EU legislation on the 2021 population and housing censuses

EXPLANATORY NOTES

2019 edition

MANUALS
AND GUIDELINES

eurostat

EU legislation on the 2021 population and housing censuses

EXPLANATORY NOTES

2019 edition

Printed by Publications Office of the European Union in Luxembourg

Manuscript completed in February 2019.

Luxembourg: Publications Office of the European Union, 2019

Theme: Population and social conditions
Collection: Manuals and guidelines

Print: ISBN 978-92-79-96919-5
ISSN 2363-197X
doi:10.2785/139261
Cat. No: KS-GQ-18-010-EN-C

PDF: ISBN 978-92-79-96918-8
ISSN 2315-0815
doi:10.2785/513433
Cat. No: KS-GQ-18-010-EN-N

Foreword

Population and housing censuses provide a precise and geographically detailed count of the population in a country. They offer information about a wide range of population characteristics at a very detailed level of cross-classification. This makes them a rich source for in-depth and flexible analysis. Censuses provide essential statistical input for formulating and evaluating policies, for administration and for research. By laying the foundation for population estimates, sample surveys and geographical data collections, population and housing censuses are the backbone of social and regional statistics.

Censuses have a long tradition in the countries of the European Union (EU). Historically, the development of census methods in each country has been shaped by numerous factors, such as: information needs; the availability of data sources and technology; data protection requirements; the burden on the respondents; and, last but not least, the financial cost of the census operation.

At the EU level, national censuses are of greater value if their results can be compared between Member States. This is why the EU has been taking continuous steps to harmonise census outputs. After the first European census programmes in 1980, 1990 and 2001, the 2011 round marked an important milestone as it was, for the first time, based on a comprehensive legal framework at the EU level. On the user side, the Census Hub was successfully launched as a central access point to the 2011 European census outputs. These innovations served the key objectives of the EU census programme, namely to disseminate more detailed data in a user-friendly way, and to make the data more comparable.

Firmly based on the experience and good practices from 2011, the 2021 round also comes at a time of fundamental transition in social statistics: administrative data sources have become the backbone of the next census in most Member States. And in most Member States, considerable progress has been achieved in recent years in the fields of methodology, institutional access to administrative data for statistical purposes, and quality assurance. These changes bring efficiency gains, leading to more powerful and more automated census systems, and are also helping Member States to address fast-evolving user expectations, where the regional dimension of EU policies is increasingly important. In this respect, the adoption of a new legal act regulating the harmonised publication of key census topics on an EU-wide 1 km^2 grid is a major innovation. It will allow for much more flexible and detailed statistical analysis at regional level, and even at cross-border level, tailored to individual research or policy needs.

The 2021 European census round will therefore be an important step. It will benefit from a well-established quality framework while piloting new methods and approaches towards a more agile, flexible and powerful system of integrated population statistics after 2021. This publication describes and explains the relevant EU legislation for the 2021 census round, with a particular focus on the new 1 km^2 grid Regulation (Chapter 4), and the updated arrangements for quality reporting (Chapter 5).

Gallo Gueye

Director

Eurostat, Directorate F

Social Statistics

Acknowledgements

Eurostat is very grateful to all those who provided expert advice in the work related to the population and housing censuses, in particular to the census experts from the statistical offices in the EU Member States and the United Nations.

On the basis of the first 2011 edition by Gregor Kyi, this publication was profoundly revised and updated for 2021 by Fabian Bach in collaboration with Maja Berber, Fabio Sartori and David Thorogood from Eurostat Unit F.2 — Population and migration. Ekkehard Petri (Unit E.4 — Regional statistics and geographical information) and Wim Kloek (Unit B.1 — Methodology; Innovation in official statistics) provided valuable inputs and comments on the new Chapter 4.

For more information:

Fabian Bach

Eurostat Unit F.2 — Population and migration

e-mail: ESTAT-Census@ec.europa.eu

Table of contents

Foreword **3**

Acknowledgements **4**

1. Introduction: EU legislation on population and housing censuses **7**

2. Technical specifications **9**

2.1. Legal and international framework 9

2.2. Technical specifications for the census topics 10

2.3. Technical specifications for the breakdowns 10

2.3.1. Breakdowns with different levels of detail 10

2.3.2. What are the breakdowns designed to disaggregate? 11

3. EU programme for the 2021 population and housing censuses **14**

3.1. Output-harmonised census data 14

3.2. Focus on EU policies 15

3.3. Structure of the 2021 hypercube programme 16

3.4. Treatment of confidential data 17

3.5. Data on the homeless 18

3.6. Metadata on specific hypercube cells: Flags 19

3.7. Metadata on definitions, sources and methodological issues 19

4. Dissemination of 2021 census data on a 1 km² grid **20**

4.1. Introduction 20

4.1.1. Background 20

4.1.2. General remarks on Regulation (EU) 2018/1799 (CIR-4) 21

4.1.3. INSPIRE issues 21

4.2. A common European 1 km² reference grid 22

4.2.1. Definition of the 1 km² reference grid 22

4.2.2. Unique 1 km² grid cell identifiers 24

4.2.3. Compiling grid data 25

4.3. The 1 km² grid dataset 26

4.3.1. Topics and breakdowns on the grid 26

4.3.2. Grid cells on which Member States have to report 27

4.3.3. The structure of the grid dataset 28

4.4. Metadata to the 1 km² grid data 29

4.4.1. Flags for the grid dataset 29

4.4.2. Topic-specific quality metadata 30

4.4.3. Quality metadata for geographic data 31

4.4.4. Further metadata provisions 32

4.5. Statistical disclosure control and output harmonisation of the 1 km² grid data at EU level 32
4.5.1. Minimising data suppression 32
4.5.2. Special features and requirements of grid data 33
4.5.3. The 'populated' flag 33
4.6. Transmission of the 1 km² grid data and metadata 35
5. Quality reporting 37
5.1. Introduction and purpose of census quality monitoring 37
5.2. Legal framework 38
5.3. Overview information 39
5.4. Data sources 40
5.4.1. Identifying the data sources for quality reporting 41
5.4.2. List of data sources 43
5.4.3. Matching of data sources to census topics 43
5.4.4. Adequacy of the data sources 43
5.4. Data sources 45
5.6. Assessment of data quality 46
5.6.1. Comparability 46
5.6.2. Timeliness and punctuality 46
5.6.3. Coherence 47
5.6.4. Coverage and accuracy 47
5.6.5. Completeness 53
5.6.6. Relevance 53
5.6.7. Geographic information — data quality 53
5.7. Access to all information relevant to the quality assessment 54
5.8. Outputs from the ESS Vision 2020 ADMIN project 54
6. Data and metadata transmission 56
6.1. Transmission of quantitative data and metadata — Census Hub 57
6.2. Transmission of textual metadata — ESS Metadata Handler 58
Annex A. Additional remarks on the topics and breakdowns in Regulation (EU) 2017/543 63
Annex B. Overview of the EU census programme under Regulation (EU) 2017/712 85
Annex C. Provisions of EU census legislation on metadata and quality reporting 90
Annex D. Transmission programme of quantitative metadata (quality hypercubes) 96
D1. How to compute QHC1: Quantitative information on the coverage of persons 98
D2. How to compute QHC2: Accuracy information on selected topics on persons 100
Annex E. Extended Euro-SDMX Metadata Structure (ESMS) for textual metadata 104
Annex F EU regulations on the 2021 population and housing census 110

1 Introduction: EU legislation on population and housing censuses

Population censuses produce the most reliable and geographically detailed count of the population. They report the benchmark figure for the 'total population' and cross-classify it in great detail for selected characteristics. This makes censuses one of the richest sources of data on the regional distribution of people living in a country and their most salient characteristics.

Censuses are expensive. In principle, the data sources that form the basis for a census should contain information on each individual statistical unit, for example on each person. Censuses entail comprehensive administrative preparation by many public agencies, such as local, regional and national authorities; public agencies that own relevant data sources; and, of course, the statistical offices. Moreover, census methods and technology are constantly changing in a sometimes arduous quest for the best data quality. These changes occur in an environment that pushes for cost savings and a 'no-heavier-than-necessary' burden on respondents. In this context, the quality of census data is becoming ever more important and must be monitored.

To produce census data, the EU Member States have developed different methods that they consider to be best suited to the administrative practices and traditions of their country. The EU legislation upholds this diversity: the framework **Regulation (EC) No 763/2008** [1] of the European Parliament and of the Council on population and housing censuses (CFR) is concerned with output harmonisation rather than input harmonisation. Member States are free to assess for themselves how to conduct their 2021 censuses and which data sources, methods and technology are best in the context of their country. This gives the Member States flexibility, in line with the principles of subsidiarity and of cost effectiveness, and in line with the competences of the statistical offices in the Member States.

However, certain conditions must be met to achieve the objective of comparability of census data from different Member States and to assess data quality. After the successful 2011 census exercise, implementing regulations for the 2011 round have been updated for the 2021 census (see Figure 1 for an overview):

1. **Regulation (EC) 2017/543** [2] (CIR-1, details in Chapter 2) contains definitions and **technical specifications for the census topics (variables) and their breakdowns** that are required to achieve Europe-wide comparability. The specifications comply with international recommendations, and have been designed to provide the best possible information value. Where appropriate, the 2021 census specifications, topics and breakdowns have been revised on the basis of the 2011 census experience. They include geographic, demographic, economic and educational characteristics of persons; international and internal migration characteristics; and household, family and housing characteristics.
2. **Regulation (EU) 2017/712** [3] (CIR-2, details in Chapter 3) lays down the reference date for the 2021 census. Moreover, it requires the data output transmitted by Member States to the Commission to comply with a **programme of statistical data** (tabulation) and with set rules on the replacement of statistical data. The content of the EU census programme serves major policy needs of the EU. The data requirements are adapted to the level of regional detail. The Regulation does not require transmission of any data that the Member States consider confidential.

   The statistical data must be supplemented by **metadata** that will facilitate interpretation of the numerical data including country-specific definitions, plus metadata on the data sources and on methodological issues. This is necessary to achieve the transparency that is a condition for valid interpretation of the data.

[1] OJ L 218, 13.8.2008, p. 14.
[2] OJ L 78, 23.3.2017, p. 13.
[3] OJ L 105, 21.4.2017, p. 1.

3. The European Commission and any user of output-harmonised census data from the EU Member States need detailed information on the quality of the censuses and their results. **Regulation (EU) 2017/881** ([4]) (CIR-3, details in Chapter 5) therefore requires transmission of a **quality report** containing a systematic description of: (i) the data sources used for census purposes in the Member States, and (ii) the quality of the census results produced from these sources. A comparably structured quality report for all EU Member States will support the exchanges of experience conducted during the 2011 round and become a reference for the development of census methodology in the future.

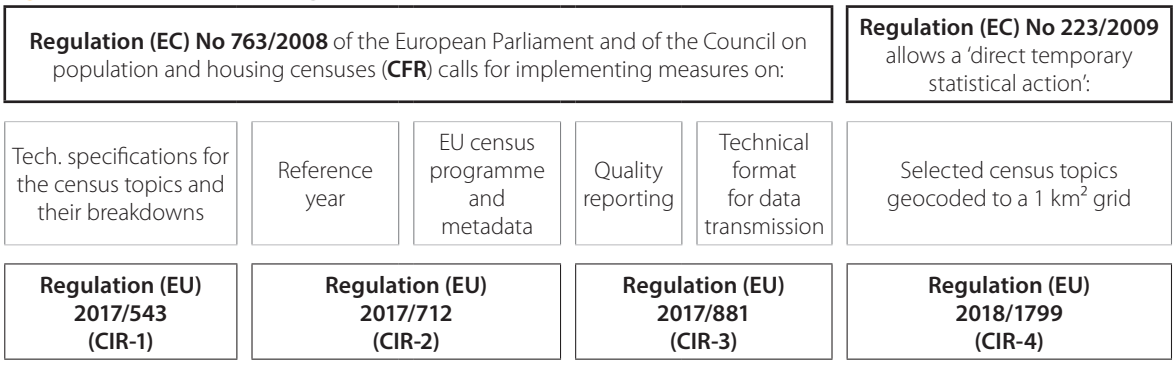
   To ensure proper transmission of the data and metadata and provide user-friendly access to this information, the **technical format to be used for transmission** must be the same for all Member States and for the Commission (Eurostat). The Regulation therefore requires the data to be transmitted in a harmonised structure and in the internationally established SDMX ([5]) format from every Member State.

In addition to these three implementing regulations based on the CFR on population and housing censuses, the Commission has adopted **Regulation (EU) 2018/1799** ([6]) on the establishment of a temporary direct statistical action for the dissemination of selected topics of the 2021 population and housing census geocoded to a 1 km^2 grid (CIR-4, details in Chapter 4). The legal basis for such a 'temporary direct statistical action' mentioned in the title is Article 14(2) of the framework **Regulation (EC) No 223/2009** ([7]) of the European Parliament and of the Council on European statistics. The Article empowers the Commission under certain provisions to take ad hoc measures to implement the European statistical programme in duly justified cases, in particular to meet unexpected needs. At the request of Member States, the harmonised dissemination of census data on a 1 km^2 grid has been identified as such an ad hoc need for the 2021 census round. CIR-4 establishes the EU-wide reference grid to be used, as well as the census topics and categories to be disseminated on that grid.

Eurostat drafted all four of the above Commission implementing regulations after intensive consultations with the national statistical offices. Between 2014 and 2017, there were 11 meetings of the Task Force on the future censuses of population and housing to advise Eurostat on technical issues ([8]). The results were discussed, amended and agreed in four Working Group meetings (September 2014, May 2015, April 2016 and June 2017). The process was accompanied by several written consultations of all the national statistical offices.

The texts of all five European Regulations that form the legal basis for the 2021 population and housing census (see Figure 1) can be found in Annex F to this publication ([9]).

Figure 1: Overview of the legal basis for the 2021 EU census



([4]) OJ L 135, 24.5.2017, p. 6.

([5]) **S**tatistical **D**ata and **M**etadata e**X**change.

([6]) OJ L 296, 22.11.2018, p. 19.

([7]) OJ L 87, 11.3.2009, p. 164.

([8]) The following countries were represented in the Task Force on the future censuses of population and housing and/or the Task Force on post-2021 census strategies: Austria, Belgium, Czechia, Denmark, France, Germany, Hungary, Ireland, Italy, Latvia, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Spain, Sweden, the United Kingdom, Turkey and Montenegro, plus the United Nations Economic Commission for Europe (UNECE) as an observer.

([9]) In this publication, regulation annexes are always referred to with their respective Roman numeral, for example 'Annex II to CIR-2', while annexes to the publication itself are arranged and referred by Roman letters, for example 'Annex A (to this publication)'.

2 Technical specifications

2.1. Legal and international framework

The annex to the CFR lists all the topics on which the EU Member States must report data. They include geographic, demographic, economic and educational characteristics of persons; international and internal migration characteristics; and household, family and housing characteristics. The annex also provides information on the level of regional detail for which data must be provided. While some topics are relevant only for NUTS level 2 aggregates, others go down to municipality level. However, the Regulation does not stipulate how the census topics are to be broken down. Nor does it specify the census topics in any further detail. Article 5(4) asks the European Commission to do this by means of an implementing regulation, which, for the 2021 round, is CIR-1.

The objective of CIR-1 is that, in every Member State, the data about the census topics should follow the same definitions and technical specifications, and the same breakdowns should be published. This is a pre-condition for Europe-wide comparability.

The specifications laid down in CIR-1 comply with international recommendations, in particular the 'Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing' (CES Recommendations). However, the Regulation is more specific in cases where the CES Recommendations have allowed countries to choose between several options. The Regulation specifies the breakdowns of the topics where the breakdowns are not given in the CES Recommendations or where more detail was required. These breakdown specifications were a precondition for the second step: defining a unified tabulation programme.

Long-standing cooperation between international and supra-national agencies, namely the United Nations Statistics Division (UNSD), the United Nations Economic Commission for Europe (UNECE) and the European Commission (Eurostat), has led to harmonisation of definitions and technical specifications stretching beyond the EU to other UNECE members or even at world level (UNSD).

The UNECE published the latest edition of the CES Recommendations (for 2020 censuses) in 2015:

https://www.unece.org/fileadmin/DAM/stats/publications/2015/ECECES41_EN.pdf.

In addition, the UNSD published the third revision of the 'Principles and Recommendations for Population and Housing Censuses' in 2015:

http://unstats.un.org/unsd/publication/seriesM/Series_M67Rev3en.pdf.

To develop international recommendations and legislation, the international and supra-national agencies cooperate closely with census experts from the national statistical offices.

2.2. Technical specifications for the census topics

To ensure that the census data received from the different EU Member States are comparable, CIR-1 sets out a common set of census topics, which follow the same technical specifications in every Member State.

The technical specifications laid down in CIR-1 provide concrete rules on which statistical units must be enumerated and to which category they must be allocated in the breakdown in question. These categories correspond largely to the specifications given in the CES Recommendations. However, the CES Recommendations contain many other valuable points of view and discussions that, by their nature, are not the subject matter for a legal text, whereas the Regulation contains several instructions not found in the CES Recommendations but which were considered necessary for practical implementation in the EU context.

Finally, note that Member States are free to add topics to those required by CIR-1. In particular, topics listed as 'non-core' in the CES Recommendations may be of interest.

2.3. Technical specifications for the breakdowns

CIR-1 lays down the breakdowns of the census topics. It specifies a number of points with a view to later use of the breakdowns in the European dissemination programme.

2.3.1. Breakdowns with different levels of detail

Many census topics can be broken down with different levels of detail. For example, the topic 'age' can be broken down into every single age (in terms of years), but five-year age groups can also be used (and even broader groups referring to particular life stages might suffice). This example demonstrates the trade-off between making more detailed information available and the size of the resulting electronic tables users can consult (called 'hypercubes', see Section 3). Excessively detailed breakdowns can lead to an exorbitant increase in the size of the tables when different census topics are cross-tabulated. This can result in very small cell values, and thus create problems with statistical confidentiality.

The level of detail of each breakdown is indicated by the suffixes listed below.

- **No suffix** means that there is only one level of breakdown detail for the topic in question. For instance, the breakdown SEX. has only one level of detail consisting of the categories 'total', 'male' and 'female'.
- The **suffix 'L'** indicates a **l**ow level of detail, for example AGE.L. consisting of broad age groups, for example 'under 15 years', '15 to 29 years', '30 to 49 years', '50 to 64 years', '65 to 84 years' and '85 years and over'.
- The **suffix 'M'** indicates a **m**edium level of detail, for example AGE.M. consisting of five-year age groups.
- The **suffix 'H'** indicates a **h**igh level of detail, for example AGE.H. consisting of a category for each single age in years.
- The **suffix 'N'** is used only in the breakdowns for 'Place of usual residence' (GEO.N.) and 'Location of place of work' (LPW.N.). It indicates that the geographical level is the whole **n**ation.
- The **suffix 'G'** is used only in the breakdowns for 'Place of usual residence' (GEO.G.) and 'Age' (AGE.G.). It indicates that the geographical level is the 1 km^2 **g**rid specified in CIR-4.

Each detailed breakdown comprises all its more aggregated siblings. For example, the breakdown AGE.H. comprises AGE.M. and AGE.L., and the breakdown AGE.M. comprises AGE.L.

The categories in the aggregated breakdowns are designed to cover the most important information. This has a practical implication: in a particular hypercube for a particular region, even if statistical information is not available for certain levels of detail (for example for reasons related to confidentiality or sampling), the aggregated information might still be available. In this example, if a user extracts data for single ages, but some cell values are blocked to protect statistical confidentiality, the user might still be able to obtain the data for the five-year age-group aggregations. The hypercubes will thus transmit the maximum amount of information that the national statistical offices are able to provide.

In an effort to reduce the complexity (and hence the response burden of the transmitting national authorities) compared to the 2011 situation, the concept of 'optional' categories was abolished for 2021. This means that all 'optional' categories in 2011 were either dropped or became mandatory in CIR-1. In general, several topic breakdowns were simplified or completely dropped, particularly those deemed less important in most Member States (see details in Annex A).

2.3.2. What are the breakdowns designed to disaggregate?

Each breakdown is designed for a specific kind of statistical unit (for example persons) and does not require a specific group of such statistical units to be applicable. In other words, all breakdowns and specifications are designed such that exactly one correct category can be identified for every possible instance of statistical units to which the breakdown applies. The implications of this are listed below.

(a) Each breakdown can disaggregate the whole total corresponding to the specified kind of statistical unit (for example total population), even if the topic in question refers only to a subgroup of that total. The advantage of this is that all breakdowns that refer to the same kind of statistical unit (for example persons) can be cross-classified. In particular, the recurring category 'not applicable' specifically fulfils the purpose of accounting for those statistical units to which the topic in question does not logically apply.
Example: The topic 'Family status' logically applies only to persons living in a family. Nevertheless, the corresponding breakdown can be applied to the total population. It is explicitly specified that persons who do not live in a family nucleus should be classified under 'not applicable'. This means that the breakdowns for the topic 'Family status' can be cross-classified with any other topic referring to persons (see more examples for this topic in Annex A).

(b) The set of available categories under each breakdown is exhaustive. More specifically, whenever there is no category 'not stated', the correct category is expected to be known (or imputed) for every single statistical unit in the population. This is because some of the topics listed in CIR-1 have been considered of such paramount importance that no missing information is allowed.
Example: The breakdowns of the topics 'sex' or 'age' (applying to persons) do not contain a category 'not stated'. This means that these breakdowns apply — and the corresponding information, in terms of available categories, should be known or inferred — for *every* person in the total population.

(c) If a breakdown can disaggregate a specified total, it can also break down each subtotal of that total. The breakdowns are complete in that they do not implicitly exclude any sub-category.
Example: Each breakdown designed for the total population can also break down any subgroup of the total population (for example the employed or the unemployed) in a meaningful way.

The topic breakdowns that can be applied, under CIR-1, to the **statistical units indicated in bold** are listed below (topics on which information must be collected or imputed for every statistical unit, in the sense of item (b) above, are marked with an asterisk*).

Persons:

- Place of usual residence (GEO.N., GEO.L., GEO.M., GEO.H., GEO.G.)*
- Location of place of work (LPW.N., LPW.L.)
- Size of the locality (LOC.)*
- Sex (SEX.)*
- Age (AGE.L., AGE.M., AGE.H., AGE.G.)*
- Legal marital status (LMS.L., LMS.H.)
- Current activity status (CAS.L., CAS.H.)
- Occupation (OCC.)
- Industry (branch of economic activity) (IND.L., IND.H.)
- Status in employment (SIE.)
- Educational attainment (EDU.)
- Country/place of birth (POB.L., POB.M., POB.H.)
- Country of citizenship (COC.L., COC.M., COC.H.)
- Ever resided abroad and year of arrival in the country (YAT., YAE.L., YAE.H.)
- Place of residence one year prior to the census (ROY.)
- Family status (FST.L., FST.M., FST.H.)
- Household status (HST.L., HST.M., HST.H.)*
- Housing arrangements (HAR.)

Family nuclei:

- Geographical area (listed under 'Place of usual residence') (GEO.N., GEO.L., GEO.M., GEO.H.)*
- Size of the locality (LOC.)*
- Type of family nucleus (TFN.L., TFN.H.)*
- Size of family nucleus (SFN.)*

Private households:

- Geographical area (listed under 'Place of usual residence') (GEO.N., GEO.L., GEO.M., GEO.H.)*
- Size of the locality (LOC.)*
- Type of private household (TPH.L., TPH.H.)*
- Size of private household (SPH.)*
- Tenure status of households (TSH.)

Living quarters:

- Geographical area (listed under 'Place of usual residence') (GEO.N., GEO.L., GEO.M., GEO.H.)*
- Size of the locality (LOC.)*
- Type of living quarter (TLQ.)

Conventional dwellings:

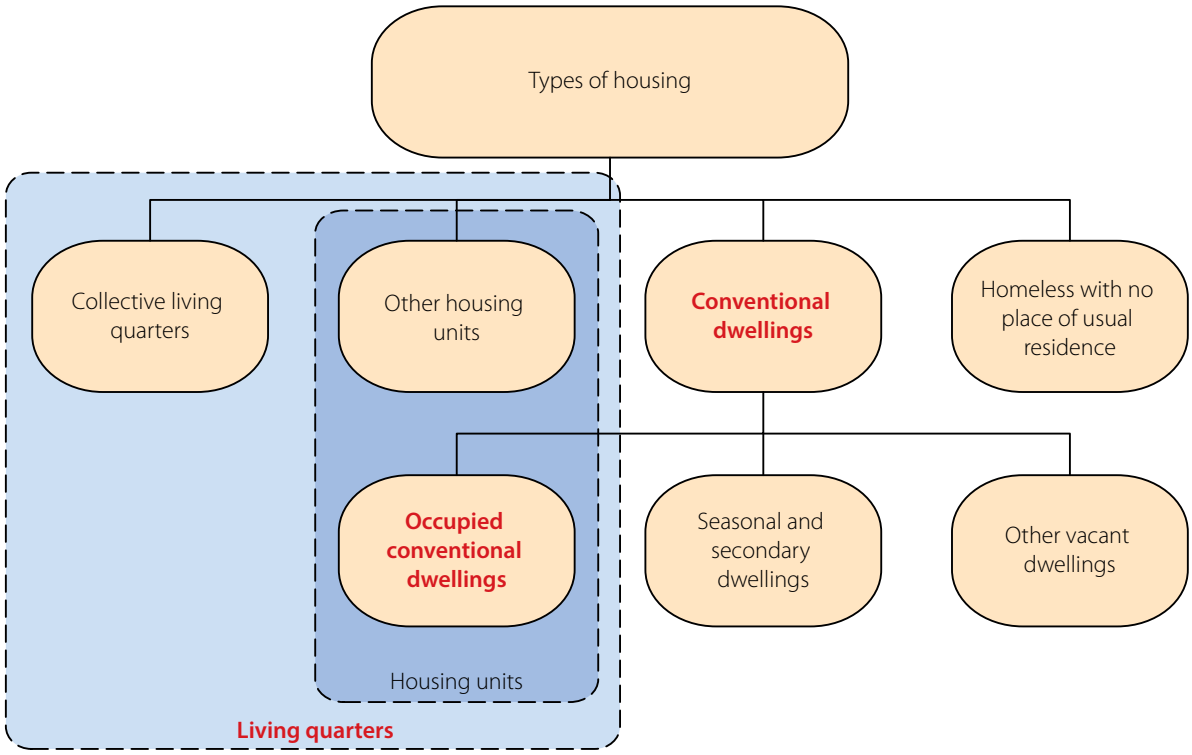
- Geographical area (listed under 'Place of usual residence') (GEO.N., GEO.L., GEO.M., GEO.H.)*
- Size of the locality (LOC.)*
- Occupancy status of conventional dwellings (OCS.)
- Type of ownership (OWS.)
- Useful floor space and/or number of rooms (UFS., NOR.)
- Water supply system (WSS.)
- Toilet facilities (TOI.)
- Bathing facilities (BAT.)
- Type of heating (TOH.)
- Dwellings by type of building (TOB.)
- Dwellings by period of construction (POC.)

Occupied conventional dwellings is a subtotal of: (i) 'living quarters' broken down by 'type of living quarter' (namely category TLQ.1.), and (ii) 'conventional dwellings' broken down by 'occupancy status' (OCS.1.), see CIR-1 and Annex A to this publication. This subtotal is used here to define a new statistical unit of its own ([10]). This means that this statistical unit inherits all breakdowns that are applicable to 'conventional dwellings' *and any subtotals* (see list above) ([11]), which are augmented by the following topics specific to this statistical unit:

- Number of occupants (NOC.)*
- Density standard (floor space or number of rooms per occupant) (DFS., DRM.)

Note also the instructive overview diagram shown in Figure 2 (based on the CES Recommendations), which illustrates the correct notion of how the statistical units 'living quarters', 'conventional dwellings' and 'occupied conventional dwellings' relate to each other.

Figure 2: Conceptual overview of housing types (taken from the CES Recommendations)



Housing types that define statistical units for the 2021 EU census are highlighted in bold red. As specified in CIR-1 and also illustrated here, any 'living quarter' must be the usual residence of at least one person.

2.3.3. Further technical specifications that apply to the breakdowns

Some breakdowns are followed by technical specifications that apply to the breakdown (instead of to the topic as such). These often concern the:

- interpretation of particular categories within a breakdown;
- order of preference for allocation of statistical units that theoretically could be classified in more than one category;
- established statistical classifications used; and
- identification and treatment of statistical units where the correct classification to one of the available breakdown categories is not obvious.

A collection of interpretation guidelines to the technical specifications, including more examples and notable changes since 2011, is given by topic in Annex A.

([10]) It is used instead of the statistical unit 'housing units' that was in use for the 2011 EU census round (see Figure 2).

([11]) Logically, it would also inherit all breakdowns applicable to 'living quarters' and any subtotals, but in fact, CIR-1 contains only a single breakdown TLQ. specific to 'living quarters', of which 'occupied conventional dwellings' is exactly one category (TLQ.1.).

3 EU programme for the 2021 population and housing censuses

3.1. Output-harmonised census data

EU legislation on population and housing censuses is 'output-oriented'. The authority and responsibility to develop appropriate census methods and technology remain with the Member States. EU legislation aims to provide census data that are comparable between the EU Member States. To achieve this, the data have to follow a European programme of statistical data and metadata.

CIR-2 sets out the EU programme of census data:

> *The legal basis for the Regulation on the programme of statistical data and metadata is Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses (the 'Census Framework Regulation' CFR). Article 5(3) authorises the Commission to adopt a programme of the statistical data and metadata to be transmitted to the Commission.*

The programme of statistical data is defined in the form of hypercubes, in other words multidimensional cross-tabulations. These contain both detailed *mesodata* and more aggregated *macrodata*. Mesodata provide an intermediate layer between microdata and macrodata ([12]). The advantage of mesodata is that they open new angles of analysis, for example by tabulating data for sufficiently small areas and/or small subgroups of the population.

The EU programme of census data is also designed in a way that aggregates the detailed mesodata step by step, both geographically and in terms of content.

- Each hypercube available for a specific geographical level (for example NUTS level 2) is also available for any more aggregated geographical level (for example NUTS level 1 and the whole nation). In this example, if, for whatever reason, certain detailed information is not available for a NUTS level 2 region, it might still be available for the corresponding NUTS level 1 region or for the country as a whole.
- Each breakdown comprises all of its more aggregated siblings that summarise the most important information (see Section 2.3.1). If, in a particular hypercube for a particular region, statistical information is not available in a certain level of detail, it might still be available for the aggregate to which the detailed category belongs.
- The two effects described above interact.

As a result, mesodata available in the EU from the 2021 census will clearly offer users greater flexibility of analysis than could be achieved with more aggregated macrodata. Experience with national censuses suggests that the pre-defined tabulation in the hypercubes will satisfy a high proportion of information requests.

For data production and statistical disclosure control, tabulation in hypercubes requires less work than processing of microdata. This is because, essentially, the information in the hypercubes has to be processed only once for numerous information requests.

([12]) See 'Terminology relating to the implementation of the vision on the production method of EU statistics', W. Radermacher, A. Baigorri, D. Delcambre, W. Kloek, H. Linden, Eurostat 2010.

3.2. Focus on EU policies

A common programme of comparable census data throughout the EU will not only support EU-wide activities but also make it easier for people acting at national (or even regional) level to see their situation as part of the larger European picture. The EU programme for the 2021 population and housing censuses has been considerably revised when compared with the 2011 case. The aim of this revision was to simplify and reduce the size of the data to be transmitted, while at the same time keeping up with major policy needs of the EU.

The programme focuses mainly on NUTS level 2 regions (total population) and NUTS level 3 regions (households and dwellings). The EU's regional cohesion policies cover large budget items (for example Structural Funds) and require detailed data to evaluate progress in the NUTS level 2 regions. Moreover, many politically relevant phenomena have a strong regional component, such as the demographic ageing of societies or economic and employment-related developments. Finally, sample surveys of the European Statistical System that also produce data for the regional level (for example the labour force survey) require a harmonised statistical framework to ensure comparability in the extrapolated survey results. The tables for the NUTS level 3 regions are further reduced and the tables for the municipalities are basic.

To tune the size of individual hypercubes, CIR-2 selects:

1. only a limited number of topics per hypercube (each hypercube focuses on one or two particular census topics and cross-tabulates these appropriately, in other words only with the other topics that have been identified as indispensable in that context – see Annex B);
2. an appropriate level of detail for the topic(s) for which CIR-1 proposes several breakdowns with different levels of detail (see Section 2.3.1).

Using these measures, CIR-2 adapts the size of the hypercubes to the level of regional detail (see also Annex B).

- The six hypercube groups at municipality level (LAU 2) are basic: the corresponding hypercubes comprise no more than 42 cells ([13]) (total population), 7 cells (households), 7 cells (families), 4 cells (living quarters) and 15 cells (conventional dwellings).
- The programme includes 13 further hypercube groups at NUTS level 3 level. Most of the contained hypercubes are reduced to well below 10 000 cells, with the exception of one 'total population' group 9, whose hypercubes have below 22 000 cells.
- All of the 20 groups for NUTS level 2 level are reduced to below 40 000 cells, with most of them well below 15 000.

In summary, the average size of the hypercubes to be transmitted for the 2021 programme was reduced very roughly by an order of magnitude compared to 2011. For example, most hypercubes at NUTS level 2 level are now smaller than 15 000 cells, while sizes of 250 000 to 300 000 were normal for the 2011 programme. This comes at the price of more individual hypercubes with fewer cross-tabulations of topics in each one, in other words reduced detail of the data. Hence, the 2021 cross-tabulations were reconsidered very carefully to achieve this simplification while keeping the resulting data fit for purpose in terms of user needs.

The significantly increased demand for geographically very detailed data is specifically addressed with the 1 km^2 grid dataset, see Section 4.

([13]) The cell numbers quoted here are based on the number of categories in the breakdowns that are part of the cross-tabulation in question, except for the GEO breakdowns whose sizes depend on the reporting country (in other words each quoted cell number must be multiplied with the number of GEO categories that exist in a country at the regional detail in question). All categories at the lowest level are counted, in other words neither totals nor subtotals are counted (because these are derived from the detail).

3.3. Structure of the 2021 hypercube programme

The programme of statistical data that the EU Member States have to transmit to the European Commission (Eurostat) is defined by the 119 hypercubes (arranged in 41 groups) that are listed in Annex I to CIR-2 ([14]), plus a single 1 km^2 grid data table defined in Annex II to CIR-4 (see Section 4). The Member States have to transmit all the data for the 119 hypercubes and the grid table ([15]).

The CFR establishes a solid foundation for supplying census data with a high resolution. The European Parliament and the Council demand that the data sources and methods that Member States use for their censuses meet, to the highest possible extent, the essential features of population and housing censuses. The Member States must make continuous efforts to enhance compliance with the essential features ([16]). These features are listed in the CFR and defined in CIR-3 ([17]) (see Section 5.2 for details).

The principle of 'individual enumeration' is of central importance for the EU programme of statistical data. It means that 'information on each statistical unit is obtained so that their characteristics can be recorded separately and cross-classified with other characteristics'. From this principle, as from the principle of 'universality' (according to which data must be provided for all statistical units within a defined territory), it follows that the programme of census data must offer much higher resolution than the results of sample surveys. The essential features of 'availability of small-area data' (data must be available for small geographical areas and small groups of statistical units) will equally lead to census data where related topics can be cross-tabulated to provide a rich basis for analysis and research.

By putting the emphasis on the essential features of population and housing censuses in the CFR, the European Parliament and the Council expect the Member States to provide detailed census data. Although the EU programme of census data laid down in Annex I to CIR-2 has been considerably revised and simplified since 2011, it remains quite comprehensive.

However, CIR-2 also upholds the principle of subsidiarity and leaves the EU Member States appropriate flexibility to assess which of their census data are reliable and fit for transmission. In doing so, the Member States can take specific problems into account that might be linked to the size of the region in question or to their reporting problems on a specific topic.

To simplify the transmission programme compared to 2011, the annex to CIR-2 now lists more individual hypercubes of smaller size each, namely 119 hypercubes sorted into 41 groups as mentioned above. On the other hand, the concept of 'principal marginal distributions' and 'secondary cells' from 2011 was dropped for 2021. This means that now all cells of all 119 hypercubes listed in CIR-2 are mandatory.

The hypercubes required for NUTS level 2 level (and above) ([18]) describe the statistical information to be transmitted for the NUTS level 2 regions, but also for the NUTS level 1 regions and the whole country. A Member State might provide more information for the more aggregated regional levels (NUTS level 1 or national) than for the NUTS level 2 regions. This should be the strategy when there is a lack of certain detailed information for some smaller (or even all) NUTS level 2 regions. At a more aggregated regional level, more observations become available, which helps to overcome problems linked to an insufficient number of observations per cell (for example error reduction, estimates and confidentiality).

Example: Hypercube group 1 'Marital status of people in households' is presented in Annex I to CIR-2 as shown in Table 1 below. Four individual hypercubes 1.1 to 1.4 are defined for group 1. Their sizes range between 249 and 4 608 cells.

([14]) CIR-2, Article 3(1).

([15]) Except statistically confidential data (see Section 3.4) and data specified in CIR-4 that are not available on the 1 km^2 grid.

([16]) CFR, Article 4(4).

([17]) CFR, Article 2(i) and CIR-3, Article 2, items (2) to (6). The essential features are 'individual enumeration', 'simultaneity', 'universality within a defined territory', 'availability of small-area data' and 'defined periodicity'.

([18]) Hypercubes containing the breakdown GEO.L.

Table 1: Example hypercube group 1 of the programme laid down in Annex I to CIR-2

No	Total	Number of cells	Breakdowns ([19])					
	Group 1		**GEO.N.**	**SEX.**	**AGE.H.**	**LMS.H.**	**HST.H.**	**FST.H.**
	Total population							
1.1		3 072	GEO.N.	SEX.	AGE.H.	LMS.H.		
1.2		3 840	GEO.N.	SEX.	AGE.H.		HST.H.	
1.3		4 608	GEO.N.	SEX.	AGE.H.			FST.H.
1.4		249	GEO.N.	SEX.		LMS.H.	HST.H.	

Comparability between 2011 and 2021 hypercubes

Although comparability — including over time — is a stated quality dimension of the European Statistical System (ESS) Code of Practice, the specific nature of (decennial) censuses makes this particularly challenging. In the 10 years between the 2011 and 2021 census rounds, changing user needs and policy priorities have led to significant differences in the underlying census methods, NUTS/LAU boundaries, concepts, definitions, classifications and disaggregations. This is an example of the well-known conflict between the quality dimensions of comparability and relevance. Moreover, as explained above, the structure of hypercubes has been substantially revised to simplify the transmission programme from 2011 to 2021. This means that most of the 2021 hypercubes only reflect sub-cubes of the 2011 programme, while others from 2011 have been dropped entirely.

Due to these technical and methodological problems — and following joint advice from the task forces on future censuses and on census IT — 2011 data will not be recalculated to meet the new geographic breakdowns or classifications that will apply for 2021 or vice versa. This implies that no concurrent (time series) extractions of combined 2011 and 2021 census data will be possible through the Census Hub (see Section 6.1).

However, better stability, and hence comparability (both geographical and over time) was one of the key incentives to establish a pan-European 1 km^2 grid dataset from the 2021 census under CIR-4 (see Chapter 4). While the detailed grid programme is expected to further develop in the future beyond 2021, this functional geographic breakdown introduces a stable reference for comparability over time, at least of the basic demographic categories.

3.4. Treatment of confidential data

Member States must suppress any numerical information if they think this is necessary to protect the statistical confidentiality of the data. CIR-2 Article 4(3) and Article 2(5) stipulate that Member States must replace a numerical cell value by the special value 'not available' if the numerical cell value must not be disclosed to protect the statistical confidentiality of the data in accordance with the Member States' statistical disclosure control. Moreover, CIR-2 Article 5(2) sets out that each affected cell must be flagged as 'confidential', see Section 3.6. Equivalent provisions exist in CIR-4 for confidential values in the 1 km^2 grid dataset, see Section 4.4.1.

All this is to make sure that no confidential information is transmitted to Eurostat under the 2021 EU census legislation, following the same standard as for the 2011 EU census round. However, it should be stated very clearly that suppression of confidential values is just one possibility for protecting such information. It is well-known that suppressing confidential values is less efficient than other methods (total information loss per confidential value), especially in high-dimensional tabulations (hypercubes) where many more non-confidential values must often be suppressed to protect confidential values from disclosure by differencing. Thus, based on the 2011 census experience, Member States were strongly encouraged to investigate more efficient protection methods for 2021, improving on the amount of disclosed information while keeping confidential information protected.

([19]) Within each table listed in the Annex to CIR-2 for a specific hypercube group, the uppermost row (in bold) lists all the breakdowns used in at least one of the hypercubes of that group, as defined in the annex to CIR-1. Each further line below (not in bold) specifies a specific hypercube, as defined in CIR-2 Article 2(2).

There is clearly a large variety of methods to protect confidential information in the census outputs. Many of these methods have already been applied in production by one or more Member States in the past. In 2016 and 2017, Eurostat hosted a dedicated project on 'Harmonised protection of census data in the ESS' to assess the situation in the Member States, and to develop recommendations for a more harmonised approach in 2021 ([20]). As a result, the project recommends two protection methods, summarised below.

(1). **Record swapping**: households containing at least one high-risk person are identified and swapped with similar, low-risk households (the geographic location of the paired households is exchanged, in other words it is a *pre-tabular* method acting directly on the microdata before aggregating output tables).
(2). **Cell key method**: random noise (±1, ±2, and so on) is added to some cell values in the output tables in a deterministic and unbiased way following a predefined probability distribution (*post-tabular* method acting on the output tables).

These two methods can be applied either separately or in combination. Both methods are variants of the 'random noise' paradigm, explicitly avoiding cell suppression or rounding. Various studies have shown that such methods can significantly reduce the overall information loss in the outputs at a comparably low disclosure risk level ([21]). In conclusion, these methods are specifically aimed at:

- overcoming the typical drawbacks of traditional methods like suppression (mentioned above) or simple rounding, and thus reducing the information loss through the treatment;
- flexibility (various parameters are available that can be adapted to particular national requirements in a Member State);
- straightforward usage, by providing implementations in SAS (original project results) ([22]) and in cooperative ESS tools (μ-/τ-Argus and sdcMicro/-Tables incl. R) ([23]) out of the box.

Member States are referred to the extensive documentation of the project results for more information (footnote 20), and strongly encouraged to consider these recommendations for the 2021 EU census round.

3.5. Data on the homeless

The topic 'housing arrangements' (HAR.) defined in CIR-1 includes information on:

- primary homeless persons (persons living in the streets without shelter);
- secondary homeless persons (persons moving frequently between temporary accommodation); and
- persons living in a non-conventional shelter (for example huts, cabins, shacks, shanties, caravans, houseboats or caves).

However, note that, as opposed to the 2011 situation, these attributes are now aggregated under a single breakdown category HAR.2. (see Annex A). CIR-2 asks Member States to provide in several hypercubes information on this category broken down by sex, age and LAU2 regional level (GEO.H.).

Member States have repeatedly pointed to the challenges involved in enumerating the homeless. However, the homeless are part of the total population, according to the relevant international recommendations and the definitions and specifications in EU legislation. Eurostat must be able to estimate the total population of each Member State on the basis of the census results. The new aggregate categories are intended to simplify reporting, while ensuring that the total count and basic characteristics of this aggregate group are included in the census population, as explicitly required in Annex II to CIR-2.

As the European Commission attaches great importance to the availability of statistics about the homeless, Eurostat will closely monitor the availability of such data in the quality evaluation of the 2021 censuses. In particular, Annex II to CIR-2 extends the table programme on the homeless (hypercubes including HAR. and/or HST.M.) by the following specific metadata provisions on the homeless:

([20]) https://ec.europa.eu/eurostat/cros/content/harmonised-protection-census-data_en.

([21]) See for example the proceedings of the 2016 International Conference on Privacy in Statistical Databases: https://link.springer.com/book/10.1007/978-3-319-45381-1.

([22]) https://ec.europa.eu/eurostat/cros/content/testing-recommendations-codes-and-instructions_en.

([23]) https://github.com/sdcTools.

The metadata shall report the number of all homeless persons. The numbers of primary homeless persons (persons living in the streets without shelter) and of secondary homeless persons (persons moving frequently between temporary accommodation) shall be shown where this distinction is possible.

A description of the methodology and data sources used to produce the data on homeless persons shall be provided.

As is the case with any other metadata requirement on topics laid down in Annex II to CIR-2, this information should be added to the topical metadata on 'household status' and 'housing arrangements'.

3.6. Metadata on specific hypercube cells: Flags

According to the definition in CIR-2 Article 2(9), Member States will have the possibility to 'flag' data items to provide transparency about specific characteristics of the data value, where CIR-2 Article 5(1) lists all available flag types for the 2021 hypercube programme ([24]).

(a) Numerical data values that have been blocked to protect statistical confidentiality (see Section 3.4) are to be marked *'confidential'* (see CIR-2 Article 5(2)).
(b) Data values that the Member States consider unreliable, according to their own statistical quality control, will be marked *'unreliable'* (see CIR-2 Article 5(3)). As is usual practice when disseminating European statistics, users should treat data flagged as unreliable with caution. Moreover, if a Member State feels strongly that some data transmitted to the Commission should not be disseminated to the public, it can ask that they should not be published. In this case, the data made available for publication would exclude the unreliable data in question, but Eurostat is entitled to request them for quality assessment only.
(c) Data values that have been *'revised after first data transmission'* must be marked as such.
(d) Member States are free to provide textual metadata for any particular hypercube cell, which must then be marked with *'see information attached'*.

CIR-2 Article 5(4) stipulates that an explanatory text must be provided for each data item that is accompanied by one or more of the flags (b) to (d).

3.7. Metadata on definitions, sources and methodological issues

Member States must report to Eurostat: the definitions of the topics used in their censuses; the source of these definitions; the method used to estimate data on the topics; or the reason for any systematic unreliability of the data.

In some cases, CIR-1 allows Member States to make choices about the specifications of the topics or certain breakdown categories. Moreover, there might be other country-specific features of the topics. Member States must report these choices and any country-specific feature of the topics. Provisions on metadata are in the following places of the 2021 EU census legislation:

- the content of the metadata that Member States have to transmit to the Commission (Eurostat) is laid down in the annex to CIR-3;
- in addition, specific topic-related metadata are required in Annex II to CIR-2;
- finally, specific metadata related to the 1 km^2 grid data are set out in Annex II to CIR-4.

Detailed explanations (including the compilation of these and all other metadata) and the unified transmission via the 'ESS Metadata Handler', are covered in Chapters 5 and 6. See also Annex C for overview tables of all legal provisions on metadata in the legislation, including where and how to report them.

([24]) The available flags for the 1 km^2 grid dataset are listed in Annex II to CIR-4; see Section 4.4.1.

4 Dissemination of 2021 census data on a 1 km² grid

4.1. Introduction

4.1.1. Background

The dissemination of population data at the 1 km² grid level is an area where there is significant user demand and which has undergone rapid development in many national statistical offices. In many situations, statistics at the level of towns and neighbourhoods will be the most relevant and most meaningful to people. These statistics will often be needed to support the local-level policy decisions that are of importance to people's day-to-day lives. For this reason, improving the availability of geocoded social statistics has been noted as a priority at the level of the whole ESS. This serves key census data users at ESS level, in particular the Directorate-General for Regional and Urban Policy (DG REGIO), which also re-stated the importance of comparable data with high spatial resolution in developing, implementing and evaluating policy — for example, to ensure efficient distribution of over EUR 350 billion from the Cohesion Policy fund.

This increased importance of developing geocoded statistics was widely noted in the years before the 2021 census. For example, Section 3.3 of the ESS Vision 2020 cited the potential value of merging geospatial data with official statistics to provide better social and environmental information. In its opinion on the draft European statistical programme for 2013-2017, the European Statistical Advisory Committee also noted the increasing importance of geospatial data combined with social and economic statistical information to support evidence-based policy making. Furthermore, the 2015 report of the fifth session of the Committee of Experts on UN Global Geospatial Information Management (UN-GIMM) highlighted the 2020/2021 round of censuses as an important opportunity for the integration of statistical and geospatial data.

The 1 km² grid collection from the 2021 EU census thus became a pilot project, aiming to exploit synergies between modern geographical information systems promoting this type of data integration and the future EU census strategy ([25]). Improving both spatial and temporal comparability is one of the core goals of the post-2021 census vision, where the successful introduction of a common European reference grid — constant in space and time — is now a significant step forward (see Section 4.2).

For implementing the new 1 km² grid data collection for 2021, the ESSnet project GEOSTAT 1 can be seen as a prototype. Based on the 2011 EU census data, the project successfully compiled and disseminated total population data geocoded to a 1 km² grid ([26]). After this promising experience, a large part of the preparation work for the 2021 EU census focused on the dissemination of a larger variety of census data geocoded to a unified European 1 km² reference grid. This represents a new development for the EU census programme, adding an entirely new product to the main programme of data dissemination from the 2021 census (Chapter 3): selected 2021 census topics geocoded to a common European 1 km² grid.

([25]) See also the paper PG2016/62/3.6 from the 62nd meeting of the Partnership Group.

([26]) http://www.efgs.info/geostat/1b/.

4.1.2. General remarks on Regulation (EU) 2018/1799 (CIR-4)

Dissemination of 1 km² grid data is not covered by the CFR. Eurostat therefore worked with the national statistical offices to prepare a dedicated act regulating this important new piece of European statistics. During this work, a *'temporary direct statistical action'* (TSA) under Article 14 of Regulation (EC) No 223/2009 on European statistics was identified as the most effective approach by the ESS members. As a result, CIR-4 was adopted on 21 November 2018 as the first legal act ever to implement European statistics by way of a TSA. CIR-4, Article 1 sets the scope:

> *A temporary direct statistical action is hereby established in order to develop, produce and disseminate selected topics of the 2021 population and housing census geocoded to a 1 km² grid [. . .].*
>
> *To this end, a unique harmonised and constant geospatial reference grid for Europe is determined consisting of cells with an area of 1 km². And the specific topics and their breakdowns are established as well as the detailed programme and metadata for the dissemination of 2021 population and housing census data geocoded to the 1 km² reference grid.*

In detail, CIR-4 establishes:

Article 2 — common definitions augmenting those in the CFR, to be applied in the context of CIR-4;
Article 3 — a common European 1 km² reference grid plus grid cell codes (see Section 4.2);
Article 4 — 2021 census topics and breakdowns for the 1 km² grid dataset (see Section 4.3.1);
Article 5 — the structure and content of the 1 km² grid dataset (see Section 4.3.3);
Article 6 — special provisions on output harmonisation (see Section 4.5);
Article 7 — the metadata belonging to the 1 km² grid dataset (see Section 4.4);
Article 8 — the reference date, aligned with CIR-2;
Articles 9 and 10 — transmission modalities (dates and technicalities), in line with the CFR and CIR-3;
Article 11 — standard quality provisions, in line with the legal practice after Regulation (EC) No 223/2009;
Article 12 — dissemination and INSPIRE issues (see Section 4.1.3).

The spirit of CIR-4 is to align as closely as possible with the other 2021 census legislation, namely the CFR plus CIR-1 to CIR-3 (cf. Figure 1). For instance, the topics and breakdowns established in Article 4 refer to CIR-1 whenever possible, and only set out necessary extensions and/or adaptations to the specifications of CIR-1 for the grid data. Similarly, the textual metadata required in Article 7 comply with the structure established in CIR-3, and only supplement specific fields with additional information needed on the grid data (see also Annex C and in particular Table C1).

One of the few notable — and unavoidable — divergences between CIR-4 and the other 2021 census legislation concerns the common definitions set out in CIR-4, Article 2 for the purpose of CIR-4, with respect to those of CIR-2, Article 2 (for the purpose of CIR-2 and CIR-3). While CIR-2 uses the term *'cell'* in a statistical sense (in other words cells in a data table), this is avoided in CIR-4 to prevent confusions with the geographical notion of a *'grid cell'*. Indeed, CIR-4, Article 2(1) links the word *'cell'* explicitly and exclusively to this geographical notion for the purpose of CIR-4. As a consequence, some definitions already present in CIR-2 had to be 'relabelled' in CIR-4 even though they refer to the same concepts: for example *'cell value'* (CIR-2) becoming *'data value'* (CIR-4), or *'numerical cell value'* (CIR-2) becoming *'numerical value'* (CIR-4).

The obvious advantage is that **each instance of the word *'cell'* in CIR-4 unambiguously refers to the geographical notion of a *'grid cell'***. We will also strictly follow this principle throughout Chapter 4 of this publication, whereas the word 'cell' may have been used more liberally in other chapters.

4.1.3. INSPIRE issues

An important aspect of CIR-4 are the legal obligations emerging from INSPIRE legislation ([27]) (fully in force from 2021), which establishes an extensive common infrastructure for the provision and exchange of spatial information in the EU. All EU census products fall under topic 10 *'Population distribution — demography'*, defined as:

> *geographical distribution of people, including population characteristics and activity levels, aggregated by grid, region, administrative unit or other analytical unit.*

([27]) 'INSPIRE' Directive 2007/2/EC (OJ L 108, 25.4.2007, p. 1) and its implementing rules.

To ensure the interoperability and harmonisation of spatial information across the EU, INSPIRE entails substantial technical requirements for public accessibility to the data. Broadly summarised, four requirements must be met for all census data:

(1) sharing of data using INSPIRE-accepted data models;
(2) the existence of a download service for the data;
(3) the existence of a view service displaying the data as a map;
(4) the provision of INSPIRE-compliant metadata on data and the services.

Member States are legally required from 2021 onwards to disseminate all statistical data on population distribution and demography — including census outputs at NUTS, LAU and grid level — in an INSPIRE-compliant manner. However, applying this requirement to the 2021 census will also be of great added value because this data collection will thus become a first real-case European pilot for the integration of statistical and geospatial information.

Because of this, **Eurostat committed to support the Member States in the fulfilment of these obligations on the 1 km² grid data**. Apart from efficiency, there will be a significant advantage for users, because all dissemination web services will be harmonised and centralised for the entire 1 km² grid data collection from all reporting countries. This makes particular sense for grid data, which are inherently suited for trans-national and cross-border analyses and applications. A single access point to harmonised census grids across the whole ESS will make these data significantly more usable.

From a technical point of view, a central dissemination of the 1 km² grid data by Eurostat in compliance with all INSPIRE requirements, thus fulfilling all respective legal obligations of the Member States, was considered the most efficient option. Eurostat makes the following legal commitment in CIR-4, Article 12(1):

> *The Commission (Eurostat) shall disseminate the 1 km² grid datasets referred to in Article 5 as well as the associated metadata referred to in Article 7.*

This provision fixes Eurostat's obligation to disseminate all national grid datasets and metadata. In addition, CIR-4, Article 12(2) reads:

> *For the purposes of this Regulation, the programme of the 1 km² grid data and metadata to be transmitted by Member States and disseminated by Eurostat corresponds to the data that Member States disseminate at national level in accordance with Directive 2007/2/EC and its implementing Regulations (EC) No 1205/2008, (EC) No 976/2009 and (EU) No 1089/2010.*

The provision identifies the data disseminated by Eurostat with the obligatory national dissemination in compliance with INSPIRE ('…*corresponds to* …'). This means that **all legal obligations on Member States emerging from INSPIRE with respect to the 1 km² grid data are fulfilled automatically by transmitting the data and metadata required in CIR-4 to Eurostat** (see Chapter 6).

4.2. A common European 1 km² reference grid

4.2.1. Definition of the 1 km² reference grid

To be able to combine geospatial information from several sources, it is essential that this information is based on the same cartographic representation of the geographic area covered. Or equivalently, the map coordinates of every piece of geocoded information must refer to the same 'reference frame'. Similarly, for the 1 km² grid data it is essential that all Member States agree on a unique map representation, or reference frame, of the entire area on which grid data must be reported: namely the European parts of the territories of all Member States ([28]).

([28]) Overseas territories are excluded, see end of this section.

The INSPIRE framework (Section 4.1.3) has also been created to answer exactly these kinds of questions. Indeed, INSPIRE stipulates such a unique reference frame in the form of a particular geographic grid to be used for any grid data of *'pan-European usage'* (and falling under INSPIRE). More specifically, Section 1.5, Item (2) of Annex IV to the INSPIRE implementing Regulation (EU) No 1089/2010 ([29]) states that for grids used as statistical units:

> *for pan-European usage, the Equal Area Grid defined in Section 2.2.1 of Annex II shall be used.*

Section 2.2.1 of Annex II, in turn, reads:

> *The grid is based on the ETRS89 Lambert Azimuthal Equal Area (ETRS89-LAEA) coordinate reference system with the centre of the projection at the point 52°N, 10°E and false easting: x0 = 4 321 000 m, false northing: y0 = 3 210 000 m.*
>
> *The origin of the grid coincides with the false origin of the ETRS89-LAEA coordinate reference system (x=0, y=0). [. . .]*
>
> *The grid is hierarchical, with resolutions of 1 m, 10 m, 100 m, 1 000 m, 10 000 m and 100 000 m. [. . .]*
>
> *The grid is designated as Grid_ETRS89-LAEA. For identification of an individual resolution level the cell size in metres is appended.*

These provisions for a 'European reference grid' definition merit some further explanations as follows.

- The starting point for the definition is the 'European Terrestrial Reference System 1989' (*ETRS89*), which fixes a unique three-dimensional reference frame for the entire Earth.
- The centre of the desired two-dimensional grid reference frame is located at 52° northern latitude and 10° eastern longitude (referring to the angular surface coordinates of the 'Geodetic Reference System 1980 ellipsoid' ETRS89-GRS80, to be precise).
- *Lambert Azimuthal Equal Area* (LAEA) is a particular projection scheme which maps every point of the Earth ellipsoid surface GRS80 ([30]) uniquely onto a point in the tangent plane touching the ellipsoid at the centre of the projection (coinciding with the later grid centre, see Figure 3). The constituting, and desired, property of this LAEA projection is that regions on the original ellipsoid surface are mapped to regions of different shape but of **equal area** in the tangent plane.
- The resulting projected plane is the desired flat reference frame. It can now be mapped with Cartesian coordinates *(x, y)* ([31]) measured in metres, where the *false easting x0* and *false northing y0* move the coordinate origin *(x=0, y=0)* away from the centre of the frame towards the southwest (into the lower-left corner of the frame), so that the centre itself is now represented by the coordinates *(x=4 321 000, y=3 210 000)* in the resulting Cartesian coordinate system.
- Extending east and north from the coordinate origin *(x=0, y=0)*, the flat reference frame can then be 'dressed' with equidistant straight grid lines, parallel to the coordinate axes and with spaces corresponding to the chosen grid resolution, in other words 1 000 m in our case (for illustration, see Figure 4 showing the reference grid with 100 000 m resolution).
- The correct name for the resulting 1 km² reference grid is *'Grid_ETRS89-LAEA1000'*.

CIR-4, Article 3(1) firstly affirms that this common European 1 km² reference grid must be used:

> *In accordance with Section 1.5 of Annex IV of Regulation (EU) No 1089/2010, the statistical 1 km² reference grid for pan-European usage shall be the Equal Area Grid 'Grid_ETRS89-LAEA1000'. [. . .]*

Then CIR-4, Article 3(1) limits the geographic extent of this reference grid:

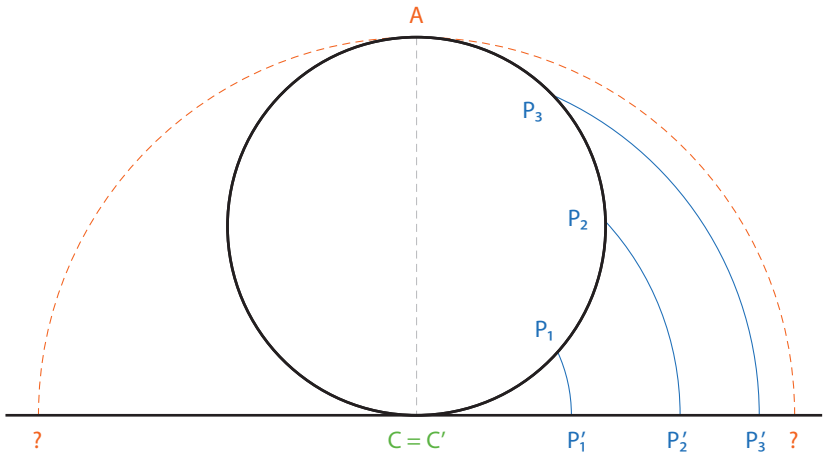
> *[. . .] The spatial extent of the reference grid in the coordinate system specified for this grid in Section 2.2.1 of Annex II of the same Regulation shall be limited to easting values between 900 000 and 7 400 000 metres and northing values between 900 000 and 5 500 000 metres for the purpose of this Regulation.*

([29]) OJ L 323, 8.12.2010, p. 11 and related amendments.

([30]) Precisely, every point except the antipodal point of the projection centre on the ellipsoid.

([31]) The *y* direction points exactly to the north from the projection centre (in other words to points >52°N and exactly 10°E in ellipsoid surface coordinates), and the *x* direction is perpendicular to it.

Figure 3: Illustration of the LAEA projection scheme (see text)



The projection centre ***C*** is mapped onto itself, whereas its antipode ***A*** is the only point on the globe (circle in the picture) which cannot be mapped in a unique way. The image also shows how the projection distorts lengths increasingly with the distance of the original point from the projection centre (the distances $C \to P_1 \to P_2 \to P_3$ are all equal on the circle, whereas the distances $C' \to P_1' \to P_2' \to P_3'$ on the projected plane are getting smaller).

The sentence defines a bounding box, which encloses the applicable geographical region of the grid for the purpose of CIR-4. Technically, this is achieved by restricting the range of applicable easting and northing values, given in flat reference frame coordinates explained above. As illustrated in Figure 4, this bounding box includes all relevant parts of continental Europe ([32]) plus adjacent islands, such as the Azores, the Canaries, the British Isles, and Iceland (among others). It is made explicit in Annex I that CIR-4 does not entail any reporting obligation on territories that are outside this bounding box, see Section 4.3.2 for details.

4.2.2. Unique 1 km² grid cell identifiers

To categorise all records in the 'census microdata database' on persons (cf. Section 5.4.1) under the geographic breakdown GEO.G. for the 1 km² grid, each grid cell needs a unique identifier, or grid cell code. CIR-4 Article 3(2-3) specifies this:

2. *In accordance with Section 1.4.1.1 of Annex IV of the same Regulation, each individual grid cell of the 1 km² reference grid shall be identified by a unique grid cell code, which is composed of the characters 'CRS3035RES1000mN'. This is followed by the northing value in metres of the grid point in the lower-left corner of the grid cell, followed by the character 'E', followed by the easting value in metres of the grid point in the lower-left corner of the grid cell.*
3. *The country code of the transmitting Member State as defined in the Interinstitutional Style Guide published by the Publications Office of the European Union, followed by the character '_', shall be prepended to the cell code of each grid cell transmitted by that Member State.*

Paragraph 2 constructs a unique grid cell identifier in accordance with INSPIRE provisions from the reference frame coordinates *(x, y)* (see Section 4.2.1) of the lower-left corner of the grid cell. The individual parts of the code string convey the following information:

- *'CRS3035'* is a unique identifier of the reference frame used for the grid (ETRS89-LAEA);
- *'RES1000 m'* gives the resolution of the grid (1 000 m or 1 km);
- *'N[y]'* encodes the so-called *northing value*, where *[y]* stands for a string containing all digits of the integer *y* coordinate in metres of the lower-left grid cell corner;
- *'E[x]'* encodes the so-called *easting value*, where *[x]* stands for a string containing all digits of the integer *x* coordinate in metres of the lower-left grid cell corner.

Paragraph 3 supplements this INSPIRE-compliant cell code with a prefix indicating the reporting country. This prefix is needed for border grid cells, on which every Member State with some territory inside the given border cell must

([32]) The continental territories of all ESS Member States, plus all candidate countries, potential candidate countries, and countries with an EU Association Agreement and a recognised prospect for membership on the date of this publication.

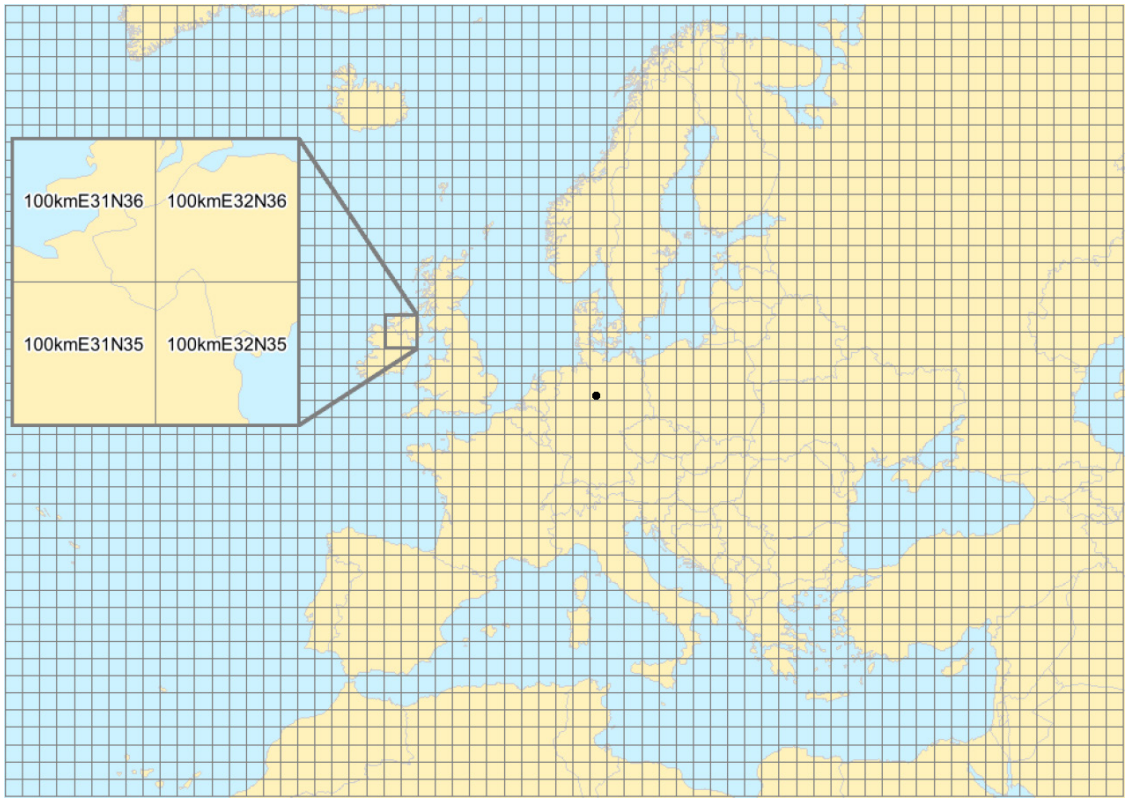
report (see Section 4.3.2). As laid down in Article 3(3), the country codes to be used must be the ones listed in the Interinstitutional Style Guide published by the Publications Office of the European Union ([33]).

To give an example, note the correctly formatted grid cell identifier, or grid cell code, for the 1 km² grid cell whose lower-left corner coincides with the projection centre of the reference frame *(x=4 321 000, y=3 210 000)* (located in Germany, cf. Figure 4):

`DE_CRS3035RES1000mN3210000E4321000`

This is in line with the correct INSPIRE format for grid cells used as statistical units (Section 1.4 of Annex IV to the INSPIRE implementing Regulation (EU) No 1089/2010, which applies for the purpose of CIR-4. Compared to the more compact format used, for example, in the 2011 GEOSTAT population grids (where it would be '1kmN3210E4321' for the example above), this new format has a quite practical advantage: it contains the coordinate values in metres, so it is immediately compatible with possible future grids at resolutions below 1 000 m.

Figure 4: Europe in the projected flat reference frame ETRS89-LAEA



The frame section on display corresponds exactly to the bounding box specified in CIR-4, Article 3(1), extending from *(x=900 000, y=900 000)* in the lower left corner to *(x=7 400 000, y=5 500 000)* in the upper right. The black dot marks the projection centre: 52°N, 10°E in ellipsoid coordinates coinciding with *(x=4 321 000, y=3 210 000)* in plane reference frame coordinates. The overlaid grid lines correspond to a resolution of 100 000 m, where four cells are magnified for illustration. In the 1 km² reference grid, each of the grid cells shown here is further divided into 100 x 100 = 10 000 cells of 1 km x 1 km size.

4.2.3. Compiling grid data

In the typical situation for the 2021 census, the 'census microdata database' on persons (cf. Section 5.4.1) will be geocoded to single points, in other words each microdata record is linked through a unique record identifier to a point address or coordinate set in a given (national) geographic reference system. Usually such geographic information is managed through some 'Geographic Information System' (GIS) software, which can straightforwardly perform the following steps:

1. transform given (national) coordinates into the required EU grid reference frame;
2. identify the corresponding 1 km² grid cell in which the transformed coordinates lie;
3. allocate the respective grid cell code (Section 4.2.2) to the microdata record in question.

([33]) http://publications.europa.eu/code/en/en-8000100.htm.

As a result, the microdata database is now categorised under a new geographic breakdown GEO.G. (details in Section 4.3.2), which can be treated like any other breakdown of a census topic for further processing, especially for the aggregation of person counts under a given combination of breakdown categories ([34]).

However, in some Member States certain grid topics (see Section 4.3.1) may not be available at single point resolution. As a consequence, the corresponding topic categorisations in the 'census microdata database' on persons will not (always) be based on individual observations for each record, but rather assigned to the records on the basis of some estimation procedure, for example disaggregation of larger area information. CIR-4 allows for such estimations but requires a detailed explanation of the applied method(s) in Annex III, under 'Metadata on the topics' (item 2). See Section 4.4.2 for an explanation of these provisions, in particular quality parameter [1], which covers this.

4.3. The 1 km² grid dataset

4.3.1. Topics and breakdowns on the grid

At an early stage of the 1 km² grid initiative, the Census Task Force identified a set of selected topics for the grid dataset and the preferred breakdown categories, considering the need for data, their potential availability and quality. It was then agreed that the following data items should be collected as a part of the 2021 EU census data collection for each 1 km² grid square:

- total population;
- sex (males, females);
- age (under 15, 15 to 64, 65 and over);
- employed persons, as far as possible ([35]);
- place of birth (in the reporting country, in another EU country, outside EU);
- usual residence 12 months before (unchanged, within reporting country, outside of the reporting country).

In effect, 12 to 13 numbers will be collected for each grid cell on which a Member State must report, representing the number of persons in each of the categories. In particular, there will be **no cross-tabulation of the different topics** (cf. Section 4.3.3).

All selected topics above are already defined in the annex to CIR-1 for the generic 2021 EU census exercise. Of the selected breakdown categories, some are defined in CIR-1, and some are ad hoc just for the grid. Hence, the approach followed in Annex I to CIR-4, where the grid topics and breakdowns are specified, is to align as closely as possible with existing provisions in CIR-1. This is initially set out in Annex I to CIR-4:

> *The technical specifications of the breakdowns for the purpose of this Regulation of the census topics specified in the Annex of [CIR-1] shall be presented as follows:*
>
> – *Each topic selected for dissemination on the 1 km² reference grid is quoted with its heading from the Annex of [CIR-1].*
> – *The technical specifications laid down in the Annex of [CIR-1] referring to that topic in general shall apply.*
> – *Thereafter, the breakdown for that topic is specified.*
> – *All breakdowns are designed to break down any total or subtotal referring to persons.*

This means that CIR-4 is restricted to persons as statistical units, and that topic specifications in CIR-1 are not touched. The sole purpose of Annex I to CIR-4 is either to select breakdown categories of these topics that already exist in CIR-1, or otherwise to create new categories for the grid, if needed. Whenever a new breakdown definition was needed, this is indicated by the letter **G.** appended to the topic acronym, which is the case only for **GEO.G.** and **AGE.G.** The specification of GEO.G. (in other words the breakdown of the geographical area for the place of usual residence by grid cells) is at the heart of CIR-4 and hence merits a dedicated explanatory section (Section 4.3.2). The other 'statistical' breakdown categories of CIR-4, including a mapping of their origin to CIR-1 or CIR-4, can be concisely summarised in Table 2.

([34]) For a detailed description of this 'grid aggregation method', including also some GIS software examples, see the relevant GEOSTAT documentation: http://www.efgs.info/wp-content/uploads/geostat/1b/GEOSTAT1B-Appendix1-production-procedures-bottom-up.pdf.

([35]) It was later agreed that information on 'employed persons' is to be collected only '*as far as possible*', which means that Member Stats are obliged to justify non-provision of such data (for example non-availability). See more detailed explanations in Section 4.3.3.

The code list **STAT.G.** in the rightmost column of this table (and defined in Annex II to CIR-4, see Section 4.3.3) simply pools all **stat**istical categories to be collected on the **g**rid for later convenience. Obviously, it is not a 'breakdown' in the sense of CIR-1, because it does not 'break down' a given population into mutually exclusive subsets according to some characteristics. Correspondingly, its categories STAT.G.1 to STAT.G.12 do not add up to the 'total population' (in other words STAT.G.0 is not the margin of STAT.G.), unlike all 'real' breakdowns specified in CIR-1. Rather, the 13 members of STAT.G. directly reflect the agreed set of 13 statistical observations initially listed in this section for collection on the 1 km² grid.

Table 2: Overview of the census topics and breakdowns selected for the 1 km² grid

Topic	Breakdown categories		Description	STAT.G.
	CIR-1	CIR-4		
GEO.		GEO.G.x.	(See Section 4.3.2)	
		GEO.G.y.		
SEX.	SEX.0.		Total population	0.
	SEX.1.		Male	1.
	SEX.2.		Female	2.
AGE.		AGE.G.1.	Under 15 years: equal to AGE.L.1. in CIR-1	3.
		AGE.G.2.	15 to 64 years: sum of AGE.L.2.-4. in CIR-1	4.
		AGE.G.3.	65 years and over: sum of AGE.L.5.-6. in CIR-1	5.
CAS.	CAS.L.1.1.		Employed persons (see details in Section 4.3.3)	6.
POB.	POB.L.1.		Place of birth in reporting country	7.
	POB.L.2.1.		Place of birth in other EU Member State	8.
	POB.L.2.2.		Place of birth elsewhere	9.
ROY.	ROY.1.		Usual residence unchanged	10.
	ROY.2.1.		Move within the reporting country	11.
	ROY.2.2.		Move from outside the reporting country	12.

4.3.2. Grid cells on which Member States have to report

The content categories of the geographic grid breakdown GEO.G. are defined in Annex I to CIR-4 under 'Topic: Place of usual residence'. The first provision states:

> *the breakdown categories of this topic on which a Member State shall report are all cells of the 1 km² reference grid specified in Article 3(1) whose area includes a part of the territory of that Member State.*

This has an unambiguous meaning in terms of the content of GEO.G. The geographic bounding box set out in CIR-4, Article 3(1) and explained in Section 4.2.1 limits the applicability of CIR-4, as explicitly stated in Article 3(1). Furthermore, it is chosen such that its boundaries coincide with 1 km² grid lines, so every single 1 km² grid cell is either completely inside or completely outside the bounding box. Thus, the bounding box defines a unique and finite list of 1 km² grid cells, namely all grid cells that are inside the bounding box (in other words inside the geographic region shown in Figure 4). The phrase *'all cells of the 1 km² reference grid specified in Article 3(1)'* above should be understood as exactly this exhaustive list of 1 km² grid cells to which CIR-4 applies.

To establish the applicable categories of GEO.G., each Member State has to select out of this exhaustive list every grid cell *'whose area includes a part of the territory of that Member State'*. The complete set of all resulting cell codes for that Member State is referred to as **GEO.G.x.** in Annex I to CIR-4, and in Table 2 of this publication. In addition, Annex I to CIR-4 provides that GEO.G.x. is:

> *complemented by a single virtual grid cell per Member State to account for persons that are not allocated.*

This adds a single category **GEO.G.y.** to the breakdown GEO.G. for each Member State, where 'y.' is merely a placeholder for the correct category code for that Member State. The correct code for each Member State, as defined in footnote 1 of the Annex, is:

```
XX_unallocated
```

where 'XX' is the country code of the reporting Member State, as laid down in CIR-4, Article 3(3) (see Section 4.2.2). The purpose of this virtual category is to report on persons that could not be geo-localised to any geographical grid cell contained in GEO.G.x.

Finally, Annex I to CIR-4 explicitly allows for appropriate estimation methods to allocate persons to GEO.G.x.:

> *If the place of usual residence of a person is unknown within the territory of the reporting Member State that is covered by the reference grid, additional scientifically-based, well-documented, and publicly available statistical estimation methods may be used to allocate this person to a specific grid cell. Persons who are not allocated to any cell of the reference grid shall be allocated to the virtual grid cell GEO.G.y. of that Member State.*

This means that Member States have a certain flexibility in treating the non-geocoded records in their 'census microdata database' on persons: they may apply estimation methods to distribute all or some of them across GEO.G.x., and/or report (remaining) non-geocoded persons under GEO.G.y. However, Annex III to CIR-4 requires under 'Metadata on the topics' (item 3) a description of how the category GEO.G.y. was used, in particular if and which estimation methods were applied to allocate non-geocoded persons, and which rules were in force to categorise persons under GEO.G.y. This is covered under quality parameter [2] of the reporting template presented in Section 4.4.2.

4.3.3. The structure of the grid dataset

Given the list of statistical categories **STAT.G.** to be collected on the grid (Section 4.3.1) and the breakdown of the geographical area into grid cells **GEO.G.** (Section 4.3.2), the final structure of the required 1 km² grid dataset is remarkably simple: it is simply the flat, two-dimensional **cross-tabulation of GEO.G. against STAT.G.**, as stipulated in Annex II to CIR-4:

> *The programme of the 1 km² grid data to be transmitted […] shall consist of one two-dimensional table that cross-tabulates the set of grid cells GEO.G. defined in Annex I against the following selection of categories from the census topic breakdowns specified in Annex I: [STAT.G.]*

This means in particular that **no topics contained in STAT.G. are cross-tabulated against each other**.

'Total population' by grid cell

One dedicated remark is in order on the 'total population' observation by grid cell, defined in CIR-4 Article 2(3):

> *'total population' means all persons of a grid cell whose usual residence is located in that grid cell.*

The breakdown SEX., defined in CIR-1 and reused for the grid, is exhaustive in the sense that it contains no category 'not stated' (cf. Section 2.3.2), and every person in the 'census microdata database' (cf. Section 5.4.1) must be allocated to one of the categories 'male' or 'female'. Therefore, its margin SEX.0. (identified with STAT.G.0. in Annex II to CIR-4), broken down by grid cell, must always equal the 'total population' of a grid cell in the sense of the definition above. As a matter of fact, this is just a special case of the general setup of the 2021 EU census legislation, where the definition in CIR-2, Article 2(1) states that:

> *'total population' of a well-defined geographical area means all persons whose usual residence, as defined in Article 2(d) of [the CFR], is located in that geographical area.*

This definition applies to all hypercubes on persons listed in Annex I to CIR-2 (where the 'Total' given in the second column is the 'Total population'). The given definition is thus imposed inside each separate GEO. category (representing a *'well-defined geographical area'*) on the margin of all breakdowns (in other words '0.' categories) of all other topics on persons in CIR-1.

'Employed persons' by grid cell

The special provision in Annex II to CIR-4 on the number of 'employed persons' (STAT.G.6., see Table 2) by grid cell merits a dedicated explanation. Footnote 2 applies to this observation, and sets out that:

> *[d]ata on the category 'employed persons' shall be transmitted as far as possible, subject to availability in the transmitting Member State.*

This exempts Member States from their legal obligation to transmit this observation if it is not available to them, or cannot be readily obtained, at 1 km² grid level. However, the formulation *'as far as possible'* ensures that, by default, available data must be transmitted, while the non-provision needs to be explicitly justified in the metadata.

4.4. Metadata to the 1 km² grid data

4.4.1. Flags for the grid dataset

Just like any 'ordinary' hypercube data item, any data item contained in the 1 km² grid dataset can also be accompanied by one or more flags. However, the set of available flag categories was adapted to the needs of the grid data in CIR-4, with respect to CIR-2 Article 5 (cf. Section 3.6). The corresponding list is laid down in Annex III to CIR-4 (in paragraph 1 under 'Metadata on data items') and contains the three flag categories set out below.

(a) *'Provisional'*: as further specified in paragraph 2 – this is a grid-specific flag, which is only applicable to 'total population' data items reported under CIR-4, Article 9(1). See Section 4.6 for a detailed explanation how to use it;
(b) *'Populated'*: as further specified in paragraph 3 – this is another grid-specific flag only applicable to 'total population' data items on the 1 km² grid. Its intention and usage is set out in CIR-4, Article 6(2) and further explained in Section 4.5.3;
(c) *'Revised'*: similar to its usage in other hypercubes – this flag is to mark all grid data items that are revised after their first transmission to Eurostat.
There is one particularity for 'total population' data items in combination with the 'provisional' flag (a) above. In general, all 'provisional' flags are to be completely removed before the 2024 transmission deadline under CIR-4, Article 9(2) (see Section 4.6). However, 'total population' data items that were originally flagged 'provisional' in 2022 under CIR-4, Article 9(1) should not simply be re-flagged as 'revised' in the transmission of all grid data under CIR-4, Article 9(1). Rather, only those 'total population' data items that were not flagged 'provisional' in 2022, but nonetheless had to be revised by 2024, should be flagged as 'revised' in the transmission. The following matrix illustrates all applicable flagging histories for a given 'total population' data item.

		'provisional' in 2022:	
		yes	no
'revised' in 2024:	yes		☑
	no	☑	☑

This approach follows a simple logic:
- A data 'revision' in the sense of the CFR and CIR-2 is a data update or correction that was not planned for at first transmission to Eurostat.
- Data considered 'provisional' entail some planned follow-up procedure to further improve their quality for a later re-submission. This well-planned procedure should not be considered a 'revision' in the above sense.

As a result, the approach leads to a more restrictive use of the 'revised' flag. This makes sense, because the idea of this flag is to track exceptional and unforeseen actions that may become necessary after first transmission, due to shortcomings that are only found later. Hence, the occurrence of this flag should not be artificially proliferated through the introduction of another flag specifically designed to indicate provisional data. In any case, paragraph 4 requires an explanatory text for each data item transmitted with the flag 'revised'.

(d) *'See information attached'*: equivalent to its usage in other hypercubes, this flag is to mark data items accompanied by some additional information in the form of free text. For consistency, paragraph 4 requires that such an explanatory text is indeed provided for each occurrence of this flag.
(e) *'Confidential'*: as further specified in paragraph 5, this flag is to mark all data items whose confidential value has been replaced by the special value 'not available' to avoid transmission of confidential information to Eurostat in accordance

with CIR-4, Article 5(2). See Section 3.4 for the general issue of confidential information in the EU census outputs, and Section 4.5.1 for particular requirements and resulting usage constraints of this flag in the 1 km² grid data.

4.4.2. Topic-specific quality metadata

Annex III to CIR-4 requires, under 'Metadata on the topics', topic-specific quality information, focusing particularly on: (i) the *reliability and accuracy* of reported grid data, (ii) any methodology used to aggregate or disaggregate grid information on the topics (person characteristics). The following table is an abridged list of quality-assessment parameters that were already in use for the 2011 GEOSTAT population grids ([36]). They have been adapted to comply with the metadata requirements on topics in Annex III to CIR-4:

[1]	Production methods	*Aggregated from point sources, from smaller grid cells, other small areas, hybrid approaches, or disaggregation methods if any* *The aggregation method may have subgroups: aggregation directly from register data, aggregation made on estimated individual data based on registers, and aggregation from very small enumeration areas. Disaggregation methods may have subgroups depending on the regional level of the source data and the ancillary data used.* **Indicator: Aggregation, disaggregation, mixed mode (total and in %)**
[2]	Accuracy of the figures	*Differences between grid data totals and totals from traditional census statistics, and the reasons for deviation between official figures and grid figures (for example due to homeless people, non-geocoded addresses, non-geocoded persons or topic information).* *In the cases of data produced by extrapolation of the sample, the measure of accuracy that results from the estimation process.* *In particular, the topic GEO. should contain information on any methodology used to allocate persons to specific grid cells, or otherwise to report them under GEO.G.y (see* Section 4.3.2*). Other topics should state if and how respective information was obtained for persons allocated to GEO.G.y.* **Indicator: Non-geocoded observations (total and in %)**
[3]	Comparability of grid data within a country	*Regional differences in quality of the data due to different production methods and source data (fully comparable as unique production method OR partially comparable – if partially comparable, then indicate the differences and the share in the total grid data for each cause of difference).* **Indicator: % of consistent and comparable data**
[4]	Temporal coherence	*Did different reference dates in any of the data sources (spatial or statistical) lead to differences in the grid data (in other words would the grid data be different if all data had the same reference date)? This might be the case if: (i) the population used for aggregation had a different reference date than the population used for disaggregation, (ii) there were differences between different regions in a country, or (iii) the address points have a different reference date than the statistical data.* **Indicator: % of population with the same reference date**

Note that the parameters [1-4] in this form are not part of CIR-4, and thus not legally binding. They are rather intended as a proposal for efficient metadata reporting to comply with Annex III to CIR-4. Ultimately there will be a single field of free text, for each topic contained in Annex I and listed in Table 2 (cf. Table C3), to address the requirements under 'Metadata on the topics' in Annex III. This text field can be structured according to points [1-4] for convenience and comparability. The suggested quantitative **indicators** above are intended to supplement the textual information.

([36]) http://www.efgs.info/wp-content/uploads/geostat/1b/GEOSTAT1B-Appendix17-GEOSTAT-grid-POP-1K-ALL-2011-QA.pdf.

4.4.3. Quality metadata for geographic data

Annex III to CIR-4 requires, under 'Reference metadata' (in its item 4), quality information on the geo-referencing of persons, covering:

> *geographic quality principles, in particular territorial coverage and comparability, positional accuracy, as well as temporal coherence and completeness of the geographic data used for geocoding.*

The quality assessment parameters in the following table cover these geographic quality principles. Again, it is an abridged list of parameters already in use for the 2011 GEOSTAT population grids ([37]), adapted to comply with the geographic metadata requirements in Annex III to CIR-4.

[1]	Geographic data used for geocoding/spatial modelling	*Brief description of the geographic features to which persons have been assigned (geocoded or disaggregated) for the census, for example geo-referenced address points, building points, building polygons, enumeration polygons.* *The source of the data (for example cadastre, municipal address registers, national mapping authorities, spatial interpolation …).*
[2]	Positional accuracy	*Positional accuracy of the data in [1], to be indicated in metres or by providing a scale number for lines and polygons.* *For example, it is important to know if the geographic data are provided from official data sources with geodetic accuracy or if they are interpolated between known points (for example between addresses at crossroads).*
[3]	Comparability	*Description of the comparability of [1] in different parts of the country (for example were different datasets of [1] used to geocode the population?).*
[4]	Completeness/coverage of geographic data	*Share of [1] that are geo-referenced and used for geo-referencing population.* *Share of population covered by [1].* *Percentage of persons covered with geo-referenced data. Share of non-geocoded addresses. Regional differences if applicable.*
[5]	Temporal coherence	*Did revisions of the geocoding infrastructure occur during the reference period of the statistical data, and if so, what were the reasons for the revisions? Last updates of the geographic data or the territorial units used for disaggregation. If applicable, extent of the updated data (%).*
[6]	Coherence with other grid systems	*If data are published on different grid systems (for example EU and national grids), independent re-aggregation of the microdata to each grid system is the preferred method. Otherwise, if one grid was produced by recasting the other grid, specify the original grid.* *If the microdata were aggregated independently into different grid systems, were the outputs checked to be coherent (for example consistent totals, subtotals, population distributions)?* *Measures implemented against disclosure by geographic differencing between different grid systems should be reported.*

Like in Section 4.4.2, the parameters [1-6] in this form are also not part of CIR-4, and thus not legally binding. They are also intended more as a proposal for efficient metadata reporting in compliance with Annex III to CIR-4. There will be a single field of free text (cf. Table C3) to address the requirements under 'Reference metadata' in Annex III (item 4 'Geographic information — data quality'), which can be structured following points [1-6] for convenience and comparability.

([37]) See footnote 36.

Referencing INSPIRE metadata

As already outlined in Section 4.2.1, typically the geo-referenced data are in a stand-alone geo-database, which is also maintained separately from the census microdata database, and merely linked to it through a unique record identifier. Usually such geo-databases are multi-purpose, and accompanied by their own INSPIRE metadata covering the quality of geographic information. In this case, the following three points apply.

- Member States may provide a unique database identifier (name or ID), plus reference (URI or similar) to INSPIRE metadata, for each geo-database used for their census grid to cover points [1], [2] and [5] above.
- In addition, points [3] and [4] would then only be needed if more than one geo-database is used: this is to give a complete picture of the resulting *comparability* and *completeness* across the entire census population, as required in CIR-4.
- Point [6] would only be needed if more than one output population grid is produced, for example one for national users and one for the EU in different projection systems.

In conclusion, **all requirements on geographic quality in Annex III to CIR-4 may be exhaustively covered by sole provision of a single geo-database identifier plus INSPIRE metadata reference** if the following three conditions are met for the reporting country:

(1) only one geo-database is used to produce the census grid data;
(2) INSPIRE metadata are available on that geo-database for referencing; and
(3) grid data are produced only on one output grid system (in other words the EU grid).

4.4.4. Further metadata provisions

Points 1 to 3 under 'Reference metadata' in Annex III to CIR-4 essentially add grid-related information to some items of the generic quality report structure laid down in the annex to CIR-3 and discussed in detail in Chapter 5. For convenience, these provisions are addressed in their respective places in the quality report where they augment generic census metadata (see Sections 5.5 and 5.6.2, and particularly Table C2, which contains the complete quality report provisions merged from the annex to CIR-3 and Annex III to CIR-4).

4.5. Statistical disclosure control and output harmonisation of the 1 km² grid data at EU level

4.5.1. Minimising data suppression

In terms of the generic 2021 census exercise, Section 3.4 already outlined recent developments in the statistical disclosure control (SDC) of confidential information, aimed at reducing the information loss while at the same time offering sufficient protection against disclosure of confidential information. For the benefit of output harmonisation and improved usage at EU level, CIR-4, Article 6(1) contains a respective provision for the grid:

> *In order to facilitate the Union-wide comparability, the output data values to be disseminated shall be harmonised. To that end, preference shall be given to numerical values over special values as far as possible.*

While this provision does not require any particular SDC measure (this remains up to the Member States), it sets out a general preference for measures allowing for the transmission of numerical values over suppression ([38]). The legal formulation *'as far as possible'* is used here, as in Annex II to CIR-4, with the same legal implication as already explained in Section 4.3.3. This means that the provision defines a default (in other words no suppression), but does not enforce it if a justification is provided laying down the particular reasons why an exemption is needed. Note that the SDC expert recommendations introduced in Section 3.4 deliberately refrain from data suppression, and are thus trivially compliant with CIR-4, Article 6(1).

([38]) The term 'suppression' is used here as a concise synonym for the procedure fixed in CIR-2 (cf. Section 3.4) and CIR-4 (cf. Section 4.4.1) of replacing confidential values by the special value 'not available', and flagging the respective data items as 'confidential'.

4.5.2. Special features and requirements of grid data

In view of the generic situation described in Section 3.4, grid data are very different from 'ordinary' hypercubes in terms of structure, scope and intended use, thus posing very specific challenges for SDC. The following three bullet points list some of the greatest challenges.

- Grids are not nested with other regional breakdowns based on administrative boundaries (NUTS, LAU), which may increase the disclosure risk of confidential information by differencing between partially overlapping small regions.
- Grid data have a special scope and purpose. Contrary to hypercubes based on administrative regions, they are not just another table of statistical data — their essential and distinguishing feature is the crossing of statistical census information with rigid and standardised geographical information. In that sense, grids may be seen as a hybrid product between traditional tables and maps. It follows that the geographic dimension of a grid plays an outstanding role: it is not just 'another dimension of a statistical table'.
- Following on from the previous point, the 'information value' of individual statistical data contained in a grid dataset (and hence the notion of 'information loss') may need to be re-evaluated, taking into account its particular purpose. Several key use cases of EU grid data rely on very accurate information about the geographic distribution of the total population. This assigns an outstanding value to the distinction of populated versus unpopulated areas (in other words grid cells), which implies a *qualitative* difference between the 'total population' by grid cell (or rather, if this number is '0', or not) and the other person characteristics collected on the grid ([39]).

Efficiently protecting confidential information along the lines of Section 3.4, while at the same time complying with the peculiarities of grid data outlined here, entails an added complexity.

For instance, during the 2011 GEOSTAT population grid exercise, the overall majority of participating Member States did not even consider 'total population' sensitive information on the 1 km² grid. Most of the remaining Member States applied simple rounding variants on small counts, but keeping the '0' counts apart to maintain the distinction between unpopulated and populated cells ([40]). However, the 2021 situation is more complicated, due to the additional person characteristics collected by grid cell. This additional information may cause more Member States to protect small 'total population' counts. Moreover, simple rounding of small 'total population' counts would likely also lead to a rounding (or suppression) of all other person characteristics.

All this was duly acknowledged and addressed by the members of the SDC project mentioned in Section 3.4. Although the recommended application of random noise methods provides an effective protection against the risk of disclosure by differencing (first point above), some challenges were noted with respect to the other two points. In particular, a certain tension was identified between the cell key method in its more common parameter setups and the requirement of accurately separating unpopulated from populated grid cells. This is because the cell key method cannot distinguish between 'total population' and other data, and there is a considerable share of small counts — typically in the range of '3' or less — being perturbed to '0' (thus turning populated grid cells into unpopulated ones) ([41]).

4.5.3. The 'populated' flag

Several approaches were discussed to alleviate the tension mentioned at the end of the previous section. Ultimately, the SDC project decided to recommend the following procedure to maintain the distinct advantages of the cell key method (cf. Section 3.4) while simultaneously meeting the requirement to accurately distinguish populated from unpopulated grid cells: introduce a new flag which marks 'total population' data items as 'populated' if and only if their observed value *before* the application of any SDC method is greater than '0'. The desired information — accurate population distribution — can thus be extracted from the metadata to the 1 km² grid dataset (in other words distribution of the 'populated' flag across the grid cells, see Figure 5 for an illustration). This does not affect the actual 'total population' data, so the cell key method can be applied to the grid dataset in a straightforward way.

([39]) This particular notion was reaffirmed by various large-scale grid users, most notably DG REGIO, already mentioned in Section 4.1.1.

([40]) Cf. the 'Confidentiality' field of the 2011 GEOSTAT quality assessment (footnote 36).

([41]) In some setups (not recommended by the SDC project, however) there are also additional perturbations in the other direction, in other words turning unpopulated grid cells into populated ones.

Table 3: Examples of possible 'total population' data value and 'populated' flag combinations for various observed values and SDC scenarios

Observed 'total population' value before any SDC treatment, according to CIR-4 Art. 2(7)	Transmitted 'populated' flag (irrespective of SDC method)	Transmitted 'total population' value in merged output table for various typical SDC scenarios			
		Cell key method (0s untreated)	Cell key method (0s treated too)	No values < 3	Rounding to base 5
0	☐	0	1	0	0
1	☑	3	0	3	0
2	☑	0	3	3	0
3	☑	3	2	3	5
4	☑	5	4	4	5
5	☑	5	6	5	5

Suppression (in other words transmission of the special value 'not available' together with the flag 'confidential') is not possible by CIR-4, Article 6(2)(a) for any of the examples in the table.

CIR-4, Article 6(2) is based on these SDC project conclusions. It ensures that the flagging procedure is applied in a harmonised way, thus ensuring that the required information is provided in a comparable manner across the whole EU. To that end, CIR-4, Article 2(7) first defines:

> *'observed value' means a numerical value that represents observed or imputed information to the best of knowledge based on all available 2021 census information, in particular before the application of any statistical disclosure control measures.*

The crucial point of this definition is that **an *'observed value'* must be determined *before* the application of any SDC measures**, for instance before processing the grid table with the cell key method. Then CIR-4, Article 6(2) itself reads as follows:

> *In order to ensure sufficiently accurate and reliable information on the spatial distribution of the total population, Member States shall respect the following requirements:*
>
> *(a) data items on total population shall not be reported as confidential;*
> *(b) data items on total population with an observed value other than '0' shall be marked with the flag 'populated'; and*
> *(c) data items on total population with an observed value '0' shall not be marked with the flag 'populated'.*

The paragraph provides an introductory justification for its existence (*'accurate and reliable information on the spatial distribution of the total population'*), before forbidding the suppression of 'total population' items in its provision (a), and finally regulating the correct usage of the 'populated' flag through provisions (b) and (c). The essential point is that **the 'populated' flag is linked uniquely to the above definition of an *'observed value'***. This establishes a legally binding, accurate distinction between populated and unpopulated grid cells, irrespective of the eventual choice and application of SDC measures, which may or may not affect 'total population' counts by grid cell. To make this even more explicit, the correct order of data processing steps, starting from the *unprotected* 'microdata database' on persons, may be outlined as follows:

1. aggregate the (unprotected) 1 km² grid table from the microdata and allocate 'populated' flags to 'total population' data items according to CIR-4, Article 6(2)(b-c) ('flag table');
2. possibly apply pre-tabular SDC methods on the microdata, for example record swapping (cf. Section 3.4);
3. aggregate the 1 km² grid table again and possibly apply post-tabular SDC methods, for example the cell key method ('output table');
4. merge the 'populated' flag information from the flag table (step 1) into the output table.

Table 3 illustrates possible data and flag values in the merged output table for various typical SDC scenarios ([42]) (see also Figure 5).

([42]) The cell key method can have different setups, some of which allow observed 0s to be changed into non-0s (for example 1, 2 — see documentation under footnote 20): This is illustrated in Table 3 by the two columns on the cell key method ('0s untreated/treated').

In conclusion, CIR-4, Article 6 in combination with the voluntary application of the recommended cell key method would have distinct **advantages from a user perspective.**

- The complete statistical data (total population plus characteristics) are made available for each grid cell, with a fixed absolute statistical variance on every data value, throughout the entire 1 km² grid data of the Member State. This means that each and every published number is affected by a design SDC uncertainty of $\pm\sqrt{V}$, where V is the variance parameter of the cell key method (typically around 0.5 – 2) ([43]). While some Member States may prefer not to publish the exact value of *V*, some reasonable indication of the noise uncertainty on the published data values should be given in the metadata.
- The accurate population distribution can be readily obtained from the distribution of the 'populated' flag across the grid cells. If the value of a 'total population' data item flagged as 'populated' equals '0', some simple heuristics, taking into account available data on characteristics in this grid cell and/or adjacent grid cells, can be applied to arrive at a fair estimate of the total population. For instance, the average of the approximate margins of the topics SEX., AGE.G., POB. and ROY. on that grid cell may convey a decent first approximation of the total population:

$$STAT.G.0 \approx \frac{1}{4}\left(\sum_{i=1}^{5} STAT.G.i + \sum_{i=7}^{12} STAT.G.i\right)$$

  cf. Table 2.

One notable drawback might be that a given grid cell with 'total population' equal to '0' but flagged as 'populated' (or vice versa, in some possible setups — see footnote 42) may be perceived as confusing or unintuitive by the user. However, it was judged that the advantages above greatly outweigh this. Moreover, it may be reasonably assumed that any serious user of the grid data would (and should!) inquire about their scope and usage. On the other hand, depending on the dissemination format, cursory data users will likely not even be aware of the 'populated' flag in the first place.

4.6. Transmission of the 1 km² grid data and metadata

CIR-4, Article 9 stipulates that the complete and validated 1 km² grid dataset plus all metadata for each reporting country must be transmitted to Eurostat **before 31 March 2024** (paragraph 2). In addition, paragraph 1 stipulates that data on the '**total population**' by grid cell (in other words a two-column table tabulating STAT.G.0 against GEO.G., plus corresponding metadata on this observation) must be transmitted **by the end of 2022**.

However, these total population data need not be the final and definite values. This means that, while the transmission of these data is obligatory, Member States are free to assess the quality of the total population data at the 2022 transmission date as compared to the expected quality at the eventual date of complete data transmission (by March 2024). If a Member State expects significant developments in the further production pr ocess between the two transmission dates, which will likely affect the quality of total population data, it may flag respective data items as '**provisional**' in 2022 (see also Section 4.4.1). While the judgement on whether and how to use this flag remains with the Member States, paragraph 2 under 'Metadata on data items' in Annex II to CIR-4 states:

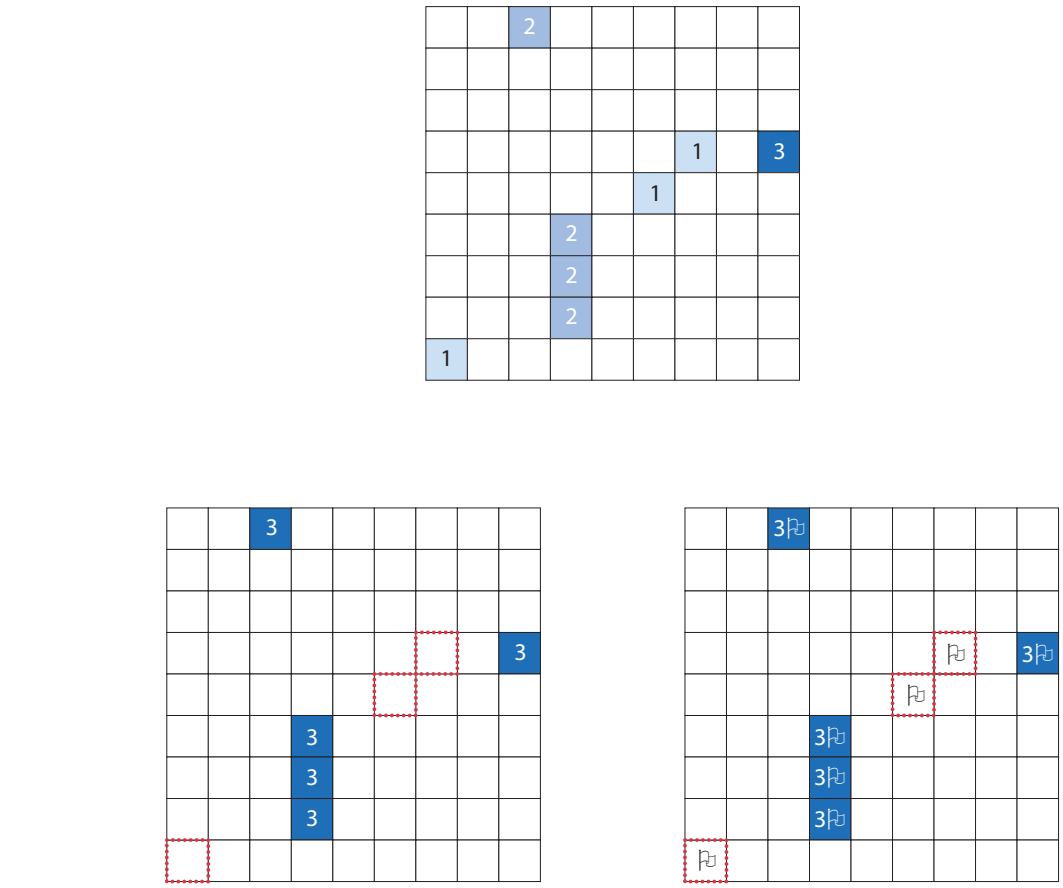
> *Only data values on 'total population' which are reported under Article 9(1) and which are not considered final data by the Member State at the time of reporting shall be accompanied by the flag 'provisional'.*

Thus, the question to be answered by each Member State at the end of 2022 is: **Do I expect notable effects on (some of) the 'total population' grid data from my further census production process until March 2024?**

Finally, CIR-4, Article 10 sets out that the same transmission tools must be used for the 1 km² data as for any other data under the 2021 EU census legislation (cf. CIR-3, Article 5). See Chapter 6 of this publication for a detailed description of these tools provided by Eurostat.

([43]) Note that record swapping (cf. Section 3.4) may further increase the SDC uncertainty on counts of person *characteristics*. Of course, all this refers only to the *SDC uncertainty*: naturally there are additional uncertainties (variance and/or bias) stemming from data collection and processing.

Figure 5: Effect of the 'populated' flag (see text) on a scarcely populated region



The upper picture shows the observed, unperturbed 'total population' data (white cells are unpopulated). The lower left picture shows a possible outcome of the cell key method (without going into detail on the method parameters), with populated areas turned into unpopulated areas marked by a red frame. The lower right picture shows how the 'populated' flag marks all originally populated areas, while 'total population' would still be reported as '0' in the red-framed cells.

(Example taken from real 2011 GEOSTAT data from Austria in a square of 9 x 9 grid cells bounded by *(x=4 693 000, y=2 744 000)* in the lower left corner and *(x=4 702 000, y=2 753 000)* in the upper right. The reported numbers may be considered as unperturbed because Austria did not report any SDC measures on total population at the time.)

5 Quality reporting

5.1. Introduction and purpose of census quality monitoring

The purpose of this chapter is to provide a comprehensive set of operational guidelines addressed to census experts in national statistical offices. This will assist them in compiling the metadata and quality reporting provided for by the EU legislation on the 2021 population and housing censuses. According to the EU legislation, Member States must transmit census metadata within 27 months from the end of the reference period, in other words by the end of March 2024 ([44]).

The EU census legislation contains several detailed requirements on these metadata and on the quality reporting. The goal of this chapter is to guide national statistical offices towards the correct interpretation of these requirements and, in particular, to clarify some important terminological issues, including how to identify the 'data sources' that must be used to report on data quality.

In addition, this chapter illustrates the channels and tools that will be set up by Eurostat to collect, process and disseminate all the metadata information associated with the 2021 census.

The metadata and the information on data quality associated with the 2021 census are of two kinds, as set out below.

(a) Quantitative metadata (which is also referred to as 'quality hypercubes' or 'data on quality'): a numeric assessment of different dimensions related to data quality, such as accuracy, completeness, coherence, and so on. **The amount of required quantitative metadata has been drastically reduced for 2021 compared to 2011 to simplify the reporting (for example, there are only 2 quality hypercubes in 2021 instead of 38 in 2011)**. Section 5.6 covers this in more detail.
(b) Qualitative, or textual, metadata: a structured set of textual information to explain the methodology used to produce census data (including those cases where different data sources have been combined to obtain the desired statistical output) plus any other generic information useful to complement the actual data.

Both kinds of metadata will be covered in detail in this chapter.

It is useful to recall the underlying philosophy that inspired the design and implementation of the quality reporting system contained in the EU census legislation. Broadly speaking, the reference scenario can be summarised in the following three bullet points.

- Member States are free to choose the data sources and census methodology that best fit their respective national scenarios.
- Nevertheless, the statistical output must meet some well-defined quality targets (in particular the 'essential features of population and housing censuses' as defined by the CFR, see Section 5.2).

([44]) Note special provisions for 1 km^2 grid data explained in Section 4.6.

- Moreover, the data produced must meet the level of detail provided for by the EU legislation, regarding both the definition of census topics and their breakdowns (see Section 2), as well as the structure of census datasets (in other words the hypercubes, see Section 3).

Within this context, there are essentially two goals of census quality reporting:

1. monitoring data production processes that can be very different from country to country, sometimes involving multisource operations (such as record linkage from different administrative sources), or data adjustments (such as record/item imputations, record weighting, and so on);
2. providing tools to evaluate the quality of the statistical output in a comparable manner, regardless of the differences in national production processes.

The rest of this chapter is structured as follows: Section 5.2 provides a schematic summary of the provisions on data quality monitoring contained in the 2021 EU census legislation. The subsequent sections, from Sections 5.3 to 5.6, directly reflect the 2021 quality report structure laid down in the annex to CIR-3.

Chapter 6 summarises the tools that will be set up by Eurostat for the collection and dissemination of all required information, including all textual and quantitative metadata. Finally, the complete metadata requirements of the 2021 census legislation are listed in Annex C, including where and how to report them using the tools described in Chapter 6.

5.2. Legal framework

Population and housing censuses are supposed to provide data of the highest quality. After the two priorities of providing a high level of regional detail and a high capacity to cross-tabulate data, the quality of the census data is an important goal of the censuses and justifies the high levels of investment. The European legislation on censuses requires Member States to report on the quality of the census results that they transmit to the European Commission (Eurostat). The Commission must then assess the quality of the data.

The recitals of the CFR refer to quality-related issues several times:

- recital 3 says that Member States should 'foster best practices';
- recital 5 says that censuses should conform to the principles of 'impartiality, transparency, reliability, relevance, cost effectiveness and confidentiality';
- recital 7 makes explicit reference to the European Statistics Code of Practice.

Quality assessment is covered by Article 6 of the CFR.

CFR, Article 6(1) introduces the applicable quality dimensions of 'relevance', 'accuracy', 'timeliness', 'punctuality', 'accessibility', 'clarity', 'comparability' and 'coherence'.

CFR, Article 6(2) requires Member States to report on the quality of the data transmitted. It refers explicitly to the five 'essential features of population and housing censuses', as listed in CFR, Article 2(i) and elaborated on in CIR-3, Articles 2(2-6):

(1) **individual enumeration**: the characteristics of each statistical unit are recorded separately, so that each characteristic can be cross-classified with others;
(2) **simultaneity**: all information refers to the same point in time (reference date);
(3) **universality within a defined territory**: data are provided for all statistical units in a defined territory (for persons in particular, data are provided for all usual residents in a defined territory);
(4) **availability of small-area data**: data are available for small geographical areas and for small subgroups of statistical units;
(5) **defined periodicity**: Member States must be able to run censuses every decade, for example by ensuring the continuity of registers.

These definitions correspond to those provided in the CES Recommendation (§§23-28).

CFR, Article 6(3) requires the Commission to define the detailed rules for — and structure of — the quality reports by means of an implementing regulation. It also requires the Commission (Eurostat) to assess the quality of the data

transmitted. This in turn implies that the quality reports from the national statistical offices must enable Eurostat to perform this assessment. To this end, the outline (headings) of the 2021 census quality reporting is laid out in the **annex to CIR-3** ([45]). The provisions on the different dimensions of data quality (see Article 6(1) above) are as follows.

(1) **Adequacy of data sources** (item 2.4) as an aspect of **relevance**: Member States must report on how their data sources meet the 'essential features of population and housing censuses' (see CFR, Article 6(2) above).
(2) **Comparability** (item 4.1): Member States must report on national definitions or practices that could impair the data comparability at EU level.
(3) **Timeliness and punctuality** (item 4.2): Member States must report on metadata about release calendars and major revisions.
(4) **Coherence** (item 4.3): Member States must report on information about any major inconsistencies between different hypercubes that have one or more topics in common.
(5) **Coverage and accuracy** (item 4.4): Member States must provide a coverage assessment of their census population bases at national level, evaluating the incidence of record imputations and deletions, as well as the estimated resulting over- and under-coverage, in other words the efforts made so that the 'census population' matches the 'target population' (quality hypercube QHC1). In the other quality hypercube (QHC2), the focus is on measuring the data sources of several selected topics as well as the incidence of imputed and missing information, in other words the degree to which the information on a topic is based on real observations.
(6) **Completeness** (item 4.5) as an aspect of **relevance**: Member States must provide information about any major methodological issue that had a significant impact on data quality and availability, and thus on comparability across the EU.
(7) **Relevance** (item 4.6): Eurostat must provide information at EU level on the actions taken to fulfil user needs, and on the extent of EU census data extractions.

CFR, Article 6(4) requires the Commission to make recommendations to ensure the quality of the data and metadata produced. This is particularly useful for continuous work on methodological harmonisation, where, in some cases, it might be appropriate to simply issue recommendations instead of launching new legislation.

In addition to the provisions on quality, the CFR also contains provisions on **eligible data sources** in Article 4. The eligible data sources that Member States can employ to produce census statistics are presented in Article 4(1). They are:

(a) conventional census;
(b) register-based census;
(c) conventional census + sample survey;
(d) register-based census + sample survey;
(e) register-based census + conventional census;
(f) register-based census + sample survey + conventional census;
(g) rolling census.

Moreover, CFR, Article 4(4) states that the chosen data sources (and methodology) must meet, to the highest possible extent, the 'essential features of population and housing censuses' (see Article 6(2) above). The same Article 4(4) requires Member States to make 'continuous efforts' to improve compliance with these essential features.

5.3. Overview information

The background information on the legal and organisational context is outlined in item 1 of the annex to CIR-3. There are no particular remarks on this item.

([45]) Note that Annex III to CIR-4 adds a few items specifically on the 1 km^2 grid data to this structure: see Section 4.4 for details.

5.4. Data sources

Census data reflect the information in the data sources used to generate them. Therefore, the quality of these sources has a direct impact on the quality of the transmitted hypercubes. This means that quality reporting on censuses must focus on the quality of the data sources.

As outlined in Section 5.2, Article 4(1) of the CFR leaves Member States effectively free to choose the data sources for the census topics. At the same time, the CFR attaches high importance to compliance with the essential features for censuses. This means that the quality reporting has to reflect both considerations:

- the Member States remain free to pool and process their raw data in the way they consider best to produce the census data required;
- the data sources must be measured against the requirements of the EU legislation on censuses, including the 'essential features of population and housing censuses'.

This is also reflected in the quite open definition of 'data source' in CIR-3, Article 2(7), aimed at the purpose — rather than the nature or origin — of a data source:

> *'Data source' means the set of data records for statistical units and/or events related to statistical units which forms a basis for the production of census data about one or more specified topics for a specified target population.*

Therefore, the quality reporting on the data sources focuses on the capacity of the data sources to contribute information in line with the requirements of the CFR. In particular, CIR-3, Article 4 requires information on the extent to which the data sources meet the following four conditions.

1. *Meet the essential features listed in point (i) of Article 2 of [the CFR]:* this is covered under item 2.4 of the annex to CIR-3 (see Section 5.4.4).
2. *Represent the target population:* this is a central element of the quantitative quality assessment covered in item 4.4 of the annex to CIR-3 (see Section 5.6.4).
3. *Comply with the relevant technical specifications in [CIR-1]:* relevant information in this respect is required under two annexes listed below.

- Annex II to CIR-2: for each topic, the metadata must name the data source(s) used to produce the statistical data on the topic and report on:
  - the definitions relating to the topic;
  - the method used to estimate data on the topic;
  - the reasons for any general unreliability of the data on the topic;
  - the handling of specified cases where CIR-1 leaves a choice between different concepts.
- Item 4.1 of the annex to CIR-3: for each topic, Member States must report on any national definition or practice that could impair the EU-wide comparability of the data (imperfect application of the CFR or CIR-1).

4. *Contribute to the provision of data for the programme of statistical data set out in [CIR-2]:* Member States must report on the general completeness of the data under item 4.5 of the annex to CIR-3 (see Section 5.6.5), and also provide a matching between data sources and topics under item 2.3 of the same annex (see Section 5.4.3).

'Administrative registers' versus 'administrative data sources'

Some clarification is needed on this terminology. The CFR, in force since 2008, establishes in Article 4(1) the term '*register-based census*' (see Section 5.2). As opposed to a '*conventional*' (or 'traditional') census, this term describes a situation where all census input microdata come from (administrative) registers not originally compiled by national statistical offices for statistical purposes. Following the CFR terminology, the term '(administrative) register' thus proliferated throughout the 2011 and 2021 implementing legislation, whereas the same concept is nowadays called 'administrative data source'. In fact, the situation today is even more complicated: while 'administrative data source' is the accepted generic term, the ESS Vision 2020 ADMIN project (see Section 5.8) glossary ([46]) defines administrative registers as special cases of unit-level administrative sources with unique identifiers and an established update policy. On the other hand, it became evident during the ADMIN project that some ESS members have still other, divergent definitions of 'administrative register' at their national level.

([46]) https://ec.europa.eu/eurostat/cros/content/draft-glossary-terms-related-vision-2020-admin-project_en; work in progress at the time of publication of these notes.

In conclusion, despite these slight tensions with the more up-to-date ADMIN terminology, **the terms 'administrative data source' and '(administrative) register' are to be understood as interchangeable for the purpose of this publication**, which reflects the intended meaning of the 2021 census legislation ([47]). A stricter use of the generic term 'administrative data source' (and possibly also of the sub-concept 'administrative register') will be strongly encouraged in the future beyond 2021.

5.4.1. Identifying the data sources for quality reporting

Overview information on all census data sources is required under item 2 of the annex to CIR-3. Population and housing censuses are becoming, in several Member States, more and more of a multisource statistical production process. For this reason, it is evident that agreeing on the precise meaning of the term 'data source' (when used in different contexts) is crucial. In particular, we need to define unambiguously the meaning of the term 'data source' whenever it appears in the EU census legislation: when a certain piece of information is to be provided 'for each data source', what exactly does this mean?

The EU census legislation uses the term 'data source' in several circumstances, sometimes with slightly different meanings. In fact, 'data source' can be understood in at least three ways, as outlined below.

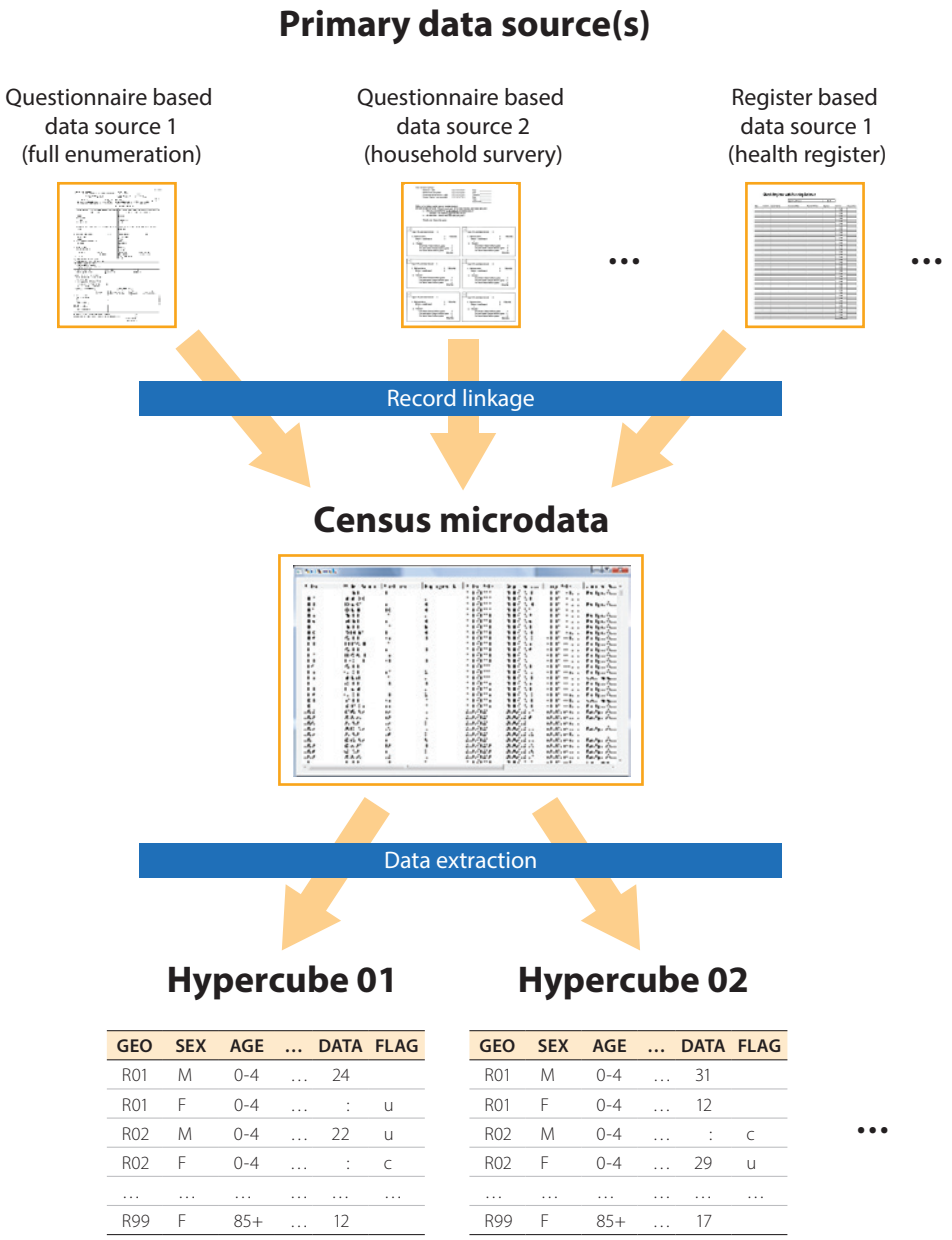
(1) It can be understood in a very general and broad sense, as 'the overall body of information that is used by a Member State as input for the statistical production process'. For instance, this is the meaning of the word 'data source' in Article 4(1) of the CFR (classification of the eligible data sources for the census), and therefore of item 2.1 of the annex to CIR-3.
(2) It can otherwise be understood to refer to a specific 'piece' or component of this information body used as input. In this sense, 'data source' indicates a particular register, questionnaire or survey used to collect the statistical information for one or more census topics. To avoid ambiguity, we will refer to these components as 'primary data sources', which can be of two kinds: either questionnaire-based (full enumeration or sample surveys) or register-based.
(3) It can finally be understood as indicating the result of the operation of record linkage, thanks to which the information contained in separate 'primary data sources' is 'brought together' at the level of each individual statistical unit covered by the census (persons, private households, families and dwellings). This 'data source', which we will indicate with the term 'census microdata database', is a typical 'intermediate output' of the statistical production process in each Member State. It is used to derive, through an operation that we will refer to as 'data extraction', the true 'final output' — at least according to the requirements of EU census legislation, in other words the cross-tabulations defined by census hypercubes.

Figure 6 illustrates the typical census workflow, covering the different meanings of the term 'data source' as used in EU census legislation. According to this framework, 'primary data sources' form the initial input to the statistical production process. They are then assembled together at the level of the individual data record (through record linkage) to constitute the 'census microdata database', which is used to derive the statistical output, in other words census hypercubes. If it is a **traditional census** (in other words the information on all census topics is derived from a comprehensive census questionnaire), the 'primary data source' and the 'census microdata database' coincide.

Table 4 provides a detailed overview of the occurrences of the term 'data source' in the EU census legislation, indicating which of the above interpretations of the term are relevant in each circumstance. In a few cases, multiple interpretations of 'data source' are possible. This is either because a piece of information has to be provided both for the 'primary data sources' and for the resulting 'census microdata database', or because the term is used in a very general sense in that circumstance, and can refer to more than one interpretation without any operational impact on the transmission of metadata.

([47]) As a side remark, the term 'administrative data source' does appear exactly once in the 2021 EU census legislation, namely in Annex II to CIR-2 (under '*Household and family topics*' – see Table C1) where it is indeed used as a synonym for 'administrative register'.

Figure 6: Conceptual overview of data sources for quality reporting



GEO	SEX	AGE	…	DATA	FLAG
R01	M	0-4	…	24	
R01	F	0-4	…	:	u
R02	M	0-4	…	22	u
R02	F	0-4	…	:	c
…	…	…	…	…	…
R99	F	85+	…	12	

GEO	SEX	AGE	…	DATA	FLAG
R01	M	0-4	…	31	
R01	F	0-4	…	12	
R02	M	0-4	…	:	c
R02	F	0-4	…	29	u
…	…	…	…	…	…
R99	F	85+	…	17	

Three additional points must be made below.

- According to CIR-3, Article 4, 'Member States shall report on any data source(s) used to […] contribute to the provision of data for the programme of statistical data set out in [CIR-2]' (in other words they must make it possible to derive the cross-tabulations defined by census hypercubes). It thus seems evident that the vast majority (if not all) of the Member States will produce exactly one 'census microdata database' for each kind of statistical unit included in the census: persons, families, private households, conventional dwellings and living quarters.
- The 'census microdata database' thus serves as the main basis for the enumeration of the census population of each statistical unit, and its existence is justified by the need to comply with the essential features of 'individual enumeration' and 'universality within a defined territory' as laid down in CFR, Article 4(4).
- Item 4.4 of the annex to CIR-3 requires two items of quantitative metadata (quality hypercubes), namely a 'coverage' assessment and an 'accuracy' assessment — both to be 'provided for person counts at national level'. This means that Member States must report on the 'census microdata database' here (see Section 5.6.4).

The previous points show that the coverage assessment, as well as the measurement of the impact of record/item imputation/weighting, does not have to be carried out for the primary data sources. It only has to be carried out for the 'census microdata database' on persons. This is particularly relevant for register-based data sources: it means that Member States **do not necessarily have to report** on the share of imputed records/items in the 'source' registers, unless this information is available. Typically, these registers are maintained by agencies or institutions other than the national statistical office, and the information on how many records/items have been imputed could be very hard — if not impossible — to obtain. However, if the national statistical office has kept track of which records/items have been imputed in the source registers (primary data sources) by the responsible agencies/institutions, the preferred option would be to include these imputations in the reporting.

5.4.2. List of data sources

Item 2.2 of the annex to CIR-3 requires a 'list of the data sources used for the 2021 census'. This is to be understood as all primary data sources and all 'census microdata databases' that were created through record linkage of primary sources. This is to provide a coherent picture of: (i) all source data used for the census, and (ii) how these sources were combined to produce census databases and ultimately extract the hypercubes.

5.4.3. Matching of data sources to census topics

Item 2.3 of the annex to CIR-3 requires a 'Matrix 'Data sources x Topics'. This should be interpreted as having two implications. Firstly, the assessment of the quality of the information provided for each census topic must be based on the 'census microdata database' (which is the same for all topics referring to the same type of statistical unit, as outlined above). Secondly, Member States must also be able **to link each census topic to exactly one 'primary data source'**. This is to evaluate the 'adequacy' of each 'primary data source' and, potentially, to understand where there is room for improvement.

5.4.4. Adequacy of the data sources

Item 2.4 of the annex to CIR-3 requires Member States to report on the 'adequacy of the data sources: extent to which the data sources meet the essential features [CFR, Article 4(4)]'. Data sources must be considered inadequate for census purposes if they are flawed by major deviations from the essential features, definitions and concepts required by the EU legislation, and if these deviations seriously impair adequate use of the census results. The hope is that this will apply only in exceptional cases. However, if major problems appear, Member States must report them honestly in the interest of an open discussion and improvement in the long run.

Table 4: Operational interpretation of the term 'data source' in 2021 EU census legislation

Reg.	Article / item	Provision	(1) overall body of information	(2) primary data source(s)	(3) census microdata database
CFR	Item (3) of the preamble	[…] the enhancement of the data sources […] should be fostered		☑	
	Art.4(1)	Member States may base the statistics on different data sources […]	☑		
	Art.4(3)	Member States shall inform the Commission […] of any changes in the chosen data sources and methodology […]		☑	
	Art.4(4)	Member States shall ensure that the data sources and the methodology […] meet […] the essential features		☑	
	Art.6(2)	[…] Member States shall report on the extent to which the chosen data sources and methodology meet the essential features […]		☑	
CIR-1	Annex topics WSS., TOI., BAT., TOH.	[…] evidence based on […] administrative data sources or from sample survey data		☑	
CIR-2	Annex II 'Metadata on the topics'	[…] name the data source(s) used to report the statistical data on the topic		☑	
	Annex II 'Metadata on the topics' — homeless	A description of the methodology and data sources used to produce the data on homeless […]		☑	
CIR-3	Art.2(7)	'data source' means the set of data records for statistical units and/or events relating to the statistical units […]	☑	☑	☑
	Art.2(10)	[…] The data records for the census population are those in the data source for the specified target population […]			☑
	Art.2(10)	[…] If a data source comprises […] data records for only a sample of the statistical units […]			☑
	Art.2(11)	[…] on which, as a result of an applied sampling methodology, the data source contains no data records			☑
	Art.2(16)	[…] the imputation of that data record into a data source			☑
	Art.2(17)	[…] ignoring/not taking into account a data record that is included in a data source […]			☑
	Art.2(18)	[…] into a data record that already exists in a data source […]			☑
	Art.2(20)	'record linkage' means the process of merging information from different data sources […]		☑	
	Art.2(21)	'unique identifier' means a variable or set of variables in the data records in a data source […] used for verifying that the data source […] includes no more than one data record for each statistical unit, and/or for record linkage		☑	
	Art.2(24)	'matching of registers' means record linkage where all matched data sources are contained in registers		☑	
	Art.4	Member States shall report on any data source(s) […]		☑	☑
	Annex item 2.1	Classification of the data sources according to […]	☑		
	Annex item 2.2	List of the data sources used for the 2011 census		☑	☑
	Annex item 2.3	Matrix 'Data sources x Topics'		☑	
	Annex item 2.4	Adequacy of data sources […]		☑	
	Annex item 3.2.4	[…] (data based on more than one type of data source) […]		☑	
	Annex item 3.2.4	[…] (types of data sources used and how information from different sources was combined, how the different sources and methods used complement and support each other and, if applicable, which parts of the population were covered by which source) […]		☑	
	Annex item 4.1	Member States shall report on the extent to which the data sources […] result in deviation from the definition of 'Current Activity Status'		☑	
	Annex item 4.4	(e) number of observed data records on the topic derived from multiple data sources		☑	
CIR-4	Annex III 'Metadata on the topics'	[…] provide metadata on each topic included in Annex I informing about the data sources […] used to obtain the data values for that topic on the 1 km^2 reference grid		☑	

5.5. Census lifecycle

The different phases of the census operation must be described under item 3 'Census lifecycle' of the annex to CIR-3. The census lifecycle refers to processes. Therefore, the data sources listed in item 3 must include all sources involved in the process, in other words all original and derived data sources (see Section 5.3). Some remarks on items, 3.2, 3.3, 3.4 and 3.5 follow below.

- Item 3.2 'Preparation and execution of data collection' distinguishes between four types of data.
  - Questionnaire-based data as defined in CIR-3, Article 2(19): item 3.2.1 covers any collection of statistical data used to produce census data that has been conducted by means of a questionnaire with reference to a pre-determined point in time.
  - Register-based data as defined in CIR-3, Article 2(23): the first three bullet points of item 3.2.2 cover all registers in the sense of 'primary data sources' (for example administrative registers, see Section 5.4.1) used to produce census data, and also the processes that affect these source registers (for example reporting delays). The last two bullet points of item 3.2.2 cover the creation of the 'census microdata databases' (Section 5.4.1) and extraction of data from them.
  - Data collected from a sample: this includes information on any census data obtained from a sample (in other words no full coverage of the census population intended), such as sampling design, estimation methodologies, biases, standard errors, and so on. Obviously, questionnaire-based data (see above) can simultaneously be sample data (for example a questionnaire-based sample survey). Because item 3 of the annex to CIR-3 concentrates on the collection/production **processes**, such data sources should therefore be addressed twice: both under item 3.2.1 (concentrating on the questionnaires) as well as under this item (concentrating on the sample aspect).
  - Data collected by combined methods: item 3.2.4 focuses on the process of producing a single 'census microdata database' for a given statistical unit by joining information from different types of primary data sources (see, for example, Figure 6). Such situations are explicitly allowed by provisions (c)-(f) of CFR, Article 4(1) (see also Section 5.2).
- Item 3.3 'Processing and evaluation' requires descriptions of methods that were applied. These descriptions must cover:
  - how primary data were processed after collection to establish the final census microdata database of each statistical unit for hypercube extraction (item 3.3.1) ([48]);
  - how the coverage was assessed for the resulting census populations contained in each microdata database, including information on the estimation of under-coverage and over-coverage and the quality of these estimates (item 3.3.2) ([49]).
  - Annex III to CIR-4 (point 1 under 'Reference metadata') adds another item 3.3.3 to the quality report, which covers:

    Additional information on generic (not topic-related) methodology applied in order to produce the 1 km² grid dataset.

    This means the generic production process for the grid dataset should be described here: for example geocoding procedure, aggregation method, disaggregation and/or estimation method(s) if applicable. Note that related topic-specific information should be reported under the topic metadata (see Section 4.4.2).

- Item 3.4 addresses dissemination issues, including the '*assurance of statistical confidentiality including statistical disclosure control*'.
  - While the choice and application of appropriate measures is within the competence of the Member States, national statistical offices are strongly encouraged to take note of, and consider applying, the recommendations for good practices compiled by the project on 'harmonised protection of census data in the ESS' ([50]). The aim of these recommendations is to support the convergence of national confidentiality measures and thus improve data quality (in particular, comparability) at the EU level. In any case, any measures applied must be reported and described under this item.
  - Annex III to CIR-4 (point 2 under 'Reference metadata') adds specific provisions related to the 1 km² grid data:

    Item 3.4. 'Dissemination' shall be supplemented by specific information about statistical disclosure control measures related to the 1 km² grid dataset. Member States shall provide the Commission (Eurostat) with information about the measures related to the harmonised protection of 1 km² grid data, in particular if they used the ESS good practices and implementation guidelines for the harmonised protection of 1 km² grid data.

([48]) Especially when combining multiple sources, note the good practices and guidelines on the use of estimation methods for the integration of administrative sources developed under the ESS Vision 2020 ADMIN project (see Section 5.8).

([49]) The Quality Guidelines for Multisource Statistics (QGMSS; see Section 5.8) are due to provide various guidelines and measures to assess coverage of data based on administrative sources.

([50]) https://ec.europa.eu/eurostat/cros/content/harmonised-protection-census-data_en; see Section 3.4.

This means that Member States must add dedicated information on the confidentiality measures applied to the grid data. They must in particular take note of the project recommendations on 1 km² grid data, and report whether and how they applied them. Ideally, this includes quantitative information about the uncertainty added to the figures in the grid table. For example, if the project recommendations were used, it should include public key parameters of the targeted record swapping and the variance applied in the cell key method. If other methods were applied, it should include similar quantitative uncertainty measures.

- Item 3.5 covers 'Measures to ensure cost effectiveness'. 'Cost effectiveness' is principle 10 of the European Statistics Code of Practice to which the CFR explicitly refers in recital 5. This principle is particularly relevant to the population and housing census, which is probably the most costly statistical exercise within the European Statistical System. However, the EU Member States follow different budgetary rules to identify the cost of a census, and the different census methods result in wide variations in the share of fixed costs against variable costs. As a result, harmonisation of figures for the total cost of the 2011 census is practically unachievable.

5.6. Assessment of data quality

5.6.1. Comparability

Item 4.1 of the annex to CIR-3 states that:

> *For each topic, Member States shall report on any deviation from the required concepts and definitions or any practice in the Member State that could impair the Union-wide comparability of the data.*

This concerns any significant imperfect application of the CFR or CIR-1.

This is essential for assessing the comparability of the data transmitted. It is equally essential for assessing the appropriateness of the primary data sources used. This is because CIR-2, Article 4 requires, among other things, that any (primary) data source must have the capacity to contribute information that complies with the relevant technical specifications laid down in CIR-1.

Note that, as a novelty for 2021, the item contains an additional provision:

> *For the topic 'Current Activity Status', Member States shall report on any estimation methods used to adjust data to meet more closely the definition set in the Annex to [CIR-1]. Member States shall report on the extent to which the data sources and any estimation methods used result in deviation from the definition of 'Current Activity Status' set in that Regulation.*

This is to address explicitly the known issues with the application of the definition of 'Current activity status' in some Member States.

5.6.2. Timeliness and punctuality

Item 4.2 of the annex to CIR-3 states that Member States must provide information about:

- date(s) of the transmission of data to the Commission (Eurostat), broken down by hypercubes ([51]);
- date(s) of major revision(s) of the transmitted data, broken down by hypercubes ([51]);
- date(s) of transmission of the metadata ([52]).

Annex III to CIR-4 (point 3 under 'Reference metadata') extends these provisions to also provide information about:

- date(s) of the transmission and possible revisions of the 1 km² grid data and metadata.

([51]) As listed in Annex I to CIR-2.

([52]) As listed in Annex II to CIR-2, in the Annex to CIR-3, and in Annex III to CIR-4.

On the basis of all this information, the timeliness and punctuality of provision of the data can be evaluated for each hypercube and set of metadata, including the 1 km^2 grid.

In the case of a major revision after the deadline set for the quality reporting in CIR-3, Article 3, Member States must report the respective date(s) within one week after the major revision.

5.6.3. Coherence

Item 4.3 of the annex to CIR-3 states that:

> *Member States shall report on any significant inconsistencies between the data transmitted in the different datasets defined in [CIR-2].*

This means that data on a given topic and its breakdowns should be consistent across all hypercubes listed in Annex I to CIR-2 that contain this topic. For instance, both hypercube 13.1 and hypercube 14.5 contain the breakdown HAR. ('Housing arrangements') as one dimension. Therefore, the aggregation over all other hypercube dimensions (GEO.L., SEX., POB.M. and YAE.H. for hypercube 13.1; and GEO.L., SEX., AGE.L., CAS.L. and ROY. for hypercube 14.5) should lead to consistent results between the two hypercubes for each single breakdown category of HAR. Obviously, such checks can easily be automatised to cover all topics of CIR-1, across all hypercubes in CIR-2.

Note that the reporting burden of national statistical offices on coherence has been substantially reduced. For the 2011 exercise, many quality hypercubes were required to underpin consistency across hypercubes, but it turned out that these tables were very rarely accessed by users after their dissemination. Therefore, the approach taken for 2021 is that such inconsistencies are assumed to be noted by the national statistical offices during their production processes. Such incidences must certainly be reported under this item, which can be done in the form of free text describing the observed incidents (textual metadata, cf. Table C1), but verified coherence of the data does not entail a particular reporting burden anymore.

5.6.4. Coverage and accuracy

The complete quantitative assessment of the quality dimension 'accuracy' (including 'coverage' as a key aspect) must be provided for data sources and selected topics that refer to natural persons. As laid down in item 4.4 of the annex to CIR-3, quantitative quality information on persons (in other words the two quality hypercubes QHC1 and QHC2) must be provided at the national level, disaggregated by SEX. and AGE.L. Table D1 in Annex D to this document illustrates the link between each quality hypercube and the specific piece of legislation it refers to.

Note that the programme of quantitative metadata has been substantially redesigned and simplified since the 2011 exercise. The number of compulsory quality hypercubes has decreased from 21 to just 2. In addition, inside the two remaining hypercubes several quality indicators were dropped completely to further simplify the compilation and provision of required information (unit/item no-information, coefficient of variation for samples, all percentage indicators).

The questions to be answered

The quantitative metadata (quality hypercubes) focus on the match between the information desired on the (ideal) target population and the information that has actually been retrieved from the data sources.

For this reason, the quantitative reporting aims to answer two fundamental questions.

1. **How well do the published data reflect the 'true' population?**
2. **To what extent are the published data based on real observations?**

These questions can be evaluated on the basis of an assessment of coverage (quality hypercube QHC1) and of the origin of information on selected topics on persons (QHC2).

5.6.4.1. QHC1 ON COVERAGE: HITTING THE TARGET POPULATION

Quality hypercube QHC1 deals with the fundamental issue of how well the census exercise in a Member State is capable of enumerating precisely the target population. It therefore requires a quantitative assessment of the impact of coverage problems (deviations between the census population and the estimated target population) and of the extent to which the census population has been affected by record imputations/deletions. This section of the report aims more at the legal provisions and how the quality reporting requirements of the CFR are served through QHC1. However, Annex D1 contains a detailed technical explanation on how to actually compile QHC1 from the information stored in the 'census microdata database' on persons.

Target population, estimated target population and census population

To report on imputation and deletion of data records and on coverage, CIR-3, Article 2(8-10) defines 'target population', 'estimated target population' and 'census population' as set out below.

- The **target population** is the ideal (or 'true') population of statistical units. It is a theoretical concept on which a census aims to report without ever achieving perfect coverage in practice:

  'Target population' means the set of all statistical units in a defined geographical area at the reference date which qualify for reporting on one or more specified topics. It includes each valid statistical unit exactly once.

  Census outputs should ideally be based on information about each statistical unit in the target population.

- The **estimated target population** is the best available estimate for the (ideal) target population. The estimate is based on the census population corrected for the errors detected by a coverage assessment:

  'Estimated target population' means the best available approximation of the target population. It consists of the census population plus under-coverage minus over-coverage.

  Mostly, only key figures will be available for the estimated target population (for example population by sex and broad age groups). To meet the requirements of item 4.4 of the annex to CIR-3, only figures for the size of the total estimated target population of persons by SEX. and AGE.L. must be provided.

- The **census population** is the population that the census data actually reflect. It is represented in the 'census microdata databases' (one per statistical unit), as explained in Section 5.4.1. The data records for the census population include all imputed records and exclude all deleted records:

  'Census population' means the set of statistical units which is factually represented by the census results on one or more specified topics for a specified target population. The data records for the census population are those in the data source for the specified target population, including all imputed records and excluding all deleted records. [. . .]

By definition, all data sources referring to persons (in a specified area and time) have the same target population and, in practice, the same estimated target population. Ideally, the census population should also be the same. In practice, however, the census population can vary from one topic to another. For example, data for one topic might be based on a fairly complete, high-quality frame, but for another topic on a frame of inferior quality. Although this is measured in QHC2 (see Section 5.6.4.2), it is assumed for QHC1 that the Member States will use all available information (in other words primary data sources) on a given statistical unit to derive a 'census microdata database' whose records reflect the estimated target population in the best possible manner (see Section 5.4.1). Hence, the census population of persons required under point (a) of QHC1 must be that contained in the complete final 'census microdata database' on persons.

Imputed and deleted data records

The essential features 'individual enumeration' and 'universality within a defined territory' imply that the census microdata database should ideally contain exactly one data record for each statistical unit of the target population. In practice, however, this will hardly ever be the case. Records might be missing or there might be too many records in the database.

To correct this, the statistical offices might choose to impute substitute records or to delete invalid records. The CFR requires Member States to report on the 'extent to which the chosen data sources and methodology meet the essential features' (Article 6(2)). This 'extent' depends partly on the number of imputed and deleted data records. It is

therefore very important to include the assessment of deletion and, particularly, of imputation of data records in the quality analysis. This information is required to answer both questions introduced at the beginning of this section.

The problem of missing or redundant data records is relevant at different stages of the census operation. Indications that data records are missing might appear during the field work, capturing, coding or editing phases of the census. Some records might also over-represent the target population because the authorities forgot to clear them from a register (for example non-deregistration or multiple registration). The standard procedure to correct coverage problems is by imputing records ([53]). Therefore, the national statistical offices must report the number of person records that are not real observations but have been imputed into the census population under point (b) of QHC1 defined in item 4.4 of the annex to CIR-3 (cf. Table D1).

They must also count under point (c) the person records that were originally in the data source but that were not included in the final census population (see Figure 7). Only actions that lead to a decrease in the size of the census population of a given statistical unit should be counted as record deletions. In many cases, this means that only records deleted from a population register or a validated address list must be counted. In the case of census microdata databases created by linkage of records from different primary data sources, it is important that all record imputations and deletions are completely inherited in the resulting census microdata databases.

Some examples:

- By mistake, all records for an apartment building are duplicated in a 'raw' address list. The mistake is detected during the enumeration and the list is corrected. The fact that the duplication is corrected does not lead to any deleted data records because the 'duplicated persons' were never considered to be part of the census population.
- In a country where personal identification numbers (PINs) are used frequently for administrative procedures, a population register contains a data record that has successfully passed a technical validation. However, the national statistical office finds out that the PIN number has not been used for any administrative procedure for the last x years. A check reveals that the person does not live at the address and that no other address for that person in the same country is known. The national statistical office assumes that the person has left the country without deregistering. The record is deleted from the census database. This record deletion is to be counted because the census population based on the population register (which is generally considered to represent the census population) had to be corrected downwards by one person.
- If a missing person is imputed into a household with, for example, four persons already present, it should not be regarded as an imputed household: The defining feature of record imputations (as opposed to item imputations, see QHC2 in Annex D) is the fact that they result in a change to the census population of that statistical unit. In this example, attributing an additional fifth person to an already existing household does not change the total number of households. It should therefore not be considered a record imputation in the census microdata database of households. However, since the number of people present in this household now results from an imputation, it does represent an *item imputation* for the variable 'SPH — Size of private household'.

Coverage assessment

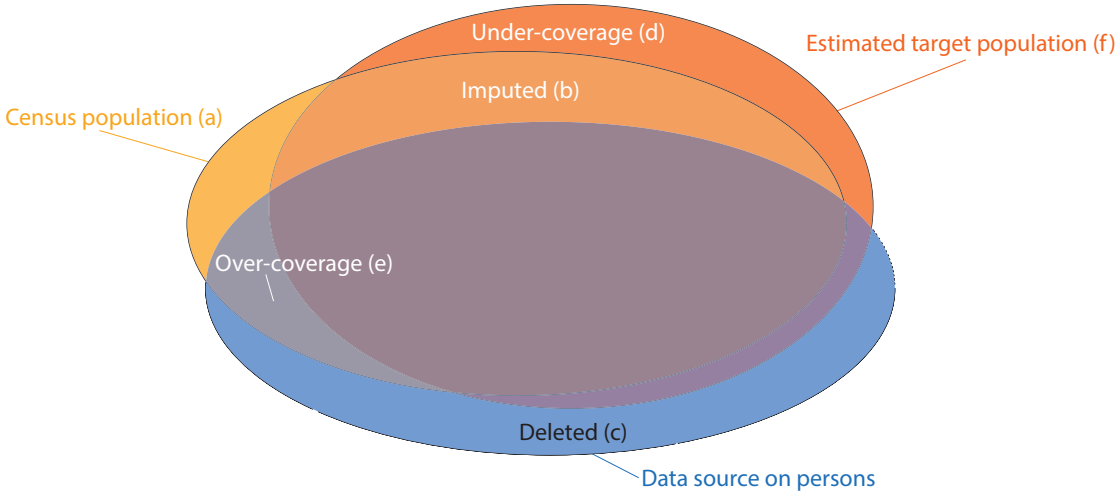
Even after processing and combining all primary data sources, the census population in the resulting microdata database will very likely not match the ideal target population. Member States must therefore conduct a coverage assessment to estimate the under-coverage or over-coverage of the total population of persons by SEX. and AGE.L. to comply with item 4.4 of the annex to CIR-3. The Regulation does not stipulate any statutory method or threshold for such assessments.

Coverage problems contravene the essential feature 'universality within a defined territory'. Information on under-coverage and over-coverage must answer, as well as possible, question 1 above ([54]). To do this, it is important that Member States report under-coverage and over-coverage separately. High under-coverage can be compensated for by high over-coverage. However, high over-coverage would lead to a situation where the published data would not reflect real (and usable) observations. As a result, a 'net balance' of coverage of 100 % can mask substantial quality problems. See Figure 7 and also the examples in Table D5.

([53]) Some statistical institutes might apply some reweighting of individual data records to improve the plausibility and/or consistency of the data, but based on previous experience such procedures are considered less relevant for the overall coverage adjustment (see Section 5.6.4.3).

([54]) How well do the published data reflect the 'true' size of the population?

Figure 7: Illustration of the census population and coverage issues addressed in QHC1



The information provided on under-coverage and over-coverage is itself an estimate and therefore affected by mistakes. As a result, a perfect answer to question 1 can never be given in practice. This is why item 3.3.2 of the annex to CIR-3 requires provision of metadata on the operation to assess under-coverage and over-coverage, including information on the quality of the estimates of under-coverage and over-coverage.

A coverage assessment can be based for instance on a post-enumeration survey or information gathered during the process. For example, over-coverage could be assessed through:

- a study of the number and kind of non-eligible records found in the data source;
- a study of duplicated personal identification numbers (PIN);
- a study of the consistency of data from different data sources on identical statistical units;
- process information on questionnaires filled in twice;
- a study focusing on population groups that are often over-counted, such as students, children of divorced parents or persons with more than one place of accommodation.

The coverage reported under points (d) and (e) of QHC1 (cf. Table D1) must refer to the census population. That means that if a country has imputed or deleted records to counter coverage problems detected during a preliminary coverage assessment, the results of this preliminary coverage assessment must not be reported here. Instead, the results for the final coverage (which could not be 'corrected for') must be reported under points (d) and (e), see Figure 7.

CIR-3, Article 2(12-15) gives a number of definitions that are relevant in this context:

> *'Coverage assessment' means a study of the difference between a specified target population and its census population.*
>
> *'Post-enumeration survey' means a survey conducted shortly after the enumeration for coverage and content assessment purposes.*
>
> *'Under-coverage' means the set of all statistical units that belong to a specified target population but are not included in the corresponding census population.*
>
> *'Over-coverage' means the set of all statistical units that are included in a census population used to report on a specified target population, but without belonging to that target population.*

5.6.4.2. QHC2 ON ACCURACY: QUANTITATIVE INFORMATION ON DATA SOURCES AND TOPICS

The purpose of QHC2 is to provide a quantitative indication, on a topic by topic basis, of the quality of the 'statistical content' of the 'census microdata database', essentially in terms of the share of the information on a topic that has been actually observed (and not imputed or missing, for instance). This section focuses on the legal provisions and how the quality reporting requirements of the CFR are served through QHC2. Annex D2 contains a detailed technical explanation how to actually compile QHC2 from the information stored in the 'census microdata database' on persons.

Identification of data records containing information on the topic

Not all data records in the census microdata database of statistical units to which a given topic applies necessarily contain observed information on that topic. Three of the reasons this might be the case are discussed in the following bullet points.

- The microdata database may have been compiled from several primary data sources, some of which might not cover the complete census population. This can lead to data records with different degrees of completeness. For example, some data records might only give details of 'sex' and 'age', whereas others contain information from other, more informative, primary sources (see Section 5.4.1).
- Unit no-information might lead to data not being available (cf. definition in CIR-3, Article 2(31)):

  'unit no-information' means the failure to collect any data from a statistical unit that is in the census population.

  Explicit information on this was dropped from all quality hypercubes for 2021, but under item 3.3.1 of the annex to CIR-3 Member States must still describe '*measures to identify or limit unit no-information*' (cf. Section 5.5).

- Even in a microdata database derived from a single primary source, there might be records with missing information on the topic.

This is why the number of data records containing observed, imputed or missing information on the topic must be reported in QHC2.

Following the logic of the reporting on the data sources, QHC2 must answer question 2 above: **To what extent are the published data on the topic based on real observations?**

For this purpose, all instances of observed information on the topic must be reported by type of data source under points (b) to (e) of QHC2 (cf. Table D1), supplemented by the size of the complementary set of statistical units in the case of sample survey under point (f) (see the following dedicated paragraph). Moreover, imputed observations on the topic must be reported under point (g) of QHC2, namely:

- all record imputations (see Section 5.6.4.1) that have an effect on the topic in question; and
- all item imputations for that topic.

For these purposes, item imputation is defined in CIR-3, Article 2(18) as the:

> *insertion of an artificial but plausible value on a specific topic into a data record that already exists in a data source but either does not contain this value or contains a value that is considered implausible.*

As already stressed in Section 5.6.4.1, in the case of microdata databases from multiple primary data sources, a record imputation into one of the primary data sources might not lead to an increase in the census population, as measured by the resulting microdata database. However, for the topics provided through that primary data source, this imputation must be counted as an item imputation.

Samples as data sources

For some topics, Member States might use data sources that are based only on statistically controlled samples of the statistical units of the target population. Clearly, such practices contravene the essential feature 'universality within a defined territory'.

QHC2 makes sure that any such action is documented in the same way as imputations in the quality measurement for the data sources and in the light of the answers to the two questions. This is achieved by the definition of 'census population' in CIR-3, Article 2(10), which states that:

> *[. . .] If a data source comprises, as a matter of methodological principle, data records for only a sample of the statistical units in its estimated target population, the census population includes the complementary set of statistical units, in addition to the statistical units in the sample.*

'Complementary set of statistical units' is defined in CIR-3, Article 2(11) as:

> *the set of those statistical units that belong to an estimated target population, but on which, as a result of an applied sampling methodology, the data source contains no data records.*

Note that there appears to be a slight tension between the two definitions, because CIR-3, Article 2(10) implies that the sample and its complementary set should add up to the 'census population', while CIR-3, Article 2(11) suggests that the sample and its complement combine to the 'estimated target population'. However, as described in detail in Section 5.6.4.1, any discrepancy between the 'census population' and the 'estimated target population' manifests itself only through a dedicated coverage assessment conducted after the actual census exercise. This in turn implies that, at the time of compiling the 'census microdata databases', the resulting set of microdata records also factually represents the best estimate available at the time for the estimated target population. It is thus consistent to require that the sample and its complementary set add up to the census population for the purpose of the accuracy assessment of the microdata database under QHC2. For instance, assuming a non-stratified random sample:

$$CS = u * \left(\frac{1}{s} - 1 \right)$$

where *CS* is the size of the complementary set of statistical units, *u* is the number of statistical units actually contained in the data source, and *s* is the relative sample size.

If information has been collected by means of a sample, Member States must provide information not only on the sample size (in QHC2), but also on the sample design, under item 3.2.3 of the annex to CIR-3 (cf. Section 5.5). Together, this information can be taken as input for the quality assessment of the data transmitted for the topic concerned.

5.6.4.3. ATTRIBUTING WEIGHTS ≠ 1 TO DATA RECORDS

Member States might wish to attribute weights not equal to 1 to individual data records or items ([55]). For instance, this might be the case if:

- Member States scale some data up or down by systematically increasing or decreasing the weights of parts of the census population;
- Member States use methods that adjust the weights of data records to improve the plausibility and/or consistency of the data.

Such practices contravene the essential feature 'individual enumeration', and in some cases they also contravene 'universality within a defined territory'. However, experience with the 2011 EU census exercise suggests that these practices are not applied in the national statistical offices in such a way as to notably affect the resulting census populations. Therefore, and to simplify the reporting, any explicit reference to reweighting has been dropped in the 2021 EU census legislation. This in turn suggests that any quantitative data transmitted in the quality hypercubes (in other words absolute person counts) should be unweighted. For comparability reasons, Member States are therefore strongly encouraged to report only unweighted record counts in the quality hypercubes. Otherwise they should document under item 3.3.1 of the annex to CIR-3 why and how they applied reweighting if they need to transmit weighted counts for some reason.

([55]) The weights can be decimals. The weights are often contained in the same database as the data records in a way that makes attribution of the weights to the records possible so that they can be retrieved during the process of calculating aggregated data.

5.6.4.4. ACCURACY: PROCESS-ORIENTED MEASUREMENTS?

The quantitative quality reporting under CIR-3 focuses on the match between the information desired on the target population and the information actually retrieved from the data sources. It does not focus on process-oriented measurements, which could count facts like the non-existence of statistical units, refusals to answer interviews, proxy interviews, changes of the status of the person, persons who replied after first/second/third contacts, records that needed to be edited in a specific way, records corrected for wrong coding, and so on.

The variety of census methods rapidly renders process-oriented measurements from different Member States incomparable. Moreover, the main purpose of process-oriented measurements is to evaluate and improve processes (which can be very useful at the level of national/regional statistical offices), whereas the main purpose of the EU quality reporting is to assess the quality of the data transmitted.

5.6.5. Completeness

Item 4.5 of the annex to CIR-3 requires Member States to provide information on the *'degree of completeness of the data in terms of the requirements of [the CFR]'*. As for the provisions on coherence (cf. item 4.3 of the annex to CIR-3 and Section 5.6.3), the reporting burden on completeness has also been substantially reduced compared to the 2011 exercise. Rather than requiring exhaustive information documenting the completeness of transmitted data in the form of quantitative metadata, CIR-3 now requires Member States to report details only about specific topics and breakdowns on which data could not be supplied as set out in CIR-1 (implementing the CFR). As for coherence, this can be reported on in the form of free text describing the observed incidents (textual metadata, cf. Table C1).

5.6.6. Relevance

Apart from completeness (see previous section), this quality dimension also covers user aspects of the data and metadata, namely:

- user needs: Description of users and their respective needs with respect to the statistical data;
- user satisfaction: Measures to determine user satisfaction.

A new feature for 2021 that was not present in 2011 is item 4.6 of the annex to CIR-3, where Eurostat commits itself to provide information at EU level on:

(1) actions taken to identify and fulfil user needs;
(2) monitoring of the extent of data extractions.

Provision 1 documents Eurostat activities to gain better knowledge on user needs, for instance the continuous coordination during the preparation of CIR-4 with DG REGIO as a core user of the resulting 1 km^2 grid data. Provision 2 gathers empirical information on the '(lack of) popularity' of the various data and metadata sets published through the 2021 EU census exercise. Both measures are intended to facilitate continuous improvements of the future EU censuses in terms of user satisfaction and fitness for purpose.

5.6.7. Geographic information — data quality

Annex III to CIR-4 (point 4 under 'Reference metadata') adds another item 4.7 to the quality report, which covers the quality of geographic information used for geo-referencing persons in the 'census microdata database'. This is mainly relevant for the quality of the 1 km^2 grid data. For convenience, it is discussed along with all other grid-related aspects in Chapter 4 (see Section 4.4.3).

5.7. Access to all information relevant to the quality assessment

Article 5 of CIR-3 states that:

4. *Member States shall provide the Commission (Eurostat), at its request, with access to any information relevant to the assessment of the quality of the data and metadata transmitted under Regulation (EU) 2017/712.*
5. *In complying with paragraph 1, Member States shall not be obliged to provide the Commission (Eurostat) with any microdata or confidential data.*

This is to comply with Article 6(3) of the CFR, which requires Eurostat to assess the quality of the census data transmitted from the Member States. To carry out this assessment, Eurostat relies entirely on the quality reporting by the Member States. Moreover, the national statistical offices not only evaluate their own processes and data, they also, in turn, depend on data provided by third parties, namely the owners of the data sources. Eurostat has the authority to take a closer look if it has the impression that the quality reporting by a Member State itself is not of sufficient quality or scope. Failure to do so can undermine the confidence users have in the whole quality assessment of the censuses in the European Statistical System.

5.8. Outputs from the ESS Vision 2020 ADMIN project

The Administrative Data Sources (ADMIN) project is one of the eight projects in the ESS Vision 2020 portfolio ([56]). Its objective is to make more use of available data sources — in particular administrative sources held by other public authorities — for official statistics. Most ESS members are moving towards an increased use of administrative data sources for statistical purposes to complement, or even replace, previous surveys. For this reason, the ADMIN project runs from 2015 until 2019 with two aims. The first aim is to help ESS members reap the benefits (decrease costs and burden, increase data availability) of using administrative data sources for the production of official statistics. The second aim is to support the quality assurance of the output produced using administrative sources, in particular the comparability of statistics required for European purposes.

One of the central goals of ADMIN is to carefully take stock of the current state of the art in the ESS, and to compile a compendium of ADMIN-related knowledge from different Member States for general reference. This work can be broadly divided into two aspects: (i) good practices reflecting conditions currently encountered in Member States, and (ii) common guidelines delineating desired target states as a synthesis of good practices as well as academic and practical knowledge on a given topic. The main outputs serving the knowledge dimension are high-level documents for common reference across the ESS, summarised in the following overview table.

Topic	Good practices and reports	Guidelines and recommendations
Access to data	Report on the legal and institutional framework for use of administrative data	
	Good practices for relations with data providers	
Methodology	**Good practices** for combining multiple sources using modelling and estimation methods	**Guidelines** on the use of estimation methods for the integration of administrative sources
Quality		Quality **guidelines** for multisource statistics (QGMSS)
	Report on the availability, kind and development of registers on persons and dwellings and derived frames used in the Member States	Quality **guidelines** for frames of social statistics (QGFSS)
		Methodological recommendations on the features of the frames used for social statistics in the Member States
Lessons learned	Summary report of specific ADMIN implementation actions ([57]) including knowledge transfer	

([56]) http://ec.europa.eu/eurostat/web/ess/about-us/ess-vision-2020/implementation-portfolio.

([57]) ADMIN work package 6 'pilots and applications' provided funding for specific projects of ESS members to develop or improve the use of administrative sources for official statistics.

The central knowledge base for publishing and sharing ADMIN information — where all documents are published as soon as they become available for publication — is the CROS portal ([58]). In fact, at the time of publication of these notes, all documents above were already available — except the QGMSS tentatively scheduled for the second quarter of 2019 (work-in-progress draft available) and the recommendations for frames (2019). Moreover, it is planned to establish all three guidelines as ESS standards ([59]). This implies that they are planned as 'evolving' documents which will not be 'final' and frozen after the end of the ADMIN project in 2019.

ADMIN relevance for EU censuses

Social statistics, and especially census/population statistics, are at the heart of the ADMIN project. This is not surprising given that most of the ESS members are using, or moving towards, register-based census systems or multisource systems that also integrate administrative sources (see Section 5.4 and Figure 6; as further explained in Section 5.5, CIR-3 explicitly provides for — and requires metadata on — register-based or multisource census data). Thus, the ADMIN project covers various core outputs that are highly relevant for census production in general, especially considering the post-2021 modernisation efforts at national and ESS level. Focusing more on the legislation for the 2021 EU census round addressed in this publication, there are important contact points with the census production lifecycle and with quality assessment.

- The legislation allows for great flexibility with data sources and methodology. However, the **good practices and guidelines for modelling/estimation methods** ([60]) using administrative data developed under ADMIN may also provide valuable inputs to further improving the census production process. In addition, the reporting on the census lifecycle (Section 5.5) can benefit from references to commonly accepted practices and guidelines.
- Also, the **quality guidelines for multisource statistics (QGMSS)** ([61]), developed under ADMIN, aim to convey a holistic picture from administrative inputs via data processing to output quality. In particular, the (output) quality dimensions discussed in Section 5.6 are extensively covered in the QGMSS, where respective guidelines and quality measures can support the quality reporting under CIR-3.
- Most of the resources related to specific implementation actions under ADMIN were invested in social statistics topics, and in particular to the preparation of the 2021 census and future census rounds. While projects typically addressed specific (national) challenges or needs, a comprehensive summary review was published to categorise the results and make them accessible and searchable by domains and keywords. This survey review especially highlighted **lessons learned and knowledge transfer** for the whole ESS ([62]). Particularly relevant topics include the estimation of census topics; quality assessment; and geocoding of the census database for the 1 km^2 grid dataset under CIR-4 (Chapter 4).

([58]) https://ec.europa.eu/eurostat/cros/content/ess-vision-2020-admin-administrative-data-sources_en.

([59]) https://ec.europa.eu/eurostat/cros/content/ess-standardisation_en.

([60]) https://ec.europa.eu/eurostat/cros/content/wp2-statistical-methods_en.

([61]) https://ec.europa.eu/eurostat/cros/content/wp3-quality_en.

([62]) https://ec.europa.eu/eurostat/cros/content/wp6-pilots-and-applications_en.

6 Data and metadata transmission

The previous chapter shows that the requirements of the EU legislation on the quality reporting for the 2021 census are rather detailed and comprehensive. As also explained in Section 5.1, the entire system has been designed so that it is simultaneously capable of:

- describing very different types of production systems, including traditional censuses, register-based censuses and censuses obtained by combining several types of data sources (questionnaire-based and/or register-based);
- providing a framework for a comparable assessment of the quality of the statistical output, regardless of the data sources and census methodology employed in each Member State.

Furthermore, the complete transmission programme consists of quantitative data (hypercubes), metadata (quality hypercubes) and textual metadata (providing quality-related information and information on the 'background' of the individual census processes in each Member State).

On the collection of all this information, CIR-3, Article 6 states that:

> *The technical format to be used for the transmission of data and metadata for the reference year 2021 shall be the Statistical Data and Metadata eXchange (SDMX) format as implemented through the Census Hub. Member States shall transmit the required data in line with the data structure definitions and related technical specifications provided by the Commission (Eurostat). Member States shall store the required data and metadata until 1 January 2035, for any later transmission requested by the Commission (Eurostat).*

In short, Article 6 stipulates that the 2021 census data must be available in SDMX format and that the Member States must use the data structure definitions provided by Eurostat. Furthermore, contrary to the 2011 situation, the obligatory use of the EU Census Hub itself is now explicitly regulated for 2021. This is the most efficient approach and does not involve any notable risk, because the system already proved itself in the 2011 exercise.

Related to the provision cited above, CIR-3, Article 2(34) also provides the following definition of 'data structure definition' ([63]):

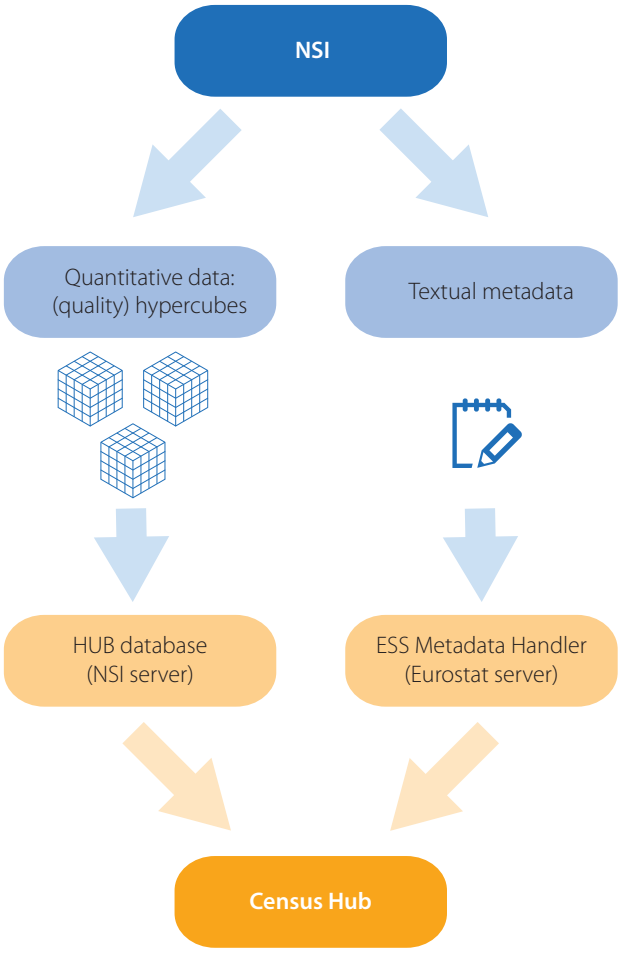
> *'Data structure definition' means a set of structural metadata associated with a data set, which includes information about how topics are associated with the measures, dimensions and attributes of a hypercube, along with breakdowns, information about the representation of data and related descriptive metadata.*

It follows from both provisions that Eurostat must provide a technical system for the harmonised collection of all required information (data and all kinds of metadata). For convenience, this system is divided into two separate channels (see Figure 8):

1. quantitative data and metadata — in other words all (quality) hypercubes — will be collected using the Census Hub;
2. all textual metadata will be collected using Eurostat's standard tool for the collection of national metadata, in other words the *ESS Metadata Handler* (ESS-MH).

([63]) In line with the SDMX Glossary (https://sdmx.org/?sdmx_news=new-sdmx-glossary-available), published in 2016 and replacing the Metadata Common Vocabulary.

Figure 8: Tools to collect and disseminate the data and metadata of the 2021 EU census



Further information on these two channels of metadata transmission is contained in the following two sections of this chapter. Moreover, Annex C provides a synoptic overview of how each individual piece of metadata required by the 2021 EU census legislation will be collected using one of the channels listed above. In particular:

- Table C1 focuses on the provisions of Annex II to CIR-2 (metadata on topics);
- Table C2 focuses on the provisions of the annex to CIR-3 (census background information and quality reporting); and
- Table C3 focuses on those provisions of Annex III to CIR-4 (metadata for the 1 km^2 grid).

6.1. Transmission of quantitative data and metadata — Census Hub

The Census Hub is the central access point provided by Eurostat for all EU census information. To comply with CIR-3, Article 6, Member States must expose the following quantitative datasets (containing numeric values) to the Census Hub:

- all 119 hypercubes listed in Annex I to CIR-2 (see Section 3.3 and Annex B);
- the two quality hypercubes QHC1 and QHC2 required under item 4.4 of the annex to CIR-3 (see Section 5.6.4 and Annex D);
- the 1 km^2 grid dataset set out in Annex II to CIR-4 (see Section 4.3).

Eurostat provides *SDMX Reference Infrastructure* (SDMX-RI) ([64]) software to all reporting national statistical offices, which should be used to map the required quantitative census information (in other words tables) onto the correct

([64]) http://ec.europa.eu/eurostat/web/sdmx-infospace/sdmx-it-tools/sdmx-ri.

SDMX data structure, and expose it to the Census Hub through a dedicated web service. The deadline for the transmission of all — data and quality — hypercubes is the **end of March 2024**, as stipulated by the 2021 EU census legislation. There is one notable exception to this deadline: CIR-4 requires transmission of 1 km^2 grid data and metadata on 'total population' by the **end of December 2022** (see Section 4.6).

As explained in Section 3.3, there are significant methodological and technical obstructions limiting the comparability between 2011 and 2021 hypercubes, which cannot be easily alleviated. Therefore, the 2021 Census Hub version will not offer functionalities allowing concurrent (time series) extractions of combined 2011 and 2021 data.

6.2. Transmission of textual metadata — ESS Metadata Handler

In addition to the quantitative data and metadata discussed in Section 6.1, the 2021 EU census legislation provides for a substantial amount of information that must be transmitted in textual form, both as census background information and in the context of quality reporting.

In addition, CIR-3, Article 6 explicitly states that '*Member States shall transmit the required data in line with the data structure definitions and related technical specifications provided by the Commission (Eurostat).*' This implies that Eurostat must provide these data structure definitions and related technical specifications for the transmission. For the textual metadata, Eurostat does this through the *Euro-SDMX Metadata Structure* (ESMS) ([65]). The ESMS is a standard for metadata transmission based on SDMX, which is in use since 2009. It was established in Regulation (EC) No 223/2009 and later revised in 2016 (version 2.0). According to the ESMS, metadata are arranged in 'concepts', which are in turn grouped into a set of 19 'areas' (for reference, see Table E1).

Nevertheless, the 2021 EU census legislation contains several provisions on textual metadata that are very specific to the census domain, and which therefore do not fit exactly into any of the ESMS v2.0 concepts. For this reason, Eurostat decided to amend slightly the ESMS for the purpose of the census. Some of the generic ESMS concepts were expanded — if needed — into multiple sub-concepts (highlighted in Table E1) to achieve optimal correspondence between the requirements of the 2021 EU census legislation and the resulting 'census extended ESMS'. A detailed summary of the complete content of the 'census extended ESMS' is contained in Annex E to this publication. Moreover, Table C1, Table C2 and Table C3 in Annex C provide the link between each specific requirement of the 2021 EU census legislation and the corresponding item(s) in the 'census extended ESMS'.

The software tool that Member States must use to transmit the ESMS metadata is the **ESS Metadata Handler** (ESS-MH) ([66]). The ESS-MH is a web application that allows Member States to input, store and transmit metadata according to the ESMS. A screenshot of the ESS-MH is displayed in Figure 9, while a comprehensive guide and link to the software is available at:

https://ec.europa.eu/eurostat/web/sdmx-infospace/sdmx-it-tools/ess-mh.

Eurostat has established a list of national local coordinators for all ESS-MH-related issues with national statistical offices. This list can be requested from ESTAT-METADATA@ec.europa.eu.

As illustrated in Figure 9, the ESS-MH presents, for each of the ESMS items present in the metadata transmission programme, a **text box** (on the right-hand side of the screen) where the data provider can type the requested textual information. Some of the ESS-MH entries are mandatory and must be filled in, while others are voluntary or not applicable and can be left empty. Some entries contain **text prefilled by Eurostat** that must remain **the same for all countries** (for example, the list of statistical units included in the census). Some other entries contain text prefilled by Eurostat that could potentially **be modified by Member States** if needed (for example, the exact definitions applied for certain census topics).

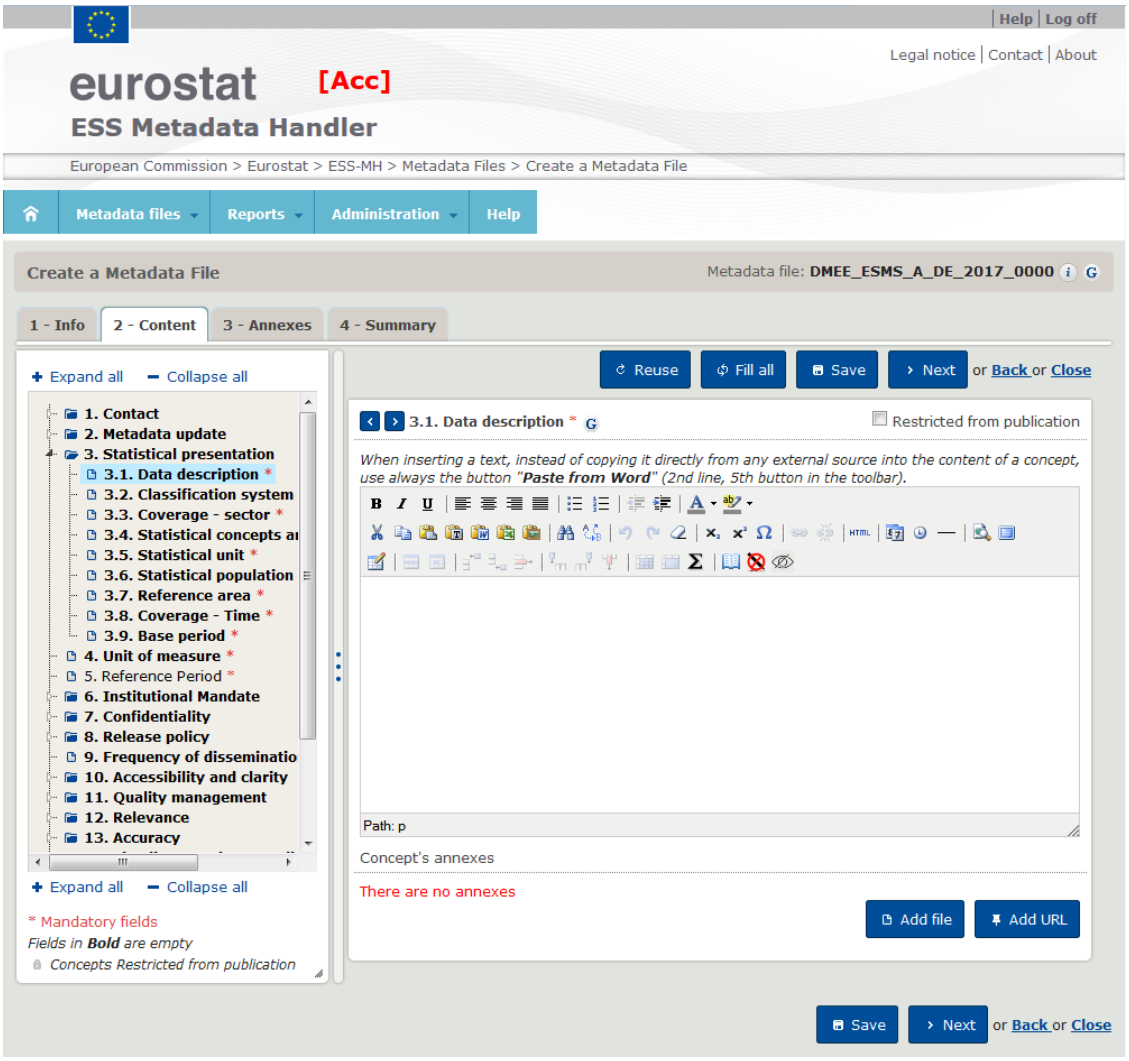
The rightmost column of Table E1 'Legal basis/notes' provides some additional legal information. However, each item in the ESS-MH itself also has an associated 'G' button (next to the metadata file name in the top-right of Figure 9). This button will show a text window providing '**G**uidelines and definitions' on how to fill this specific metadata entry,

([65]) http://ec.europa.eu/eurostat/data/metadata/metadata-structure.

([66]) Formerly called the *National Reference Metadata Editor* (NRME). It is in use by Eurostat since 2010 to collect national metadata in several statistical domains.

covering both content and style (for example the recommended length of the text to be provided). This 'G' pop-up window will be used by Eurostat, when setting up the ESS-MH for census metadata, to provide a set of information notes to help Member States in the correct interpretation of each of the metadata entries listed in Table E1.

Figure 9: Screenshot of the ESS Metadata Handler (ESS-MH)



Member States can choose to provide census textual metadata in their **national language**, rather than in English. In this case, the workflow will include the following steps:

(1) metadata are provided in national language by the Member State;
(2) metadata are translated into English by Eurostat;
(3) the Member State verifies and approves the proposed translation into English;
(4) metadata are published in English through the Census Hub.

Finally, the deadline for the transmission of the ESMS metadata is the **end of March 2024**, as provided for by the 2021 EU census legislation ([67]).

([67]) With some exceptions: for instance, for data revisions occurring after the end of March 2024, the metadata item '17.2 Data revision — practice' can be edited when new data revisions are made. However, CIR-4 requires transmission of grid data and metadata on 'total population' as early as the **end of December 2022** (see Section 4.6).

Annexes

A Additional remarks on the topics and breakdowns in Regulation (EU) 2017/543

This annex contains a list of all topics and breakdowns covered in CIR-1 (and CIR-4), where additional specific guidelines on their definition and application may be helpful based on the 2011 experience. Specific topic-related metadata required in Annex II to CIR-2 are also mentioned where applicable. In addition to general remarks and explanations, each topic may contain dedicated paragraphs with:

- a list of *changes with respect to 2011*;
- *further examples* based on questions from the 2011 round.

When specific codes for the breakdown categories (as defined in CIR-1) are mentioned explicitly, they will always be given with the lowest-detail breakdown that includes the category. The inclusion of that category in all higher-detail breakdowns is then implicitly understood, as explained in Section 2.3.1. For instance, the age group 'under 15 years' would be referenced with the code AGE.L.1., where the existence of the codes AGE.M.1. and AGE.H.1. is implicit.

Finally, note that this annex is not intended to simply repeat or rephrase sufficiently clear definitions and specifications contained in CIR-1. The idea is rather to add clarifications and/or rephrase CIR-1 provisions whenever needed to make the intended interpretation clearer. As in the previous sections of this publication, verbatim citations of original legal provisions are *set in italic*, and used specifically to support certain interpretations of specifications or examples. There are also many more cross-references in this annex than in the original annex to CIR-1. This may make the work with various concepts used across a multitude of topics more efficient (see also the concept glossary in Table A1 at the end of this annex). Thus, the idea is to use this annex as a supplement to the original CIR-1 text when applying the specifications.

Topic: Place of usual residence/geographical area (GEO)

The EU census legislation follows the CES Recommendations (§393) on the definition of 'place of usual residence'. More specifically, CFR, Article 2(d) states:

> *The following persons alone shall be considered to be usual residents of the geographical area in question:*
>
> *i) those who have lived in their place of usual residence for a continuous period of at least 12 months before the reference date; or*
>
> *ii) those who arrived in their place of usual residence during the 12 months before the reference date with the intention of staying there for at least one year.*

In addition, Article 2(d) states that, where these circumstances for the usual residence *'cannot be established, 'usual residence' shall mean the place of legal or registered residence'*. Priority is thus given to the concept of usual residence as the place where a person normally spends the daily period of rest over the other two concepts (legal residence or registered residence). In any case, Annex II to CIR-2 requires topical metadata explaining in which way the definition of 'usual residence' has been applied, in particular to what extent the preferred concept based on the 12-month criterion had to be replaced by the legal or registered residence concept.

The annex to CIR-1 further specifies the interpretation of this definition for various special cases (please see items (a)-(m) of the topic specifications). Furthermore, Annex II to CIR-2 requires topical metadata on any country-specific application of these special cases.

Finally, note that the breakdowns of this topic can be applied to *any* type of statistical unit that is used in the 2021 census programme. The breakdown categories are to be interpreted as the 'place of usual residence' in the case of persons or family nuclei, or alternatively as the 'geographical area (in which the unit is located)' for any other kind of statistical unit (such as dwellings, living quarters, and so on).

Changes with respect to 2011

- New special cases (h)-(m) were added to give additional interpretation guidance, namely:
  (h) merchant seamen and fishermen;
  (i) persons irregularly staying or undocumented/asylum seekers/refugees or applicants;
  (j) children born in the 12 months before the census;
  (k) persons staying exactly 1 year;
  (l) foreign military, naval and diplomatic personnel located in the country;
  (m) national military, naval and diplomatic personnel located outside of the country.
- Annex I to CIR-4 introduces a new breakdown GEO.G., which represents the European 1 km^2 reference grid 'Grid_ETRS89-LAEA_1000' (CRS:3035) ([68]) for the dissemination of selected 2021 census topics on a grid under CIR-4. The correct code format for the GEO.G. breakdown categories ('grid cell identifiers') is regulated in Article 3(2-3) of CIR-4 (see detailed explanation, including an example, in Section 4.2.2). An explanation of the special code GEO.G.y. (virtual grid cell for persons with their usual residence in the reporting country, but who cannot be allocated to any geographic grid cell) is given in Section 4.3.2.

Further examples

1. *Persons other than usually resident can be enumerated* (coverage will be larger than the usually resident population): Enumeration itself does not imply the inclusion in the total usually resident population, and countries may have national interests in specific categories of persons. However, it is important that the overall quality of the enumeration of the usually resident population is not affected by the extension of the exercise to other categories. It is also of paramount importance that the total usually resident population includes only the categories listed in the relevant EU regulations. Countries extending the census coverage should therefore put in place all the necessary measures to avoid mistakes in the computations and to minimise the negative effects of the additional surveys. It is recommended that the information collected on non-usual residents be kept to a minimum.
2. *International officials:* Unlike foreign diplomats, naval personnel and military forces, who — according to the CES Recommendations §176 — are not part of the total usually resident population of a country, international officials must be included in the total usually resident population, provided that they meet the requirements on the place of usual residence (12 months criterion).
3. *Tertiary students studying abroad but returning home every weekend:* Special case (c) of the topic specifications in CIR-1 sets out that these students must *'consider their term-time address to be their place of usual residence'*, with the only possible exception where *'the place of education is within the country'* (see example 5). This is in line with the CES Recommendations. Therefore, tertiary students studying abroad should always be classified in the country where they are studying, in line with item (c) of the specifications and irrespective of the treatment of workers abroad in item (a).
4. *Tertiary students studying within the country and returning home every weekend:* Special case (c) of the topic specifications in CIR-1 states that the place of usual residence of such students may exceptionally be their family home. Such exceptions may be defined within the reporting country, but in any case Annex II to CIR-2 requires that the applied rule be reported in the metadata on this topic.
5. *Persons who are temporarily present somewhere other than their own place of usual residence:* The topic specifications in CIR-1 state that *'persons who are enumerated but do not meet the criteria for usual residence in the place of enumeration […] are considered temporarily present and are therefore not counted in the total usual resident population'*. This sentence obviously refers to the place of enumeration and not to the entire country. It may therefore be read as follows:

   Persons who are enumerated but do not meet the criteria for usual residence in the place of enumeration […] are considered temporarily present and are therefore not counted in the total usual resident population [of the place of enumeration].

([68]) http://spatialreference.org/ref/epsg/etrs89-etrs-laea/.

Of course, the persons who are temporarily present in the place of enumeration must be included in the total (national) usual resident population only if they are usually resident in some other part of the country; otherwise they are indeed excluded.
The same logic applies to another specification of this topic: *'persons living or expected to live outside the place of enumeration for one year or more shall not be considered temporarily absent and shall therefore be excluded from the total population'*. In this case, total population also refers to the place of enumeration, and thus the provision should be read as: *'… and shall therefore be excluded from the total population [of the place of enumeration]'*.
In conclusion, the place of enumeration can indeed be different from the place of usual residence. Inclusion in — or exclusion from — the total population depends strictly on the question if the place of usual residence is in any part of the reporting country.

6. *Persons in military service for a period of less than 1 year:* These persons do not comply with (i) nor with (ii) of the definition of 'place of usual residence' cited above. Therefore, their place of usual residence remains their home address.
7. *National population registers with a time criterion of less than 1 year (for example 3 or 6 months):* The CFR leaves to each country the choice of the methodology that best suits its national census, provided that the essential features of the census are met. Therefore, the choice of using the national population register for the census purposes is fully legitimate.
   As explained above, the CFR provides a definition of 'place of usual residence' for the EU censuses, to which CIR-1 refers explicitly. However, as also mentioned above, only if this definition cannot be applied, CIR-1 allows for two alternative concepts (legal residence and registered residence). If a population register is used in combination with a traditional census survey, it should be possible to identify the place of usual residence. If there are such technical difficulties that the usual residence cannot be identified, then one of the other two concepts can be adopted. However, in accordance with Annex II to CIR-2, explanations should be provided in the metadata to be transmitted.

Topic: Location of place of work (LPW)

If in doubt, countries should take into account the CES Recommendations (§441), and in particular the recommendation that *'persons who do not have a fixed place of work but who report to a fixed address at the beginning of their work period (for example bus drivers, airline pilots and stewards, operators of street market stalls which are not removed at the end of the workday) should provide information on this address'*. This recommendation was also included in CIR-1.

Note that, although the topic breakdowns do not contain a category 'Not stated', the two new categories 'Unknown place of work (unknown if inside or outside the Member State)' (LPW.N.4.) and 'Unknown place of work in the Member State' (LPW.L.1.y.) essentially have the same effect. Unknown information does not need to be imputed for every person; rather the special arrangement makes it possible to include the partial information that the unknown work location is inside the reporting country (see also example 2).

Changes with respect to 2011

- Inclusion of the CES Recommendations (§441) into the topic specifications.
- Inclusion of three new breakdown categories 'No fixed place of work (inside or outside the Member State)' (LPW.N.3.), 'Unknown place of work (unknown if inside or outside the Member State)' (LPW.N.4.) and 'Unknown place of work in the Member State' (LPW.L.1.y.) to simplify the allocation of individuals (imputation not always necessary).

Further examples

1. *Individuals with no fixed place of work, for example construction workers or travelling salesmen:* As explained in CIR-1, the correct allocation depends on whether a fixed address exists to which these persons report at the beginning of their work period. Only if this is not the case, should they be allocated to 'No fixed place of work' (LPW.N.3.). For instance, if a construction worker is employed by a company on whose premises he/she starts the shift, the address of these premises is the correct 'Location of place of work'. On the other hand, if he/she starts the work periods in changing places, then the category 'No fixed place of work' applies.
2. *Individuals having left the place of work entry blank or entered a place that could not be coded:* The correct allocation depends whether the work location is known to be inside the reporting country. If yes, they may be included in the new category 'Unknown place of work in the Member State', otherwise the category 'Unknown place of work (unknown if inside or outside the Member State)' applies.

Topic: Locality (LOC)

CIR-1 defines the following concepts needed to specify the breakdown LOC. of this topic:

- 'locality'
- 'population of a locality'.

In addition, CIR-1 sets out that any scattered building that cannot be assigned to a locality is to be treated like a locality itself (although failing the formal definition). This means that it is to be allocated to a LOC. category according to the number of its usual residents.

Like the breakdowns of 'place of usual residence' (GEO, see above), the breakdown LOC. of this topic can be applied to any type of statistical unit used in the 2021 census programme.

Topic: Sex (SEX)

The information on this topic must be collected for every person in the total population (there is no category 'Not stated' or similar). If the information is missing, the most appropriate imputation method should be used in the reporting country to impute it.

Topic: Age (AGE)

The information on this topic must be collected for every person in the total population (there is no category 'Not stated' or similar). If the information is missing, the most appropriate imputation method should be used in the reporting country to impute it.

Changes with respect to 2011

CIR-4 adds a new age breakdown for the 1 km^2 grid, which is even less detailed than AGE.L. established in CIR-1. The new breakdown AGE.G. contains only three categories: 'Under 15 years' (AGE.G.1.), '15 to 64 years' (AGE.G.2.) and '65 years and over' (AGE.G.3).

Topic: Legal marital status (LMS)

For the 2021 census round, CIR-1 sets out that the concepts of marriage and registered partnership must be treated as equivalent statuses. These equivalent statuses are characterised through the existence of a legal framework within the reporting country, which stipulates that the two partners:

(h) commit themselves to legal conjugal obligations; and
(i) cannot commit themselves into multiple partnerships with different partners.

The equivalence of marriage and registered partnership means in particular that no distinction between the two is required for the breakdowns. The categories of the topic breakdowns LMS.L. and LMS.H. have been updated to reflect this (see specific changes).

The correct classification under the new categories (in other words the distinction between LMS.L.1. 'never married and never in a registered partnership' and either of LMS.L.2., LMS.L.3, or LMS.L.4) should be made based on the relevant national legal basis in the reporting country. However, note that CIR-2 requires in its Annex II that the metadata on this topic should provide information on the relevant legal basis in the reporting country (for example concerning opposite-sex/same-sex marriages, minimum age for marriages, registered partnerships, or divorce).

If same-sex partnerships are not legally recognised in the reporting country, the respective categories can be reported as '0'.

Changes with respect to 2011

- For the 2021 census round, the distinction between the concepts 'marriage' and 'registered partnership' has been lifted. This means that marriages and registered partnerships are treated as equivalent statuses in all breakdown categories included in CIR-1 (a redesign of all categories compared to 2011 legislation).
- In line with the general decision to abolish optional categories to simplify the reporting, there are no optional categories under this topic. In particular, the categories 'In an opposite-sex marriage or registered partnership' (LMS.H.2.1.) and 'In a same-sex marriage or registered partnership' (LMS.H.2.2.) are now compulsory. (In countries where same-sex partnerships are not legally recognised, '0' can be reported in the respective category.)
- There now two distinct breakdowns, LMS.L. and LMS.H., where only the latter includes the distinction between opposite-sex/same-sex partnerships.

Topic: Current activity status (CAS)

CIR-1 defines the following concepts used for the breakdown categories:

- 'current activity status'
- 'labour force'
- 'employed'
- 'unemployed'
- 'formal job attachment'
- 'student'.

CIR-1 further specifies the correct classification of 'self-employed persons' (as defined in the CES Recommendations §541) and 'contributing family workers' (for a definition, see topic 'status in employment' in CIR-1).

For the sake of uniqueness, CIR-1 also specifies an order of priority, both for the assignment of a single activity status ('employed' before 'unemployed' before 'outside of the labour force'), and for the categories breaking down 'outside of the labour force' (CAS.L.2.). As a general rule, the categories in the breakdowns are *strictly* ordered by priority. This means that, in case of doubt, the topmost applicable category in the list is *always* the correct one (see also examples below).

The definitions applied for the 2021 EU census are in line with the ILO definitions ([69]) and CES Recommendations (§496ff.). However, the difficulties of collecting information based on these definitions using different types of census — and the resulting potential incomparability between countries — are well-known and cannot readily be solved. At the least, CIR-2 requires in its Annex II specific metadata information on country-specific rules/methods for this topic (register definitions, minimum age for economic activity, persons with more than one job, and so on).

Changes with respect to 2011

- The concept 'currently economically active population' has been renamed to 'labour force' (but the same definition has been kept).
- The correct classification of 'contributing family workers' has been elaborated on in CIR-1 (and it is slightly different from the 2011 situation). Contributing family workers must be assigned to the 'employed' category (CAS.L.1.1.) if they fulfil the conditions on working time within the reference period for the 'employed'. The category 'employed' (CAS.L.1.1.) thus becomes less restrictive than in 2011.
- The sub-categories of the category 'unemployed' (CAS.L.1.2.) have been abolished.
- The distinction between 'homemakers' and 'others' has been lifted. This means that the collective category has been renamed to 'others' (CAS.H.2.4.), and its optional sub-categories have been abolished.
- CIR-1 now clarifies that the new category 'others' (CAS.H.2.4.) must be applicable to all persons 'outside of the labour force' (CAS.L.2.) to which none of the foregoing categories (CAS.H.2.1., CAS.H.2.2., CAS.H.2.3.) applies. This is consistent with the general priority rule explained above (see also examples).

([69]) http://www.ilo.org/global/statistics-and-databases/statistics-overview-and-topics/lang--en/index.htm.

Further examples

1. *Recipients of pensions or capital income:* In line with the general priority rule stated above, persons receiving a pension or capital income must be allocated to the respective category CAS.H.2.2. only if they are of — or above — the national minimum age for economic activity, in other words if CAS.H.2.1. does not apply.
2. *Students:* In line with the general priority rule stated above, students must be allocated to the respective category CAS.H.2.3. only if they are of — or above — the national minimum age for economic activity and do not receive any pension or capital income, in other words if neither CAS.H.2.1. nor CAS.H.2.2. apply.
3. *Unemployed persons:* In line with the general priority rule stated above, persons can only be 'unemployed' (CAS.L.1.2.) if they fulfil all three defining criteria given in CIR-1, namely they are 'without work', 'currently available for work' and 'seeking work'. If any of these criteria is not satisfied (and the person does not fulfil the definition of 'employed', of course), that person must be allocated to the correct category under 'outside of the labour force' (CAS.L.2.).

Topic: Occupation (OCC)

This topic collects information on the type of work done for a job, in other words information describing the main tasks and duties of the work. The sole breakdown OCC. of this topic in CIR-1 follows the ISCO ([70]) classification. However, CIR-1 also states that the classification in force on 1 January 2021 must be used if it deviates from the one on which the table in CIR-1 is based.

If a person has more than one job, the main job must be identified, according to CIR-1, using the following criteria:

(1) the time spent on the job or, if not available;
(2) the income received.

If a person has several jobs, CIR-1 requires the *same* (main) job to be used for the allocation to the topics 'occupation' (OCC), 'industry' (IND, see below) and 'status in employment' (SIE, see below). Furthermore, Annex II to CIR-2 requires information in the topical metadata on the way and extent to which these criteria were applied to identify the 'main job' in the reporting country.

The correct allocation of a person to their 'occupation' is closely aligned with the category 'employed' (CAS.L.1.1.) of the topic 'current activity status' (CAS, see above):

- all 'employed' persons (CAS.L.1.1.) must be allocated to exactly one of the 'occupation' breakdown categories OCC.1.-11.;
- all persons who are not 'employed', but whose stated current activity status is CAS.L.1.2. or CAS.L.2. should be allocated to 'not applicable' (OCC.12.);
- the 'occupation' of persons whose current activity status is 'not stated' (CAS.L.3.) should equally be 'not stated' (OCC.11.).

Changes with respect to 2011

Compared to 2011, the reporting on this topic has been simplified and aligned with the 'employed' concept of the topic 'current activity status' (CAS, see above).

Topic: Industry (IND)

The specifications for this topic on a person's job are aligned with the specifications for the topics 'occupation' (OCC) and 'status in employment' (SIE). This means in particular that the 'main job' (of a person with several jobs) identified for the purpose of allocating an OCC. category (see in OCC above) must also be used for this topic (and for SIE).

([70]) http://www.ilo.org/public/english/bureau/stat/isco/.

CIR-1 also states that the NACE ([71]) classification in force on 1 January 2021 must be used if it deviates from the one used for the categories IND.H.1.-10. as listed in CIR-1.

The correct allocation of a person to their 'industry' follows the same logic as in the case of 'occupation':

- 'employed' persons (CAS.L.1.1.) must be allocated to exactly one of the 'industry' breakdown categories IND.H.1.-11.;
- persons who are not 'employed', but whose current activity status is stated as CAS.L.1.2. or CAS.L.2., should be allocated to 'not applicable' (IND.L.12.);
- the 'industry' of persons whose current activity status is 'not stated' (CAS.L.3.) should equally be 'not stated' (IND.L.11.).

Changes with respect to 2011

For the topics 'occupation' (OCC, see above) and 'status in employment' (SIE, see below), the reporting on this topic has been simplified and aligned with the 'employed' concept of the topic 'current activity status' (CAS, see above).

Topic: Status in employment (SIE)

The specifications for this topic on a person's job are aligned with the specifications for 'occupation' (OCC) and 'industry' (IND). This means in particular that the 'main job' (of a person with several jobs) identified for the purpose of allocating an OCC. category (see above in OCC) must also be used for this topic (and for IND).

Furthermore, CIR-1 defines the following concepts used for the breakdown categories:

- 'employee'
- 'paid employment'
- 'employer'
- 'own-account worker'
- 'contributing family worker'
- 'member of a producers' cooperative'.

The correct allocation of a person to their 'status in employment' follows the same logic as in the case of 'occupation':

- 'employed' persons (CAS.L.1.1.) must be allocated to exactly one of the 'status in employment' breakdown categories SIE.1.-5.;
- persons who are not 'employed' but whose current activity status is stated as CAS.L.1.2. or CAS.L.2. should be allocated to 'not applicable' (SIE.6.);
- the 'status in employment' of persons whose current activity status is 'not stated' (CAS.L.3.) should equally be 'not stated' (SIE.5.).

In addition, CIR-1 sets out that 'contributing family workers' and 'members of a producers' cooperative' must be allocated to 'other members of the labour force' (SIE.4.).

Finally, if a person is found to be both 'employer' and 'employee', he/she must be allocated to one of the two groups, according to CIR-1 using the following criteria:

(1) the time spent on the job or, if not available;
(2) the income received.

This is consistent with the identification criteria for a person's 'main job' (see above in OCC). In any case, CIR-2 requires topical metadata on the country-specific application of the rules and specifications laid down in CIR-1 on this topic.

([71]) http://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF.

Changes with respect to 2011

- The reporting on the topics 'occupation' (OCC, see above) and 'industry' (IND, see above) has been simplified and aligned with the 'employed' concept of the topic 'current activity status' (CAS, see above).
- The 2011 breakdown category 'Others ('Contributing family workers' and 'Members of producers' cooperatives')' has been renamed to 'Other members of the labour force' (SIE.4.), while the category definition remains the same.
- In line with the general decision to abolish optional categories to simplify the reporting, there are no optional categories anymore under this topic, in other words a distinction between 'contributing family workers' and 'members of a producers' cooperative' is not required anymore.

Topic: Educational attainment (EDU)

The sole breakdown EDU. of this topic in CIR-1 follows the ISCED 2011 ([72]) classification. However, CIR-1 also states that the classification in force on 1 January 2021 must be used if it deviates from ISCED 2011.

Changes with respect to 2011

The categories of the breakdown EDU. were updated from ISCED 1997 to the ISCED 2011 classification.

Topic: Country/place of birth (POB)

CIR-1 sets out that *'Information on the 'Place of birth' shall be collected according to the place of usual residence of the mother at the time of the birth, or, if not available, the place in which the birth took place'*. However, if the latter concept is used, CIR-2 requires respective information to be added to the metadata on this topic.

Moreover, if no information — or only incomplete information — is available based on the international boundaries at the time of the census, CIR-2 requires information on the methods used in the reporting country to allocate affected persons.

Changes with respect to 2011

- Croatia (POB.H.2.1.11.) is now listed under 'Other EU Member State' (POB.L.2.1.).
- Turkey (POB.H.2.2.1.17.) is now listed under 'Other European country' (POB.M.2.2.1.). This is in line with the standard geographical classification applied by Eurostat, as Turkey is an EU candidate country.
- Saint Martin is replaced by Saint Martin (FR) (POB.H.2.2.3.37.) and St Maarten (NL) (POB.H.2.2.3.38.).
- Bermuda (POB.H.2.2.4.04.) and Saint Pierre and Miquelon (POB.H.2.2.4.05.) are now listed under 'North America' (POB.M.2.2.4).
- French Southern Territories (POB.H.2.2.6.06.) is now listed under 'Oceania' (POB.M.2.2.6.).
- Added to the list of categories under the breakdown POB.H.: South Sudan (POB.H.2.2.2.48.), Western Sahara (POB.H.2.2.2.54.), Curaçao (POB.H.2.2.3.16.), Palestine (POB.H.2.2.5.46.), and Cook Islands (NZ) (POB.H.2.2.6.03.).
- Removed from the list of categories under the breakdown POB.H.: Mayotte and Netherlands Antilles.
- The special categories 'Other country in [*continent*]' (POB.H.2.2.1.26., POB.H.2.2.2.57., POB.H.2.2.3.45., POB.H.2.2.4.06., POB.H.2.2.5.48., POB.H.2.2.6.21.) were added to the breakdown POB.H.
- The breakdown category 'Information not classifiable according to current borders (optional)' and its sub-categories were dropped from the category 'Other' (POB.L.3.). This is in line with the general approach to have no more optional categories for 2021.

On the relation between the new special categories 'Other country in [*continent*]' and the generic category 'Other' (POB.L.3.): these categories must be used such that the reported data reflect all available information on the place of birth in the most accurate manner. For instance, if the place of birth of a person can be inferred to lie within the geographic boundaries of Europe, but cannot be allocated to any of the categories POB.H.2.1.X. or POB.H.2.2.1.1.-25., this person must be allocated to 'Other country in Europe' (POB.H.2.2.1.26). The category 'Other' (POB.L.3.) only applies to persons where the stated information does not allow an unambiguous allocation to any continent (see examples below). Only if no information was stated at all should the person be allocated to 'Not stated' (POB.L.4.).

([72]) http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf.

Further examples

In 2011, the category 'Other' contained optional sub-categories to account for *information not classifiable according to current borders*, such as former entities like the Soviet Union or Czechoslovakia. For 2021, these optional sub-categories were dropped, but persons must nonetheless be allocated to the correct continent if possible. For example, while a person born in the Soviet Union (with no further information available) would still have to be allocated to 'Other' (POB.L.3.), persons born in Czechoslovakia or Yugoslavia should be reported under 'Other country in Europe' (POB.H.2.2.1.26.). The same reasoning applies to any other former entity at the date 1 January 2021 that can be uniquely related to a single continent.

Topic: Country of citizenship (COC)

There is a comprehensive list of countries of citizenship included in the EU regulation, and priority criteria are given for the classification of a person holding more than one citizenship. In particular, a person with two or more citizenships must be allocated to only one country of citizenship, to be determined in the following order of precedence:

1. reporting country; or
2. if the person does not have the citizenship of the reporting country: other EU Member State; or
3. if the person does not have the citizenship of another EU Member State: other country outside the EU.

In cases of dual citizenship where both are EU countries but neither is the reporting country, Member States must determine which country of citizenship is to be allocated. Although not specified in the legislation, this approach, whereby the reporting Member State must determine which citizenship is allocated, must also apply where a person has dual citizenship of more than one non-EU country.

The list of categories included in CIR-1 under this topic is intended only for statistical purposes. It does not imply the expression of any opinion whatsoever on the part of the European Commission concerning the legal status, the authorities, or the delimitation of the frontiers or boundaries of any country, territory, city or area.

Whenever the special category 'recognised non-citizens' (COC.H.2.2.1.20.) is applicable to one or more persons in the total population of the reporting country, CIR-2 requires relevant information on this in the metadata on the topic.

Changes with respect to 2011

- CIR-1 now clarifies explicitly that the category COC.H.2.1.XX. referring to the reporting country itself under the category 'Citizenship not of reporting country, but other EU Member State' (COC.L.2.1.) does not apply to the reporting country.
- If the reporting country is not an EU Member State, the category 'Citizenship not of reporting country, but other EU Member State' (COC.L.2.1.) is renamed to 'Citizenship of an EU Member State'.
- Croatia (COC.H.2.1.11.) is now listed under 'Citizenship not of reporting country, but other EU Member State' (COC.L.2.1.).
- Turkey (COC.H.2.2.1.17.) is now listed under 'other European country' (COC.M.2.2.1.). This is in line with the standard geographical classification applied by Eurostat, as Turkey is an EU candidate country.
- Added to the list of categories under the breakdown COC.H.: Kosovo (COC.H.2.2.1.06.) ([73]), South Sudan (COC.H.2.2.2.47.), Curaçao (COC.H.2.2.3.13.) St Maarten (NL) (COC.H.2.2.3.31.), and Palestine (COC.H.2.2.5.46.).
- Removed from the list of categories under the breakdown COC.H.: Anguilla, Bermuda, British Virgin Islands, Cayman Islands, Falkland Islands (Malvinas), French Polynesia, French Southern Territories, Gibraltar, Guernsey, Isle of Man, Jersey, Montserrat, Netherlands Antilles, New Caledonia, Saint Barthelemy, Saint Martin, Saint Pierre and Miquelon, Sark, Turks and Caicos Islands and Wallis and Futuna Islands.
- The special categories 'Other country in [*continent*]' (COC.H.2.2.1.21., COC.H.2.2.2.56., COC.H.2.2.3.37., COC.H.2.2.4.03., COC.H.2.2.5.48., COC.H.2.2.6.15.) were added to the breakdown COC.H (see also the dedicated example on this topic).

([73]) This is without prejudice to positions on status, and is in line with UNSCR 1244/1999 and the ICJ Opinion on the Kosovo declaration of independence.

Further examples

Classification of persons holding only the citizenship of countries not listed in CIR-1: This occurs because those citizenships are the particular legal bond between an individual and an entity which is recognised by some countries as a sovereign state, but not by others. Citizens of those entities residing in a foreign country may be recognised as such or not, depending on the recognition foreign policy of the country of residence. To allow the classification of persons with the citizenship of such territories, CIR-1 includes in the breakdown COC.H. the additional codes 'Other countries in Europe/Africa/Asia/etc.', to be used whenever relevant.

Topic: Ever resided abroad and year of arrival in the country (YAT, YAE)

CIR-1 sets out that the most recent arrival in the reporting country should be considered, rather than the first arrival. The topic therefore does not provide information on interrupted stays.

The topic contains a less-detailed breakdown YAT., which focuses on more recent international migration since 2010, as well as two more-detailed breakdowns YAE.L. and YAE.H., which focus on international migration since 1980.

Changes with respect to 2011

- The title of the breakdown YAT. 'Year of arrival in the country since 2010' (was 2000) has been adapted to the census reference year 2021.
- The categories 'ever resided abroad and arrived in 2010 or after' (YAT.1.) and 'resided abroad and arrived in 2009 or before, or never resided abroad' (YAT.2.) have been adapted to the census reference year 2021.
- To simplify the reporting on this topic, the sub-categories breaking YAE.H. down into single years have been abolished for all years between 1980 and 2004.
- There are no optional sub-categories anymore to further break down 'resided abroad and arrived in 1979 or before, or never resided abroad' (YAE.L.2.), in other words a distinction between 'resided abroad and arrived in 1979 or before' and 'never resided abroad' is not required anymore.

Topic: Place of usual residence one year prior to the census (ROY)

The topic contains a single breakdown ROY. CIR-1 set out several specifications to clarify the correct allocation to the available breakdown categories:

- the reference period for the topic starts exactly one year before the reference date of the census, as reported by the country in compliance with CIR-2, and ends with that reference date;
- children aged less than one year must be allocated to 'not applicable' (ROY.4.);
- if there has been more than one move in the reference period (see above), the most recent move to the current place of usual residence should be considered;
- a move within the same LAU2 area must also be considered as a 'move within the same NUTS 3 area as the current usual residence' (ROY.2.1.1.).

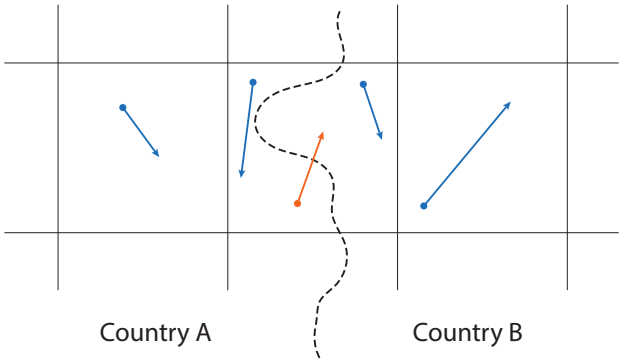
Changes with respect to 2011

To be consistent with other data collections, CIR-1 now clarifies that *'[a] move within the same LAU2 area shall also be considered as a move within the same NUTS 3 area.'*

Moreover, for the sake of internal consistency with this specification, CIR-4 adds the following rule for the reporting on the 1 km^2 grid: *'A move within the same grid cell shall be reported either as a 'Move within the reporting country' (ROY.2.1.) or as a 'Move from outside the reporting country' (ROY.2.2.) as appropriate.'* This means that the correct

category is ROY.2.1. whenever the person moved within the reporting country (regardless of whether the move was within the same grid cell or across different grid cells). However, the code ROY.2.2. is to be used whenever the person moved across a country border into the reporting country (regardless of whether the move was within the same grid cell or across different grid cells) — see Figure 10.

Figure 10: Illustration of the correct reporting of ROY. on the 1 km^2 grid, for the special case where the move happened to be within the same geographical grid cell



The picture sketches a border region between countries A and B (border shown as bold dashed line), where the correct category for all moves indicated by a blue arrow is 'Move within the reporting country' (ROY.2.1.) — the two on the left reported by country A, the two on the right by country B. Only the move shown as a red arrow is to be reported as 'Move from outside the reporting country' (ROY.2.2.) by country B.

Topic: Family status (FST)

The topic contains three breakdowns FST.L., FST.M. and FST.H., where CIR-1 defines the following concepts used for the breakdown categories:

- 'family nucleus'
- 'child (son/daughter)'
- 'partners'
- 'consensual union'.

The rules specified in CIR-1 for these concepts facilitate the correct classification of each person's 'family status'. These rules also form the basis for establishing the total statistical population of 'family nuclei' in the reporting country (see topics 'type/size of family nucleus' below). In particular, for the purpose of allocating a unique breakdown category — and hence establishing the family nuclei in a household — CIR-1 explicitly states:

> *A son or daughter who lives with a spouse, with a registered partner, with a partner in a consensual union, or with one or more own children, is not considered to be a child.*

This entails two simple rules:

1. the concept 'partners' (FST.L.1.) has a higher priority than the concept 'son/daughter' (FST.L.3.);
2. only persons of the youngest generation in a household can be classified as 'sons/daughters';

see examples below.

Furthermore, the concepts 'married' and 'registered partnership' are treated equivalently in the 2021 census round: they are both classified as 'persons in a married couple or registered partnership' (FST.M.1.1.), and no distinction between these concepts is needed.

Changes with respect to 2011

- The concepts 'married' and 'registered partnership' have been merged into the category 'persons in a married couple or registered partnership' (FST.M.1.1.).
- For a 'child (son/daughter)' who spends equal amounts of time with each of his/her separated parents, the legal or registered residence of the child may now also be used to identify a unique household/family nucleus where this child is to be enumerated.
- In line with the general decision to abolish optional categories to simplify the reporting, there are no optional categories anymore in this topic. In particular:
  - the sub-categories 'persons in an opposite-sex marriage or registered partnership' (FST.H.1.1.1.) and 'persons in a same-sex marriage or registered partnership' (FST.H.1.1.2.) are now compulsory (in countries where same-sex partnerships are not legally recognised, '0' can be reported in the respective category.);
  - the distinction between opposite-sex partners and same-sex partners has been abolished for 'partners in a consensual union' (FST.M.1.2.);
  - the sub-categories 'not of lone parent' (FST.M.3.1.) and 'of lone parent' (FST.M.3.2.) breaking down 'sons/daughters' (FST.L.3.) are now compulsory.

Further examples

1. *A lone mother living with her children and her sister:* The persons composing the household cannot be classified as belonging to the same family nucleus. The relevant definition is in CIR-1 under the topic 'family status' (FST):

   [. . .] a family comprises a couple without children or a couple with one or more children, or a lone parent with one or more children. This family concept limits relationships between children and adults to direct (first-degree) relationships, that is, between parents and children.

   This means that the lone mother and her sister cannot be considered as being members of the same family nucleus, and each person must be classified as follows: the mother as 'lone parent' (FST.L.2.), her children as 'sons/daughters of lone parent' (FST.M.3.2.), and her sister as 'not applicable — not in a family nucleus' (FST.L.5.).

2. *A married couple living in the same household with their child and his/her consensual union partner:* Both the child and the child's partner must be classified as a 'partners in a consensual union' (FST.M.1.2.) in accordance with the priority rule stated above, while the married parents must be classified as 'persons in an opposite-sex married couple or registered partnership' (FST.H.1.1.1.).
3. However, consider the case of *one parent living in the same household with their child and his/her consensual union partner:* While nothing changes for the child and his/her partner in this situation, the parent must not be classified as 'lone parent', but rather as 'not applicable — not in a family nucleus' (FST.L.5.). This may seem illogical, as the parent actually lives with a resident child. However, this child lives with his/her partner, and therefore must not be considered a child (see rule 1 above). The parent must therefore not be considered a parent in the sense of the census definition (see CIR-1).
4. Finally, consider the case of a *three-generation household consisting of a mother living with her children and her father:* According to rule 2 above, only the children (youngest generation) are to be classified as 'sons/daughters' (FST.L.3.), while their mother must be classified as 'lone parent' (FST.L.2). The mother is thus not a child in the sense of CIR-1, and therefore her father in turn cannot be a 'lone parent' — he is instead classified as 'not applicable — not in a family nucleus' (FST.L.5.).

Note that, as explained above, the specifications of this topic form the basis for the correct identification of 'family nuclei'. They will therefore be reused below to further elaborate the consequences of the family nucleus assignment whenever relevant (for the topics 'type of family nucleus', 'household status' and 'type of private household').

Topic: Type of family nucleus (TFN)

The CIR-1 definition of a 'family nucleus' given under the topic 'family status' (FST, see above) must be applied to establish the total statistical population of 'family nuclei' in the reporting country, which is broken down by the breakdowns TFN.L. and TFN.H. included in this topic. Moreover, just like the topic 'legal marital status' (LMS, see above), this topic now treats the concepts 'married partnership' and 'registered partnership' equivalently, in other words no more distinction is required between the two.

The information on this topic must be collected for every family nucleus in the data source (there is no category 'not stated' or similar). If the information is missing, the most appropriate method should be used in the reporting country to impute it. However, a simple imputation method might suffice in this circumstance. Often what the users are interested in is the share of 'lone parent' families as given by the sum of 'lone father families' (TFN.L.3.) and 'lone mother families' (TFN.L.4.). Therefore, the imputation of the gender of the lone parent will probably not lead to a significant distortion of the message in the data.

As already explained in the topic 'legal marital status' (LMS, see above), in countries where same-sex couples are not legally recognised the respective categories can be reported as '0'.

Changes with respect to 2011

- The two categories 'married couple families' and 'registered partnership couple families' (and all respective sub-categories) have been merged into a single category 'married or registered partnership couple families' (TFN.L.1.) and respective sub-categories.
- The distinction between 'opposite-sex couple families' and 'same-sex couple families' has been abolished for all types of 'consensual union couple families' (TFN.L.2.).

In line with the general decision to abolish optional categories to simplify the reporting, all categories remaining in the breakdown TFN.H. are now compulsory.

Further examples

The examples below revisit those from 'family status' (FST, see above).

1. *A lone mother living with her children and her sister:* As explained in FST example 1, the persons composing the household cannot be classified as belonging to the same family nucleus. While the mother and her children form a family nucleus 'lone mother family' (TFN.L.4.), the sister must not be included in the underlying statistical population of 'family nuclei' broken down here. This is because she does not constitute a family nucleus in the sense of CIR-1 (her correct family status is 'not applicable — not in a family nucleus' (FST.L.5.)).
2. *Married couple living in the same household with their child and his/her consensual union partner:* Both couples each form together an individual family nucleus of two persons (see definition of 'family nucleus' in CIR-1). The family nucleus of the child and his/her partner must be classified as 'consensual union couples without resident children' (TFN.H.2.1.). But note that the family nucleus of the married parents must be classified as 'married or registered partnership couple families without resident children' (TFN.H.1.1.). This may seem illogical, since there is actually a resident child. However, this child lives with his/her partner, and therefore must not be considered a child.
3. *One parent living in the same household with their child and his/her consensual union partner:* While nothing changes for the child and his/her partner compared to example 2, the single parent must not be classified as 'lone father/mother family'. In fact, the parent must not be included in the underlying statistical population of 'family nuclei' broken down here, as he/she does not constitute a family nucleus in the sense of CIR-1 (his/her correct family status is 'not applicable — not in a family nucleus' (FST.L.5.), see FST example 3 above).
4. *The three-generation household consisting of a mother living with her children and her father:* The mother and her children form a 'lone mother family' (TFN.L.4.) as in example 1 above. However, following the reasoning of example 3, the mother's father is not a part of the underlying statistical population of 'family nuclei' and hence the breakdown TFN does not apply.

Topic: Size of family nucleus (SFN)

The CIR-1 definition of a 'family nucleus' given under the topic 'family status' (FST, see above) must be applied to establish the total statistical population of 'family nuclei' in the reporting country, which is broken down by the breakdown SFN. included in this topic.

The information on this topic must be collected for every family nucleus in the data source (there is no category 'not stated' or similar). If the information is missing, the most appropriate method should be used in the reporting country to impute it.

Changes with respect to 2011

- The categories breaking down '6 to 10 persons' have been abolished.
- There is now only one breakdown SFN., the detail of which corresponds to SFN.M. of 2011.

Further examples

People living alone: They do not constitute a 'family nucleus' and hence this topic does not apply to them (they should not be part of the total statistical population of 'family nuclei'). In fact, persons living alone are not, technically speaking, a family. They must therefore not be included in the count of family nuclei. In the data on persons, these persons living alone must be classified with family status 'not applicable — not in a family nucleus' (FST.L.5.).

Topic: Household status (HST)

The topic contains three breakdowns HST.L., HST.M. and HST.H., where CIR-1 defines the following concepts used for the breakdown categories:

- 'private household'
- 'housekeeping', a concept that includes the sub-concepts:
  - 'one-person household'
  - 'multiperson household'.
- 'household-dwelling'
- 'non-family household'
- 'institutional household'
- 'homeless persons', consisting of:
  - 'primary homeless' persons
  - 'secondary homeless' persons.

Note that the concept 'private household' is also the basis for establishing the total statistical population of 'private households' in the reporting country. In particular, CIR-1 stipulates that the reporting country must give preference to the 'housekeeping' concept over the 'household-dwelling' concept in identifying 'private households'. In any case, CIR-2 requires topical metadata explaining which concept was used to establish the total statistical population of private households in the reporting country.

Furthermore, the CIR-1 definition of a 'family nucleus' given under the topic 'family status' (FST, see above) must be used to identify the correct type and number of families living in a household, and hence the correct 'household status' of each person in that household (see examples).

Homeless persons

For the 2021 census round, CIR-1 combines primary and secondary homeless together with other persons of unknown household status into the category 'persons not living in a private household (including homeless persons), but category not stated' (HST.M.2.2.). A distinction between these concepts is therefore not legally required. This approach of merging population subgroups that are hard to enumerate into one breakdown category makes reporting easier, while ensuring that these subgroups are included in the total population.

In fact, Annex II of CIR-2 requires the total population of the reporting country to include all homeless persons (classified in HST.M.2.2.). It also requires the number of all (primary plus secondary) homeless persons to be reported in the topical metadata. In addition, CIR-2 requests individual numbers for primary and secondary homeless *'where this distinction is possible'*. Finally, CIR-2 requires the methodology applied to produce data on the homeless to be described in the topical metadata.

Note that the definition for the concept 'institutional households' given in CIR-1 is important to sharply delimit 'persons living in an institutional household' (HST.M.2.1.) from HST.M.2.2. (see above).

Changes with respect to 2011

This topic was largely revised and simplified for the 2021 round compared to 2011. This means that the number of breakdown categories under this topic has now been considerably reduced. In particular:

- all sub-categories of 'persons in a family nucleus' (HST.M.1.1.) were abolished;
- all sub-categories of persons not in a family nucleus and 'not living alone' (HST.H.1.2.2.) were abolished;
- optional sub-categories of 'persons in an institutional household' (HST.M.2.1.) were abolished, and the definition of 'institutional household' was added instead;
- 'primary homeless persons' were merged with 'secondary homeless persons' and 'persons not living in a private household, but category not stated' into one single category 'persons not living in a private household (including homeless persons), but category not stated' (HST.M.2.2., see further explanations above).

Further examples

The following four examples revisit the ones for 'family status' (FST, see above) and consider, in particular, their application under the topic 'type of family nucleus' (TFN, see above).

1. *A lone mother living with her children and her sister:* While the mother and her children are all 'persons in a family nucleus' (HST.M.1.1.), the sister must be classified as a 'person not in a family nucleus' (same reasoning as for family status FST.L.5. above) who is 'not living alone' (HST.H.1.2.2.).
2. *A married couple living in the same household with their child and his/her consensual union partner:* All four persons are 'persons in a family nucleus' (HST.M.1.1.).
3. However, consider the case of *one parent living in the same household with their child and his/her consensual union partner:* While nothing changes for the child and his/her partner compared to example 2, the single parent must be classified as a 'person not in a family nucleus' who is 'not living alone' (HST.H.1.2.2.) (compare with example 1).
4. *A three-generation household consisting of a mother living with her children and her father:* Following the allocation to family nuclei in FST example 4, and adapting the reasoning of example 1 above, the mother and her children are 'persons in a family nucleus' (HST.M.1.1.). However, the mother's father must be classified as a 'person not in a family nucleus' (same as family status FST.L.5. above) who is 'not living alone' (HST.H.1.2.2.).

Topic: Type of private household (TPH)

The CIR-1 definition of a 'private household' given under the topic 'household status' (HST, see above) must be applied to establish the total statistical population of 'private households' in the reporting country, which is broken down by the breakdowns TPH.L. and TPH.H. included in this topic. In addition, CIR-2 requires topical metadata explaining which concept ('housekeeping' or 'household-dwelling', see explanations under HST above) was used to establish the total statistical population of private households in the reporting country.

Furthermore, the CIR-1 definition of a 'family nucleus' given under the topic 'family status' (FST, see above) must be used to identify the correct type and number of families living in that household, and hence the correct breakdown category of this topic (see examples).

The information on this topic must be collected for every private household in the data source (there is no category 'not stated' or similar). If the information is missing, the most appropriate method should be used in the reporting country to impute it.

This topic treats the concepts 'married couple household', 'registered partnership household' and 'consensual union couple household' equivalently, in other words no distinction is required between them. In fact, the topic specifications in CIR-1 state explicitly: *'"Couple households" shall include married couple households, registered partnership households and consensual union couple households.'*

Note that CIR-2 also requires topical metadata explaining how the relationships between household members are identified in the reporting country (for example relationship matrix, from administrative registers, or based on a statistical model).

Changes with respect to 2011

The breakdowns for this topic have been largely simplified for 2021:

- the concepts of married couple/registered partnership/consensual union couple households (and all respective sub-categories) have been merged into one category 'couple households' (TPH.H.2.1.) and respective sub-categories;
- the optional sub-categories 'opposite-sex couple households' and 'same-sex couple households' have been abolished.

Further examples

Let us reconsider once more the four examples introduced under the topic 'family status' (FST, see above), and in particular the correct composition of family nuclei assigned to these examples under the topic 'type of family nucleus' (TFN, see above). None of these four examples can be identified as belonging to just one single family in the sense of CIR-1. Therefore the correct breakdown category for the private households in all these examples is 'two-or-more-family household' (TPH.L.3.).

Topic: Size of private household (SPH)

The CIR-1 definition of a 'private household' given under the topic 'household status' (HST, see above) must be applied to establish the total statistical population of 'private households' in the reporting country, which is broken down by the breakdown SPH. included in this topic. In addition, CIR-2 requires topical metadata explaining which concept ('housekeeping' or 'household-dwelling', see explanations under HST above) was used to establish the total statistical population of private households in the reporting country.

The information on this topic must be collected for every private household in the data source (there is no category 'not stated' or similar). If the information is missing, the most appropriate method should be used in the reporting country to impute it.

Changes with respect to 2011

- The categories breaking down '6 to 10 persons' have been abolished.
- There is now only breakdown SPH., the detail of which corresponds to SPH.M. of 2011.

Topic: Housing arrangements (HAR)

The topic contains one breakdown HAR., where CIR-1 defines the following concepts used for the breakdown categories:

- 'occupant'
- 'conventional dwelling' (including the sub-concepts 'separate' and 'independent')
- 'other housing unit'
- 'collective living quarter'
- 'living quarter'
- 'housing unit'.

All the concept definitions and logical relations between the concepts given in CIR-1 follow the CES Recommendations (§866ff., see also Section 2.3.2 and in particular Figure 2 of this publication). Note that these definitions in CIR-1 are also the basis for establishing the total statistical populations of 'conventional dwellings' and 'living quarters' in the reporting country, to be broken down by several other topics addressed further below.

Homeless persons

In line with the simplifications made in the topic 'household status' (HST, see above), the 'homeless' are also not required here anymore as a separate compulsory category. Note however the specific information obligations on 'homeless': Annex II of CIR-2 requires the total population of the reporting country to include all homeless persons (classified in HAR.2.). It also requires the number of all (primary plus secondary) homeless persons to be reported in the topical metadata. In addition, CIR-2 requests individual numbers for primary and secondary homeless *'where this distinction is possible'*. Finally, CIR-2 requires the applied methodology to produce data on the homeless to be described in the topical metadata.

Changes with respect to 2011

The sub-categories breaking down the category 'occupants living in another housing unit and the homeless' (HAR.2.) have been abolished.

Topic: Tenure status of households (TSH)

The CIR-1 definition of a 'private household' given under the topic 'household status' (HST, see above) must be applied to establish the total statistical population of 'private households' in the reporting country, which is broken down by the breakdown TSH. included in this topic. In addition, CIR-2 requires topical metadata explaining which concept ('housekeeping' or 'household-dwelling', see explanations under HST above) was used to establish the total statistical population of private households in the reporting country.

Topic: Type of living quarters (TLQ)

The CIR-1 definition of a 'living quarter' given under the topic 'housing arrangements' (HAR, see above) must be applied to establish the total statistical population of 'living quarters' in the reporting country, which is broken down by the breakdown TLQ. included in this topic. See also Figure 2 of this publication.

Furthermore, the CIR-1 definitions of 'other housing unit' and 'collective living quarters' given under the topic 'housing arrangements' (HAR, see above), as well as the definition of 'occupied conventional dwelling' given under the topic 'occupancy status of conventional dwellings' (OCS, see below) are used here to define the breakdown categories of TLQ.

Topic: Occupancy status of conventional dwellings (OCS)

The CIR-1 definition of a 'conventional dwelling' given under the topic 'housing arrangements' (HAR, see above) must be applied to establish the total statistical population of 'conventional dwellings' in the reporting country, which is broken down by the breakdown OCS. included in this topic. See also Figure 2 of this publication.

CIR-1 defines the following concepts, which are used for the breakdown categories:

- 'occupied conventional dwelling'
- 'unoccupied conventional dwelling'.

Furthermore, CIR-1 lists several special cases which must be classified as 'unoccupied conventional dwellings' (OCS.2.):

- dwellings reserved for seasonal or secondary use
- vacant dwellings
- conventional dwellings with persons present but not included in the census.

Note that the breakdown category 'occupied conventional dwelling' (OCS.1.) is also used to define the total statistical population 'occupied conventional dwellings' in the reporting country, to be broken down by the dedicated topics 'number of occupants' (NOC) and 'density standard' (DFS/DRM) (see further details below).

Changes with respect to 2011

- A list of special cases to be classified under the category 'unoccupied conventional dwellings' (OCS.2.) was added (see above).
- The optional sub-categories of 'unoccupied conventional dwellings' (OCS.2.) was abolished.

Further examples:

Buildings owned/occupied by foreign countries (for example private residences of personnel of foreign embassies): One might consider these buildings as being not part of the territory of the state (extra-territoriality), and therefore not to be enumerated by the reporting country. However, these buildings do belong to the housing stock of the hosting (in other words reporting) country, and once a foreign representation leaves that country these buildings would suddenly 'appear' in its national housing stock. It is also true that certain categories of persons are to be excluded from the total usually resident population (like foreign military, naval and diplomatic personnel located in the country) but their coverage by the census is ensured in the country of origin (see points (l) and (m) of the specifications on 'place of usual residence' (GEO)). However, this is certainly not the case for the buildings, and thus the buildings would not be covered in any national census. Therefore, buildings possessed/occupied by foreign authorities must be enumerated, although it is clear that these units may be more difficult to cover.

Finally, as follows from CIR-1 through the special case 'conventional dwellings with persons present but not included in the census' (see above), the correct breakdown category for such buildings is 'unoccupied conventional dwellings' (OCS.2.).

Topic: Type of ownership (OWS)

The CIR-1 definition of a 'conventional dwelling' given under the topic 'housing arrangements' (HAR, see above) must be applied to establish the total statistical population of 'conventional dwellings' in the reporting country, which is broken down by the breakdown OWS. included in this topic. See also Figure 2 of this publication.

CIR-1 defines the following concepts used for the breakdown categories:

- 'owner-occupied dwelling'
- 'rented dwelling'.

A dwelling with known ownership, but which does not fulfil the definition of 'owner-occupied dwelling' nor of 'rented dwelling', should be classified as 'dwellings in other types of ownership' (OWS.3.). The category 'not stated' (OWS.4.) is reserved for genuinely unknown/unstated ownership, while the category 'not applicable' (OWS.5.) is for 'unoccupied dwellings'.

Changes with respect to 2011

The breakdown category 'dwellings in cooperative ownership' was abolished.

Topic: Number of occupants (NOC)

The CIR-1 definition of an 'occupied conventional dwelling' given under the topic 'occupancy status of conventional dwellings' (OCS, see above) must be applied (in other words select its category OCS.1.) to establish the total statistical population of 'occupied conventional dwellings' in the reporting country, which is broken down by the breakdown NOC. included in this topic. See also Figure 2 of this publication.

The information on this topic must be collected for every occupied conventional dwelling in the data source (there is no category 'not stated' or similar). If the information is missing, the most appropriate method should be used in the reporting country to impute it.

Topic: Useful floor space and/or number of rooms (UFS/NOR)

The CIR-1 definition of a 'conventional dwelling' given under the topic 'housing arrangements' (HAR, see above) must be applied to establish the total statistical population of 'conventional dwellings' in the reporting country, which is broken down by the breakdowns UFS. and/or NOR. included in this topic. See also Figure 2 of this publication.

CIR-1 defines the following concepts used for the breakdowns:

- 'room'
- 'useful floor space'.

Based on these concepts, CIR-1 establishes two different breakdowns, 'useful floor space' (UFS.) and 'number of rooms' (NOR.), where NOR. is to be applied only if it is not possible to apply UFS.

Topic: Density standard (DFS/DRM)

The CIR-1 definition of an 'occupied conventional dwelling' given under the topic 'occupancy status of conventional dwellings' (OCS, see above) must be applied (in other words select its category OCS.1.) to establish the total statistical population of 'occupied conventional dwellings' in the reporting country, which is broken down by the breakdowns DFS. and/or DRM. included in this topic. See also Figure 2 of this publication.

This topic relates the topic 'number of occupants' (NOC, see above) to the topic 'useful floor space and/or number of rooms' (USF/NOR), where the useful floor space ([74]) divided by the number of occupants ([74]) must be classified under the breakdown 'density standard (floor space)' (DFS.). If this is not possible because information on the useful floor space is not available, the number of rooms ([74]) divided by the number of occupants ([74]) must be classified under the breakdown 'density standard (number of rooms)' (DRM.). This priority rule is consistent with the rule defined for the topic 'useful floor space and/or number of rooms' (DFS/DRM, see above).

Topic: Water supply system (WSS)

The topic specifications in CIR-1 now explicitly state that, in countries where sufficient evidence indicates that virtually all conventional dwellings have piped water, these countries are allowed to report all conventional dwellings in their data source under the breakdown category 'Piped water in conventional dwelling' (WSS.1.). However, countries making use of this exemption must certify this assumption and provide explanations in the metadata for this topic.

Changes with respect to 2011

The reporting obligation on this topic has been relaxed, given sufficient evidence that such facilities exist in virtually all conventional dwellings of the reporting country (see above).

([74]) Exact value, in other words not yet categorised under the respective topic breakdown.

Topic: Toilet facilities (TOI)

The topic specifications in CIR-1 now explicitly state that, in countries where sufficient evidence indicates that virtually all conventional dwellings have toilet facilities, these countries are allowed to report all conventional dwellings in their data source under the breakdown category 'Flush toilet in conventional dwelling' (TOI.1.). However, countries making use of this exemption must certify this assumption and provide explanations in the metadata for this topic.

Changes with respect to 2011

The reporting obligation on this topic has been relaxed, given sufficient evidence that such facilities exist in virtually all conventional dwellings of the reporting country (see above).

Topic: Bathing facilities (BAT)

The topic specifications in CIR-1 now explicitly state that, in countries where sufficient evidence indicates that virtually all conventional dwellings have bathing facilities, these countries are allowed to report all conventional dwellings in their data source under the breakdown category 'Fixed bath or shower in conventional dwelling' (BAT.1.). However, countries making use of this exemption must certify this assumption and provide explanations in the metadata for this topic.

Changes with respect to 2011

The reporting obligation on this topic has been relaxed, given sufficient evidence that such facilities exist in virtually all conventional dwellings of the reporting country (see above).

Topic: Type of heating (TOH)

The topic specifications in CIR-1 now explicitly state that, in countries where sufficient evidence indicates that virtually all conventional dwellings have central heating, these countries are allowed to report all conventional dwellings in their data source under the breakdown category 'central heating' (TOH.1.). However, countries making use of this exemption must certify this assumption and provide explanations in the metadata for this topic.

Whenever the allocation to the correct breakdown category is unclear for a given conventional dwelling, the guiding principle for the distinction between 'central heating' (TOH.1.) and 'no central heating' (TOH.2.) is given by the CES Recommendations (§962): '*A housing unit is considered as centrally heated if heating is provided either from a community heating centre or from an installation built in the building or in the housing unit, established for heating purposes, without regard to the source of energy.*' See also the examples below.

Changes with respect to 2011

The reporting obligation on this topic has been relaxed, given sufficient evidence that such facilities exist in virtually all conventional dwellings of the reporting country (see above).

Further examples

Wood stove/open fire place/wall-mounted gas fires/wall-mounted electric heaters: These facilities do not fulfil the recommended CES definition cited above, as they can only provide heating room by room and thus do not represent a central installation for the purpose of heating the entire dwelling. They should therefore be reported under 'no central heating' (TOH.2.).

Installed systems of electric storage heaters: Although such systems are firmly installed in a dwelling for the sole purpose of heating, they usually comprise a set of at least one stand-alone unit per heated room, instead of a central installation. Therefore, they do not strictly fulfil the CES Recommendation cited above for 'central heating'. They should therefore be reported under 'no central heating' (TOH.2).

Topic: Dwellings by type of building (TOB)

The CIR-1 definition of a 'conventional dwelling' given under the topic 'housing arrangements' (HAR, see above) must be applied to establish the total statistical population of 'conventional dwellings' in the reporting country, which is broken down by the breakdown TOB. included in this topic. See also Figure 2 of this publication.

Moreover, CIR-1 defines the concept 'non-residential building' to facilitate the correct allocation of a conventional dwelling to the breakdown category 'conventional dwellings in non-residential dwellings' (TOB.2.).

Changes with respect to 2011

A concept definition for 'non-residential building' has been added to CIR-1 (see above).

Topic: Dwellings by period of construction (POC)

The CIR-1 definition of a 'conventional dwelling' given under the topic 'housing arrangements' (HAR, see above) must be applied to establish the total statistical population of 'conventional dwellings' in the reporting country, which is broken down by the breakdown TOB. included in this topic. See also Figure 2 of this publication.

Changes with respect to 2011

The overall number of breakdown categories was reduced for the 2021 census round, although the breakdown now covers a longer period than in 2011. This implies that each category now covers a longer time period on average (reduced detail compared to 2011).

Glossary of concept definitions

For convenience, this glossary contains a table with all concepts defined somewhere in the EU census 2021 legislation. The table is ordered by concept name as used in CIR-1 and listed in this publication, and shows where to find the respective definition (legal base or CIR-1 topic acronym).

Table A1 — Glossary of concept definitions

Concept name	Definition
Building	CFR, Art. 2
Central heating	TOH
Child (son/daughter)	FST
Citizenship	COC
Collective living quarter	HAR
Consensual union	FST
Contributing family worker	SIE
Conventional dwelling	HAR
Couple	FST
Couple household	TPH
Current activity status	CAS
Educational attainment	EDU
Employed	CAS
Employee	SIE
Employer	SIE
EU Member State	POB
Family nucleus	FST
Formal job attachment	CAS
Homeless person	HST
Household-dwelling concept	HST
Housekeeping concept	HST
Housing	CFR, Art. 2
Housing unit	HAR
Industry (branch of economic activity)	IND
Institutional household	HST
Labour force	CAS
(Legal) marital status	LMS
Living quarter	HAR
Locality	LOC
Lone parent	FST
Member of a producers' cooperative	SIE
Multiperson household	HST
Non-family household	HST
Non-residential building	TOB
Occupant	HAR
Occupation	OCC
Occupied conventional dwelling	OCS
One-person household	HST
Other housing unit	HAR
Own-account worker	SIE
Owner-occupied dwelling	OWS
Paid employment	SIE
Partner	FST
Place of usual residence one year prior to the census	ROY
Population	CFR, Art. 2
Population of a locality	LOC
Primary homeless person	HST
Private household	HST
Reference date	CFR, Art. 2
Registered partnership	LMS
Rented dwelling	OWS
Room	UFS/NOR
Secondary homeless person	HST
Self-employment	CES §541
Skip-generation household	FSZ
Student	CAS
Total population (of a geographical area)	CIR-2 Art. 2
Unemployed	CAS
Unoccupied conventional dwelling	OCS
Useful floor space	UFS/NOR
Usual resident (place of usual residence)	CFR, Art. 2
Year of arrival	YAT/YAE

B Overview of the EU census programme under Regulation (EU) 2017/712

This overview contains four sheets:

- Hypercube groups about persons — most detailed regional level: NUTS level 2;
- Hypercube groups about persons — most detailed regional level: below NUTS level 2;
- Hypercube groups about private households and families — all regional levels;
- Hypercube groups about housing — all regional levels.

The hypercubes about the housing arrangements of persons (hypercubes 38, 39 and 40) are presented in the overview table on 'housing'.

The overview tables indicate the hypercubes in the columns and the topics those hypercubes contain in the rows.

There are three kinds of marker, as set out in the bullet points below.

- Crosses in bold and red: these indicate the principal focus of the hypercube. For example, the main focus of hypercube 1 is the household status of persons.
- Crosses in bold and black: these indicate the particular angles from which the principal focus of the hypercube is looked at. For example, hypercube 1 looks at the legal marital status of persons with a particular household status (for example of lone parents).
- Crosses not in bold and black: these indicate the 'structure topics' helpful to analyse the information. For example, in hypercube 1 the information on persons by their household status and legal marital status is broken down by their sex, five-year age groups, current activity status, country of birth and citizenship.

Table B1: Hypercube groups about persons — most detailed regional level: NUTS level 2 ([75])

Topic	Breakdowns	Hypercube group number																				
		1	4	5	6	7	15	16	17	18	19	20	21	22	23	24	28	29	30	31	32
GEO. Place of usual residence	x = NUTS level 2, NUTS level 1, nation (GEO.L.) N = only nation (GEO.N.)	N	x	x	x	N	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
SEX. Sex	total / male / female	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
AGE.M. Age	five-year age groups	x	x	x	x	x	x	x	x		x	x	**x**	x	x	x	x	x	x	x	x
AGE.H. Age	single years of age	**x**	**x**																		
ROY. Residence one year before census	unchanged / same country / same region							x	**x**								**x**				
LPW. Location of place of work	N = only national level, in other words persons working abroad (LPW.N.)				**N**	**x**						**N**									
LOC. Size of the locality	13 size classes																				
LMS.L. Legal marital status	never married / married / widowed / divorced (or equivalent for registered partnership)	x											**x**			**x**		x	x		
LMS.H. Legal marital status	never married / married or equiv. (opposite-/same-sex) / widowed / divorced (equiv. for reg. partnership)	x																			
HST.L. Household status	in private household / not in private household	x											x	x	x	x	x	x	x	x	x
HST.M. Household status	in private household: in family / not in family not in private household: in institutional household / not in institutional household incl. homeless	x											x	x		**x**	**x**	x	x	**x**	**x**
HST.H. Household status	in private household: in family / not in family (living alone / not alone) not in private household: in institutional household / not in institutional household incl. homeless	**x**											**x**	**x**							
FST.L. Family status	partner / lone parent / child	x											x		**x**	**x**	x	x	x	**x**	**x**
FST.M. Family status	partner (married or reg. partnership / consensual union) / lone parent / child (not lone / lone parent)	x											**x**				**x**				
FST.H. Family status	partner (opposite-/same-sex married or equiv. / cons. union) / lone parent / child (not lone / lone parent)	**x**																			
EDU. Educational attainment	nine ISCED 2011 categories		**x**	**x**	**x**		**x**			**x**	**x**		x	**x**	**x**					x	x
CAS.L. Current economic activity status	labour force (employed / unemployed) / outside of the LF		**x**				**x**						x		**x**	**x**		**x**	**x**		
CAS.H. Current economic activity status	labour force (employed / unemployed) / outside of the LF (subgroups)												**x**								
OCC. Occupation	type of work by 10 major ISCO groups		**x**	**x**	**x**			**x**													
IND. Industry	branch of economic activity by 21 NACE sections			**x**	x	**x**			**x**	**x**											
SIE. Status in employment	employees / employers / own-account workers / other LF member			**x**	x	**x**				**x**				**x**						**x**	**x**
POB.L. Place of birth	in country / not in country (EU / non-EU)						x	**x**		**x**	**x**	**x**							**x**		**x**	
COC.L. Citizenship	in country / not in country (EU / non-EU)						x	x	**x**	x		**x**								**x**		**x**
YAE.L. Year of arrival	since 1980 (five-year groups)						x	**x**	**x**		x											
YAE.H. Year of arrival	since 1980 (single years from 2005)						x				**x**											

([75]) Excluding hypercubes about 'Housing arrangements' of persons, which are presented in the table on 'Hypercubes about housing'.

Table B2: Hypercube groups about persons — most detailed regional level: below NUTS level 2 ([76])

Topic	Breakdowns	Hypercube group number								
		3	8	9	10	11	12	25	26	27
GEO.M. Place of usual residence	NUTS level 3, NUTS level 2, NUTS level 1, nation	x	x	x	x	x	x	x	x	x
GEO.H. Place of usual residence	municipalities	x	x							
SEX. Sex	total / male / female	x	x	x	x	x	x	x	x	x
AGE.M. Age	five-year age groups	x		x	x	x	x	x	x	x
ROY. Residence one year before census	unchanged / same country / same region						x			
LMS.L. Legal marital status	never married / married / widowed / divorced (or equivalent for reg. partnership)	x						x		
HST.M. Household status	in private household (in family / not in family) / not in private household (in inst. household / not in inst. household incl. homeless)	x						x	x	x
FST.L. Family status	partner / lone parent / child								x	x
FST.M. Family status	partner (subgroups) / lone parent / child (subgroups)									x
CAS.L. Current economic activity status	labour force (employed/unemployed) / outside of the LF				x					
SIE. Status in employment	employees / employers / own-account workers / other LF member						x			
POB.L. Place of birth	in country / not in country (EU / non-EU)		x	x	x		x	x	x	
POB.M. Place of birth	continents			x			x			
POB.H. Place of birth	countries			x						
COC.L. Citizenship	in country / not in country (EU / non-EU)		x	x	x	x	x	x	x	
COC.M. Citizenship	continents					x	x			
COC.H. Citizenship	countries					x				
YAT. Year of arrival	since 2000				x					
YAE.L. Year of arrival	since 1980 (five-year groups)			x		x	x			x
YAE.H. Year of arrival	since 1980 (single years from 2005)			x						

([76]) Excluding hypercubes about 'Housing arrangements' of persons, which are presented in the table on 'Hypercubes about housing'.

Table B3: Hypercube groups about private households and families

Topic	Breakdowns	Hypercube group number			
		33	34	35	36
GEO.M. Regional detail	NUTS level 3, NUTS level 2, NUTS level 1, nation (H = number of private households, F = number of families in the geographical area)	**H**	**F**	H	F
GEO.H. Regional detail	municipalities			**H**	**F**
TPH.L. Type of private household	non-family household (one person/multi-person) / one-family household / two-or-more-family household	x		**x**	
TPH.H. Type of private household	non-family household (one person/multi-person) / one-family household (kind of couple and presence of children) / two-or-more-family household	**x**			
SPH. Size of private household	1, 2, 3, 4, 5, 6-10, 11+	x		**x**	
TSH. Tenure status of household	housing unit owned / rented / other	**x**			
TFN.L. Type of family nucleus	by kind of couple		x		**x**
TFN.H. Type of family nucleus	by kind of couple and presence of children		**x**		
SFN. Size of family nucleus	1, 2, 3, 4, 5, 6-10, 11+		x		**x**

Table B4: Hypercube groups about housing ([77])

Topic	Breakdowns	Hypercube group number							
		2	13	14	37	38	39	40	41
GEO.L. Regional detail	NUTS level 2, NUTS level 1, nation (P = Population at their place of usual residence, D = conventional dwellings, O = occupied conventional dwellings, L = living quarters in the geographical area)	P	P	**P**	D	D	O	**O**	L
GEO.M. Regional detail	NUTS level 3	**P**	**P**		**D**	D	**O**		L
GEO.H. Regional detail	municipalities					**D**			**L**
SEX. Sex	total / male / female	x	x	x					
AGE.M. Age	five-year age groups	x	x	x					
ROY. Residence one year before census	unchanged / same country / same region		**x**	x					
LOC. Size of the locality	13 size classes	**x**							
LMS.L. Legal marital status	never married / married / widowed / divorced (or equivalent for registered partnership)	x							
HST.H. Household status	in private household: in family / not in family (living alone / not alone) not in private household: in institutional household / not in institutional household incl. homeless	**x**							
FST.H. Family status	partner (opposite-/same-sex married or equiv. / cons. union) / lone parent / child (not lone / lone parent)	**x**							
HAR. Housing arrangement	occupants living in a conventional dwelling / collective living quarter / other housing unit (incl. homeless)	**x**	**x**	**x**					
CAS.H. Current economic activity status	active: employed / unemployed / not active			**x**					
POB.L. Place of birth	in country / not in country (EU / non-EU)		x	x					
POB.M. Place of birth	continents		**x**						
COC.L. Citizenship	in country / not in country (EU / non-EU)		x	**x**					
COC.M. Citizenship	continents		**x**						
YAE.L. Year of arrival	since 1980 (five-year groups)		x	**x**					
YAE.H. Year of arrival	since 1980 (single years from 2005)		**x**						
TLQ. Type of living quarter	occupied conventional dwellings / other housing units (for example hut, caravan) / collective living quarters								**x**
OCS. Occupancy status of conventional dwelling	occupied / unoccupied (seasonal use / vacant)				**x**	**x**			
OWS. Type of ownership	owner-occupied / rented / other						**x**		
NOC. Number of occupants	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11+						**x**		
UFS./NOR. Useful space	nine size classes for floor space (UFS.) or number of rooms (NOR.)						**x**		
DFS./DRM. Density standard (useful space per occupant)	eight classes for floor space per occupant (DFS.) or number of rooms per occupant (DRM.)						**x**		
WSS. Water supply system: piped water	yes / no							**x**	
TOI. Toilet facilities: flush toilet	yes / no							**x**	
BAT. Bathing facilities: fixed bath or shower	yes / no							**x**	
TOH. Type of heating: central heating	yes / no							**x**	
TOB. Dwellings by type of building	conventional dwellings in residential buildings (1/2/3+ dwelling buildings) / in non-residential buildings				**x**	**x**	**x**		
POC. Dwelling by period of construction	eight classes: before 1919, 1919-1945, 1946-1960, 1961-1980, 1981-2000, 2001-2010, 2011-2015, 2016 and later				**x**				

([77]) Including hypercubes about 'Housing arrangements' of persons.

C Provisions of EU census legislation on metadata and quality reporting

Table C1: Annex II to Regulation 2017/712 (CIR-2)

Requirements of Annex II to Regulation 2017/712	Collected through
Metadata on the topics referred to in Article 6 Member States shall transmit to the Commission (Eurostat) textual metadata on the definitions referring to the census topics. For each topic, the metadata shall:	–
— name the data source(s) used to report the statistical data on the topic;	ESMS items 18.1.3.x
— report on the methodology used to estimate data on the topic;	ESMS items 13.1.x
— report on the reasons for any unreliability of the data on the topic.	ESMS items 13.1.x
In addition, Member States shall provide the following metadata: ESMS items 13.1.x (estimation methodology and possible unreliability) **Place of usual residence** The metadata shall explain in which way the definition of 'usual residence' of Article 2(d) of Regulation (EC) No 763/2008 has been applied, in particular to what extent the legal or registered residence has been reported as a substitute for the usual residence according to the 12 months criterion, as well as a clear definition of the concept adopted for the resident population. The metadata shall report if third level students whose term-time address is not the one of their family home have been considered to have their usual residence at their family home. The metadata shall report on any other country-specific application of the rules for the 'special cases' listed in the technical specifications for the topic 'Place of usual residence' in the Annex to Implementing Regulation (EU) 2017/543. **Homeless** The data on total population shall include all primary homeless persons (persons living in the streets without shelter) and all secondary homeless persons (persons moving frequently between temporary accommodation). The metadata shall report the number of all homeless persons. The numbers of primary homeless persons (persons living in the streets without shelter) and of secondary homeless persons (persons moving frequently between temporary accommodation) shall be shown where this distinction is possible. A description of the methodology and data sources used to produce the data on homeless persons shall be provided. **Legal marital status/partnerships** The metadata shall report on the relevant legal basis in the Member State concerning opposite-sex and same-sex marriages, the minimum age for marriages, opposite-sex and same-sex registered partnerships, and the possibility to divorce or legally separate.	ESMS items 3.4.x (definitions and concepts) + ESMS items 13.1.x (estimation methodology and possible unreliability)

Table C1 (continued): Annex II to Regulation 2017/712 (CIR-2)

Requirements of Annex II to Regulation 2017/712	Collected through
Economic topics The metadata shall report on any country-specific application of the rules listed in the technical specifications for the topic 'Current activity status' in the Annex to Implementing Regulation (EU) 2017/543. The metadata shall report whether the current activity status has been reported on the basis of registers, and, if this is the case, on the relevant definitions used in this register. The metadata shall report on the national minimum age for economic activity in the country, and the relevant legal basis. Where the census in the Member State identifies persons doing more than one job, the metadata shall describe the method used to allocate them to their main job (for example, on the basis of time spent on the job, income received). The metadata shall report on any country-specific application of the rules listed in the technical specifications for the topic 'Status in employment' in the Annex to Implementing Regulation (EU) 2017/543. Where the census in the Member State identifies person who are both, employer and employee, the metadata shall describe the method used to allocate them to one of the two categories.	
Country/place of birth For censuses for which no or incomplete information is available on the country of birth according to international boundaries existing at the time of the census, the metadata shall inform about the methodology used to allocate persons within the breakdown of the topic 'Country/place of birth'. The metadata shall report if information on the place where the birth took place was used as a substitute for the place of usual residence of the mother at the time of the birth.	
Country of citizenship In countries where a part of the population are persons who are 'Recognised Non-Citizens' (that is persons who are neither citizens of any country nor stateless and who have some but not all of the rights and duties associated with citizenship), the metadata shall provide relevant information.	ESMS items 3.4.x (definitions and concepts) + ESMS items 13.1.x (estimation methodology and possible unreliability)
Place of usual residence one year prior to the census Where the census in the Member State collects information on the topic 'Previous place of usual residence and date of arrival in the current place', the metadata shall describe any methodology used to report on the place of usual residence one year prior to the census.	
Household and family topics The metadata shall specify whether the census in the Member State applies the 'housekeeping' or the 'household-dwelling' concept to identify private households. The metadata shall report on the method used to generate households and families. The metadata shall report on the way the relationships between household members are identified (e.g. relationship matrix; relation to reference person). If these data are obtained from administrative registers, it shall be reported whether information on the relationship between household and family members is recorded in and obtained directly from the administrative source(s), or whether this information is based on a statistical model.	
Type of ownership The metadata shall explain and provide examples of the ownership types under national property laws or customs that have been classified under 'Dwellings in other types of ownership'.	
Useful floor space and/or number of rooms of housing unit, density standard The metadata shall report on the application of the concept of either 'useful floor space', or 'number of rooms' as appropriate, and on the definition adopted for the corresponding measurement of the density standard.	

Table C2: Structure and content of the 2021 census quality reports: Annex to Regulation 2017/881 (CIR-3) merged with relevant provisions of Annex III to Regulation 2018/1799 (CIR-4, fields with grey background)

Requirements	Collected through
1. OVERVIEW	–
1.1. Legal background	ESMS item 6.1
1.2. Bodies responsible	ESMS item 6.1.1
2. DATA SOURCES	–
2.1. Classification of the data sources according to Article 4(1) of Regulation (EC) No 763/2008	ESMS items 18.1.2.x
2.2. List of the data sources used for the 2021 census	ESMS items 18.1.1.x
2.3. 'Data sources x topics' matrix	ESMS item 18.1.3
2.4. Adequacy of data sources: extent to which they meet the essential features (Article 4(4) of Regulation (EC) No 763/2008) 2.4.1. *Individual enumeration;* 2.4.2. *Simultaneity;* 2.4.3. *Universality within the defined territory;* 2.4.4. *Availability of small-area data;* 2.4.5. *Defined periodicity.*	ESMS item 18.1.4
3. CENSUS LIFE CYCLE	–
3.1. Reference date according to Article 3 of Commission Regulation (EU) 2017/712	ESMS item 5
3.2. Preparation and execution of data collection	–
3.2.1. Questionnaire-based data — design and testing of questionnaires (including copies of all final questionnaires), — preparation of any address lists, preparation of the field work, mapping, publicity, — data collection (including field work), — legal obligation to collect information, incentives for providing truthful information or possible reasons for providing false information.	ESMS item 18.3.1
3.2.2. Register-based data — creation of new registers from the year 2011 onwards (where applicable), — redesign of existing registers from 2011 onwards (including changes in the contents of registers, adaptation of the census population, adaptation of definitions and/or technical specifications) (where applicable), — maintenance of the registers (for each register used for the 2021 census), including — content of the register (statistical units and information on them, any record editing and/or item and record imputation in the register), — administrative responsibilities — legal obligation to register information, incentives for providing truthful information and possible reasons for providing false information, — delays in reporting, in particular legal/official delays, data registration delays, late reporting, — evaluation of and clearance for non-registration, non-deregistration, multiple registration, — any major register revisions or updating of records that affects the 2021 census data, periodicity of register revisions, — usage, including 'statistical usage of the register other than for the census' and 'usage of the register other than for statistical purposes (e.g. administrative purposes)', — matching and linking of registers (including unique identifier(s) used for record linkage), — data extraction.	ESMS item 18.3.2

Table C2 (continued): Structure and content of the 2021 census quality reports: Annex to Regulation 2017/881 (CIR-3) merged with relevant provisions of Annex III to Regulation 2018/1799 (CIR-4, fields with grey background)

Requirements	Collected through
3.2.3. Data collected by means of a sample For topics for which information has been collected by means of a sample, the metadata shall also contain descriptions of: — the sampling design, — methodologies used for any estimations, models or imputations, — possible biases in the estimation due to methodologies applied, — formulae and algorithms used to calculate the standard error.	ESMS item 18.3.3
3.2.4. Data collected by combined methods (data based on more than one type of data source) For topics for which information has been collected by combined methods, the metadata shall also contain: — a description of the methods (types of data sources used and how information from different sources was combined, how the different sources and methods used complement and support each other and, if applicable, which parts of the population were covered by which source), — any other quality issues relating to the process of using combined methods.	ESMS item 18.3.4
3.3. Processing and evaluation	–
3.3.1. Data processing (including capturing, coding, identifying variable(s), record editing, record imputation, record deletion, estimation, record linkage including identifying variable(s) used for the record linkage, generation of households and families, measures to identify or limit unit-no-information);	ESMS item 18.5
3.3.2. Coverage assessment activities, methodology to treat non-response, post-enumeration survey(s) (where applicable), final data validation: method of assessing under- and over-coverage, including information on the quality of the under- and over-coverage estimates.	ESMS items 11.2, 11.2.1, 11.2.2, 18.4
3.3.3. Additional information on generic (not topic-related) methodology applied in order to produce the 1 km^2 grid dataset.	ESMS item 18.5.1
3.4. Dissemination (dissemination channels, assurance of statistical confidentiality including statistical disclosure control)	ESMS items 10.1, 10.2, 10.3, 10.4, 10.5 and 7.2
Specific information about statistical disclosure control measures related to the 1 km^2 grid dataset: Member States shall provide the Commission (Eurostat) with information about the measures related to the harmonised protection of 1 km^2 grid data, in particular if they used the ESS good practices and implementation guidelines for the harmonised protection of 1 km^2 grid data.	ESMS item 7.2.1
3.5. Measures to ensure cost effectiveness	ESMS item 16
4. ASSESSMENT OF DATA QUALITY	**–**
4.1. Comparability For each topic, Member States shall report on any deviation from the required concepts and definitions or any practice in the Member State that could impair the Union-wide comparability of the data. For the topic 'Current Activity Status', Member States shall report on any estimation methods used to adjust data to meet more closely the definition set in the Annex to Commission Implementing Regulation (EU) 2017/543. Member States shall report on the extent to which the data sources and any estimation methods used result in deviation from the definition of 'Current Activity Status' set in that Regulation.	ESMS item 16.1
4.2. Timeliness and punctuality The following information shall be provided at national level:	–
— date(s) of the transmission of data to the Commission (Eurostat), broken down by hypercubes,	ESMS item 14.1
— date(s) of major revision(s) of the transmitted data, broken down by hypercubes,	ESMS item 14.2
— date(s) of transmission of the metadata,	ESMS item 2.2
— date(s) of the transmission and possible revisions of the 1 km^2 grid data and metadata.	ESMS items 14.1, 14.2 and 2.2 (see above)

Table C2 (continued): Structure and content of the 2021 census quality reports: Annex to Regulation 2017/881 (CIR-3) merged with relevant provisions of Annex III to Regulation 2018/1799 (CIR-4, fields with grey background)

Requirements	Collected through
4.3. Coherence Member States shall report on any significant inconsistencies between the data transmitted in the different datasets defined in Commission Regulation (EU) 2017/712.	ESMS item 15.3
4.4. Coverage and accuracy	–
To indicate coverage, the following **absolute values** shall be provided for person counts at national level and shall be disaggregated by sex and broad age groups as defined in Commission Implementing Regulation (EU) 2017/543: (a) census population; (b) number of all record imputations; (c) number of all record deletions; (d) under-coverage (estimated); (e) over-coverage (estimated); (f) estimated target population.	QHC1
For the assessment of accuracy the following absolute values shall be provided for person counts at national level and shall be disaggregated by sex and broad age groups as defined in Commission Implementing Regulation (EU) 2017/543: (a) census population; (b) number of observed data records on the topic derived from traditional census; (c) number of observed data records on the topic derived from administrative registers; (d) number of observed data records on the topic derived from sample surveys; (e) number of observed data records on the topic derived from multiple data sources; (f) complementary set of statistical units on the topic (for samples); (g) number of imputed observations on the topic; (h) number of records with missing information on the topic.	QHC2
The above absolute values for the assessment of accuracy shall be provided for the following census topics: (a) legal Marital status (LMS); (b) family status (FST); (c) household status (HST); (d) current activity status (CAS); (e) occupation (OCC); (f) industry (IND); (g) status in employment (SIE); (h) location of place of work (LPW); (i) educational attainment (EDU); (j) country/Place of birth (POB); (k) country of citizenship (COC); (l) year of arrival in the country since 2010 (YAT); (m) year of arrival in the country since 1980 (YAE); (n) place of usual residence one year prior to the census (ROY); (o) housing arrangements (HAR).	QHC2
4.5. Completeness Member States shall report on the degree of completeness of the data in terms of the requirements of Regulation (EC) No 763/2008. They shall give details of any census topics or associated breakdowns for which data are not supplied.	ESMS item 12.3
4.6. Relevance Information on the following shall be provided at Union level:	–
(a) actions taken to identify and fulfil user needs;	ESMS item 12.1
(b) monitoring of the extent of data extractions.	ESMS item 12.2
4.7. Geographic information — data quality covering geographic quality principles, in particular territorial coverage and comparability, positional accuracy, as well as temporal oherence and completeness of the geographic data used for geocoding.	ESMS item 13.1.2

Table C3: Annex III to Regulation 2018/1799 (CIR-4)

Requirements	Collected through
Metadata on data items 1. Where applicable, Member States shall add the following flags to a data item: (a) 'provisional'; (b) 'populated'; (c) 'revised'; (d) 'see information attached'; (e) 'confidential'. 2. Only data values on 'total population' which are reported under Article 9(1) and which are not considered final data by the Member State at the time of reporting shall be accompanied by the flag 'provisional'. 3. The flag 'populated' shall be applicable only to 'total population' data items under the provisions specified in Article 6(2). 4. For each data value accompanied by at least one of the flags 'revised' or 'see information attached' an explanatory text shall be provided. 5. Each data item whose confidential value has been replaced by the special value 'not available' shall be marked with the flag 'confidential'.	1 km² grid dataset
Metadata on the topics In addition to the metadata on the topics transmitted to the Commission (Eurostat) under Article 6 of [CIR-2], Member States shall provide metadata on each topic included in Annex I informing about the data sources and methodology used to obtain the data values for that topic on the 1 km² reference grid. In particular, the metadata shall contain: — information on the reliability and accuracy of the reported data values; — a description of any methodology used to estimate the data values on the 1 km² reference grid, including reliability and accuracy of the resulting data values; — a description of any methodology used to allocate persons to specific grid cells under the topic 'place of usual residence', including information on the characteristics of persons under the category GEO.G.y.	ESMS items 13.1.1.x (cf. Table C1)
Reference metadata The metadata information and structure laid down in the Annex of [CIR-3] shall be supplemented for the purpose of this Regulation by the following items specifically referring to the 1 km² grid data:	cf. Table C2
— Item 3.3. 'Processing and evaluation' shall be supplemented by the additional sub-item 3.3.3. 'Additional information on generic (not topic-related) methodology applied in order to produce the 1 km² grid dataset'.	ESMS item 18.5.1
— Item 3.4. 'Dissemination' shall be supplemented by specific information about statistical disclosure control measures related to the 1 km² grid dataset. Member States shall provide the Commission (Eurostat) with information about the measures related to the harmonised protection of 1 km² grid data, in particular if they used the ESS good practices and implementation guidelines for the harmonised protection of 1 km² grid data.	ESMS item 7.2.1
— Item 4.2. 'Timeliness and punctuality' shall be supplemented by specific calendar date(s) of the transmission and possible revisions of the 1 km² grid data and metadata.	ESMS items 15.1, 19.2
— Item 4. 'Assessment of data quality' shall be supplemented by the additional sub-item 4.7 'Geographic information — data quality' covering geographic quality principles, in particular territorial coverage and comparability, positional accuracy, as well as temporal coherence and completeness of the geographic data used for geocoding.	ESMS item 13.1.2

D Transmission programme of quantitative metadata (quality hypercubes)

Table D1: Legal basis and data structure of quality hypercubes QHC1 and QHC2

QHC	Legal basis according to item 4.4 of the annex to CIR-3	Breakdowns
QHC1	*To indicate coverage, the following **absolute values** shall be provided for person counts at national level and shall be disaggregated by sex and broad age groups as defined in Commission Implementing Regulation (EU) 2017/543:*	**SEX.** **AGE.L.**
	(a) census population; *(b) number of all record imputations;* *(c) number of all record deletions;* *(d) under-coverage (estimated);* *(e) over-coverage (estimated);* *(f) estimated target population.*	**COV_IND.**
QHC2	*For the assessment of accuracy the following **absolute values** shall be provided for person counts at national level and shall be disaggregated by sex and broad age groups as defined in Commission Implementing Regulation (EU) 2017/543:*	**SEX.** **AGE.L.**
	(a) census population; *(b) number of observed data records on the topic derived from traditional census;* *(c) number of observed data records on the topic derived from administrative registers;* *(d) number of observed data records on the topic derived from sample surveys;* *(e) number of observed data records on the topic derived from multiple data sources;* *(f) complementary set of statistical units on the topic (for samples);* *(g) number of imputed observations on the topic;* *(h) number of records with missing information on the topic.*	**TOPIC_IND.**
	The above absolute values for the assessment of accuracy shall be provided for the following census topics: *(a) legal Marital status (LMS);* *(b) family status (FST);* *(c) household status (HST);* *(d) current activity status (CAS);* *(e) occupation (OCC);* *(f) industry (IND);* *(g) status in employment (SIE);* *(h) location of place of work (LPW);* *(i) educational attainment (EDU);* *(j) country/Place of birth (POB);* *(k) country of citizenship (COC);* *(l) year of arrival in the country since 2010 (YAT);* *(m) year of arrival in the country since 1980 (YAE);* *(n) place of usual residence one year prior to the census (ROY);* *(o) housing arrangements (HAR).*	**TOPIC.**

Table D2: Code list for the COV_IND dimension in quality hypercube QHC1

Code	Description (absolute values only, see Table D1)
NBR_POP_CENSUS	Census population
NBR_IMPU_REC	Number of all record imputations
NBR_DEL_REC	Number of all record deletions
NBR_UNDERCOV	Under-coverage (estimated)
NBR_OVERCOV	Over-coverage (estimated)
NBR_POP_ETP	Estimated target population: absolute value

Table D3: Code list for the TOPIC_IND dimension in quality hypercube QHC2

Code	Description (absolute values only, see Table D1)
NBR_POP_CENSUS	Census population
NBR_OBS_TRAD	Number of observed data records on the topic derived from traditional census
NBR_OBS_REGISTER	Number of observed data records on the topic derived from administrative registers
NBR_OBS_SURVEY	Number of observed data records on the topic derived from sample surveys
NBR_OBS_MULTI	Number of observed data records on the topic derived from multiple data sources
NBR_COMPSET	Complementary set of statistical units on the topic (for samples)
NBR_IMPU_TOPIC	Number of imputed observations on the topic
NBR_MISS_TOPIC	Number of records with missing information on the topic

Table D4: Code list for the TOPIC dimension in quality hypercube QHC2

Code	Description
GEO	Usual residence
SEX	Sex
AGE	Age
LMS	Marital status
FST	Family status
HST	Household status
CAS	Current activity status
OCC	Occupation
IND	Industry
SIE	Status in employment
LPW	Place of work
EDU	Educational attainment
POB	Place of birth
COC	Country of citizenship
YAT	Year of arrival in the country since 2010
YAE	Year of arrival in the country since 1980
ROY	Residence one year prior to the census
HAR	Housing arrangements

D1. How to compute QHC1: Quantitative information on the coverage of persons

As outlined in Table D1, QHC1 has three dimensions:

- **SEX.**
- **AGE.L.**
- **COV_IND.**

where COV_IND. is the list of coverage indicators that must be reported to measure how well the census population relates to the estimated target population. For the code list of the dimension COV_IND., see Table D2. A detailed set of guidelines on each of these coverage indicators is provided in the following table, where the column 'Point' refers to the list of coverage indicators in the annex to CIR-3, item 4.4 (cf. Table D1).

Point	Coverage indicator	Explanation
(a)	Census population	Absolute number of records included in the 'census microdata database' on persons (after record imputation/deletion).
(b)	Number of all record imputations	Absolute number of records in the 'census microdata database' on persons that have been imputed.
(c)	Number of all record deletions	Number of records in the 'census microdata database' on persons that have been deleted.
(d)	Under-coverage (estimated)	Estimate of under-coverage as resulting from the chosen method of coverage assessment (see above). This is meant to be the final under-coverage, generally conducted after the publication of census results (in other words the residual under-coverage after record imputation).
(e)	Over-coverage (estimated)	Estimate of over-coverage as resulting from the chosen method of coverage assessment. As for item (d), this is meant to be the final over-coverage (in other words after record deletion).
(f)	Estimated target population	Equals the census population (a) plus the under-coverage (d) minus the over-coverage (e) in the respective SEX. and AGE.L. class.

As mentioned earlier, the indicators included in QHC1 must be computed on the basis of the 'census microdata database' on persons, the data source that — through record linking — is obtained by merging together the different primary data sources on persons of the census, as explained in Section 5.4.1.

We provide here some examples of how to compute the indicators of QHC1. For the sake of simplicity, these calculations will be illustrated using plain SQL on the microdata database. The correct translations of these queries into equivalents of the statistical software used in each Member State (R, SAS, and so on) will be obvious. To write these queries, we assume that the table containing the census microdata on persons has two additional columns with the variables REC_IMPUTED and REC_DELETED as listed below.

- `REC _ IMPUTED = {"Y", "N"}` is a flag indicating whether the record was fully imputed or not. This does not mean 'whether the record includes some imputed items or not' but rather 'whether the entire record was added to the data source as a result of record imputation'. We recall that any record imputation increases the size of the census population, so it is just the 'record additions' that have to be counted as 'record imputations'.
- `REC _ DELETED = {"Y", "N"}` is a flag indicating whether the record was deleted (in other words excluded from the census population) or not. We assume therefore that our data source (the 'census microdata database' on persons) also includes records that were deleted and that must not be included in the population count. If this is not the case (meaning that the deleted records are physically removed from the 'census microdata database'), Member States must keep track of how many record deletions occurred, by SEX. and AGE.L. at national level, so that it is still possible to compute the number of all record deletions, as required by the annex to CIR-3. It is clear that a record cannot be considered as imputed and deleted at the same time. Therefore REC_IMPUTED and REC_DELETED cannot be both 'Y' for the same microdata record.

The indicators included in QHC1 must be provided for the breakdowns SEX. and AGE.L. Therefore we assume that the 'census microdata database' on persons includes respective columns SEX and AGE_L containing the breakdown categories to which each person record belongs.

The 2021 census legislation contains no explicit provisions on record weighting, so for the sake of simplicity we assume here only unweighted absolute numbers of physical records in the microdata database to be reported (see Section 5.6.4.3 for more details). Obviously, if record weighting plays a significant role for the optimal coverage of the target population in a Member State, the respective national statistical office should adapt its queries for QHC1 accordingly. The national statistical office should also add explanatory textual metadata to item 3.3.1. of the annex to CIR-3.

The example queries for the coverage indicators in QHC1 are:

(d) census population:

The census population is the count of all records in the 'census microdata database' on persons, excluding the records marked as 'deleted':

```
SELECT SEX, AGE_L, COUNT(*)
FROM CENSUS_DB
WHERE REC_DELETED = "N"
GROUP BY SEX, AGE_L;
```

(e) number of all record imputations:

Assuming unweighted absolute counts to be reported (cf. Section 5.6.4.3), this is the simple count of the records flagged as imputed:

```
SELECT SEX, AGE_L, COUNT(*)
FROM CENSUS_DB
WHERE REC_IMPUTED = "Y"
GROUP BY SEX, AGE_L;
```

(f) number of all record deletions:

Again, if there is no weighting, this is the count of the records flagged as deleted:

```
SELECT SEX, AGE_L, COUNT(*)
FROM CENSUS_DB
WHERE REC_DELETED = "Y"
GROUP BY SEX, AGE_L;
```

(g) under-coverage and (e) over-coverage:

As explained earlier, these result from the coverage assessment that is performed after the publication of census data, using the method that is chosen in the Member State to perform this task (for example a post-enumeration survey). If in a Member State:

- all coverage problems have been corrected by record imputations/deletions, or
- the post-enumeration survey does not provide significant results

then these indicators must be '0', and the estimated target population coincides with the census population. We recall that item 3.3.2 of the annex to CIR-3 requires that Member States report on the 'method of assessing under- and over-coverage, including information on the quality of the under- and over-coverage estimates'.

(f) estimated target population:

The estimated target population is computed starting from the census population (a), if estimates of under- and over-coverage are available. In this case, the estimated target population is the sum of the census population plus the under-coverage minus the over-coverage.

Table D5: Some examples how the content of QHC1 could look

		High under-coverage	Imputations	High over-coverage	Deletions	Under- and over-coverage
(a)	**Census population**	700	1 000	1 300	1 000	1 000
(b)	**Imputed records**	0	300	0	0	0
(c)	**Deleted records**	0	0	0	300	0
(d)	**Estimated under-coverage**	300	0	0	0	300
(e)	**Estimated over-coverage**	0	0	300	0	300
(f)	**Estimated target population**	1 000	1 000	1 000	1 000	1 000
	Net coverage	70 %	100 %	130 %	100 %	100 %

D2. How to compute QHC2: Accuracy information on selected topics on persons

As outlined in Table D1, QHC2 has the four dimensions listed below.

- SEX.
- AGE.L.
- TOPIC.: As explained before, the quantitative assessment of data quality should be based on the 'census microdata database' on persons. However, in the case of hypercube QHC2, the quality assessment must be performed on a column-by-column basis, in other words separately for each census topic. The code list for the dimension TOPIC. is presented in Table D4, and coincides with the breakdown codes of the selected topics as laid down in CIR-1.
- TOPIC_IND.: This is the list of accuracy indicators that must be reported to measure how much of the information on a given census topic is based on real observations. A detailed set of guidelines on each of these indicators is provided in the following table, where the column 'Point' refers to the points listed in the annex to CIR-3 under item 4.4 (for the corresponding code list, see Table D3).

Point	Accuracy indicator	Explanation
(a)	Census population	Assuming that the reporting is carried out on one single 'census microdata database' on persons, the 'census population' to be reported here is the same for all topics on persons in question. However, to benchmark the other accuracy indicators it is convenient to report the census population separately for each topic.
(b)	Number of observed data records on the topic derived from traditional census	This is the number of topic observations obtained through a traditional census. Note that, even if the traditional census constitutes the 'census microdata database' forming the basis for reporting QHC2, this number is not necessarily equal to the census population. This is because information on the topic in question might be missing — either due to non-responses on the topic, or (less likely) due to insufficient coverage of the topic in the census questionnaire.
(c)	Number of observed data records on the topic derived from administrative registers	This indicator reflects information on the topic originating from an administrative register, or from several administrative registers that were matched, in the sense of CIR-3, Article 2(24), in the process of compiling the 'census microdata database' on persons.
(d)	Number of observed data records on the topic derived from sample surveys	If information on the topic in question stems from a (series of) sample survey(s), this number reflects the total sample size.

Point	Accuracy indicator	Explanation
(e)	Number of observed data records on the topic derived from multiple data sources	Observed values on the topic in question should be counted to this number only if the information on this topic originates from several different types of 'primary data sources' that were linked, in the sense of CIR-3, Article 2(20), in the process of compiling the 'census microdata database' on persons. However, this categorisation must be made on a value-by-value basis. **Only those individual values where information on the topic from more than one primary data source was combined to obtain them — for example through averaging or similar estimation techniques — should be counted here**. In other words, only if an individual value is definitely based on real observations of the topic but cannot be uniquely counted among any of the other 'observed' indicators (b) — (d), should it be counted here. **Example**: if information on the topic was collected through a register in combination with a sample survey, then all topic values originating only from the register should be counted under (c), while all values coming solely from the survey are counted under (d). Only those individual values where topic information from the register and the survey was combined to obtain them should be counted here.
(f)	Complementary set of statistical units on the topic (for samples)	If information on the topic in question stems from a (series of) sample survey(s), this number gives the size of the complementary set of statistical units of the sample(s), with respect to the census population, as defined in CIR-3, Article 2(10-11) and further described in Section 5.6.4.2, under 'Samples as data sources'. In short: non-zero indicators (f) + (d) should always sum to the census population (a); see 'Samples as data sources' in Section 5.6.4.2.
(g)	Number of imputed observations on the topic	This is the total number of values on the topic that were imputed into the 'census microdata database' on persons, either through 'record imputation' of the entire person record or through 'item imputation' of the topic information to an observed person record. This includes cases where an existing but implausible topic value was replaced by a more plausible one, or where a missing topic value was inferred through estimation methods using additional information available about that person record and/or information from other person records.
(h)	Number of records with missing information on the topic	This is the total number of values on the topic that were missing in the 'census microdata database' on persons and could not be imputed in the sense of indicator (g) above. This indicator should be the sum of all instances of the category 'not stated' under the topic in question (by SEX. and AGE.L.).

The accuracy indicators in QHC2 require that Member States provide, separately for each requested topic, an assessment of the accuracy of the information present in their data source, in other words in the 'census microdata database' on persons. One of the ways to achieve this is to add, for each topic, an additional column to the microdata database. This additional column should indicate, for each person record in the database, the 'status' of the information collected for that topic.

Suppose, for instance, that the topic under study is 'EDU — educational attainment'. The 'census microdata database' could then contain a 'status' variable called EDU_STATUS. Assume further that the possible values for EDU_STATUS directly reflect the available TOPIC_IND categories (cf. table above and also Table D3) listed in the bullet points below.

- 'OBS_TRAD', 'OBS_REGISTER', 'OBS_SURVEY' and 'OBS_MULTI': these categories imply that the value for the topic in question was observed, in other words correctly recorded in the corresponding 'primary data source(s)' from which the information on that topic was drawn (traditional census, administrative register, sample survey or a combination of primary data sources). Note in particular that to be able to provide this information, **the correct status category must be inherited from the 'primary data source(s)'** in the process of compiling the 'census microdata database' on persons (see Section 5.4.1).

- 'COMPSET': the value for the topic in question belongs to the complementary set of a sample survey that was conducted to gain information on that topic, in other words this category should only occur for topics which are measured on a sample basis in the Member State.
- 'IMPU_TOPIC': the value for the topic in question was imputed, in other words not directly observed. 'Imputed' here means either 'item-imputed' into an observed data record, or imputed as a consequence of the fact that the whole record was itself 'record-imputed' (cf. table below). 'Item imputations' should also comprise those cases in which the value stored for that topic comes from an operation of 'data correction' (in other words the value recorded in the 'primary data source' is found to be manifestly wrong, and a new value is stored instead, consistent with the values of other census topics).
- 'MISS_TOPIC': the value for the topic in question is missing, in other words corresponds to the category 'not stated' for the topic in question. Note, however, that a number of topics (for example SEX. and AGE.) **do not have the category 'not stated'** in their code list. In these cases, Member States are **obliged to impute any instance of missing information** for that topic, so that the TOPIC_IND. category cannot be 'MISS_TOPIC'.

Note that the available categories are mutually exclusive by construction, in other words each record on a given topic can only have exactly one of the categories.

The status category 'IMPU_TOPIC' can be used to identify whether the topic in question was 'item-imputed' in the sense of CIR-3, Article 2(18). This is different from the variable 'REC_IMPUTED' introduced in Section 5.6.4.1, which indicates whether the whole record was imputed (CIR-3, Article 2(16)). We recall that 'record imputation' is an operation that results in an increase of the census population, while an item imputation in itself does not change the size of the census population. All possible combinations of record and item imputation are summarised in this table:

Item imputation	Record imputation	
	REC_IMPUTED = "N"	REC_IMPUTED = "Y"
"OBS_[…]"	☑	
"IMPU_TOPIC"	☑*	☑
"MISS_TOPIC"	☑	☑

where the asterisk marks the sole combination fulfilling the definition of 'item imputation' in the sense of CIR-3, Article 2(18). Obviously, a topic cannot be 'observed' if the entire record was imputed, while all other combinations are possible.

The example queries for the coverage indicators in QHC2 are (cf. Table D1):

(a) census population:

As shown in Annex D1, the census population is the number of records in the 'census microdata database' on persons (excluding the 'deleted' ones). It is thus constant for all topics. This redundant information about the 'census population' is included in QHC2 for all selected topics in the TOPIC dimension because it is a very useful reference for the user to compare the values of the other TOPIC_IND indicators in QHC2.

(b) to (h):

Assuming the existence of a 'status' column as described above for each topic included in the TOPIC dimension in the 'census microdata database' on persons, the query to extract all remaining accuracy indicators of QHC2 is straightforward, for example for TOPIC.EDU.:

```
SELECT SEX, AGE_L, EDU_STATUS, COUNT(*)
FROM CENSUS_DB
WHERE REC_DELETED = "N"
GROUP BY SEX, AGE_L, EDU_STATUS;
```

Table D6: Some examples how the content of QHC2 could look

		No sizeable problems	High imputation	Sample survey
(a)	**Census population**	1 000	1 000	1 000
(b)	**Observed data records on the topic (traditional census)**	995	0	0
(c)	**Observed data records on the topic (admin. registers)**	0	**600**	0
(d)	**Observed data records on the topic (sample surveys)**	0	0	**145**
(e)	**Observed data records on the topic (multiple sources)**	0	0	0
(f)	**Complementary set of statistical units on the topic**	0	0	**850**
(g)	**Imputed observations on the topic**	0	**300**	0
(h)	**Records with missing information on the topic**	**5**	**100**	**5**
	% census population	0.5 %	10.0 %	0.5 %

E Extended Euro-SDMX Metadata Structure (ESMS) for textual metadata

The following Table E1 provides an overview of the concepts of the current version v2.0 of the Euro-SDMX Metadata Structure (ESMS) ([78]), as implemented in the ESS Metadata Handler (ESS-MH, see Section 6.2). Moreover, it shows how these concepts are matched to the textual metadata required for the quality reporting under the 2021 census legislation, including notes and/or legal basis where applicable.

As explained in Section 6.2, some rather specific census-related metadata concepts could not be accommodated in the generic ESMS: these were added ad hoc to the structure to facilitate compliance with the legislation, hence 'extended ESMS' in Table E1.

([78]) http://ec.europa.eu/eurostat/data/metadata/metadata-structure.

Table E1: Extended ESMS for the collection of census textual metadata (main concept 'areas' with yellow background, census-specific sub-concepts with light grey background)

	Concept code	Concept name	Breakdown	Representation	Legal basis / Notes
1	CONTACT	Contact		–	This is just a header
1.1	CONTACT_ORGANISATION	Contact organisation		Text	
1.2	ORGANISATION_UNIT	Contact organisation unit		Text	
1.3	CONTACT_NAME	Contact name		Text	
1.4	CONTACT_FUNCT	Contact person function		Text	
1.5	CONTACT_MAIL	Contact mail address		Text	
1.6	CONTACT_EMAIL	Contact email address		Email	
1.7	CONTACT_PHONE	Contact phone number		Telephone	
1.8	CONTACT_FAX	Contact fax number		Fax	Optional
2	META_UPDATE	Metadata update		–	This is just a header
2.1	META_CERTIFIED	Metadata last certified		Date	Date of metadata latest certification
2.2	META_POSTED	Metadata last posted		Date	Date inserted by the system, requested by the annex to CIR-3, item 4.2 and by Annex III to CIR-4
2.3	META_LAST_UPDATE	Metadata last update		Date	Date inserted by the system
3	STAT_PRES	Statistical presentation		–	This is just a header
3.1	DATA_DESCR	Data description		Text (prefilled)	Prefilled text by Eurostat to provide a description of the main characteristics of the data
3.2	CLASS_SYSTEM	Classification system		Text (prefilled)	Prefilled text by Eurostat (ISCED, and so on)
3.3	COVERAGE_SECTOR	Sector coverage		–	Not applicable
3.4	STAT_CONC_DEF	Statistical concepts and definitions	**By TOPIC**	Text	Boxes corresponding to each topic with prefilled text by Eurostat containing the default definition. If needed, Member States can amend the default definition. For some topics, additional information requested by Annex II to CIR-2 and Annex III to CIR-4 is to be submitted here
3.5	STAT_UNIT	Statistical unit		Text (prefilled)	Prefilled text by Eurostat
3.6	STAT_POP	Statistical population		Text (prefilled)	Prefilled text by Eurostat to explain which is the relevant statistical population
3.7	REF_AREA	Reference area		Text	Should explain what is the territorial coverage of census data. Possibly Eurostat should prepare a prefilled text that should serve as a default answer. This question is more relevant for countries (such as France) with lots of semi-independent territories

Table E1 (continued): Extended ESMS for the collection of census textual metadata (main concept 'areas' with yellow background, census-specific sub-concepts with light grey background)

	Concept code	Concept name	Breakdown	Representation	Legal basis / Notes
3.8	COVERAGE_TIME	Time coverage		– (or prefilled text)	Either not applicable or prefilled text like 'the data reflect the situation in the country at the reference date, see item 5'
3.9	BASE_PER	Base period		–	Not applicable
4	UNIT_MEASURE	Unit of measure		Text (prefilled)	Prefilled text by Eurostat: 'counts of statistical units'
5	REF_PERIOD	Reference period		Date	CIR-3-A item 3.1: census reference date
6	INST_MANDATE	Institutional mandate		–	This is just a header
6.1	INST_MAN_LA_OA	Legal acts and other agreements		Text	Legal background as requested by CIR-3-A item 1.1
6.1.1	INST_MAN_RESP	Bodies responsible		Text	Bodies responsible as requested by CIR-3-A item 1.2
6 .2	INST_MAN_SHAR	Data sharing		Text (prefilled)	Prefilled text by Eurostat to explain the ownership status of the data disseminated via the Census Hub
7	CONF	Confidentiality		–	This is just a header
7.1	CONF_POLICY	Confidentiality — policy		Text	Description of the national legislation concerning statistical disclosure control
7.2	CONF_DATA_TR	Confidentiality — data treatment		Text	Information requested by CIR-3-A item 3.4
7.2.1	CONF_DATA_TR_GRID	Confidentiality — 1 km^2 grid data		Text	Information on the confidentiality treatment of grid data, as requested by Annex III to CIR-4
8	REL_POLICY	Release policy		–	This is just a header
8.1	REL_CAL_POLICY	Release calendar		Text (prefilled)	Prefilled text by Eurostat explaining the provisions of the census legislation concerning the data transmission to Eurostat
8.2	REL_CAL_ACCESS	Release calendar access		–	Not applicable
8.3	REL_POL_US_AC	User access		–	Not applicable
9	FREQ_DISS	Frequency of dissemination		Text (prefilled)	Prefilled text by Eurostat: 'every decade'
10	ACCESSIBILITY_CLARITY	Accessibility and clarity		–	This is just a header

Table E1 (continued): Extended ESMS for the collection of census textual metadata (main concept 'areas' with yellow background, census-specific sub-concepts with light grey background)

	Concept code	Concept name	Breakdown	Representation	Legal basis / Notes
10.1	NEWS_REL	News release		Text	If applicable, national press release(s) concerning the 2021 census, as part of the requirements of CIR-3-A item 3.4
10.2	PUBLICATIONS	Publications		Text	National publications concerning the 2021 census, as part of the requirements of the annex to CIR-3, item 3.4
10.3	ONLINE_DB	On-line database		Text	National on-line database on the 2021 census, as part of the requirements of the annex to CIR-3, item 3.4
10.4	MICRO_DAT_ACC	Microdata access		Text	Whether census data are accessible as microdata, as part of the requirements of the annex to CIR-3, item 3.4
10.5	DISS_OTHER	Other		Text	Other dissemination channels, as part of the requirements of the annex to CIR-3, item 3.4
10.6	DOC_METHOD	Documentation on methodology		Text	Prefilled text by Eurostat guiding to the ESS-MH files on the Census Hub
10.7	QUALITY_DOC	Quality documentation		Text	Prefilled text by Eurostat
11	QUALITY_MGMNT	Quality management		–	This is just a header
11.1	QUALITY_ASSURE	Quality assurance		Text	To be provided on a voluntary basis, where applicable
11.2	QUALITY_ASSMNT	Quality assessment		Text	Procedures to ensure quality assessment, as requested by the annex to CIR-3, item 3.3.2
11.2.1	COVERAGE_ASSMNT	Coverage assessment		Text	Procedures to assess coverage, as requested by the annex to CIR-3, item 3.3.2
11.2.2	POST_ENUM_SURVEY	Post-enumeration survey(s)		Text	If applicable, post-enumeration survey(s) as requested by the annex to CIR-3, item 3.3.2
12	RELEVANCE	Relevance		–	This is just a header
12.1	USER_NEEDS	User needs		Text (prefilled)	Prefilled text by Eurostat about the actions taken to identify and fulfil user needs, as requested by the annex to CIR-3, item 4.6
12.2	USER_SAT	User satisfaction		Text (prefilled)	Prefilled text by Eurostat about the tools to monitor data extractions and receive feedback from users, as requested by the annex to CIR-3, item 4.6
12 .3	COMPLETENESS	Completeness		Text	Degree of completeness of the data in terms of the requirements of the CFR, as requested by the annex to CIR-3, item 4.5

Table E1 (continued): Extended ESMS for the collection of census textual metadata (main concept 'areas' with yellow background, census-specific sub-concepts with light grey background)

	Concept code	Concept name	Breakdown	Representation	Legal basis / Notes
13	ACCURACY	Accuracy and reliability		–	This is just a header
13.1	ACCURACY_OVERALL	Overall accuracy	**By TOPIC**	Text	Reasons for unreliability of the data on the topic, as requested by Annex II to CIR-2 and Annex III to CIR-4
13.1.1	ACCURACY_GRID	Accuracy of grid data on topics	**By TOPIC**	Text	Information on the accuracy of grid data by topic, as requested by Annex III to CIR-4
13.1.2	GEO_QUALITY	Geographic information — data quality		Text	Information on geographic quality principles, as requested by Annex III to CIR-4
13.2	SAMPLING_ERR	Sampling error		–	Not applicable
13.3	NONSAMPLING_ERR	Non-sampling error		–	Not applicable
14	TIMELINESS_PUNCT	Timeliness and punctuality		–	This is just a header
14.1	TIMELINESS	Timeliness		Text (table)	Calendar date(s) by hypercube of the transmission of census data, as requested by the annex to CIR-3, item 4.2 and Annex III to CIR-4
14.2	PUNCTUALITY	Punctuality		–	Not applicable
15	COHER_COMPAR	Coherence and comparability		–	This is just a header
15.1	COMPAR_GEO	Comparability — geographical		Text	National practices which could impair the EU-wide comparability of the data, as requested by the annex to CIR-3, item 4.1
15.2	COMPAR_TIME	Comparability — over time		–	Not applicable
15.3	COHER_X_DOM	Coherence — cross domain		Text (prefilled)	Prefilled text by Eurostat explain how census data relate to other social statistics
15.3	COHER_INTERNAL	Coherence — internal		Text	Information on any significant inconsistencies between the hypercubes, as requested by the annex to CIR-3, item 4.3
16	COST_BURDEN	Cost and burden		Text	Measures to ensure cost effectiveness, as requested by the annex to CIR-3, item 3.5
17	DATA_REV	Data revision		–	This is just a header
17.1	REV_POLICY	Data revision — policy		Text	If applicable, to be provided on a voluntary basis
17.2	REV_PRACTICE	Data revision — practice		Text (table)	Calendar date(s) of major revision(s) by hypercube, as requested by the annex to CIR-3, item 4.2 and Annex III to CIR-4

Table E1 (continued): Extended ESMS for the collection of census textual metadata (main concept 'areas' with yellow background, census-specific sub-concepts with light grey background)

	Concept code	Concept name	Breakdown	Representation	Legal basis / Notes
18	STAT_PROCESS	Statistical processing		–	This is just a header
18.1	SOURCE_TYPE	Source data		–	This is just a header
18.1.1	SOURCE_LIST	List of data sources	**By STAT_UNIT**	Text	List of census data sources as requested by the annex to CIR-3, item 2.2
18.1.2	SOURCE_CLASS	Classification of data sources	**By STAT_UNIT**	Codelist	Classification of the data sources as requested by the annex to CIR-3, item 2.1
18.1.3	SOURCE_X_TOPICS	'Data sources x topics' matrix		Matrix of tick boxes	Requested by the annex to CIR-3, item 2.3
18.1.4	SOURCE_ADEQUACY	Adequacy of data sources		Matrix of tick boxes	Requested by the annex to CIR-3, item 2.4
18 .2	FREQ_COLL	Frequency of data collection		Text (prefilled)	Prefilled text by Eurostat: 'every decade'
18.3	COLL_METHOD	Data collection		–	This is just a header
18.3.1	COLL_METHOD_QUEST	Questionnaire-based data		Text	Information on questionnaire-based data as requested by the annex to CIR-3, item 3.2.1
18.3.2	COLL_METHOD_REG	Register-based data		Text	Information on register-based data as requested by the annex to CIR-3, item 3.2.2
18.3.3	COLL_METHOD_SAMPLE	Sample data		Text	Information on sample data as requested by the annex to CIR-3, item 3.2.3
18.3.4	COLL_METHOD_MULT	Data from combined methods		Text	Information on data from combined methods as requested by the annex to CIR-3, item 3.2.4
18.4	DATA_VALIDATION	Data validation		Text	Procedures implemented for final data validation, as requested by the annex to CIR-3, item 3.3.2
18.5	DATA_COMP	Data compilation		Text	Steps involved in data processing, as requested by the annex to CIR-3, item 3.3.1
18.5.1	DATA_COMP_GRID	Compilation of 1 km^2 grid data			Methodology applied to produce the 1 km^2 grid dataset, as requested by Annex III to CIR-4
18.6	ADJUSTMENT	Adjustment		–	Not applicable
19	COMMENT_DSET	Comment		Text	Voluntary

F EU regulations on the 2021 population and housing census

Acronym	Complete regulation title and legal reference
CFR	Regulation (EC) No 763/2008 of the European Parliament and of the Council of 9 July 2008 on population and housing censuses (OJ L 218, 13.8.2008, p. 14)
CIR-1	Commission Implementing Regulation (EU) 2017/543 of 22 March 2017 implementing Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses as regards the technical specifications of the topics and of their breakdowns (OJ L 78, 23.3.2017, p. 13)
CIR-2	Commission Regulation (EU) 2017/712 of 20 April 2017 establishing the reference year and the programme of the statistical data and metadata for population and housing censuses provided for by Regulation (EC) No 763/2008 of the European Parliament and of the Council (OJ L 105, 21.4.2017, p. 1)
CIR-3	Commission Implementing Regulation (EU) 2017/881 of 23 May 2017 implementing Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses, as regards the modalities and structure of the quality reports and the technical format for data transmission, and amending Regulation (EU) No 1151/2010 (OJ L 135, 24.5.2017, p. 6)
CIR-4	Commission Implementing Regulation (EU) 2018/1799 of 21 November 2018 on the establishment of a temporary direct statistical action for the dissemination of selected topics of the 2021 population and housing census geocoded to a 1 km^2 grid (OJ L 296, 22.11.2018, p. 19)

REGULATION (EC) No 763/2008 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

of 9 July 2008

on population and housing censuses

(Text with EEA relevance)

THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION,

Having regard to the Treaty establishing the European Community, and in particular Article 285(1) thereof,

Having regard to the proposal from the Commission,

Acting in accordance with the procedure laid down in Article 251 of the Treaty ([1]),

Whereas:

(1) The Commission (Eurostat) needs to be in possession of sufficiently reliable, detailed and comparable data on the population and housing, in order to enable the Community to fulfil the tasks assigned to it, in particular by Articles 2 and 3 of the Treaty. Sufficient comparability must be ensured at Community level as regards methodology, definitions and the programme of the statistical data and the metadata.

(2) Periodic statistical data on the population and the main family, social, economic and housing characteristics of persons are necessary for the study and definition of regional, social and environmental policies affecting particular sectors of the Community. In particular, there is a need to collect detailed information on housing in support of various Community activities, such as the promotion of social inclusion and the monitoring of social cohesion at regional level, or the protection of the environment and the promotion of energy efficiency.

(3) In view of methodological and technological developments, best practices should be identified and the enhancement of the data sources and methodologies used for censuses in the Member States should be fostered.

(4) In order to ensure the comparability of the data provided by the Member States and for reliable overviews to be drawn up at Community level, the data used should refer to the same reference year.

([1]) Opinion of the European Parliament of 20 February 2008 (not yet published in the Official Journal) and Council Decision of 23 June 2008.

(5) In accordance with Council Regulation (EC) No 322/97 of 17 February 1997 on Community Statistics ([2]), which constitutes the reference framework for the provisions of this Regulation, it is necessary for the collection of statistics to conform to the principles of impartiality, in particular objectivity and scientific independence, as well as transparency, reliability, relevance, cost-effectiveness and statistical confidentiality.

(6) The transmission of data subject to statistical confidentiality is governed by Regulation (EC) No 322/97 and Council Regulation (Euratom, EEC) No 1588/90 of 11 June 1990 on the transmission of data subject to statistical confidentiality to the Statistical Office of the European Communities ([3]). Measures that are taken in accordance with those Regulations ensure the physical and logical protection of confidential data and that no unlawful disclosure or non-statistical use occurs when Community statistics are produced and disseminated.

(7) In the production and dissemination of Community statistics under this Regulation, the national and Community statistical authorities should take account of the principles set out in the European Statistics Code of Practice adopted on 24 February 2005 by the Statistical Programme Committee, established by Council Decision 89/382/EEC, Euratom ([4]) and attached to the Recommendation of the Commission on the independence, integrity and accountability of the national and Community statistical authorities.

(8) Since the objectives of this Regulation, namely the collection and compilation of comparable and comprehensive Community statistics on population and housing, cannot be sufficiently achieved by the Member States, due to the absence of common statistical features and quality requirements as well as a lack of methodological transparency, and can therefore, by way of a common statistical framework, be better achieved at Community level, the Community may adopt measures, in accordance with the principle of subsidiarity as set out in Article 5 of the Treaty. In accordance with the principle of proportionality, as set out in that Article, this Regulation does not go beyond what is necessary in order to achieve those objectives.

([2]) OJ L 52, 22.2.1997, p. 1. Regulation as amended by Regulation (EC) No 1882/2003 of the European Parliament and of the Council (OJ L 284, 31.10.2003, p. 1).

([3]) OJ L 151, 15.6.1990, p. 1. Regulation as last amended by Regulation (EC) No 1882/2003.

([4]) OJ L 181, 28.6.1989, p. 47.

(9) The measures necessary for the implementation of this Regulation should be adopted in accordance with Council Decision 1999/468/EC of 28 June 1999 laying down the procedures for the exercise of implementing powers conferred on the Commission ([1]).

(10) In particular, the Commission should be empowered to establish the conditions for the establishment of subsequent reference years and the adoption of the programme of the statistical data and the metadata. Since those measures are of general scope and are designed to amend non-essential elements of this Regulation, *inter alia*, by supplementing it with new non-essential elements, they must be adopted in accordance with the regulatory procedure with scrutiny provided for in Article 5a of Decision 1999/468/EC.

(11) The Statistical Programme Committee has been consulted in accordance with Article 3 of Decision 89/382/EEC, Euratom,

HAVE ADOPTED THIS REGULATION:

Article 1

Subject matter

This Regulation establishes common rules for the decennial provision of comprehensive data on population and housing.

Article 2

Definitions

For the purpose of this Regulation, the following definitions shall apply:

(a) 'population' shall mean the national, regional and local population at its usual residence at the reference date;

(b) 'housing' shall mean living quarters and buildings as well as housing arrangements and the relationship between the population and living quarters at the national, regional and local levels at the reference date;

(c) 'buildings' shall mean permanent buildings that contain living quarters designed for human habitation, or conventional dwellings that are reserved for seasonal or secondary use or that are vacant;

(d) 'usual residence' shall mean the place where a person normally spends the daily period of rest, regardless of temporary absences for purposes of recreation, holidays, visits to friends and relatives, business, medical treatment or religious pilgrimage.

The following persons alone shall be considered to be usual residents of the geographical area in question:

(i) those who have lived in their place of usual residence for a continuous period of at least 12 months before the reference date; or

(ii) those who arrived in their place of usual residence during the 12 months before the reference date with the intention of staying there for at least one year.

Where the circumstances described in point (i) or (ii) cannot be established, 'usual residence' shall mean the place of legal or registered residence;

(e) 'reference date' shall mean the date to which the data of the respective Member State refer, in accordance with Article 5(1);

(f) 'national' shall mean on the territory of a Member State;

(g) 'regional' shall mean at NUTS level 1, NUTS level 2 or NUTS level 3, as defined in the classification of territorial units for statistics (NUTS), established by Regulation (EC) No 1059/2003 of the European Parliament and of the Council ([2]) in its version applicable at the reference date;

(h) 'local' shall mean at Local Administrative Units level 2 (LAU level 2);

(i) 'essential features of population and housing censuses' shall mean individual enumeration, simultaneity, universality within a defined territory, availability of small-area data and defined periodicity.

Article 3

Data submission

Member States shall submit to the Commission (Eurostat) data on the population covering determined demographic, social and economic characteristics of persons, families and households, as well as on housing at a national, regional and local level, as set out in the Annex.

Article 4

Data sources

1. Member States may base the statistics on different data sources, in particular on:

(a) conventional censuses;

(b) register-based censuses;

(c) a combination of conventional censuses and sample surveys;

(d) a combination of register-based censuses and sample surveys;

(e) a combination of register-based censuses and conventional censuses;

([1]) OJ L 184, 17.7.1999, p. 23. Decision as amended by Decision 2006/512/EC (OJ L 200, 22.7.2006, p. 11).

([2]) OJ L 154, 21.6.2003, p. 1. Regulation as last amended by Regulation (EC) No 176/2008 of the European Parliament and of the Council (OJ L 61, 5.3.2008, p. 1).

(f) a combination of register-based censuses, sample surveys and conventional censuses; and

(g) appropriate surveys with rotating samples (rolling censuses).

2. Member States shall take all measures necessary to meet the requirements of data protection. The Member States' own data protection provisions shall not be affected by this Regulation.

3. Member States shall inform the Commission (Eurostat) of any revision or correction of the statistics supplied under this Regulation, as well as of any changes in the chosen data sources and methodology, no later than one month before the release of the revised data.

4. Member States shall ensure that the data sources and the methodology used to satisfy the requirements of this Regulation meet, to the highest possible extent, the essential features of population and housing censuses, as defined in Article 2(i). They shall make continuous efforts to enhance compliance with those essential features.

Article 5

Data transmission

1. Each Member State shall determine a reference date. The reference date shall fall in a year specified on the basis of this Regulation (reference year). The first reference year shall be 2011. The Commission (Eurostat) shall establish subsequent reference years in accordance with the regulatory procedure with scrutiny referred to in Article 8(3). Reference years shall fall during the beginning of every decade.

2. Member States shall provide the Commission (Eurostat) with final, validated and aggregated data and with metadata, as required by this Regulation, within 27 months of the end of the reference year.

3. The Commission (Eurostat) shall adopt a programme of the statistical data and of the metadata to be transmitted to fulfil the requirements of this Regulation, in accordance with the regulatory procedure with scrutiny referred to in Article 8(3).

4. The Commission (Eurostat) shall adopt the technical specifications of the topics as required by this Regulation as well as of their breakdowns, in accordance with the regulatory procedure referred to in Article 8(2).

5. Member States shall transmit to the Commission (Eurostat) the validated data and metadata in electronic form. The Commission (Eurostat) shall adopt the appropriate technical format to be used for the transmission of the required data, in accordance with the regulatory procedure referred to in Article 8(2).

6. In the event of a revision or correction in accordance with Article 4(3), Member States shall transmit the modified data to the Commission (Eurostat) no later than on the date of release of the revised data.

Article 6

Quality assessment

1. For the purpose of this Regulation, the following quality assessment dimensions shall apply to the data to be transmitted:

— 'relevance' shall refer to the degree to which statistics meet the current and potential needs of users,

— 'accuracy' shall refer to the closeness of estimates to the unknown true values,

— 'timeliness' and 'punctuality' shall refer to the delay between the reference period and the availability of results,

— 'accessibility' and 'clarity' shall refer to the conditions under and modalities by which users can obtain, use and interpret data,

— 'comparability' shall refer to the measurement of the impact of differences in applied statistical concepts and measurement tools and procedures when statistics are compared between geographical areas, sectoral domains, or over time, and

— 'coherence' shall refer to the adequacy of the data to be reliably combined in different ways and for various uses.

2. Member States shall provide the Commission (Eurostat) with a report on the quality of the data transmitted. In this context, Member States shall report on the extent to which the chosen data sources and methodology meet the essential features of population and housing censuses as defined in Article 2(i).

3. In applying the quality assessment dimensions laid down in paragraph 1 to the data covered by this Regulation, the modalities and structure of the quality reports shall be defined in accordance with the regulatory procedure referred to in Article 8(2). The Commission (Eurostat) shall assess the quality of the data transmitted.

4. The Commission (Eurostat), in cooperation with the competent authorities of the Member States, shall provide methodological recommendations designed to ensure the quality of the data and metadata produced, acknowledging, in particular, the Conference of European Statisticians Recommendations for the 2010 Censuses of Population and Housing.

Article 7

Implementing measures

1. The following measures necessary for the implementation of this Regulation shall be adopted in accordance with the regulatory procedure referred to in Article 8(2):

(a) technical specifications of the topics as required by this Regulation as well as of their breakdowns as provided for in Article 5(4);

(b) the establishment of the appropriate technical format as provided for in Article 5(5); and

(c) modalities and structure of the quality reports as provided for in Article 6(3).

2. The following measures necessary for the implementation of this Regulation, designed to amend non-essential elements of this Regulation by supplementing it, shall be adopted in accordance with the regulatory procedure with scrutiny referred to in Article 8(3):

(a) the establishment of the reference years, as provided for in Article 5(1); and

(b) the adoption of the programme of the statistical data and the metadata, as provided for in Article 5(3).

3. Consideration shall be given to the principles that the benefits of the measures taken must outweigh their costs and that additional costs and burdens must remain within a reasonable limit.

Article 8

Committee procedure

1. The Commission shall be assisted by the Statistical Programme Committee.

2. Where reference is made to this paragraph, Articles 5 and 7 of Decision 1999/468/EC shall apply, having regard to the provisions of Article 8 thereof.

The period laid down in Article 5(6) of Decision 1999/468/EC shall be set at three months.

3. Where reference is made to this paragraph, Article 5a(1) to (4) and Article 7 of Decision 1999/468/EC shall apply, having regard to the provisions of Article 8 thereof.

Article 9

Entry into force

This Regulation shall enter into force on the 20th day following its publication in the *Official Journal of the European Union*.

This Regulation shall be binding in its entirety and directly applicable in all Member States.

Done at Strasbourg, 9 July 2008.

For the European Parliament	*For the Council*
The President	*The President*
H.-G. PÖTTERING	J.-P. JOUYET

ANNEX

Topics to be covered in Population and Housing Censuses

1. Population topics

1.1. Obligatory topics for the geographical levels: NUTS 3, LAU 2

1.1.1. Non-derived topics

— Place of usual residence,

— sex,

— age,

— legal marital status,

— country/place of birth,

— country of citizenship,

— previous place of usual residence and date of arrival in the current place; or place of usual residence one year prior to the census,

— relationships between household members

1.1.2. Derived topics

— Total population,

— locality,

— household status,

— family status,

— type of family nucleus,

— size of family nucleus,

— type of private household,

— size of private household

1.2. Obligatory topics for the geographical levels: national level, NUTS 1, NUTS 2

1.2.1. Non-derived topics

— Place of usual residence,

— location of place of work,

— sex,

— age,

— legal marital status,

— current activity status,

— occupation,

— industry (branch of economic activity),

— status in employment,

— educational attainment,

— country/place of birth,

— country of citizenship,

— ever resided abroad and year of arrival in the country (from 1980),

— previous place of usual residence and date of arrival in the current place; or place of usual residence one year prior to the census,

— relationships between household members,

— tenure status of households

1.2.2. Derived topics

— Total population,

— locality,

— household status,

— family status,

— type of family nucleus,

— size of family nucleus,

— type of private household,

— size of private household

2. Housing topics

2.1. Obligatory topics for the geographical levels: NUTS 3, LAU 2

2.1.1. Non-derived topics

— Type of living quarters,

— location of living quarters,

— occupancy status of conventional dwellings,

— number of occupants,

— useful floor space and/or number of rooms of housing units,

— dwellings by type of building,

— dwellings by period of construction

2.1.2. Derived topics

— Density standard

2.2. Obligatory topics for the geographical levels: national level, NUTS 1, NUTS 2

2.2.1. Non-derived topics

— Housing arrangements,

— type of living quarters,

— location of living quarters,

— occupancy status of conventional dwellings,

— type of ownership,

— number of occupants,

— useful floor space and/or number of rooms of housing units,

— water supply system,

— toilet facilities,

— bathing facilities,

— type of heating,

— dwellings by type of building,

— dwellings by period of construction

2.2.2. Derived topics

— Density standard

COMMISSION IMPLEMENTING REGULATION (EU) 2017/543

of 22 March 2017

laying down rules for the application of Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses as regards the technical specifications of the topics and of their breakdowns

(Text with EEA relevance)

THE EUROPEAN COMMISSION,

Having regard to the Treaty on the Functioning of the European Union,

Having regard to Regulation (EC) No 763/2008 of the European Parliament and of the Council of 9 July 2008 on population and housing censuses ([1]), and in particular Article 5(4) thereof,

Whereas:

(1) To ensure that data from the population and housing censuses conducted in the Member States are comparable, and to allow reliable Union-wide overviews to be drawn up, the census topics need to be established and broken down in the same way in all Member States. Technical specifications for those topics and their breakdowns should therefore be adopted.

(2) The measures provided for in this Regulation are in accordance with the opinion of the European Statistical System Committee,

HAS ADOPTED THIS REGULATION:

Article 1

The technical specifications for the census topics and their breakdowns which are to be applied to the data to be sent to the Commission for the reference year 2021 are listed in the Annex.

Article 2

This Regulation shall enter into force on the twentieth day following that of its publication in the *Official Journal of the European Union*.

This Regulation shall be binding in its entirety and directly applicable in all Member States.

Done at Brussels, 22 March 2017.

For the Commission

The President

Jean-Claude JUNCKER

([1]) OJ L 218, 13.8.2008, p. 14.

ANNEX

Technical specifications of the census topics and their breakdowns

The technical specifications are presented as follows:

— Each topic is identified by a heading.

— The heading of the topic may be followed by technical specifications referring to that topic in general.

— Thereafter, the breakdown(s) for the topic is (are) specified. Some topics have more than one breakdown, each with different levels of detail. Where this is the case, 'H' identifies breakdowns with the highest level of detail, 'M' identifies breakdowns with a medium level of detail, and 'L' identifies breakdowns with the lowest level of detail and 'N' identifies the breakdown that refers to the national level.

— The totals to which the breakdowns apply are identified. Each breakdown may be followed by further technical specifications that relate specifically to that breakdown.

Topic: Place of usual residence

In applying the definition of 'usual residence' given in Article 2(d) of Regulation (EC) No 763/2008, Member States shall treat special cases as follows:

(a) Where a person regularly lives in more than one residence during the year, the residence where he/she spends the majority of the year shall be taken as his/her place of usual residence regardless of whether this is located elsewhere within the country or abroad. However, a person who works away from home during the week and who returns to the family home at weekends shall consider the family home to be his/her place of usual residence regardless of whether his/her place of work is elsewhere in the country or abroad.

(b) Primary and secondary school pupils and students who are away from home during the school term and regardless of how often they return to their family home, shall consider their family home to be their place of usual residence (regardless of whether they are pursuing their education elsewhere in the country or abroad).

(c) Tertiary students who are away from home while at college or university shall consider their term-time address to be their place of usual residence regardless of whether this is an institution (such as a boarding school) or a private residence and regardless of whether they are pursuing their education elsewhere in the country or abroad. Exceptionally, where the place of education is within the country, the place of usual residence may be considered to be the family home.

(d) An institution shall be taken as the place of usual residence of all its residents who at the time of the census have spent, or are likely to spend, 12 months or more living there.

(e) The general rule in relation to where most of the daily period of rest is spent applies to persons doing compulsory military service and to members of the armed forces who live in military barracks or camps.

(f) The place of enumeration shall be taken as the place of usual residence of homeless or roofless persons, nomads, vagrants and persons with no concept of usual residence.

(g) A child who alternates between two places of residence (for instance if his or her parents are divorced) shall consider the one where he or she spends the majority of the time as his or her place of usual residence. Where an equal amount of time is spent with both parents the place of usual residence shall be the one where the child is found at the time on census night or, alternatively, the household where the child has his or her legal or registered residence.

(h) Merchant seamen and fishermen usually resident in the country but at sea at the time of the census (including those who have no place of residence other than their quarters aboard ship) shall be included.

(i) Persons who may be irregularly staying or undocumented, as well as asylum seekers and persons who have applied for, or been granted, refugee status or similar types of international protections, provided that they meet the criteria for the usual residence in the country shall be included. The intention is not to distinguish these persons separately, but rather to ensure that they are not missed from the enumeration.

(j) Children born in the 12 months before the census reference time and whose families are usually resident in the country at the census reference time shall be included.

(k) Persons whose stay in the country (actual and/or intended) is exactly one year shall be included.

Military, naval and diplomatic personnel and their families

(l) Foreign military, naval and diplomatic personnel and their families, located in the country, regardless of their duration of stay shall be excluded from the usually resident population of a country.

(m) Where the duration of residence outside of the country can be established for national military, naval and diplomatic service personnel and their families located outside the country, the following shall be applied:

— if they are residing abroad for less than 12 months and they are intending to return to the place of departure they shall be allocated within the country in accordance with the rules for usual residence. In particular, they could be allocated to (by decreasing order of priority):

(i) the family home address within the country, if any, or

(ii) the duty station within the country to which they were attached before leaving.

— If they are residing abroad for at least 12 months or if they are not intending to return to the place of departure (although returning in the country within a 12-month period), they shall be attributed to a 'virtual place' (extra-region) of the country of departure.

On the basis of the definition of the place of usual residence, persons usually resident in the place of enumeration but absent, or expected to be absent, at the time of the census for less than one year shall be considered as temporarily absent persons and thus included in the total population. In contrast, persons living or expected to live outside the place of enumeration for one year or more shall not be considered temporarily absent and shall therefore be excluded from the total population. This is regardless of the length of visits that they may pay to their families from time to time.

Persons who are enumerated but do not meet the criteria for usual residence in the place of enumeration, i.e. do not live or do not expect to live in the place of enumeration for a continuous period of at least 12 months, are considered temporarily present and are therefore not counted in the total usual resident population.

Geographical area [1]					**GEO.N.**	**GEO.L.**	**GEO.M.**	**GEO.H.**
0.	Total (in the territory of the Member State)				0.	0.	0.	0.
x.	All NUTS 1 regions in the Member State					x.	x.	x.
	x.x.	All NUTS 2 regions in the Member State				x.x.	x.x.	x.x.
		x.x.x.	All NUTS 3 regions in the Member State				x.x.x.	x.x.x.
			x.x.x.x.	All LAU 2 regions in the Member State				x.x.x.x.

[1] The codes 'x.', 'x.x.' and 'x.x.x.' depend on the NUTS classification, the code 'x.x.x.x.' on the LAU classification, valid for the Member State on 1 January 2021. The annotation 'N' identifies the breakdown that refers to the national level.

The breakdowns for 'Geographical area' are designed to break down any total or subtotal referring to persons (Place of usual residence). They can also be used to regionally break down any total to which neither the topic 'Place of usual residence' nor 'Location of place of work' applies.

For the breakdowns for 'Geographical area' the versions of the classification of territorial units for statistics (NUTS [1]) and of the classification for Local Administrative Units (LAU) valid on 1 January 2021 shall be used.

Topic: Location of place of work

The location of the place of work is the geographical area in which a currently employed person does his/her job.

[1] In accordance with Regulation (EC) No 1059/2003, all Member States' statistics that are transmitted to the Commission and that are to be broken down by territorial units should use the NUTS classification. Consequently, in order to establish comparable regional statistics, data on the territorial units should be provided in accordance with the NUTS classification. (Regulation (EC) No 1059/2003 of the European Parliament and of the Council of 26 May 2003 on the establishment of a common classification of territorial units for statistics (NUTS) (OJ L 154, 21.6.2003, p. 1).

The place of work of those mostly working at home is the same as their usual residence. The term 'working' refers to work done as an 'employed person' as defined under the topic 'Current activity status'. 'Mostly' working at home means that the person spends all or most of the time working at home, and less, or no, time in a place of work other than at home.

For the breakdowns for 'Location of place of work' the version of the classification of territorial units for statistics (NUTS) valid on 1 January 2021 shall be used.

Information on persons who do not have a fixed place of work but who report to a fixed address at the beginning of their work period (for example bus drivers, airline crew, operators of street market stalls that are not removed at the end of the workday) should refer to that address. This group may also include individuals who travel to work, on a regular basis, across the border to a neighbouring country. Breakdown 'No fixed place of work (inside or outside the Member State)' includes all persons without fixed place of work but will also refer to persons such as sailors, fishermen and offshore workers for whom it may not be possible to allocate the place of work.

Location of place of work [1]				**LPW.N.**	**LPW.L.**
0.	Total			0.	0.
1.	In the territory of the Member State			1.	1.
	1.x.	All NUTS 1 regions in the Member State			1.x.
		1.x.x.	All NUTS 2 regions in the Member State		1.x.x.
	1.y.	Unknown place of work in the Member State			1.y.
2.	Not in the territory of the Member State			2.	2.
3.	No fixed place of work (inside or outside the Member State)			3.	3.
4.	Unknown place of work (unknown if inside or outside the Member State)			4.	4.
5.	Not applicable (not working)			5.	5.

[1] The codes '1.x.' and '1.x.x.' depend on the NUTS classification valid for the Member State on 1 January 2021. The annotation 'N' identifies the breakdown that refers to the national level.

The breakdowns for 'Location of place of work' are designed to break down any total or subtotal referring to persons.

Topic: Locality

A locality is defined as a distinct population cluster, that is an area defined by population living in neighbouring or contiguous buildings. Such buildings may either:

(a) form a continuous built-up area with a clearly recognisable street formation; or

(b) though not part of such a built-up area, comprise a group of buildings to which a locally recognized place name is uniquely attached; or

(c) though not meeting either of the above two criteria, constitute a group of buildings, none of which is separated from its nearest neighbour by more than 200 meters.

In applying this definition, certain land-use categories shall not be regarded as breaking the continuity of a built-up area. These categories include: industrial and commercial buildings and facilities, public parks, playgrounds and gardens, football fields and other sports facilities, bridged rivers, railway lines, canals, parking lots and other transport infrastructure, churchyards and cemeteries.

LAU 2 regions with a total population of less than 2 000 can be assumed to be one locality.

Population of a locality is defined as persons having their usual residence in that locality.

A scattered building is to be allocated to the category that represents the number of persons that have their usual residence in the building.

Size of the locality		LOC.
0.	Total	0.
1.	1 000 000 and more persons	1.
2.	500 000 — 999 999 persons	2.
3.	200 000 — 499 999 persons	3.
4.	100 000 — 199 999 persons	4.
5.	50 000 — 99 999 persons	5.
6.	20 000 — 49 999 persons	6.
7.	10 000 — 19 999 persons	7.
8.	5 000 — 9 999 persons	8.
9.	2 000 — 4 999 persons	9.
10.	1 000 — 1 999 persons	10.
11.	500 — 999 persons	11.
12.	200 — 499 persons	12.
13.	less than 200 persons	13.

The breakdown 'Size of the locality' is designed to break down any total or subtotal of units that can be located in 'localities', including any total or subtotal referring to persons.

Topic: Sex

Sex		SEX.
0.	Total	0.
1.	Male	1.
2.	Female	2.

The breakdown 'Sex' is designed to break down any total or subtotal referring to persons.

Topic: Age

The age reached in completed years of age at the reference date shall be reported.

Age				AGE.L.	AGE.M.	AGE.H.
0.	Total			0.	0.	0.
1.	under 15 years			1.	1.	1.
	1.1.	under 5 years			1.1.	1.1.
		1.1.1	under 1 year			1.1.1
		1.1.2.	1 year			1.1.2.

Age				AGE.L.	AGE.M.	AGE.H.
		1.1.3.	2 years			1.1.3.
		1.1.4.	3 years			1.1.4.
		1.1.5.	4 years			1.1.5.
	1.2.	5 to 9 years			1.2.	1.2.
		1.2.1.	5 years			1.2.1.
		1.2.2.	6 years			1.2.2.
		1.2.3.	7 years			1.2.3.
		1.2.4.	8 years			1.2.4.
		1.2.5.	9 years			1.2.5.
	1.3.	10 to 14 years			1.3.	1.3.
		1.3.1.	10 years			1.3.1.
		1.3.2.	11 years			1.3.2.
		1.3.3.	12 years			1.3.3.
		1.3.4.	13 years			1.3.4.
		1.3.5.	14 years			1.3.5.
2.	15 to 29 years			2.	2.	2.
	2.1.	15 to 19 years			2.1.	2.1.
		2.1.1.	15 years			2.1.1.
		2.1.2.	16 years			2.1.2.
		2.1.3.	17 years			2.1.3.
		2.1.4.	18 years			2.1.4.
		2.1.5.	19 years			2.1.5.
	2.2.	20 to 24 years			2.2.	2.2.
		2.2.1.	20 years			2.2.1.
		2.2.2.	21 years			2.2.2.
		2.2.3.	22 years			2.2.3.
		2.2.4.	23 years			2.2.4.
		2.2.5.	24 years			2.2.5.
	2.3.	25 to 29 years			2.3.	2.3.
		2.3.1.	25 years			2.3.1.
		2.3.2.	26 years			2.3.2.
		2.3.3.	27 years			2.3.3.

Age				AGE.L.	AGE.M.	AGE.H.
		2.3.4.	28 years			2.3.4.
		2.3.5.	29 years			2.3.5.
3.	30 to 49 years			3.	3.	3.
	3.1.	30 to 34 years			3.1.	3.1.
		3.1.1.	30 years			3.1.1.
		3.1.2.	31 years			3.1.2.
		3.1.3.	32 years			3.1.3.
		3.1.4.	33 years			3.1.4.
		3.1.5.	34 years			3.1.5.
	3.2.	35 to 39 years			3.2.	3.2.
		3.2.1.	35 years			3.2.1.
		3.2.2.	36 years			3.2.2.
		3.2.3.	37 years			3.2.3.
		3.2.4.	38 years			3.2.4.
		3.2.5.	39 years			3.2.5.
	3.3.	40 to 44 years			3.3.	3.3.
		3.3.1.	40 years			3.3.1.
		3.3.2.	41 years			3.3.2.
		3.3.3.	42 years			3.3.3.
		3.3.4.	43 years			3.3.4.
		3.3.5.	44 years			3.3.5.
	3.4.	45 to 49 years			3.4.	3.4.
		3.4.1.	45 years			3.4.1.
		3.4.2.	46 years			3.4.2.
		3.4.3.	47 years			3.4.3.
		3.4.4.	48 years			3.4.4.
		3.4.5.	49 years			3.4.5.
4.	50 to 64 years			4.	4.	4.
	4.1.	50 to 54 years			4.1.	4.1.
		4.1.1.	50 years			4.1.1.
		4.1.2.	51 years			4.1.2.
		4.1.3.	52 years			4.1.3.

Age				AGE.L.	AGE.M.	AGE.H.
		4.1.4.	53 years			4.1.4.
		4.1.5.	54 years			4.1.5.
	4.2.	55 to 59 years			4.2.	4.2.
		4.2.1.	55 years			4.2.1.
		4.2.2.	56 years			4.2.2.
		4.2.3.	57 years			4.2.3.
		4.2.4.	58 years			4.2.4.
		4.2.5.	59 years			4.2.5.
	4.3.	60 to 64 years			4.3.	4.3.
		4.3.1.	60 years			4.3.1.
		4.3.2.	61 years			4.3.2.
		4.3.3.	62 years			4.3.3.
		4.3.4.	63 years			4.3.4.
		4.3.5.	64 years			4.3.5.
5.	65 to 84 years			5.	5.	5.
	5.1.	65 to 69 years			5.1.	5.1.
		5.1.1.	65 years			5.1.1.
		5.1.2.	66 years			5.1.2.
		5.1.3.	67 years			5.1.3.
		5.1.4.	68 years			5.1.4.
		5.1.5.	69 years			5.1.5.
	5.2.	70 to 74 years			5.2.	5.2.
		5.2.1.	70 years			5.2.1.
		5.2.2.	71 years			5.2.2.
		5.2.3.	72 years			5.2.3.
		5.2.4.	73 years			5.2.4.
		5.2.5.	74 years			5.2.5.
	5.3.	75 to 79 years			5.3.	5.3.
		5.3.1.	75 years			5.3.1.
		5.3.2.	76 years			5.3.2.
		5.3.3.	77 years			5.3.3.
		5.3.4.	78 years			5.3.4.
		5.3.5.	79 years			5.3.5.

Age				AGE.L.	AGE.M.	AGE.H.
	5.4.	80 to 84 years			5.4.	5.4.
		5.4.1.	80 years			5.4.1.
		5.4.2.	81 years			5.4.2.
		5.4.3.	82 years			5.4.3.
		5.4.4.	83 years			5.4.4.
		5.4.5.	84 years			5.4.5.
6.	85 years and over			6.	6.	6.
	6.1.	85 to 89 years			6.1.	6.1.
		6.1.1.	85 years			6.1.1.
		6.1.2.	86 years			6.1.2.
		6.1.3.	87 years			6.1.3.
		6.1.4.	88 years			6.1.4.
		6.1.5.	89 years			6.1.5.
	6.2.	90 to 94 years			6.2.	6.2.
		6.2.1.	90 years			6.2.1.
		6.2.2.	91 years			6.2.2.
		6.2.3.	92 years			6.2.3.
		6.2.4.	93 years			6.2.4.
		6.2.5.	94 years			6.2.5.
	6.3.	95 to 99 years			6.3.	6.3.
		6.3.1.	95 years			6.3.1.
		6.3.2.	96 years			6.3.2.
		6.3.3.	97 years			6.3.3.
		6.3.4.	98 years			6.3.4.
		6.3.5.	99 years			6.3.5.
	6.4.	100 years and over			6.4.	6.4.

The breakdowns for 'Age' are designed to break down any total or subtotal referring to persons.

Topic: Legal Marital Status

Marital status is defined as the (legal) conjugal status of an individual in relation to the marriage laws (or customs) of the country (that is the *de jure* status).

A person shall be classified according to his/her most recently acquired legal marital status at the reference date.

In Member States where the legislation includes provisions for married partners or partners in registered partnership to be 'legally separated', such 'legally separated' persons shall be classified under 'Married or in registered partnership' (LMS.L. 2. and LMS.H. 2.).

Legal Marital Status			LMS.L.	LMS.H.
0.	Total		0.	0.
1.	Never married and never in a registered partnership		1.	1.
2.	Married or in registered partnership		2.	2.
	2.1.	In an opposite-sex marriage or registered partnership		2.1.
	2.2.	In a same-sex marriage or registered partnership		2.2.
3.	Widowed or registered partnership ended with the death of partner (and not re-married or in a registered partnership)		3.	3.
4.	Divorced or registered partnership legally dissolved (and not remarried or in a registered partnership)		4.	4.
5.	Not stated		5.	5.

The breakdown 'Legal marital status' is designed to break down any total or subtotal referring to persons.

Topic: Current activity status

'Current activity status' is the current relationship of a person to economic activity, based on a reference period of one week, which may be either a specified, recent, fixed, calendar week, or the last complete calendar week, or the last seven days prior to enumeration.

The 'labour force' comprises all persons who fulfil the requirements for inclusion among the employed or the unemployed.

'*Employed*' persons comprise all persons aged 15 years or over who during the reference week:

(a) performed at least one hour of work for pay or profit, in cash or in kind, or

(b) were temporarily absent from a job in which they had already worked and to which they maintained a formal attachment, or from a self-employment activity.

Employees temporarily not at work shall be considered as in paid employment provided they had a formal job attachment. The possible reasons for such temporary absences are:

(a) illness or injury; or

(b) holiday or vacation; or

(c) strike or lock-out; or

(d) educational or training leave; or

(e) maternity or parental leave; or

(f) reduction in economic activity; or

(g) temporary disorganization or suspension of work due to such reasons as bad weather, mechanical or electrical breakdown, or shortage of raw materials or fuels; or

(h) other temporary absence with or without leave.

The formal job attachment shall be determined on the basis of one or more of the following criteria:

(a) a continued receipt of wage or salary; or

(b) an assurance of return to work following the end of the contingency, or an agreement as to the date of return; or

(c) the elapsed duration of absence from the job which, wherever relevant, may be that duration for which workers can receive compensation benefits without obligations to accept other jobs.

Self-employed persons shall be considered as 'employed' if they have worked as such during the reference week or if they are temporarily absent from work and their enterprise meanwhile continues to exist.

Contributing family workers shall be considered to be 'employed' at work on the same basis as other employed persons; that is irrespective of the number of hours worked during the reference period. Similarly, persons who perform tasks or duties of an employee job held by a family member living in the same, or in another, household shall also be classified as employed.

The '*unemployed*' comprise all persons aged 15 years or over who were:

(a) 'without work', that is, were not in wage employment or self-employment during the reference week; and

(b) 'currently available for work', that is, were available for wage employment or self-employment during the reference week and for two weeks after that; and

(c) 'seeking work', that is, had taken specific steps to seek wage employment or self-employment within four weeks ending with the reference week.

The 'Others' comprise persons outside of the labour force who are receiving public aid or private support, and all other persons not falling into any of the above categories.

In ascribing a single activity status to each person, priority shall be given to the status of 'Employed' in preference to 'Unemployed', and to the status of 'Unemployed' in preference to 'Outside of the labour force'.

In ascribing a single activity status to each person currently outside of the labour force, priority shall be given to the status of 'Persons below the national minimum age for economic activity' in preference to 'Pension or capital income recipients', to the status of 'Pension or capital income recipients' in preference to 'Students', and of 'Students' in preference to 'Others'.

The category 'Students' (CAS.H.2.3.) shall thus comprise secondary and tertiary students who:

— have attained the national minimum age for economic activity or above, and

— are outside of the labour force, and

— are not recipients of a pension or of capital income..

Current activity status			**CAS.L.**	**CAS.H.**
0.	Total		0.	0.
1.	Labour force		1.	1.
	1.1.	Employed	1.1.	1.1.
	1.2.	Unemployed	1.2.	1.2.
2.	Outside of the labour force		2.	2.
	2.1.	Persons below the national minimum age for economic activity		2.1.
	2.2.	Pension or capital income recipients		2.2.
	2.3.	Students		2.3.
	2.4.	Others		2.4.
3.	Not stated		3.	3.

The breakdowns for 'Current activity status' are designed to break down any total or subtotal referring to persons.

Topic: Occupation

'Occupation' refers to the type of work done in a job. 'Type of work' is described by the main tasks and duties of the work.

The allocation of a person within the breakdowns of the topics 'Occupation', 'Industry' and 'Status in employment' shall be based on the same job. Persons doing more than one job shall be allocated an occupation based on their main job, which is to be identified according to:

(1) the time spent on the job or, if not available,

(2) the income received.

Persons aged 15 or over that were employed (i.e. had the 'Current activity status — CAS of "Employed"' (CAS.L. and CAS.H. 1.1)) during the reference week shall be classified under only one category of OCC.1. to OCC.11.

Persons under the age of 15 years, as well as persons aged 15 or over that were:

— unemployed during the reference week ('Current activity status — "Unemployed"' (CAS.L. 1.2)) or that were

— outside of the labour force ('Current activity status — "outside of the labour force"' (CAS.L. and CAS.H.2.)) shall be classified under 'Not applicable' (OCC.12.).

If the denomination of categories of the ISCO classification in force on 1 January 2021 deviates from that listed in the categories OCC.2. to OCC.11., the denomination of the ISCO classification in force on 1 January 2021 shall be used.

	Occupation	**OCC.**
0.	Total	0.
1.	Managers	1.
2.	Professionals	2.
3.	Technicians and associate professionals	3.
4.	Clerical support workers	4.
5.	Service and sales workers	5.
6.	Skilled agricultural, forestry, and fishery workers	6.
7.	Craft and related trades workers	7.
8.	Plant and machine operators, and assemblers	8.
9.	Elementary occupations	9.
10.	Armed forces occupations	10.
11.	Not stated	11.
12.	Not applicable	12.

The breakdown 'Occupation' is designed to break down any total or subtotal referring to persons.

Topic: Industry

Industry (branch of economic activity) refers to the kind of production or activity of the establishment or similar unit in which the job of an employed person is located. For persons who are recruited and employed by one enterprise but who actually have their place of work in another enterprise ('agency workers', 'seconded workers') the industry (branch of economic activity) of the establishment or similar unit where the place of work actually is shall be reported.

The allocation of a person within the breakdowns of the topics '*Occupation*', '*Industry*' and '*Status in employment*' shall be based on the same job. Persons doing more than one job shall be allocated an industry (branch of economic activity) based on their main job which is to be identified according to:

— the time spent on the job or, if not available,

— the income received.

Persons aged 15 or over that were employed (i.e. had the 'Current activity status — CAS' of 'employed' (CAS.L. and CAS.H. 1.1)) during the reference week shall be classified under only one category of IND.1. to IND.11.

Persons under the age of 15 years, as well as persons aged 15 or over that were:

— unemployed during the reference week ('Current activity status — "unemployed"' (CAS.L. and CAS.H.1.2)) or that were

— outside of the labour force ('Current activity status — "outside of the labour force"' (CAS.L. and CAS.H.2.)) shall be classified under 'Not applicable' (IND.12.).

Industry			IND.L.	IND.H.
0.	Total		0.	0.
1.	Agriculture, forestry and fishing		1.	1.
2	Manufacturing, mining and quarrying, and other industry		2.	2.
	2.1.	Mining and quarrying		2.1.
	2.2.	Manufacturing		2.2.
	2.3.	Electricity, gas, steam and air conditioning supply		2.3.
	2.4	Water supply; sewerage, waste management and remediation activities		2.4.
3.	Construction		3.	3.
4.	Wholesale and retail trade, transportation and storage, accommodation and food service activities		4.	4.
	4.1	Wholesale and retail trade; repair of motor vehicles and motorcycles		4.1.
	4.2	Transportation and storage		4.2.
	4.3	Accommodation and food service activities		4.3.
5.	Information and communication		5.	5.
6.	Financial and insurance activities		6.	6.
7.	Real estate activities		7.	7.
8.	Professional, scientific, technical, administrative and support service activities		8.	8.
	8.1.	Professional, scientific and technical activities		8.1.
	8.2.	Administrative and support service activities		8.2.
9.	Public administration, defence, education, human health and social work activities		9.	9.
	9.1.	Public administration and defence; compulsory social security		9.1.
	9.2.	Education		9.2.
	9.3.	Human health and social work activities		9.3.

Industry			IND.L.	IND.H.
10.	Other services		10.	10.
	10.1.	Arts, entertainment and recreation		10.1.
	10.2.	Other service activities		10.2.
	10.3.	Activities of households as employers; undifferentiated goods- and services producing activities of households for own use		10.3.
	10.4.	Activities of extraterritorial organisations and bodies		10.4.
11.	Not stated		11.	11.
12.	Not applicable		12.	12.

If the denomination of categories of the NACE classification in force on 1 January 2021 deviates from that listed in the categories IND.1. to IND.10., the denomination of the NACE classification in force on 1 January 2021 shall be used.

The breakdown 'Industry (branch of economic activity)' is designed to break down any total or subtotal referring to persons.

Topic: Status in employment

An 'employee' is a person who works in a 'paid employment' job, that is a job where the explicit or implicit contract of employment gives the incumbent a basic remuneration, which is independent of the revenue of the unit for which he/she works (this unit may be a corporation, a non-profit institution, government unit or a household). Persons in 'paid employment' jobs are typically remunerated by wages and salaries, but may be paid by commission from sales, by piece rates, bonuses or in-kind payment such as food, housing or training. Some or all of the tools, capital equipment, information systems and/or premises used by the incumbent may be owned by others, and the incumbent may work under direct supervision of, or according to strict guidelines set by, the owner(s) or persons in the owner's employment.

An 'employer' is a person who, working on his or her own account or with a small number of partners, holds a 'self-employment' job and, in this capacity, on a continuous basis (including the reference week) has engaged one or more persons to work for him/her as 'employees'. The incumbent makes the operational decisions affecting the enterprise, or delegates such decisions while retaining responsibility for the welfare of the enterprise.

If a person is both employer and employee, he/she shall be allocated to only one group according to:

— the time spent on the job or, if not available,

— the income received.

An 'own-account worker' is a person who, working on his/her own account or with one or a few partners, holds a 'self-employment job' and has not engaged, on a continuous basis (including the reference week), any 'employees'.

'Other employed persons' includes persons who are 'contributing family workers' and 'members of producers' cooperatives'.

A 'contributing family worker' is a person who

— holds a 'self-employment' job in a market-oriented establishment operated by a related person, living in the same household, and

— cannot be regarded as a partner (that is an employer or own-account worker) because the degree of commitment to the operation of the establishment, in terms of working time or other factors to be determined by national circumstances, is not at a level comparable to that of the head of the establishment.

A 'member of a producers' cooperative' is a person who holds a 'self-employment' job in an establishment organised as a cooperative, in which each member takes part on an equal footing with other members in determining the organisation of production, sales and/or other work, the investments and the distribution of the proceeds among the members.

The allocation of a person within the breakdowns of the topics 'Occupation', 'Industry' and 'Status in employment' shall be based on the same job. Persons doing more than one job shall be allocated a status in employment based on their main job, which is to be identified according to:

— the time spent on the job or, if not available,

— the income received.

Persons aged 15 or over that were employed (i.e. had the 'Current activity status — CAS' of 'employed' (CAS.L. and CAS.H.1.1)) during the reference week shall be classified under only one category of SIE.1. to SIE.5., according to their status in employment.

Persons under the age of 15 years, as well as persons aged 15 or over that were:

— unemployed during the reference week ('Current activity status — "unemployed"' (CAS.L.1.2. and CAS.H. 1.2)) or that were

— outside of the labour force ('Current activity status — "outside of the labour force"' (CAS.L.2. and CAS.H.2.)) shall be classified under 'Not applicable' (SIE.6.).

Status in employment		SIE.
0.	Total	0.
1.	Employees	1.
2.	Employers	2.
3.	Own-account workers	3.
4.	Other employed persons	4.
5.	Not stated	5.
6.	Not applicable	6.

The breakdown 'Status in employment' is designed to break down any total or subtotal referring to persons.

Topic: Educational attainment

Educational attainment refers to the highest level successfully completed in the educational system of the country where the education was received. All education which is relevant to the completion of a level shall be taken into account even if this was provided outside schools and universities.

Persons aged 15 years or over shall be classified under only one of the categories from EDU.1. to EDU.10., according to their educational attainment (highest completed level). Persons under the age of 15 years shall be classified under 'Not applicable' (EDU.11.).

EDU.1. refers to persons not having successfully completed ISCED level 1. This includes individuals who: have never attended an education programme; have attended some early childhood education (defined as ISCED level 0 in the classification of education programmes); or have attended some primary education but have not successfully completed ISCED level 1.

If the denomination of categories of the ISCED classification in force on 1 January 2021 deviates from that listed in the categories EDU.2. to EDU.9., the denomination of the ISCED classification in force on 1 January 2021 shall be used.

Educational attainment (highest completed level)		EDU.
0.	Total	0.
1.	ISCED level 0: Less than primary education	1.
2.	ISCED level 1: Primary education	2.
3.	ISCED level 2: Lower secondary education	3.

Educational attainment (highest completed level)		EDU.
4.	ISCED level 3: Upper secondary education	4.
5.	ISCED level 4: Post-secondary non-tertiary education	5.
6.	ISCED level 5: Short-cycle tertiary education	6.
7.	ISCED level 6: Bachelor's or equivalent level;	7.
8.	ISCED level 7: Master's or equivalent level;	8.
9.	ISCED level 8: Doctoral or equivalent level.	9.
10.	Not stated (of the persons aged 15 years or over)	10.
11.	Not applicable (persons under 15 years of age)	11.

The breakdown 'Educational attainment (highest completed level)' is designed to break down any total or subtotal referring to persons.

Topic: Country/place of birth

Information on the 'Place of birth' shall be collected according to the place of usual residence of the mother at the time of the birth, or, if not available, the place in which the birth took place.

Information on the country of birth shall be collected on the basis of international boundaries existing on 1 January 2021.

'EU Member State' means a country that is a member of the European Union on 1 January 2021.

Country/place of birth				POB.L.	POB.M.	POB.H.
0.	Total			0.	0.	0.
1.	Place of birth in reporting country			1.	1.	1.
2.	Place of birth not in reporting country			2.	2.	2.
	2.1.	Other EU Member State		2.1.	2.1.	2.1.
		2.1.01.	Belgium			2.1.01.
		2.1.02.	Bulgaria			2.1.02.
		2.1.03.	Czech Republic			2.1.03.
		2.1.04.	Denmark			2.1.04.
		2.1.05.	Germany			2.1.05.
		2.1.06.	Estonia			2.1.06.
		2.1.07.	Ireland			2.1.07.
		2.1.08.	Greece			2.1.08.
		2.1.09.	Spain			2.1.09.
		2.1.10.	France			2.1.10.
		2.1.11.	Croatia			2.1.11.
		2.1.12.	Italy			2.1.12.
		2.1.13.	Cyprus			2.1.13.

Country/place of birth					POB.L.	POB.M.	POB.H.
		2.1.14.	Latvia				2.1.14.
		2.1.15.	Lithuania				2.1.15.
		2.1.16.	Luxembourg				2.1.16.
		2.1.17.	Hungary				2.1.17.
		2.1.18.	Malta				2.1.18.
		2.1.19.	Netherlands				2.1.19.
		2.1.20.	Austria				2.1.20.
		2.1.21.	Poland				2.1.21.
		2.1.22.	Portugal				2.1.22.
		2.1.23.	Romania				2.1.23.
		2.1.24.	Slovenia				2.1.24.
		2.1.25.	Slovakia				2.1.25.
		2.1.26.	Finland				2.1.26.
		2.1.27.	Sweden				2.1.27.
		2.1.28.	United Kingdom				2.1.28.
	2.2.	Elsewhere			2.2.	2.2.	2.2.
		2.2.1.	Elsewhere within Europe			2.2.1.	2.2.1.
			2.2.1.01.	Albania			2.2.1.01.
			2.2.1.02.	Andorra			2.2.1.02.
			2.2.1.03.	Belarus			2.2.1.03.
			2.2.1.04.	Former Yugoslav Republic of Macedonia, The			2.2.1.04.
			2.2.1.05.	Iceland			2.2.1.05.
			2.2.1.06.	Kosovo (*)			2.2.1.06.
			2.2.1.07.	Liechtenstein			2.2.1.07.
			2.2.1.08.	Moldova			2.2.1.08.
			2.2.1.09.	Monaco			2.2.1.09.
			2.2.1.10..	Montenegro			2.2.1.10..
			2.2.1.11.	Norway			2.2.1.11.
			2.2.1.12.	Bosnia and Herzegovina			2.2.1.12.
			2.2.1.13.	Russian Federation			2.2.1.13.
			2.2.1.14.	San Marino			2.2.1.14.
			2.2.1.15.	Serbia			2.2.1.15.

Country/place of birth					POB.L.	POB.M.	POB.H.
			2.2.1.16.	Switzerland			2.2.1.16.
			2.2.1.17.	Turkey			2.2.1.17.
			2.2.1.18.	Ukraine			2.2.1.18.
			2.2.1.19.	Vatican City State			2.2.1.19.
			2.2.1.20.	Faroe Islands			2.2.1.20.
			2.2.1.21.	Gibraltar			2.2.1.21.
			2.2.1.22.	Guernsey			2.2.1.22.
			2.2.1.23.	Isle of Man			2.2.1.23.
			2.2.1.24.	Jersey			2.2.1.24.
			2.2.1.25.	Sark			2.2.1.25.
			2.2.1.26.	Other country in Europe			2.2.1.26.
		2.2.2.	Africa			2.2.2.	2.2.2.
			2.2.2.01.	Algeria			2.2.2.01.
			2.2.2.02.	Angola			2.2.2.02.
			2.2.2.03.	Benin			2.2.2.03.
			2.2.2.04.	Botswana			2.2.2.04.
			2.2.2.05.	Burkina Faso			2.2.2.05.
			2.2.2.06.	Burundi			2.2.2.06.
			2.2.2.07.	Cameroon			2.2.2.07.
			2.2.2.08.	Cape Verde			2.2.2.08.
			2.2.2.09.	Central African Republic			2.2.2.09.
			2.2.2.10.	Chad			2.2.2.10.
			2.2.2.11.	Comoros			2.2.2.11.
			2.2.2.12.	Congo			2.2.2.12.
			2.2.2.13.	Côte d'Ivoire			2.2.2.13.
			2.2.2.14.	Democratic Republic of the Congo			2.2.2.14.
			2.2.2.15.	Djibouti			2.2.2.15.
			2.2.2.16.	Egypt			2.2.2.16.
			2.2.2.17.	Equatorial Guinea			2.2.2.17.
			2.2.2.18.	Eritrea			2.2.2.18.
			2.2.2.19.	Ethiopia			2.2.2.19.
			2.2.2.20.	Gabon			2.2.2.20.

Country/place of birth					POB.L.	POB.M.	POB.H.
			2.2.2.21.	Gambia			2.2.2.21.
			2.2.2.22.	Ghana			2.2.2.22.
			2.2.2.23.	Guinea			2.2.2.23.
			2.2.2.24.	Guinea-Bissau			2.2.2.24.
			2.2.2.25.	Kenya			2.2.2.25.
			2.2.2.26.	Lesotho			2.2.2.26.
			2.2.2.27.	Liberia			2.2.2.27.
			2.2.2.28.	Libya			2.2.2.28.
			2.2.2.29.	Madagascar			2.2.2.29.
			2.2.2.30.	Malawi			2.2.2.30.
			2.2.2.31.	Mali			2.2.2.31.
			2.2.2.32.	Mauritania			2.2.2.32.
			2.2.2.33.	Mauritius			2.2.2.33.
			2.2.2.34.	Morocco			2.2.2.34.
			2.2.2.35.	Mozambique			2.2.2.35.
			2.2.2.36.	Namibia			2.2.2.36.
			2.2.2.37.	Niger			2.2.2.37.
			2.2.2.38.	Nigeria			2.2.2.38.
			2.2.2.39.	Rwanda			2.2.2.39.
			2.2.2.40.	Saint Helena			2.2.2.40.
			2.2.2.41.	Sao Tome and Principe			2.2.2.41.
			2.2.2.42.	Senegal			2.2.2.42.
			2.2.2.43.	Seychelles			2.2.2.43.
			2.2.2.44.	Sierra Leone			2.2.2.44.
			2.2.2.45.	Somalia			2.2.2.45.
			2.2.2.46.	South Africa			2.2.2.46.
			2.2.2.47.	Sudan			2.2.2.47.
			2.2.2.48.	South Sudan			2.2.2.48.
			2.2.2.49.	Swaziland			2.2.2.49.
			2.2.2.50.	Togo			2.2.2.50.
			2.2.2.51.	Tunisia			2.2.2.51.
			2.2.2.52.	Uganda			2.2.2.52.

Country/place of birth					POB.L.	POB.M.	POB.H.
			2.2.2.53.	Tanzania			2.2.2.53.
			2.2.2.54.	Western Sahara			2.2.2.54.
			2.2.2.55.	Zambia			2.2.2.55.
			2.2.2.56.	Zimbabwe			2.2.2.56.
			2.2.2.57.	Other country in Africa			2.2.2.57.
		2.2.3.	Caribbean, South or Central America			2.2.3.	2.2.3.
			2.2.3.01.	Anguilla			2.2.3.01.
			2.2.3.02.	Antigua and Barbuda			2.2.3.02.
			2.2.3.03.	Argentina			2.2.3.03.
			2.2.3.04.	Aruba			2.2.3.04.
			2.2.3.05.	Bahamas			2.2.3.05.
			2.2.3.06.	Barbados			2.2.3.06.
			2.2.3.07.	Belize			2.2.3.07.
			2.2.3.08.	Bolivia			2.2.3.08.
			2.2.3.09.	Brazil			2.2.3.09.
			2.2.3.10.	British Virgin Islands			2.2.3.10.
			2.2.3.11.	Cayman Islands			2.2.3.11.
			2.2.3.12.	Chile			2.2.3.12.
			2.2.3.13.	Colombia			2.2.3.13.
			2.2.3.14.	Costa Rica			2.2.3.14.
			2.2.3.15.	Cuba			2.2.3.15.
			2.2.3.16.	Curaçao			2.2.3.16.
			2.2.3.17.	Dominica			2.2.3.17.
			2.2.3.18.	Dominican Republic			2.2.3.18.
			2.2.3.19.	Ecuador			2.2.3.19.
			2.2.3.20.	El Salvador			2.2.3.20.
			2.2.3.21.	Falkland Islands (Malvinas)			2.2.3.21.
			2.2.3.22.	Grenada			2.2.3.22.
			2.2.3.23.	Guatemala			2.2.3.23.
			2.2.3.24.	Guyana			2.2.3.24.
			2.2.3.25.	Haiti			2.2.3.25.
			2.2.3.26.	Honduras			2.2.3.26.

Country/place of birth					POB.L.	POB.M.	POB.H.
			2.2.3.27.	Jamaica			2.2.3.27.
			2.2.3.28.	Mexico			2.2.3.28.
			2.2.3.29.	Montserrat			2.2.3.29.
			2.2.3.30.	Nicaragua			2.2.3.30.
			2.2.3.31.	Panama			2.2.3.31.
			2.2.3.32.	Paraguay			2.2.3.32.
			2.2.3.33.	Peru			2.2.3.33.
			2.2.3.34.	Saint Barthelemy			2.2.3.34.
			2.2.3.35.	Saint Kitts and Nevis			2.2.3.35.
			2.2.3.36.	Saint Lucia			2.2.3.36.
			2.2.3.37.	Saint Martin (FR)			2.2.3.37.
			2.2.3.38.	St Maarten (NL)			2.2.3.38.
			2.2.3.39.	Saint Vincent and the Grenadines			2.2.3.39.
			2.2.3.40.	Suriname			2.2.3.40.
			2.2.3.41.	Trinidad and Tobago			2.2.3.41.
			2.2.3.42.	Turks and Caicos Islands			2.2.3.42.
			2.2.3.43.	Uruguay			2.2.3.43.
			2.2.3.44.	Venezuela			2.2.3.44.
			2.2.3.45.	Other country in the Caribbean, South or Central America			2.2.3.45.
		2.2.4.	North America			2.2.4.	2.2.4.
			2.2.4.01.	Canada			2.2.4.01.
			2.2.4.02.	Greenland			2.2.4.02.
			2.2.4.03.	United States of America			2.2.4.03.
			2.2.4.04.	Bermuda			2.2.4.04.
			2.2.4.05.	Saint Pierre and Miquelon			2.2.4.05.
			2.2.4.06.	Other country in North America			2.2.4.06.
		2.2.5.	Asia			2.2.5.	2.2.5.
			2.2.5.01.	Afghanistan			2.2.5.01.
			2.2.5.02.	Armenia			2.2.5.02.
			2.2.5.03.	Azerbaijan			2.2.5.03.
			2.2.5.04.	Bahrain			2.2.5.04.
			2.2.5.05.	Bangladesh			2.2.5.05.

Country/place of birth					POB.L.	POB.M.	POB.H.
			2.2.5.06.	Bhutan			2.2.5.06.
			2.2.5.07.	Brunei Darussalam			2.2.5.07.
			2.2.5.08.	Cambodia			2.2.5.08.
			2.2.5.09.	China			2.2.5.09.
			2.2.5.10.	Georgia			2.2.5.10.
			2.2.5.11.	India			2.2.5.11.
			2.2.5.12.	Indonesia			2.2.5.12.
			2.2.5.13.	Iraq			2.2.5.13.
			2.2.5.14.	Iran			2.2.5.14.
			2.2.5.15.	Israel			2.2.5.15.
			2.2.5.16.	Japan			2.2.5.16.
			2.2.5.17.	Jordan			2.2.5.17.
			2.2.5.18.	Kazakhstan			2.2.5.18.
			2.2.5.19.	North Korea			2.2.5.19.
			2.2.5.20.	South Korea			2.2.5.20.
			2.2.5.21.	Kuwait			2.2.5.21.
			2.2.5.22.	Kyrgyzstan			2.2.5.22.
			2.2.5.23.	Laos			2.2.5.23.
			2.2.5.24.	Lebanon			2.2.5.24.
			2.2.5.25.	Malaysia			2.2.5.25.
			2.2.5.26.	Maldives			2.2.5.26.
			2.2.5.27.	Mongolia			2.2.5.27.
			2.2.5.28.	Myanmar			2.2.5.28.
			2.2.5.29.	Nepal			2.2.5.29.
			2.2.5.30.	Oman			2.2.5.30.
			2.2.5.31.	Pakistan			2.2.5.31.
			2.2.5.32.	Philippines			2.2.5.32.
			2.2.5.33.	Qatar			2.2.5.33.
			2.2.5.34.	Saudi Arabia			2.2.5.34.
			2.2.5.35.	Singapore			2.2.5.35.
			2.2.5.36.	Sri Lanka			2.2.5.36.
			2.2.5.37.	Syria			2.2.5.37.

Country/place of birth					POB.L.	POB.M.	POB.H.
			2.2.5.38.	Taiwan			2.2.5.38.
			2.2.5.39.	Tajikistan			2.2.5.39.
			2.2.5.40.	Thailand			2.2.5.40.
			2.2.5.41.	Timor-Leste			2.2.5.41.
			2.2.5.42.	Turkmenistan			2.2.5.42.
			2.2.5.43.	United Arab Emirates			2.2.5.43.
			2.2.5.44.	Uzbekistan			2.2.5.44.
			2.2.5.45.	Vietnam			2.2.5.45.
			2.2.5.46.	Palestine			2.2.5.46.
			2.2.5.47.	Yemen			2.2.5.47.
			2.2.5.48.	Other country in Asia			2.2.5.48.
		2.2.6.	Oceania			2.2.6.	2.2.6.
			2.2.6.01.	Australia			2.2.6.01.
			2.2.6.02.	Federated States of Micronesia			2.2.6.02.
			2.2.6.03.	Cook Islands (NZ)			2.2.6.03.
			2.2.6.04.	Fiji			2.2.6.04.
			2.2.6.05.	French Polynesia			2.2.6.05.
			2.2.6.06.	French Southern Territories			2.2.6.06.
			2.2.6.07.	Kiribati			2.2.6.07.
			2.2.6.08.	Marshall Islands			2.2.6.08.
			2.2.6.09.	Nauru			2.2.6.09.
			2.2.6.10.	New Caledonia			2.2.6.10.
			2.2.6.11.	New Zealand			2.2.6.11.
			2.2.6.12.	Palau			2.2.6.12.
			2.2.6.13.	Papua New Guinea			2.2.6.13.
			2.2.6.14.	Samoa			2.2.6.14.
			2.2.6.15.	Solomon Islands			2.2.6.15.
			2.2.6.16.	Tonga			2.2.6.16.
			2.2.6.17.	Tuvalu			2.2.6.17.
			2.2.6.18.	Pitcairn			2.2.6.18.
			2.2.6.19.	Vanuatu			2.2.6.19.

Country/place of birth					POB.L.	POB.M.	POB.H.
			2.2.6.20.	Wallis and Futuna Islands			2.2.6.20.
			2.2.6.21.	Other country in Oceania			2.2.6.21.
3.	Other				3.	3.	3.
4.	Not stated				4.	4.	4.

(*) This designation is without prejudice to positions on status, and is in line with UNSCR 1244/1999 and the ICJ Opinion on the Kosovo declaration of independence.

The breakdowns for 'Country/place of birth' are designed to break down any total or subtotal referring to persons.

The list of countries in the breakdown 'Country/place of birth' shall only apply for statistical purposes.

Topic: Country of citizenship

Citizenship is defined as the particular legal bond between an individual and his/her State, acquired by birth or naturalisation, whether by declaration, option, marriage or other means according to the national legislation.

A person with two or more citizenships shall be allocated to only one country of citizenship, to be determined in the following order of precedence:

1. reporting country; or

2. if the person does not have the citizenship of the reporting country: other EU Member State; or

3. if the person does not have the citizenship of another EU Member State: other country outside the European Union.

Where there are cases of dual citizenship where both countries are within the European Union but neither is the reporting country, Member States shall determine which country of citizenship is to be allocated.

'EU Member State' means a country that is a member of the European Union on 1 January 2021.

For reporting countries that are EU Member States, the sub-category of the category 'Citizenship not of reporting country, but other EU Member State' (COC.H.2.1.) that refers to their Member State does not apply. For reporting countries that are not EU Member States, the category 'Citizenship not of reporting country, but other EU Member State' (COC.L.2.1. COC.M.2.1. COC.H.2.1.) shall be changed to 'Citizenship of an EU Member State'.

Persons who are neither citizens of any country nor stateless and who have some but not all of the rights and duties associated with citizenship shall be classified under 'Recognised Non-Citizens' (COC.H. 2.2.1.20.).

Country of citizenship				COC.L.	COC.M.	COC.H.
0.	Total			0.	0.	0.
1.	Citizenship of reporting country			1.	1.	1.
2.	Citizenship not of reporting country			2.	2.	2.
	2.1.	Citizenship not of reporting country, but other EU Member State		2.1.	2.1.	2.1.
		2.1.01.	Belgium			2.1.01.
		2.1.02.	Bulgaria			2.1.02.
		2.1.03.	Czech Republic			2.1.03.
		2.1.04.	Denmark			2.1.04.
		2.1.05.	Germany			2.1.05.

Country of citizenship					COC.L.	COC.M.	COC.H.
		2.1.06.	Estonia				2.1.06.
		2.1.07.	Ireland				2.1.07.
		2.1.08.	Greece				2.1.08.
		2.1.09.	Spain				2.1.09.
		2.1.10.	France				2.1.10.
		2.1.11.	Croatia				2.1.11.
		2.1.12.	Italy				2.1.12.
		2.1.13.	Cyprus				2.1.13.
		2.1.14.	Latvia				2.1.14.
		2.1.15.	Lithuania				2.1.15.
		2.1.16.	Luxembourg				2.1.16.
		2.1.17.	Hungary				2.1.17.
		2.1.18.	Malta				2.1.18.
		2.1.19.	Netherlands				2.1.19.
		2.1.20.	Austria				2.1.20.
		2.1.21.	Poland				2.1.21.
		2.1.22.	Portugal				2.1.22.
		2.1.23.	Romania				2.1.23.
		2.1.24.	Slovenia				2.1.24.
		2.1.25.	Slovakia				2.1.25.
		2.1.26.	Finland				2.1.26.
		2.1.27.	Sweden				2.1.27.
		2.1.28.	United Kingdom				2.1.28.
	2.2.	Citizenship of country not member of the EU			2.2.	2.2.	2.2.
		2.2.1.	Other European country			2.2.1.	2.2.1.
			2.2.1.01.	Albania			2.2.1.01.
			2.2.1.02.	Andorra			2.2.1.02.
			2.2.1.03.	Belarus			2.2.1.03.
			2.2.1.04.	Former Yugoslav Republic of Macedonia, The			2.2.1.04.
			2.2.1.05.	Iceland			2.2.1.05.
			2.2.1.06.	Kosovo (*)			2.2.1.06.
			2.2.1.07.	Liechtenstein			2.2.1.07.

Country of citizenship					COC.L.	COC.M.	COC.H.
			2.2.1.08.	Moldova			2.2.1.08.
			2.2.1.09.	Monaco			2.2.1.09.
			2.2.1.10.	Montenegro			2.2.1.10.
			2.2.1.11.	Norway			2.2.1.11.
			2.2.1.12.	Bosnia and Herzegovina			2.2.1.12.
			2.2.1.13.	Russian Federation			2.2.1.13.
			2.2.1.14.	San Marino			2.2.1.14.
			2.2.1.15.	Serbia			2.2.1.15.
			2.2.1.16.	Switzerland			2.2.1.16.
			2.2.1.17.	Turkey			2.2.1.17.
			2.2.1.18.	Ukraine			2.2.1.18.
			2.2.1.19.	Vatican City State			2.2.1.19.
			2.2.1.20.	Recognised Non-Citizens			2.2.1.20.
			2.2.1.21.	Other country in Europe			2.2.1.21.
		2.2.2.	Country in Africa			2.2.2.	2.2.2.
			2.2.2.01.	Algeria			2.2.2.01.
			2.2.2.02.	Angola			2.2.2.02.
			2.2.2.03.	Benin			2.2.2.03.
			2.2.2.04.	Botswana			2.2.2.04.
			2.2.2.05.	Burkina Faso			2.2.2.05.
			2.2.2.06.	Burundi			2.2.2.06.
			2.2.2.07.	Cameroon			2.2.2.07.
			2.2.2.08.	Cape Verde			2.2.2.08.
			2.2.2.09.	Central African Republic			2.2.2.09.
			2.2.2.10.	Chad			2.2.2.10.
			2.2.2.11.	Comoros			2.2.2.11.
			2.2.2.12.	Congo			2.2.2.12.
			2.2.2.13.	Côte d'Ivoire			2.2.2.13.
			2.2.2.14.	Democratic Republic of the Congo			2.2.2.14.
			2.2.2.15.	Djibouti			2.2.2.15.
			2.2.2.16.	Egypt			2.2.2.16.
			2.2.2.17.	Equatorial Guinea			2.2.2.17.

Country of citizenship					COC.L.	COC.M.	COC.H.
			2.2.2.18.	Eritrea			2.2.2.18.
			2.2.2.19.	Ethiopia			2.2.2.19.
			2.2.2.20.	Gabon			2.2.2.20.
			2.2.2.21.	Gambia			2.2.2.21.
			2.2.2.22.	Ghana			2.2.2.22.
			2.2.2.23.	Guinea			2.2.2.23.
			2.2.2.24.	Guinea-Bissau			2.2.2.24.
			2.2.2.25.	Kenya			2.2.2.25.
			2.2.2.26.	Lesotho			2.2.2.26.
			2.2.2.27.	Liberia			2.2.2.27.
			2.2.2.28.	Libya			2.2.2.28.
			2.2.2.29.	Madagascar			2.2.2.29.
			2.2.2.30.	Malawi			2.2.2.30.
			2.2.2.31.	Mali			2.2.2.31.
			2.2.2.32.	Mauritania			2.2.2.32.
			2.2.2.33.	Mauritius			2.2.2.33.
			2.2.2.34.	Morocco			2.2.2.34.
			2.2.2.35.	Mozambique			2.2.2.35.
			2.2.2.36.	Namibia			2.2.2.36.
			2.2.2.37.	Niger			2.2.2.37.
			2.2.2.38.	Nigeria			2.2.2.38.
			2.2.2.39.	Rwanda			2.2.2.39.
			2.2.2.40.	Sao Tome and Principe			2.2.2.40.
			2.2.2.41.	Senegal			2.2.2.41.
			2.2.2.42.	Seychelles			2.2.2.42.
			2.2.2.43.	Sierra Leone			2.2.2.43.
			2.2.2.44.	Somalia			2.2.2.44.
			2.2.2.45.	South Africa			2.2.2.45.
			2.2.2.46.	Sudan			2.2.2.46.
			2.2.2.47.	South Sudan			2.2.2.47.
			2.2.2.48.	Swaziland			2.2.2.48.
			2.2.2.49.	Togo			2.2.2.49.

Country of citizenship					COC.L.	COC.M.	COC.H.
			2.2.2.50.	Tunisia			2.2.2.50.
			2.2.2.51.	Uganda			2.2.2.51.
			2.2.2.52.	Tanzania			2.2.2.52.
			2.2.2.53.	Western Sahara			2.2.2.53.
			2.2.2.54.	Zambia			2.2.2.54.
			2.2.2.55.	Zimbabwe			2.2.2.55.
			2.2.2.56.	Other country in Africa			2.2.2.56.
		2.2.3.	Country in the Caribbean, South or Central America			2.2.3.	2.2.3.
			2.2.3.01.	Antigua and Barbuda			2.2.3.01.
			2.2.3.02.	Argentina			2.2.3.02.
			2.2.3.03.	Aruba			2.2.3.03.
			2.2.3.04.	Bahamas			2.2.3.04.
			2.2.3.05.	Barbados			2.2.3.05.
			2.2.3.06.	Belize			2.2.3.06.
			2.2.3.07.	Bolivia			2.2.3.07.
			2.2.3.08.	Brazil			2.2.3.08.
			2.2.3.09.	Chile			2.2.3.09.
			2.2.3.10.	Colombia			2.2.3.10.
			2.2.3.11.	Costa Rica			2.2.3.11.
			2.2.3.12.	Cuba			2.2.3.12.
			2.2.3.13.	Curaçao			2.2.3.13.
			2.2.3.14.	Dominica			2.2.3.14.
			2.2.3.15.	Dominican Republic			2.2.3.15.
			2.2.3.16.	Ecuador			2.2.3.16.
			2.2.3.17.	El Salvador			2.2.3.17.
			2.2.3.18.	Grenada			2.2.3.18.
			2.2.3.19.	Guatemala			2.2.3.19.
			2.2.3.20.	Guyana			2.2.3.20.
			2.2.3.21.	Haiti			2.2.3.21.
			2.2.3.22.	Honduras			2.2.3.22.
			2.2.3.23.	Jamaica			2.2.3.23.
			2.2.3.24.	Mexico			2.2.3.24.

Country of citizenship					COC.L.	COC.M.	COC.H.
			2.2.3.25.	Nicaragua			2.2.3.25.
			2.2.3.26.	Panama			2.2.3.26.
			2.2.3.27.	Paraguay			2.2.3.27.
			2.2.3.28.	Peru			2.2.3.28.
			2.2.3.29.	Saint Kitts and Nevis			2.2.3.29.
			2.2.3.30.	Saint Lucia			2.2.3.30.
			2.2.3.31.	St Maarten (NL)			2.2.3.31.
			2.2.3.32.	Saint Vincent and the Grenadines			2.2.3.32.
			2.2.3.33.	Suriname			2.2.3.33.
			2.2.3.34.	Trinidad and Tobago			2.2.3.34.
			2.2.3.35.	Uruguay			2.2.3.35.
			2.2.3.36.	Venezuela			2.2.3.36.
			2.2.3.37.	Other country in the Caribbean, South or Central America			2.2.3.37.
		2.2.4.	Country in North America			2.2.4.	2.2.4.
			2.2.4.01.	Canada			2.2.4.01.
			2.2.4.02.	United States of America			2.2.4.02.
			2.2.4.03.	Other country in North America			2.2.4.03.
		2.2.5.	Country in Asia			2.2.5.	2.2.5.
			2.2.5.01.	Afghanistan			2.2.5.01.
			2.2.5.02.	Armenia			2.2.5.02.
			2.2.5.03.	Azerbaijan			2.2.5.03.
			2.2.5.04.	Bahrain			2.2.5.04.
			2.2.5.05.	Bangladesh			2.2.5.05.
			2.2.5.06.	Bhutan			2.2.5.06.
			2.2.5.07.	Brunei Darussalam			2.2.5.07.
			2.2.5.08.	Cambodia			2.2.5.08.
			2.2.5.09.	China			2.2.5.09.
			2.2.5.10.	Georgia			2.2.5.10.
			2.2.5.11.	India			2.2.5.11.
			2.2.5.12.	Indonesia			2.2.5.12.
			2.2.5.13.	Iraq			2.2.5.13.
			2.2.5.14.	Iran			2.2.5.14.

Country of citizenship					COC.L.	COC.M.	COC.H.
			2.2.5.15.	Israel			2.2.5.15.
			2.2.5.16.	Japan			2.2.5.16.
			2.2.5.17.	Jordan			2.2.5.17.
			2.2.5.18.	Kazakhstan			2.2.5.18.
			2.2.5.19.	North Korea			2.2.5.19.
			2.2.5.20.	South Korea			2.2.5.20.
			2.2.5.21.	Kuwait			2.2.5.21.
			2.2.5.22.	Kyrgyzstan			2.2.5.22.
			2.2.5.23.	Laos			2.2.5.23.
			2.2.5.24.	Lebanon			2.2.5.24.
			2.2.5.25.	Malaysia			2.2.5.25.
			2.2.5.26.	Maldives			2.2.5.26.
			2.2.5.27.	Mongolia			2.2.5.27.
			2.2.5.28.	Myanmar/Burma			2.2.5.28.
			2.2.5.29.	Nepal			2.2.5.29.
			2.2.5.30.	Oman			2.2.5.30.
			2.2.5.31.	Pakistan			2.2.5.31.
			2.2.5.32.	Philippines			2.2.5.32.
			2.2.5.33.	Qatar			2.2.5.33.
			2.2.5.34.	Saudi Arabia			2.2.5.34.
			2.2.5.35.	Singapore			2.2.5.35.
			2.2.5.36.	Sri Lanka			2.2.5.36.
			2.2.5.37.	Syria			2.2.5.37.
			2.2.5.38.	Taiwan			2.2.5.38.
			2.2.5.39.	Tajikistan			2.2.5.39.
			2.2.5.40.	Thailand			2.2.5.40.
			2.2.5.41.	Timor-Leste			2.2.5.41.
			2.2.5.42.	Turkmenistan			2.2.5.42.
			2.2.5.43.	United Arab Emirates			2.2.5.43.
			2.2.5.44.	Uzbekistan			2.2.5.44.
			2.2.5.45.	Vietnam			2.2.5.45.
			2.2.5.46.	Palestine			2.2.5.46.
			2.2.5.47.	Yemen			2.2.5.47.

Country of citizenship					COC.L.	COC.M.	COC.H.
			2.2.5.48.	Other country in Asia			2.2.5.48.
		2.2.6.	Country in Oceania			2.2.6.	2.2.6.
			2.2.6.01.	Australia			2.2.6.01.
			2.2.6.02.	Federated States of Micronesia			2.2.6.02.
			2.2.6.03.	Fiji			2.2.6.03.
			2.2.6.04.	Kiribati			2.2.6.04.
			2.2.6.05.	Marshall Islands			2.2.6.05.
			2.2.6.06.	Nauru			2.2.6.06.
			2.2.6.07.	New Zealand			2.2.6.07.
			2.2.6.08.	Palau			2.2.6.08.
			2.2.6.09.	Papua New Guinea			2.2.6.09.
			2.2.6.10.	Samoa			2.2.6.10.
			2.2.6.11.	Solomon Islands			2.2.6.11.
			2.2.6.12.	Tonga			2.2.6.12.
			2.2.6.13.	Tuvalu			2.2.6.13.
			2.2.6.14.	Vanuatu			2.2.6.14.
			2.2.6.15.	Other country in Oceania			2.2.6.15.
3.	Stateless				3.	3.	3.
4.	Not stated				4.	4.	4.

(*) This designation is without prejudice to positions on status, and is in line with UNSCR 1244/1999 and the ICJ Opinion on the Kosovo declaration of independence.

The breakdowns for 'Country of citizenship' are designed to break down any total or subtotal referring to persons.

The list of countries in the breakdown 'Country of citizenship' shall only apply for statistical purposes.

Topic: Ever resided abroad and year of arrival in the country (from 1980)

The year of arrival shall be the calendar year in which a person most recently established usual residence in the country. The year of the most recent arrival in the country shall be reported rather than the year of first arrival (i.e. the topic 'Year of arrival in the country' does not provide information on interrupted stays).

The breakdown 'Year of arrival in the country since 2010' focuses on more recent international migration since 2010.

Year of arrival in the country since 2010		YAT.
0.	Total	0.
1.	Ever resided abroad and arrived in 2010 or after	1.
2.	Resided abroad and arrived in 2009 or before, or never resided abroad	2.
3.	Not stated	3.

The breakdown 'Year of arrival in the country since 2010' is designed to break down any total or subtotal referring to persons.

The breakdowns for 'Year of arrival in the country since 1980' focus on international migration since 1980.

The data for 2021 shall refer to the time span between 1 January 2021 and the reference date.

Year of arrival in the country since 1980				YAE.L.	YAE.H.
0.	Total			0.	0.
1.	Ever resided abroad and arrived in 1980 or after			1.	1.
	1.1.	2020 to 2021		1.1.	1.1.
		1.1.1.	2021		1.1.1.
		1.1.2.	2020		1.1.2.
	1.2.	2015 to 2019		1.2.	1.2.
		1.2.1.	2019		1.2.1.
		1.2.2.	2018		1.2.2.
		1.2.3.	2017		1.2.3.
		1.2.4.	2016		1.2.4.
		1.2.5.	2015		1.2.5.
	1.3.	2010 to 2014		1.3.	1.3.
		1.3.1.	2014		1.3.1.
		1.3.2.	2013		1.3.2.
		1.3.3.	2012		1.3.3.
		1.3.4.	2011		1.3.4.
		1.3.5.	2010		1.3.5.
	1.4.	2005 to 2009		1.4.	1.4.
		1.4.1.	2009		1.4.1.
		1.4.2.	2008		1.4.2.
		1.4.3.	2007		1.4.3.
		1.4.4.	2006		1.4.4.
		1.4.5.	2005		1.4.5.
	1.5.	2000 to 2004		1.5.	1.5.
	1.6.	1995 to 1999		1.6.	1.6.
	1.7.	1990 to 1994		1.7.	1.7.
	1.8.	1985 to 1989		1.8.	1.8.
	1.9.	1980 to 1984		1.9.	1.9.
2.	Resided abroad and arrived 1979 or before, or never resided abroad			2.	2.
3.	Not stated			3.	3.

The breakdowns for 'Year of arrival in the country since 1980' are designed to break down any total or subtotal referring to persons.

Topic: Previous place of usual residence and date of arrival in the current place; or Place of usual residence one year prior to the census

The relationship between the current place of usual residence and the place of usual residence one year prior to the census shall be reported.

In the breakdown 'Place of usual residence one year prior to the census' any change of residence shall refer to the time span between one year prior to the reference date and the reference date. A move within the same LAU2 area shall also be considered as a move within the same NUTS 3 area.

Children under one year of age shall be classified under 'Not applicable' (ROY.4.).

Countries collecting information on the topic 'Previous place of usual residence and date of arrival in the current place' shall classify all persons that have changed their usual residence more than once within the year prior to the reference date according to their previous place of usual residence, i.e. the place of usual residence from which they moved to their current place of usual residence.

Place of usual residence one year prior to the census				ROY.
0.	Total			0.
1.	Usual residence unchanged			1.
2.	Usual residence changed			2.
	2.1.	Move within the reporting country		2.1.
		2.1.1.	Usual residence one year prior to the census within the same NUTS 3 area as the current usual residence	2.1.1.
		2.1.2.	Usual residence one year prior to the census outside the NUTS 3 area of the current usual residence	2.1.2.
	2.2.	Move from outside the reporting country		2.2.
3.	Not stated			3.
4.	Not applicable			4.

The breakdown 'Place of usual residence one year prior to the census' is designed to break down any total or subtotal referring to persons.

Topic: Family status

The family nucleus is defined in a narrow sense; that is as two or more persons who belong to the same household and who are related as husband and wife, as partners in a registered partnership, as partners in a consensual union, or as parent and child. Thus a family comprises a couple without children or a couple with one or more children, or a lone parent with one or more children. This family concept limits relationships between children and adults to direct (first-degree) relationships, that is, between parents and children.

Child (son/daughter) refers to a blood, step- or adopted son or daughter (regardless of age or marital status) who has usual residence in the household of at least one of the parents, and who has no partner or own children in the same household. Foster children shall not be included. A son or daughter who lives with a spouse, with a registered partner, with a partner in a consensual union, or with one or more own children, is not considered to be a child. A child who alternates between two households (for instance if his or her parents are divorced) shall consider the one where he or she spends the majority of the time as his or her household. Where an equal amount of time is spent with both parents the household shall be the one where the child is found at the time of census night or, alternatively, the household where the child has his or her legal or registered residence.

'Partners' shall include married couples, couples in registered partnerships, and couples who live in a consensual union. 'Registered partnership' is defined as in the technical specifications for the topic 'Legal marital status'.

Two persons are considered to be partners in a 'consensual union' when they

— belong to the same household, and

— have a marriage-like relationship with each other, and

— are not married to or in a registered partnership with each other.

'Skip-generation households' (households consisting of a grandparent or grandparents and one or more grandchildren, but no parent of those grandchildren) are not included in the definition of a family.

Family status				FST.L.	FST.M.	FST.H.
0.	Total			0.	0.	0.
1.	Partners			1.	1.	1.
	1.1.	Persons in a married couple or registered partnership			1.1.	1.1.
		1.1.1.	Persons in an opposite-sex married couple or registered partnership			1.1.1.
		1.1.2.	Persons in a same-sex married couple or registered partnership			1.1.2.
	1.2.	Partners in a consensual union			1.2.	1.2.
2.	Lone parents			2.	2.	2.
3.	Sons/daughters			3.	3.	3.
	3.1.	Not of lone parent			3.1.	3.1.
	3.2.	Of lone parent			3.2.	3.2.
4.	Not stated			4.	4.	4.
5.	Not applicable — not in a family nucleus			5.	5.	5.

The breakdowns for 'Family status' are designed to break down any total or subtotal referring to persons.

Topic: Type of family nucleus

The specifications for family concepts and the definitions of the terms 'family nucleus', 'child', 'couple' and 'consensual union' provided for the topic 'Family status' also apply for the topic 'Type of family nucleus'.

Type of family nucleus				TFN.L.	TFN.H.
0.	Total			0.	0.
1.	Married or Registered partnership couple families			1.	1.
	1.1.	Married or Registered partnership couple families without resident children			1.1.
		1.1.1.	Husband/wife couple families		1.1.1.
		1.1.2.	Married or Registered partnership same-sex couple families		1.1.2.

Type of family nucleus				TFN.L.	TFN.H.
	1.2.	Married or Registered partnership couple families with at least one resident child under 25			1.2.
		1.2.1.	Husband/wife couple families		1.2.1.
		1.2.2.	Married or Registered partnership same-sex couple families		1.2.2.
	1.3.	Married or Registered partnership couple families, youngest resident son/daughter 25 or older			1.3.
		1.3.1.	Husband/wife couple families		1.3.1.
		1.3.2.	Married or Registered partnership same-sex couple families		1.3.2.
2.	Consensual union couple families			2.	2.
	2.1.	Consensual union couples without resident children			2.1.
	2.2.	Consensual union couples with at least one resident child under 25			2.2.
	2.3.	Consensual union couples, youngest resident son/daughter 25 or older			2.3.
3.	Lone father families			3.	3.
	3.1.	Lone father families with at least one resident child under 25			3.1.
	3.2.	Lone father families, youngest resident son/daughter 25 or older			3.2.
4.	Lone mother families			4.	4.
	4.1.	Lone mother families with at least one resident child under 25			4.1.
	4.2.	Lone mother families, youngest resident son/daughter 25 or older			4.2.

The breakdowns for 'Type of family nucleus' are designed to break down the total of 'family nuclei', and any subtotals.

Topic: Size of family nucleus

The definition of the term 'family nucleus' provided for the topic 'Family status' also applies to the topic 'Size of family nucleus'.

Size of family nucleus			SFN.
0.	Total		0.
1.	2 persons		1.
2.	3 to 5 persons		2.
	2.1.	3 persons	2.1.
	2.2.	4 persons	2.2.

Size of family nucleus			SFN.
	2.3.	5 persons	2.3.
3.	6 and more persons		3.
	3.1.	6 to 10 persons	3.1.
	3.2.	11 and more persons	3.2.

The breakdowns for 'Size of family nucleus' are designed to break down the total of 'family nuclei', and any subtotals.

Topic: Household status

Member States shall apply the 'housekeeping concept' to identify private households, or, if not possible, the 'household-dwelling' concept.

1. Housekeeping concept

According to the housekeeping concept, a private household is either:

(a) A one-person household, that is a person who lives alone in a separate housing unit or who occupies, as a lodger, a separate room (or rooms) of a housing unit but does not join with any of the other occupants of the housing unit to form part of a multiperson household as defined below; or

(b) A multiperson household, that is a group of two or more persons who combine to occupy the whole or part of a housing unit and to provide themselves with food and possibly other essentials for living. Members of the group may pool their incomes to a greater or lesser extent.

2. Household-dwelling concept

The household-dwelling concept considers all persons living in a housing unit to be members of the same household, such that there is one household per occupied housing unit. In the household dwelling concept, then, the number of occupied housing units and the number of households occupying them is equal, and the locations of the housing units and households are identical.

The category 'Persons living in a private household' comprises 'Persons in a family nucleus' (HST.M. and HST.H.1.1.) and 'Persons not in a family nucleus' (HST.M. and HST.H. 1.2.). The category 'Persons in a family nucleus' comprises all persons who belong to a private household that contains a family nucleus of which they are a member. 'Persons not in a family nucleus' comprises all persons who either belong to a non-family household or to a family household without being member of any family nucleus in that household.

A non-family household can be a one-person household (person is 'Living alone' (HST.H.1.2.1.)) or a multiperson household without any family nucleus. The category 'Not living alone' (HST.H.1.2.2.) comprises persons that live either in a multiperson household without any family nucleus or in a family household without being member of any family nucleus in that household.

An *institutional household* comprises persons whose need for shelter and subsistence are being provided by an institution. An institution is understood to be a legal body for the purpose of long-term inhabitation and provision of services to a group of persons. Institutions usually have common facilities shared by the occupants (baths, lounges, eating facilities, dormitories and so forth).

'Homeless persons' (HST.M. 2.2. and HST.H.2.2.) are persons living in the streets without a shelter that would fall within the scope of living quarters (primary homelessness) or persons moving frequently between temporary accommodation (secondary homelessness).

Household status			HST.L.	HST.M.	HST.H.
0.	Total		0.	0.	0.
1.	Persons living in a private household		1.	1.	1.
	1.1.	Persons in a family nucleus		1.1.	1.1.

Household status				HST.L.	HST.M.	HST.H.
	1.2.	Persons not in a family nucleus			1.2.	1.2.
		1.2.1.	Living alone			1.2.1.
		1.2.2.	Not living alone			1.2.2.
	1.3.	Persons living in a private household, but category not stated			1.3.	1.3.
2.	Persons not living in a private household			2.	2.	2.
	2.1.	Persons in an institutional household			2.1.	2.1.
	2.2.	Persons not living in a private household (including homeless persons), but category not stated			2.2.	2.2.

The breakdowns for 'Household status' are designed to break down any total or subtotal referring to persons.

Topic: Type of private household

The specifications for the household concepts provided for the topic 'Household status' also apply to the topic 'Type of private household'.

'Couple households' shall include married couple households, registered partnership households and consensual union couple households.

Type of private household				TPH.L.	TPH.H.
0.	Total			0.	0.
1.	Non-family households			1.	1.
	1.1.	One-person households		1.1.	1.1.
	1.2.	Multiperson households		1.2.	1.2.
2.	One-family households			2.	2.
	2.1.	Couple households			2.1.
		2.1.1.	Couples without resident children		2.1.1.
		2.1.2.	Couples with at least one resident child under 25		2.1.2.
		2.1.3.	Couples, youngest resident son/daughter 25 or older		2.1.3.
	2.2.	Lone father households			2.2.
		2.2.1.	Lone father households with at least one resident child under 25		2.2.1.
		2.2.2.	Lone father households, youngest resident son/daughter 25 or older		2.2.2.

Type of private household				TPH.L.	TPH.H.
	2.3.	Lone mother households			2.3.
		2.3.1.	Lone mother households with at least one resident child under 25		2.3.1.
		2.3.2.	Lone mother households, youngest resident son/daughter 25 or older		2.3.2.
3.	Two-or-more-family households			3.	3.

The breakdowns for 'Type of private household' are designed to break down the total of 'private households', and any subtotals.

Topic: Size of private household

The specifications for the household concepts provided for the topic 'Household status' also apply to the topic 'Size of private household'.

Size of private household			SPH.
0.	Total		0.
1.	1 person		1.
2.	2 persons		2.
3.	3 to 5 persons		3.
	3.1.	3 persons	3.1.
	3.2.	4 persons	3.2.
	3.3.	5 persons	3.3.
4.	6 to 10 persons		4.
5.	11 or more persons		5.

The breakdowns for 'Size of private household' are designed to break down the total of 'private households', and any subtotals

Topic: Housing arrangements

The topic 'Housing arrangements' covers the whole population and refers to the type of housing in which a person usually resides at the time of the census. This covers all persons who are usual residents in different types of living quarters, or who do not have a usual residence and stay temporarily in some type of living quarters, or who are roofless, sleeping rough or in emergency shelters, when the census is taken.

Occupants are persons with their usual residence in the places listed in the respective category.

'Conventional dwellings' are structurally separate and independent premises at fixed locations which are designed for permanent human habitation and are, at the reference date,

(a) used as a residence, or

(b) vacant, or

(c) reserved for seasonal or secondary use.

'Separate' means surrounded by walls and covered by a roof or ceiling so that one or more persons can isolate themselves. 'Independent' means having direct access from a street or a staircase, passage, gallery or grounds.

'Other housing units' are huts, cabins, shacks, shanties, caravans, houseboats, barns, mills, caves or any other shelter used for human habitation at the time of the census, irrespective if it was designed for human habitation.

'Collective living quarters' are premises which are designed for habitation by large groups of individuals or several households and which are used as the usual residence of at least one person at the time of the census.

'Occupied conventional dwellings', other housing units and collective living quarters together represent 'living quarters'. Any 'living quarter' must be the usual residence of at least one person.

The sum of occupied conventional dwellings and other housing units represents 'housing units'.

The homeless (persons who are not usual residents in any living quarter category) are persons living in the streets without a shelter that would fall within the scope of living quarters (primary homelessness) or persons moving frequently between temporary accommodation (secondary homelessness).

Housing arrangements			HAR.
0.	Total		0.
1.	Occupants living in a conventional dwelling or in a collective living quarter		1.
	1.1.	Occupants living in a conventional dwelling	1.1.
	1.2.	Occupants living in a collective living quarter	1.2.
2.	Occupants living in an other housing unit and the homeless		2.
3.	Not stated		3.

The breakdowns for 'Housing arrangements' are designed to break down any total or subtotal referring to persons.

Topic: Tenure status of households

The topic 'Tenure status of households' refers to the arrangements under which a private household occupies all or part of a housing unit.

Households that are in the process of paying off a mortgage on the housing unit in which they live or purchasing their housing unit over time under other financial arrangements shall be classified under category 'Households of which at least one member is the owner of all or part of the housing unit' (TSH.1.).

Households of which at least one member is the owner of the housing unit and at least one member tenant of all or part of the housing unit shall be classified under category 'Households of which at least one member is the owner of all or part of the housing unit' (TSH.1.).

Tenure status of households		TSH.
0.	Total	0.
1.	Households of which at least one member is the owner of all or part of the housing unit	1.
2.	Households of which at least one member is a tenant of all or part of the housing unit (and no other member is the owner)	2.
3.	Households occupying all or part of a housing unit under some other form of tenure	3.
4.	Not stated	4.

The breakdowns for 'Tenure status of households' is designed to break down the total of 'private households', and any subtotals.

Topic: Type of living quarters

A living quarter is housing which is the usual residence of one or more persons. The terms 'Conventional dwellings', 'Other housing units' and 'Collective living quarters' are defined as under the topic 'Housing arrangements'.

	Type of living quarter	TLQ.
0.	Total	0.
1.	Occupied conventional dwellings	1.
2.	Other housing units	2.
3.	Collective living quarters	3.
4.	Not stated	4.

The breakdown 'Type of living quarter' is designed to break down the total of 'living quarters', and any subtotals.

Topic: Occupancy status of conventional dwellings

'Occupied conventional dwellings' are conventional dwellings which are the usual residence of one or more persons at the time of the census. 'Unoccupied conventional dwellings' are conventional dwellings which are not the usual residence of any person at the time of the census.

Dwellings reserved for seasonal or secondary use, vacant dwellings, as well as conventional dwellings with persons present but not included in the census shall be classified under the category 'Unoccupied conventional dwellings' (OCS.2.).

	Occupancy status of conventional dwelling	OCS.
0.	Total	0.
1.	Occupied conventional dwellings	1.
2.	Unoccupied conventional dwellings	2.
3.	Not stated	3.

The breakdowns for 'Occupancy status of conventional dwelling' are designed to break down the total of 'conventional dwellings', and any subtotals.

Topic: Type of ownership (tenure arrangements under which the dwelling is occupied)

The topic 'Type of ownership' refers to the ownership of the dwelling and not to that of the land on which the dwelling stands. It is intended to show the tenure arrangements under which the dwelling is occupied.

'Owner-occupied dwellings' are those where at least one occupant of the dwelling owns parts or the whole of the dwelling.

'Rented dwellings' are those where at least one occupant pays a rent for the occupation of the dwelling, and where no occupant owns parts or the whole of the dwelling.

Unoccupied conventional dwellings shall be classified under 'Not applicable' (OWS.5.).

	Type of ownership	OWS.
0.	Total	0.
1.	Owner-occupied dwellings	1.
2.	Rented dwellings	2.

Type of ownership		OWS.
3.	Dwellings in other types of ownership	3.
4.	Not stated	4.
5.	Not applicable	5.

The breakdown 'Type of ownership' is designed to break down the total of 'conventional dwellings', and any subtotals.

Topic: Number of occupants

The number of occupants of a housing unit is the number of people for whom the housing unit is the usual residence.

Number of occupants				NOC.
0.	Total			0.
1.	1 person			1.
2.	2 persons			2.
3.	3 to 5 persons			3.
	3.1.	3 persons		3.1.
	3.2.	4 persons		3.2.
	3.3	5 persons		3.3.
4.	6 and more persons			4.
	4.1.	6 to 10 persons		4.1.
		4.1.1.	6 persons	4.1.1.
		4.1.2.	7 persons	4.1.2.
		4.1.3.	8 persons	4.1.3.
		4.1.4.	9 persons	4.1.4.
		4.1.5.	10 persons	4.1.5.
	4.2.	11 or more persons		4.2.

The breakdowns for 'Number of occupants' are designed to break down the total of "occupied conventional dwellings', and any subtotals.

Topic: Useful floor space and/or Number of rooms of housing units

Useful floor space is defined as:

— the floor space measured inside the outer walls excluding non-habitable cellars and attics and, in multi-dwelling buildings, all common spaces; or

— the total floor space of rooms falling under the concept of 'room'.

A 'room' is defined as a space in a housing unit enclosed by walls reaching from the floor to the ceiling or roof, of a size large enough to hold a bed for an adult (4 square metres at least) and at least 2 metres high over the major area of the ceiling.

The Member States shall report on the 'useful floor space' or, if this is not possible, on the 'number of rooms'.

	Useful floor space	**UFS.**
0.	Total	0.
1.	Under 30 square metres	1.
2.	30 — less than 40 square metres	2.
3.	40 — less than 50 square metres	3.
4.	50 — less than 60 square metres	4.
5.	60 — less than 80 square metres	5.
6.	80 — less than 100 square metres	6.
7.	100 — less than 120 square metres	7.
8.	120 — less than 150 square metres	8.
9.	150 square metres and over	9.
10.	Not stated	10.

The breakdown 'Useful floor space' is designed to break down total of 'conventional dwellings', and any subtotals.

	Number of rooms	**NOR.**
0.	Total	0.
1.	1 room	1.
2.	2 rooms	2.
3.	3 rooms	3.
4.	4 rooms	4.
5.	5 rooms	5.
6.	6 rooms	6.
7.	7 rooms	7.
8.	8 rooms	8.
9.	9 rooms and more	9.
10.	Not stated	10.

The breakdown 'Number of rooms' is designed to break down the total of 'conventional dwellings', and any subtotals.

Topic: Density standard

The topic 'Density standard' relates the useful floor space in square meters or the number of rooms to the number of occupants, as specified under the topic 'Number of occupants'. Member States shall report on the density standard measured by the 'useful floor space', or, if not possible, by the 'number of rooms'.

	Density standard (floor space)	**DFS.**
0.	Total	0.
1.	Under 10 square metres per occupant	1.
2.	10 — less than 15 square metres per occupant	2.
3.	15 — less than 20 square metres per occupant	3.
4.	20 — less than 30 square metres per occupant	4.
5.	30 — less than 40 square metres per occupant	5.
6.	40 — less than 60 square metres per occupant	6.
7.	60 — less than 80 square metres per occupant	7.
8.	80 square metres and over per occupant	8.
9.	Not stated	9.

The breakdown 'Density standard (floor space)' is designed to break down the total of 'occupied conventional dwellings', and any subtotals.

	Density standard (number of rooms)	**DRM.**
0.	Total	0.
1.	Under 0,5 room per occupant	1.
2.	0,5 — less than 1,0 room per occupant	2.
3.	1,0 — less than 1,25 rooms per occupant	3.
4.	1,25 — less than 1,5 rooms per occupant	4.
5.	1,5 — less than 2,0 rooms per occupant	5.
6.	2,0 — less than 2,5 rooms per occupant	6.
7.	2,5 — less than 3,0 rooms per occupant	7.
8.	3,0 and more rooms per occupant	8.
9.	Not stated	9.

The breakdown 'Density standard (number of rooms)' is designed to break down the total of 'occupied conventional dwellings', and any subtotals.

Topic: Water supply system

It is recognised that, for some Member States, where there is evidence based on previous censuses, administrative data sources or from sample survey data, it may be assumed that virtually all conventional dwellings have 'Piped water in conventional dwelling'. Therefore, for these Member States, all conventional dwellings may be coded as WSS.1 — 'Piped water in conventional dwelling'. When Member States adopt this option, they shall certify this assumption and explain it in the metadata.

	Water supply system	**WSS.**
0.	Total	0.
1.	Piped water in conventional dwelling	1.
2.	No piped water in conventional dwelling	2.
3.	Not stated	3.

The breakdown 'Water supply system' is designed to break down the total of 'conventional dwellings', and any subtotals.

Topic: Toilet facilities

It is recognised that, for some Member States, where there is evidence based on previous censuses, administrative data sources or from sample survey data, it may be assumed that virtually all conventional dwellings have 'Toilet facilities'. Therefore, for these Member States, all conventional dwellings may be coded as TOI.1 — 'Flush toilet in conventional dwelling'. When Member States adopt this option, they shall certify this assumption and explain it in the metadata.

	Toilet facilities	**TOI.**
0.	Total	0.
1.	Flush toilet in conventional dwelling	1.
2.	No flush toilet in conventional dwelling	2.
3.	Not stated	3.

The breakdown 'Toilet facilities' is designed to break down the total of 'conventional dwellings', and any subtotals.

Topic: Bathing facilities

A bathing facility is any facility designed to wash the whole body and includes shower facilities.

It is recognised that, for some Member States, where there is evidence based on previous censuses, administrative data sources or from sample survey data, it may be assumed that virtually all conventional dwellings have 'Bathing facilities' . Therefore, for these Member States, all conventional dwellings may be coded as BAT.1. — 'Fixed bath or shower in conventional dwelling' . When Member States adopt this option, they shall certify this assumption and explain it in the metadata.

	Bathing facilities	**BAT.**
0.	Total	0.
1.	Fixed bath or shower in conventional dwelling	1.
2.	No fixed bath or shower in conventional dwelling	2.
3.	Not stated	3.

The breakdown 'Bathing facilities' is designed to break down the total of 'conventional dwellings', and any subtotals.

Topic: Type of heating

Conventional dwelling is considered as centrally heated if heating is provided either from a community heating centre or from an installation built in the building or in the conventional dwelling, established for heating purposes, without regard to the source of energy.

It is recognised that, for some Member States, where there is evidence based on previous censuses, administrative data sources or from sample survey data, it may be assumed that virtually all conventional dwellings have 'Central heating'. Therefore, for these Member States, all conventional dwellings may be coded as TOH.1 — 'Central heating'. When Member States adopt this option, they shall certify this assumption and explain it in the metadata.

Type of heating		TOH.
0.	Total	0.
1.	Central heating	1.
2.	No central heating	2.
3.	Not stated	3.

The breakdown 'Type of heating' is designed to break down the total of 'conventional dwellings', and any subtotals.

Topic: Dwellings by type of building

The topic 'Dwellings by type of building' refers to the number of dwellings in the building in which the dwelling is placed.

A building is considered as a Non-residential building when its purpose is predominantly non-residential (commercial building, office building, factory) but contains very few dwellings, e. g. for the property caretaker or an employee dwelling.

Dwellings by type of building			TOB.
0.	Total		0.
1.	Conventional dwellings in residential buildings		1.
	1.1.	Conventional dwellings in one-dwelling buildings	1.1.
	1.2.	Conventional dwellings in two-dwelling buildings	1.2.
	1.3.	Conventional dwellings in three or more dwelling buildings	1.3.
2.	Conventional dwellings in non-residential buildings		2.
3.	Not stated		3.

The breakdown 'Dwellings by type of building' is designed to break down the total of 'conventional dwellings', and any subtotals.

Topic: Dwellings by period of construction

The topic 'Dwellings by period of construction' refers to the year when the building in which the dwelling is placed was completed.

Dwellings by period of construction		POC.
0.	Total	0.
1.	Before 1919	1.
2.	1919-1945	2.
3.	1946-1960	3.

Dwellings by period of construction		POC.
4.	1961-1980	4.
5.	1981-2000	5.
6.	2001-2010	6.
7.	2011-2015	7.
8.	2016 and later	8.
9.	Not stated	9.

The breakdown 'Dwellings by period of construction' is designed to break down the total of 'conventional dwellings', and any subtotals.

II

(Non-legislative acts)

REGULATIONS

COMMISSION REGULATION (EU) 2017/712

of 20 April 2017

establishing the reference year and the programme of the statistical data and metadata for population and housing censuses provided for by Regulation (EC) No 763/2008 of the European Parliament and of the Council

(Text with EEA relevance)

THE EUROPEAN COMMISSION,

Having regard to the Treaty on the Functioning of the European Union,

Having regard to Regulation (EC) No 763/2008 of the European Parliament and the Council of 9 July 2008 on population and housing censuses ([1]), and in particular Article 5(1) and (3) thereof,

Whereas:

(1) Pursuant to Article 5(1) of Regulation (EC) No 763/2008, the Commission should define a reference year. The reference date selected by each Member State for the population and housing censuses data to be transmitted to the Commission should fall in that year.

(2) Pursuant to Article 5(3) of Regulation (EC) No 763/2008, the Commission should adopt a programme of the statistical data and of the metadata for the population and housing censuses to be transmitted to the Commission.

(3) In order to ensure data from the population and housing censuses conducted in the Member States are comparable, and to allow reliable Union-wide overviews to be drawn up, this programme should be the same in all Member States.

(4) In particular, it is necessary to define the content, format and structure of hypercubes which should be the same in all Member States, the special cell values and flags that the Member States can use in these hypercubes as well as the metadata on the topics.

(5) Commission Implementing Regulation (EU) 2017/543 ([2]) lays down the technical specifications for the census topics and their breakdowns to be applied to the data to be sent to the Commission for the reference year 2021.

(6) The measures provided for in this Regulation are in accordance with the opinion of the European Statistical System Committee,

([1]) OJ L 218, 13.8.2008, p. 14.

([2]) Commission Implementing Regulation (EU) 2017/543 of 22 March 2017 laying down rules for the application of Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses as regards the technical specifications of the topics and of their breakdowns (OJ L 78, 23.3.2017, p. 13).

HAS ADOPTED THIS REGULATION:

Article 1

Subject matter

This Regulation establishes the programme of statistical data and the metadata for the population and housing censuses to be transmitted to the Commission (Eurostat) for the reference year 2021.

Article 2

Definitions

The definitions laid down in Regulation (EC) No 763/2008 and the specifications set out in the Annex to Commission Implementing Regulation (EU) 2017/543 shall apply. The following definitions shall also apply for the purpose of this Regulation:

(1) 'total population' of a well-defined geographical area means all persons whose usual residence, as defined in Article 2(d) of Regulation (EC) No 763/2008, is located in that geographical area;

(2) 'hypercube' means a multidimensional cross tabulation of breakdowns which contains a cell value for the measurement of each category of each breakdown cross-tabulated by each category of any other breakdown used in that hypercube;

(3) 'cell value' means the information provided in a hypercube cell. A cell value can be either a 'numerical cell value' or a 'special cell value';

(4) 'numerical cell value' means a numerical value that is transmitted in a cell in order to provide the statistical information on the observation for that cell;

(5) 'confidential cell value' means a numerical cell value which in order to protect the statistical confidentiality of the data must not be divulged, according to the Member States' protective measures against disclosure of statistical data;

(6) 'non-confidential cell value' means a numerical cell value which is not a confidential cell value;

(7) 'unreliable cell value' means a numerical cell value which is unreliable according to the Member States' quality control;

(8) 'special cell value' means a symbol that is transmitted in a hypercube cell instead of a numerical cell value;

(9) 'flag' means a code that can accompany a particular cell value to describe a specific characteristic of that cell value.

Article 3

Reference date

Each Member State shall determine a reference date falling in 2021 for the population and housing census data to be transmitted to the Commission (Eurostat). Member States shall inform the Commission (Eurostat) by 31 December 2019 of the reference date selected.

Article 4

Programme of the statistical data

1. The programme of the statistical data to be transmitted to the Commission (Eurostat) for the reference year 2021 shall consist of the hypercubes listed in Annex I.

2. Member States shall provide the special cell value 'not applicable' only in the following cases:

(a) when a cell refers to the category 'not applicable' of at least one breakdown; or

(b) when a cell describes an observation that does not exist in the Member State.

3. Member States shall replace any confidential cell value by the special cell value 'not available'.

4. At the request of a Member State the Commission (Eurostat) shall refrain from making public any unreliable cell value provided by that Member State.

Article 5

Metadata on the cell values

1. Where applicable, Member States shall add the following flags to a hypercube cell:

(a) 'confidential';

(b) 'unreliable';

(c) 'revised after first data transmission';

(d) 'see information attached'.

2. Each cell whose confidential cell value has been replaced by the special value 'not available' shall be marked with the flag 'confidential'.

3. Each cell whose numerical cell value is unreliable shall be marked with the flag 'unreliable'.

4. For each cell accompanied by at least one of the flags 'unreliable', 'revised after first data transmission' or 'see information attached' an explanatory text shall be provided.

Article 6

Metadata on the topics

Member States shall provide the Commission (Eurostat) with the metadata on the topics as laid out in Annex II.

Article 7

Entry into force

This Regulation shall enter into force on the twentieth day following that of its publication in the *Official Journal of the European Union*.

This Regulation shall be binding in its entirety and directly applicable in all Member States.

Done at Brussels, 20 April 2017.

For the Commission
The President
Jean-Claude JUNCKER

ANNEX I

Programme of the statistical data (hypercubes) for the reference year 2021

No	Total	Breakdowns							
	Group 1 Total population	**GEO.N.**	**SEX.**	**AGE.H.**	**LMS.H.**	**HST.H.**	**FST.H.**		
1.1		GEO.N.	SEX.	AGE.H.	LMS.H.				
1.2		GEO.N.	SEX.	AGE.H.		HST.H.			
1.3		GEO.N.	SEX.	AGE.H.			FST.H.		
1.4		GEO.N.	SEX.		LMS.H.	HST.H.			
	Group 2 Total population	**GEO.M.**	**SEX.**	**AGE.M.**	**LMS.L.**	**HST.H.**	**FST.H.**	**HAR.**	**LOC.**
2.1		GEO.M.	SEX.	AGE.L.	LMS.L.		FST.H.		
2.2		GEO.M.	SEX.	AGE.L.		HST.H.		HAR.	
2.3		GEO.M.	SEX.	AGE.M.				HAR.	LOC.
	Group 3 Total population	**GEO.H.**	**SEX.**	**AGE.M.**	**HST.M.**	**LMS.L.**			
3.1		GEO.H.	SEX.	AGE.M.					
3.2		GEO.H.	SEX.		HST.M.				
3.3		GEO.H.	SEX.			LMS.L.			
	Group 4 Total population	**GEO.L.**	**SEX.**	**AGE.H.**	**CAS.H.**	**OCC.**	**EDU.**		
4.1		GEO.L.	SEX.	AGE.H.	CAS.H.				
4.2		GEO.L.	SEX.	AGE.H.		OCC.			
4.3		GEO.L.	SEX.	AGE.H.			EDU.		
	Group 5 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**OCC.**	**IND.L.**	**SIE.**	**EDU.**	
5.1		GEO.L.	SEX.	AGE.M.	OCC.	IND.L.			
5.2		GEO.L.	SEX.	AGE.M.	OCC.		SIE.		
5.3		GEO.L.	SEX.	AGE.M.	OCC.			EDU.	
5.4		GEO.L.	SEX.	AGE.L.			SIE.	EDU.	
5.5		GEO.N.	SEX.		OCC.	IND.L.		EDU.	
5.6		GEO.L.	SEX.	AGE.M.		IND.L.	SIE.		
5.7		GEO.L.	SEX.	AGE.L.		IND.L.		EDU.	
	Group 6 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**LPW.N.**	**OCC.**	**IND.L.**	**SIE.**	**EDU.**
6.1		GEO.L.	SEX.	AGE.M.	LPW.N.	OCC.			
6.2		GEO.L.	SEX.	AGE.M.	LPW.N.				EDU.
6.3		GEO.L.	SEX.		LPW.N.		IND.L.	SIE.	

No	Total	Breakdowns							
	Group 7 Total population	**GEO.N.**	**SEX.**	**AGE.M.**	**LPW.L.**	**IND.L.**	**SIE.**		
7.1		GEO.N.	SEX.	AGE.M.	LPW.L.	IND.L.			
7.2		GEO.N.	SEX.	AGE.M.	LPW.L.		SIE.		
	Group 8 Total population	**GEO.H.**	**SEX.**	**COC.L.**	**POB.L.**				
8.1		GEO.H.	SEX.	COC.L.					
8.2		GEO.H.	SEX.		POB.L.				
	Group 9 Total population	**GEO.M.**	**SEX.**	**AGE.M.**	**COC.L.**	**POB.H.**	**YAE.H.**		
9.1		GEO.N.	SEX.	AGE.M.	COC.L.	POB.H.			
9.2		GEO.M.	SEX.	AGE.M.			YAE.H.		
9.3		GEO.M.	SEX.	AGE.M.		POB.H.			
9.4		GEO.M.	SEX.			POB.H.	YAE.H.		
	Group 10 Total population	**GEO.M.**	**SEX.**	**AGE.M.**	**CAS.L.**	**COC.L.**	**POB.L.**	**YAT.**	
10.1		GEO.M.	SEX.	AGE.M.		COC.L.		YAT.	
10.2		GEO.M.	SEX.	AGE.M.			POB.L.	YAT.	
10.3		GEO.L.	SEX.	AGE.M.	CAS.L.	COC.L.		YAT.	
	Group 11 Total population	**GEO.M.**	**SEX.**	**AGE.M.**	**COC.H.**	**YAE.L.**			
11.1		GEO.M.	SEX.	AGE.M.	COC.H.				
11.2		GEO.M.	SEX.		COC.H.	YAE.L.			
	Group 12 Total population	**GEO.M.**	**SEX.**	**AGE.M.**	**COC.M.**	**POB.M.**	**YAE.L.**	**SIE.**	**ROY.**
12.1		GEO.M.	SEX.	AGE.M.			YAE.L.		ROY.
12.2		GEO.M.	SEX.	AGE.M.		POB.M.			ROY.
12.3		GEO.L.	SEX.		COC.M.	POB.M.			ROY.
12.4		GEO.L.	SEX.	AGE.M.				SIE.	ROY.
	Group 13 Total population	**GEO.M.**	**SEX.**	**AGE.M.**	**COC.M.**	**POB.M.**	**YAE.H.**	**ROY.**	**HAR.**
13.1		GEO.L.	SEX.			POB.M.	YAE.H.		HAR.
13.2		GEO.M.	SEX.	AGE.M.				ROY.	HAR.
13.3		GEO.M.		AGE.M.		POB.M.			HAR.
13.4		GEO.M.		AGE.M.	COC.M.				HAR.
13.5		GEO.L.			COC.M.	POB.M.	YAE.H.		

No	Total	Breakdowns								
	Group 14 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**CAS.H.**	**COC.L.**	**POB.L.**	**YAE.L.**	**ROY.**	**HAR.**
14.1		GEO.L.	SEX.	AGE.M.	CAS.H.	COC.L.				
14.2		GEO.L.	SEX.	AGE.M.	CAS.H.		POB.L.			
14.3		GEO.L.	SEX.	AGE.M.	CAS.H.			YAE.L.		
14.4		GEO.L.	SEX.	AGE.M.	CAS.H.				ROY.	
14.5		GEO.L.	SEX.	AGE.L.	CAS.L.				ROY.	HAR.
	Group 15 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**CAS.L.**	**EDU.**	**COC.L.**	**POB.L.**	**YAE.H.**	
15.1		GEO.L.	SEX.	AGE.L.	CAS.L.	EDU.		POB.L.		
15.2		GEO.L.	SEX.		CAS.L.	EDU.			YAE.H.	
15.3		GEO.L.	SEX.		CAS.L.		COC.L.		YAE.H.	
15.4		GEO.L.	SEX.	AGE.M.	CAS.L.		COC.L.	POB.L.		
	Group 16 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**OCC.**	**COC.L.**	**POB.L.**	**YAE.L.**	**ROY.**	
16.1		GEO.L.	SEX.	AGE.M.	OCC.	COC.L.				
16.2		GEO.L.	SEX.	AGE.M.	OCC.		POB.L.			
16.3		GEO.L.	SEX.	AGE.M.	OCC.			YAE.L.		
16.4		GEO.L.	SEX.	AGE.M.	OCC.				ROY.	
16.5		GEO.L.	SEX.		OCC.		POB.L.	YAE.L.		
	Group 17 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**IND.H.**	**COC.L.**	**YAE.L.**	**ROY.**		
17.1		GEO.L.	SEX.	AGE.M.	IND.H.	COC.L.				
17.2		GEO.N.	SEX.	AGE.M.	IND.H.		YAE.L.			
17.3		GEO.L.	SEX.	AGE.M.	IND.H.			ROY.		
	Group 18 Total population	**GEO.L.**	**SEX.**	**IND.H.**	**SIE.**	**EDU.**	**COC.L.**	**POB.L.**		
18.1		GEO.L.	SEX.	IND.H.	SIE.			POB.L.		
18.2		GEO.L.	SEX.	IND.H.		EDU.		POB.L.		
18.3		GEO.L.	SEX.	IND.L.			COC.L.	POB.L.		
	Group19 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**EDU.**	**POB.L.**	**YAE.H.**			
19.1		GEO.L.	SEX.	AGE.M.	EDU.	POB.L.				
19.2		GEO.L.	SEX.	AGE.M.	EDU.		YAE.L.			
19.3		GEO.L.	SEX.		EDU.	POB.L.	YAE.H.			
	Group 20 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**LPW.N.**	**COC.L.**	**POB.L.**			
20.1		GEO.L.	SEX.	AGE.M.	LPW.N.	COC.L.				
20.2		GEO.L.	SEX.	AGE.M.	LPW.N.		POB.L.			

No	Total	Breakdowns							
	Group 21 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**LMS.L.**	**FST.M.**	**HST.H.**	**CAS.H.**	**EDU.**
21.1		GEO.L.	SEX.	AGE.M.	LMS.L.			CAS.H.	
21.2		GEO.L.	SEX.	AGE.M.	LMS.L.				EDU.
21.3		GEO.L.	SEX.	AGE.M.		FST.M.		CAS.H.	
21.4		GEO.L.	SEX.	AGE.M.		FST.M.			EDU.
21.5		GEO.L.	SEX.	AGE.M.			HST.H.	CAS.H.	
	Group 22 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**HST.H.**	**EDU.**	**SIE.**		
22.1		GEO.L.	SEX.	AGE.M.	HST.H.	EDU.			
22.2		GEO.L.	SEX.	AGE.M.	HST.H.		SIE.		
	Group 23 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**FST.L.**	**HST.L.**	**CAS.L.**	**EDU.**	
23.1		GEO.N.	SEX.	AGE.M.		HST.L.	CAS.L.	EDU.	
23.2		GEO.L.	SEX.	AGE.M.	FST.L.		CAS.L.	EDU.	
	Group 24 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**LMS.L.**	**FST.L.**	**HST.M.**	**CAS.L.**	
24.1		GEO.N.	SEX.	AGE.M.	LMS.L.	FST.L.		CAS.L.	
24.2		GEO.L.	SEX.	AGE.M.	LMS.L.		HST.M.	CAS.L.	
	Group 25 Total population	**GEO.M.**	**SEX.**	**AGE.M.**	**LMS.L.**	**HST.M.**	**COC.L.**	**POB.L.**	
25.1		GEO.M.	SEX.	AGE.M.	LMS.L.			POB.L.	
25.2		GEO.L.	SEX.	AGE.M.	LMS.L.	HST.M.	COC.L.		
25.3		GEO.L.	SEX.	AGE.M.	LMS.L.	HST.M.		POB.L.	
	Group 26 Total population	**GEO.M.**	**SEX.**	**AGE.M.**	**FST.L.**	**HST.M.**	**COC.L.**	**POB.L.**	
26.1		GEO.M.	SEX.	AGE.M.	FST.L.		COC.L.		
26.2		GEO.M.	SEX.	AGE.M.		HST.M.		POB.L.	
26.3		GEO.L.	SEX.	AGE.M.		HST.M.	COC.L.	POB.L.	
	Group 27 Total population	**GEO.M.**	**SEX.**	**AGE.L.**	**FST.M.**	**HST.M.**	**YAE.L.**		
27.1		GEO.M.	SEX.	AGE.L.	FST.M.		YAE.L.		
27.2		GEO.M.	SEX.	AGE.L.		HST.M.	YAE.L.		
	Group 28 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**FST.M.**	**HST.M.**	**ROY.**		
28.1		GEO.L.	SEX.	AGE.M.	FST.M.		ROY.		
28.2		GEO.L.	SEX.	AGE.M.		HST.M.	ROY.		

No	Total	Breakdowns							
	Group 29 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**LMS.L.**	**FST.L.**	**HST.M.**	**CAS.L.**	**POB.L.**
29.1		GEO.L.	SEX.	AGE.M.	LMS.L.			CAS.L.	POB.L.
29.2		GEO.L.	SEX.	AGE.M.		FST.L.		CAS.L.	POB.L.
29.3		GEO.L.	SEX.	AGE.M.			HST.M.	CAS.L.	POB.L.
	Group 30 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**LMS.L.**	**FST.L.**	**HST.M.**	**CAS.L.**	**COC.L.**
30.1		GEO.L.	SEX.	AGE.M.	LMS.L.			CAS.L.	COC.L.
30.2		GEO.L.	SEX.	AGE.M.		FST.L.		CAS.L.	COC.L.
30.3		GEO.L.	SEX.	AGE.M.			HST.M.	CAS.L.	COC.L.
	Group 31 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**FST.L.**	**HST.M.**	**SIE.**	**EDU.**	**POB.L.**
31.1		GEO.L.	SEX.	AGE.M.		HST.M.	SIE.		POB.L.
31.2		GEO.L.	SEX.	AGE.M.	FST.L.		SIE.		POB.L.
31.3		GEO.L.	SEX.			HST.M.		EDU.	POB.L.
	Group 32 Total population	**GEO.L.**	**SEX.**	**AGE.M.**	**FST.L.**	**HST.M.**	**SIE.**	**EDU.**	**COC.L.**
32.1		GEO.L.	SEX.	AGE.M.		HST.M.	SIE.		COC.L.
32.2		GEO.L.	SEX.	AGE.M.	FST.L.		SIE.		COC.L.
32.3		GEO.L.	SEX.			HST.M.		EDU.	COC.L.
	Group 33 Number of all private households	**GEO.M.**	**TPH.H.**	**SPH.**	**TSH.**				
33.1		GEO.M.	TPH.H.	SPH.	TSH.				
	Group 34 Number of all families	**GEO.M.**	**TFN.H.**	**SFN.**					
34.1		GEO.M.	TFN.H.	SFN.					
	Group 35 Number of all private households	**GEO.H.**	**TPH.L.**	**SPH.**					
35.1		GEO.H.	TPH.L.						
35.2		GEO.H.		SPH.					
	Group 36 Number of all families	**GEO.H.**	**TFN.L.**	**SFN.**					
36.1		GEO.H.	TFN.L.						
36.2		GEO.H.		SFN.					

No	Total	Breakdowns					
	Group 37 Number of all conventional dwellings	**GEO.M.**	**TOB.**	**OCS.**	**POC.**		
37.1		GEO.M.	TOB.	OCS.	POC.		
	Group 38 Number of all conventional dwellings	**GEO.H.**	**TOB.**	**OCS.**			
38.1		GEO.H.	TOB.	OCS.			
	Group 39 Number of all occupied conventional dwellings	**GEO.M.**	**TOB.**	**(UFS.or NOR)**	**(DFS.or DRM)**	**OWS.**	**NOC.**
39.1		GEO.L.	TOB.			OWS.	NOC.
39.2		GEO.M.	TOB.	(UFS.or NOR)			NOC.
39.3		GEO.M.	TOB.		(DFS.or DRM)		NOC.
	Group 40 Number of all occupied conventional dwellings	**GEO.L.**	**WSS.**	**TOI.**	**BAT.**	**TOH.**	
40.1		GEO.L.	WSS.				
40.2		GEO.L.		TOI.			
40.3		GEO.L.			BAT.		
40.4		GEO.L.				TOH.	
	Group 41 Number of all living quarters	**GEO.H.**	**TLQ.**				
41.1		GEO.H.	TLQ.				

ANNEX II

Metadata on the topics referred to in Article 6

Member States shall transmit to the Commission (Eurostat) textual metadata on the definitions referring to the census topics.

For each topic, the metadata shall:

— name the data source(s) used to report the statistical data on the topic;

— report on the methodology used to estimate data on the topic;

— report on the reasons for any unreliability of the data on the topic.

In addition, Member States shall provide the following metadata:

Place of usual residence

The metadata shall explain in which way the definition of 'usual residence' of Article 2(d) of Regulation (EC) No 763/2008 has been applied, in particular to what extent the legal or registered residence has been reported as a substitute for the usual residence according to the 12 months criterion, as well as a clear definition of the concept adopted for the resident population.

The metadata shall report if third level students whose term-time address is not the one of their family home have been considered to have their usual residence at their family home.

The metadata shall report on any other country-specific application of the rules for the 'special cases' listed in the technical specifications for the topic 'Place of usual residence' in the Annex to Implementing Regulation (EU) 2017/543.

Homeless

The data on total population shall include all primary homeless persons (persons living in the streets without shelter) and all secondary homeless persons (persons moving frequently between temporary accommodation).

The metadata shall report the number of all homeless persons. The numbers of primary homeless persons (persons living in the streets without shelter) and of secondary homeless persons (persons moving frequently between temporary accommodation) shall be shown where this distinction is possible.

A description of the methodology and data sources used to produce the data on homeless persons shall be provided.

Legal marital status/partnerships

The metadata shall report on the relevant legal basis in the Member State concerning opposite-sex and same-sex marriages, the minimum age for marriages, opposite-sex and same-sex registered partnerships, and the possibility to divorce or legally separate.

Economic topics

The metadata shall report on any country-specific application of the rules listed in the technical specifications for the topic 'Current activity status' in the Annex to Implementing Regulation (EU) 2017/543. The metadata shall report whether the current activity status has been reported on the basis of registers, and, if this is the case, on the relevant definitions used in this register.

The metadata shall report on the national minimum age for economic activity in the country, and the relevant legal basis.

Where the census in the Member State identifies persons doing more than one job, the metadata shall describe the method used to allocate them to their main job (for example, on the basis of time spent on the job, income received).

The metadata shall report on any country-specific application of the rules listed in the technical specifications for the topic 'Status in employment' in the Annex to Implementing Regulation (EU) 2017/543. Where the census in the Member State identifies person who are both, employer and employee, the metadata shall describe the method used to allocate them to one of the two categories.

Country/place of birth

For censuses for which no or incomplete information is available on the country of birth according to international boundaries existing at the time of the census, the metadata shall inform about the methodology used to allocate persons within the breakdown of the topic 'Country/place of birth'.

The metadata shall report if information on the place where the birth took place was used as a substitute for the place of usual residence of the mother at the time of the birth.

Country of citizenship

In countries where a part of the population are persons who are 'Recognised Non-Citizens' (that is persons who are neither citizens of any country nor stateless and who have some but not all of the rights and duties associated with citizenship), the metadata shall provide relevant information.

Place of usual residence one year prior to the census

Where the census in the Member State collects information on the topic 'Previous place of usual residence and date of arrival in the current place', the metadata shall describe any methodology used to report on the place of usual residence one year prior to the census.

Household and family topics

The metadata shall specify whether the census in the Member State applies the 'housekeeping' or the 'household-dwelling' concept to identify private households. The metadata shall report on the method used to generate households and families.

The metadata shall report on the way the relationships between household members are identified (e.g. relationship matrix; relation to reference person). If these data are obtained from administrative registers, it shall be reported whether information on the relationship between household and family members is recorded in and obtained directly from the administrative source(s), or whether this information is based on a statistical model.

Type of ownership

The metadata shall explain and provide examples of the ownership types under national property laws or customs that have been classified under 'Dwellings in other types of ownership'.

Useful floor space and/or number of rooms of housing unit, density standard

The metadata shall report on the application of the concept of either 'useful floor space', or 'number of rooms' as appropriate, and on the definition adopted for the corresponding measurement of the density standard.

COMMISSION IMPLEMENTING REGULATION (EU) 2017/881

of 23 May 2017

implementing Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses, as regards the modalities and structure of the quality reports and the technical format for data transmission, and amending Regulation (EU) No 1151/2010

(Text with EEA relevance)

THE EUROPEAN COMMISSION,

Having regard to the Treaty on the Functioning of the European Union,

Having regard to Regulation (EC) No 763/2008 of the European Parliament and of the Council of 9 July 2008 on population and housing censuses ([1]), and in particular Articles 5(5) and 6(3) thereof,

Whereas:

(1) Regulation (EC) No 763/2008 establishes common rules for the decennial provision of comprehensive data on population and housing.

(2) Under Commission Regulation (EU) 2017/712 ([2]), the next population and housing census should relate to the reference year 2021.

(3) Commission Regulation (EU) No 1151/2010 ([3]) laid down the modalities and structure of the quality reports and the technical format for data transmission on population and housing censuses for the reference year 2011.

(4) For the purposes of the next population and housing census 2021, and in order to assess the quality of the data transmitted to the Commission (Eurostat), it is necessary to lay down new modalities and structure for the quality reports and the technical format for data transmission.

(5) In accordance with Article 5(5) of Regulation (EC) No 763/2008, Member States should transmit their validated data and metadata in electronic form, in an appropriate technical format to be adopted by the Commission. The Statistical Data and Metadata eXchange (SDMX) initiative on statistical and technical standards for the exchange and sharing of data and metadata, on which Census Hub is based, was launched by the Bank of International Settlements, the European Central Bank, the Commission (Eurostat), the International Monetary Fund, the Organisation for Economic Cooperation and Development (OECD), the United Nations and the World Bank. For the exchange of official statistics, SDMX and the Census Hub provide statistical, technical and transmission standards. A technical format in accordance with those standards should therefore be introduced.

(6) Regulation (EU) No 1151/2010 requires Member States to store the population and housing census data for the 2011 reference year until 1 January 2025. In order to allow users to make comparisons between the two censuses, the 2011 census data should be available until 1 January 2035 in parallel with the 2021 data.

(7) Regulation (EU) No 1151/2010 should therefore be amended accordingly.

(8) The measures provided for in this Regulation are in accordance with the opinion of the European Statistical System Committee,

([1]) OJ L 218, 13.8.2008, p. 14.

([2]) Commission Regulation (EU) 2017/712 of 20 April 2017 establishing the reference year and the programme of the statistical data and metadata for population and housing censuses provided for by Regulation (EC) No 763/2008 of the European Parliament and of the Council (OJ L 105, 21.4.2017, p. 1).

([3]) Commission Regulation (EU) No 1151/2010 of 8 December 2010 implementing Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses, as regards the modalities and structure of the quality reports and the technical format for data transmission (OJ L 324, 9.12.2010, p. 1).

HAS ADOPTED THIS REGULATION:

Article 1

Subject matter

This Regulation lays down the modalities and structure of Member States' reports on the quality of the data that they transmit to the Commission (Eurostat) from their population and housing censuses for the reference year 2021, and the technical format for data transmission.

Article 2

Definitions

The definitions and technical specifications set out in Regulation (EC) No 763/2008, Commission Implementing Regulation (EU) 2017/543 ([1]) and Regulation (EU) 2017/712 shall apply for the purpose of this Regulation.

The following definitions shall also apply:

1. 'statistical unit' means the basic observation unit, namely a natural person, household, family, living quarter or conventional dwelling;

2. 'individual enumeration' means the collection of information on each statistical unit so that its characteristics can be recorded separately and cross-classified with other characteristics;

3. 'simultaneity' means that the information obtained in a census refers to the same point in time (reference date);

4. 'universality within a defined territory' means that data are provided for all statistical units within a precisely defined territory. Where the statistical units are persons, 'universality within a defined territory' means that data are provided based on information for all persons who have their usual residence in the defined territory (total population);

5. 'availability of small-area data' means the availability of data for small geographical areas and small groups of statistical units;

6. 'defined periodicity' means the regular conducting of censuses at the beginning of every decade, including the continuity of registers;

7. 'data source' means the set of data records for statistical units and/or events relating to the statistical units which forms a direct basis for the production of census data about one or more specified topics for a specified target population;

8. 'target population' means the set of all statistical units in a defined geographical area at the reference date which qualify for reporting on one or more specified topics. It includes each valid statistical unit only once;

9. 'estimated target population' means the best available approximation of the target population. It consists of the census population plus under-coverage minus over-coverage;

10. 'census population' means the set of statistical units that is factually represented by the census results on one or more specified topics for a specified target population. The data records for the census population are those in the data source for the specified target population, including all imputed records and excluding all deleted records. If a data source comprises, as a matter of methodological principle, data records for only a sample of the statistical units in its estimated target population, the census population includes the complementary set of statistical units, in addition to the statistical units in the sample;

11. 'complementary set of statistical units' means the set of those statistical units that belong to an estimated target population, but on which, as a result of an applied sampling methodology, the data source contains no data records;

([1]) Commission Implementing Regulation (EU) 2017/543 of 22 March 2017 laying down rules for the application of Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses as regards the technical specifications of the topics and of their breakdowns (OJ L 78, 23.3.2017, p. 13).

12. 'coverage assessment' means a study of the difference between a specified target population and its census population;

13. 'post-enumeration survey' means a survey conducted shortly after the enumeration for coverage and content assessment purposes;

14. 'under-coverage' means the set of all statistical units that belong to a specified target population but are not included in the corresponding census population;

15. 'over-coverage' means the set of all statistical units that are included in a census population used to report on a specified target population, but without belonging to that target population;

16. 'record imputation' means the assignment of an artificial but plausible data record to exactly one geographical area at the most detailed geographical level for which census data are produced, and the imputation of that data record into a data source;

17. 'record deletion' means the act of deleting or ignoring/not taking into account a data record that is included in a data source used to report on a specified target population but does not report any valid information on any statistical unit within that target population;

18. 'item imputation' means the insertion of an artificial but plausible value on a specific topic into a data record that already exists in a data source but either does not contain this value or contains a value that is considered implausible;

19. 'questionnaire-based data' means data originally obtained from respondents by means of a questionnaire in the context of a collection of statistical data that refer to a specified point in time;

20. 'record linkage' means the process of merging information from different data sources by comparing records for individual statistical units and merging information referring to the same statistical unit;

21. 'unique identifier' means a variable or set of variables in the data records in a data source or any list of statistical units which is used for

    — verifying that the data source (or list of statistical units) includes no more than one data record for each statistical unit, and/or

    — record linkage;

22. 'register' means a repository of information about statistical units which is directly updated in the course of events affecting the statistical units;

23. 'register-based data' means data in or from a register;

24. 'matching of registers' means record linkage where all matched data sources are contained in registers;

25. 'data extraction' means the process of retrieving census information from information contained in a register and relating to individual statistical units;

26. 'coding' means the process of converting information into codes representing classes within a classification scheme;

27. 'capturing' means the process by which collected data are imported into a form suitable for further processing;

28. 'record editing' means the process of checking and modifying data records to make them plausible while at the same time preserving major parts of them;

29. 'generation of a household' means the identification of a private household according to the household-dwelling concept as defined in the Annex to Implementing Regulation (EU) 2017/543 under the topic 'Household status';

30. 'generation of a family' means the identification of a family based on information on whether the persons live in the same household, but with no or incomplete information on family relationships between them. The term 'family' is specified in 'family nucleus' as defined in the Annex to Implementing Regulation (EU) 2017/543 under the topic 'Family status';

31. 'unit no-information' means the failure to collect any data from a statistical unit that is in the census population;

32. 'statistical disclosure control' means the methods and processes applied in order to minimise the risk of disclosing information on individual statistical units when releasing statistical information;

33. 'estimation' means the calculation of statistical estimates using a mathematical formula and/or algorithm applied to the available data;

34. 'data structure definition' means a set of structural metadata associated with a data set, which includes information about how topics are associated with the measures, dimensions and attributes of a hypercube, along with breakdowns, information about the representation of data and related descriptive metadata.

Article 3

Metadata and quality reporting

Member States shall report to the Commission (Eurostat), by 31 March 2024, the background information and quality-related data and metadata specified in the Annex to this Regulation, with reference to their population and housing censuses for the reference year 2021 and to the data and metadata transmitted to the Commission (Eurostat) under Regulation (EU) 2017/712.

Article 4

Data sources

Member States shall report on any data source(s) used for the collection of information needed to fulfil the requirements of Regulation (EC) No 763/2008, in particular to:

(a) meet the essential features listed in point (i) of Article 2 of Regulation (EC) No 763/2008;

(b) represent the target population;

(c) comply with the relevant technical specifications set out in Implementing Regulation (EU) 2017/543; and

(d) contribute to the provision of data for the programme of statistical data set out in Regulation (EU) 2017/712.

Article 5

Access to relevant information

1. Member States shall provide the Commission (Eurostat), at its request, with access to any information relevant to the assessment of the quality of the data and metadata transmitted under Regulation (EU) 2017/712.

2. In complying with paragraph 1, Member States shall not be obliged to provide the Commission (Eurostat) with any microdata or confidential data.

Article 6

Technical format for data transmission

The technical format to be used for the transmission of data and metadata for the reference year 2021 shall be the Statistical Data and Metadata eXchange (SDMX) format as implemented through the Census Hub. Member States shall transmit the required data in line with the data structure definitions and related technical specifications provided by the Commission (Eurostat). Member States shall store the required data and metadata until 1 January 2035, for any later transmission requested by the Commission (Eurostat).

Article 7

Amendment to Regulation (EU) No 1151/2010

In Article 6 of Regulation (EU) No 1151/2010, the third sentence is replaced by the following:

'Member States shall store the data and metadata for the 2011 reference year until 1 January 2035. Member States shall not be obliged to make changes or revisions to these data after 1 January 2025. Member States choosing to do so shall inform the Commission (Eurostat) about the changes or revisions before they are implemented.'

Article 8

Entry into force

This Regulation shall enter into force on the twentieth day following its publication in the *Official Journal of the European Union*.

This Regulation shall be binding in its entirety and directly applicable in all Member States.

Done at Brussels, 23 May 2017.

For the Commission
The President
Jean-Claude JUNCKER

ANNEX

Content and structure of the quality reports for data transmission

The **textual** and **quantitative** metadata information on the population and housing censuses conducted in the Member States for the reference year 2021 comprises the following sections:

1. **OVERVIEW**

1.1. Legal background

1.2. Bodies responsible

2. **DATA SOURCES**

2.1. Classification of data sources in accordance with Article 4(1) of Regulation (EC) No 763/2008.

2.2. List of the data sources used for the 2021 census

2.3. 'Data sources x topics' matrix

2.4. Adequacy of data sources: extent to which they meet the essential features (Article 4(4) of Regulation (EC) No 763/2008)

2.4.1. *Individual enumeration;*

2.4.2. *Simultaneity;*

2.4.3. *Universality within the defined territory;*

2.4.4. *Availability of small-area data;*

2.4.5. *Defined periodicity.*

3. **CENSUS LIFECYCLE**

3.1. **Reference date**

Member States shall provide the Commission (Eurostat) with the reference date according to Article 3 of Commission Regulation (EU) 2017/712.

3.2. **Preparation and execution of data collection**

3.2.1. *Questionnaire-based data*

— design and testing of questionnaires (including copies of all final questionnaires),

— preparation of any address lists, preparation of the field work, mapping, publicity,

— data collection (including field work),

— legal obligation to collect information, incentives for providing truthful information or possible reasons for providing false information.

3.2.2. *Register-based data*

— creation of new registers from 2011 onwards (where applicable),

— redesign of existing registers from 2011 onwards (including changes in the contents of registers, adaptation of the census population, adaptation of definitions and/or technical specifications) (where applicable),

— maintenance of the registers (for each register used for the 2021 census), including

  — content of the register (statistical units and information on them, any record editing and/or item and record imputation in the register),

  — administrative responsibilities,

  — legal obligation to register information, incentives for providing truthful information and possible reasons for providing false information,

  — delays in reporting, in particular legal/official delays, data registration delays, late reporting,

  — evaluation of and clearance for non-registration, non-deregistration, multiple registration,

  — any major register revisions or updating of records that affects the 2021 census data, periodicity of register revisions,

  — usage, including 'statistical usage of the register other than for the census' and 'usage of the register other than for statistical purposes (e.g. administrative purposes)',

— matching and linking of registers (including unique identifier(s) used for record linkage),

— data extraction.

3.2.3. *Data collected by means of a sample*

For topics for which information has been collected by means of a sample, the metadata shall also contain descriptions of:

— the sampling design,

— methodologies used for any estimations, models or imputations,

— possible biases in the estimation due to methodologies applied,

— formulae and algorithms used to calculate the standard error.

3.2.4. *Data collected by combined methods (data based on more than one type of data source)*

For topics for which information has been collected by combined methods, the metadata shall also contain:

— a description of the methods (types of data sources used and how information from different sources was combined, how the different sources and methods used complement and support each other and, if applicable, which parts of the population were covered by which source),

— any other quality issues relating to the process of using combined methods.

3.3. **Processing and evaluation**

3.3.1. Data processing (including capturing, coding, identifying variable(s), record editing, record imputation, record deletion, estimation, record linkage including identifying variable(s) used for the record linkage, generation of households and families, measures to identify or limit unit-no-information);

3.3.2. Coverage assessment activities, methodology to treat non-response, post-enumeration survey(s) (where applicable), final data validation: method of assessing under- and over-coverage, including information on the quality of the under- and over-coverage estimates.

3.4. **Dissemination** (dissemination channels, ensuring statistical confidentiality including statistical disclosure control)

3.5. **Measures to ensure cost-effectiveness**

4. **ASSESSMENT OF DATA QUALITY**

4.1. **Comparability**

For each topic, Member States shall report on any deviation from the required concepts and definitions or any practice in the Member State that could impair the Union-wide comparability of the data.

For the topic 'Current Activity Status', Member States shall report on any estimation methods used to adjust data to meet more closely the definition set in the Annex to Commission Implementing Regulation (EU) 2017/543. Member States shall report on the extent to which the data sources and any estimation methods used result in deviation from the definition of 'Current Activity Status' set in that Regulation.

4.2. **Timeliness and punctuality**

The following information shall be provided at national level:

— date(s) of the transmission of data to the Commission (Eurostat), broken down by hypercubes,

— date(s) of major revision(s) of the transmitted data, broken down by hypercubes,

— date(s) of transmission of the metadata.

In the event of major revisions on or after 1 April 2024, Member States shall report the respective date(s) separately to the Commission (Eurostat) within a week.

4.3. **Coherence**

Member States shall report on any significant inconsistencies between the data transmitted in the different datasets defined in Commission Regulation (EU) 2017/712.

4.4. **Coverage and accuracy**

To indicate coverage, the following absolute values shall be provided for person counts at national level and shall be disaggregated by sex and broad age groups as defined in Commission Implementing Regulation (EU) 2017/543:

(a) census population;

(b) number of all record imputations;

(c) number of all record deletions;

(d) under-coverage (estimated);

(e) over-coverage (estimated);

(f) estimated target population.

For the assessment of accuracy the following **absolute values** shall be provided for person counts at national level and shall be disaggregated by sex and broad age groups as defined in Commission Implementing Regulation (EU) 2017/543:

(a) census population;

(b) number of observed data records on the topic derived from traditional census;

(c) number of observed data records on the topic derived from administrative registers;

(d) number of observed data records on the topic derived from sample surveys;

(e) number of observed data records on the topic derived from multiple data sources;

(f) complementary set of statistical units on the topic (for samples);

(g) number of imputed observations on the topic;

(h) number of records with missing information on the topic.

The above absolute values for the assessment of accuracy shall be provided for the following census topics:

(a) Legal Marital status (LMS);

(b) Family status (FST);

(c) Household status (HST);

(d) Current activity status (CAS);

(e) Occupation (OCC);

(f) Industry (IND);

(g) Status in employment (SIE);

(h) Location of place of work (LPW);

(i) Educational attainment (EDU);

(j) Country/Place of birth (POB);

(k) Country of citizenship (COC);

(l) Year of arrival in the country since 2010 (YAT);

(m) Year of arrival in the country since 1980 (YAE);

(n) Place of usual residence one year prior to the census (ROY);

(o) Housing arrangements (HAR).

4.5. **Completeness**

Member States shall report on the degree of completeness of the data in terms of the requirements of Regulation (EC) No 763/2008. They shall give details of any census topics or associated breakdowns for which data are not supplied.

4.6. **Relevance**

Information on the following shall be provided at Union level:

(a) actions taken to identify and fulfil user needs;

(b) monitoring of the extent of data extractions.

COMMISSION IMPLEMENTING REGULATION (EU) 2018/1799

of 21 November 2018

on the establishment of a temporary direct statistical action for the dissemination of selected topics of the 2021 population and housing census geocoded to a 1 km² grid

(Text with EEA relevance)

THE EUROPEAN COMMISSION,

Having regard to the Treaty on the Functioning of the European Union,

Having regard to Regulation (EC) No 223/2009 of 11 March 2009 on European statistics and repealing Regulation (EC, Euratom) No 1101/2008 of the European Parliament and of the Council on the transmission of data subject to statistical confidentiality to the Statistical Office of the European Communities, Council Regulation (EC) No 322/97 on Community Statistics, and Council Decision 89/382/EEC, Euratom establishing a Committee on the Statistical Programmes of the European Communities ([1]), and in particular Article 14(2) thereof,

Whereas:

(1) The dissemination of Union-wide harmonised census topics on a constant area grid, in particular on a 1 km² grid, is a key European statistical output for future policy-making and census strategies by the Member States.

(2) Pursuant to Article 14(1)(b) of Regulation (EC) No 223/2009, the Commission may decide in specific and duly justified cases, in order to meet unexpected needs, on a temporary direct statistical action which shall fulfil conditions laid down in Article 14(2) of that Regulation.

(3) This temporary direct statistical action should provide for data collection covering one reference year. All Member States should be able to produce univariate census data geocoded to a 1 km² grid at the reference date for the 2021 population and housing census; the Union should also make financial contributions to National Statistical Institutes and other national authorities to cover the incremental costs incurred by them. This action is supported by a cost-effectiveness analysis and an estimate of overall incremental production costs provided by the Commission.

(4) This action is justified by a common need across the Union for reliable, accurate and comparable information on population distribution with sufficient spatial resolution, founded on harmonised output requirements and intended in particular for pan-European regional policy-making.

(5) Harmonised, spatially resolved demographic information across the Union is available, and the objective is to disseminate one dataset per Member State, containing selected topics of the 2021 population and housing census geocoded to a 1 km² grid. There is no additional burden on respondents as all necessary information will be obtained from the 2021 census data.

(6) In particular, in order to achieve comparable harmonised outputs across the Union, a Union-wide constant area grid consisting of 1 km² cells needs to be determined. Furthermore, the specific topics and their breakdowns as well as the detailed programme to be disseminated on this 1 km² grid need to be established. Finally, it is necessary to specify the required spatial and statistical metadata for such a dataset.

(7) Directive 2007/2/EC of the European Parliament and of the Council ([2]) and related Commission Implementing Regulations lay down the required metadata ([3]), data format ([4]), and network services ([5]) for the dissemination of spatial data. In particular, point 1 of Annex III covers possible statistical grid systems for spatial data dissemination and, according to point 10 of Annex III, is applicable to spatial datasets under the theme 'Population distribution — demography'.

([1]) OJ L 87, 31.3.2009, p. 164.
([2]) Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an infrastructure for spatial information in the European Community (INSPIRE) (OJ L 108, 25.4.2007, p. 1).
([3]) Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata (OJ L 326, 4.12.2008, p. 12).
([4]) Commission Regulation (EU) No 1089/2010 of 23 November 2010 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards interoperability of spatial data sets and services (OJ L 323, 8.12.2010, p. 11).
([5]) Commission Regulation (EC) No 976/2009 of 19 October 2009 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards the Network Services (OJ L 274, 20.10.2009, p. 9).

(8) Regulation (EC) No 763/2008 of the European Parliament and the Council ([1]) and related Commission Implementing Regulations establish common rules for transmission of the 2021 census data, in particular the reference year and the required metadata ([2]), the technical specifications of the census topics and their breakdowns ([3]), and the technical format ([4]).

(9) Member States should transmit their validated data and metadata in electronic form, in an appropriate technical format to be provided by the Commission. The Statistical Data and Metadata eXchange (SDMX) initiative on statistical and technical standards for the exchange and sharing of data and metadata, on which the Census Hub is based, was launched by the Bank of International Settlements, the European Central Bank, the Commission (Eurostat), the International Monetary Fund, the Organisation for Economic Cooperation and Development (OECD), the United Nations and the World Bank. For the exchange of official statistics, SDMX and the Census Hub provide statistical, technical and transmission standards. A technical format in accordance with those standards should therefore be introduced.

(10) The Commission (Eurostat) hosted a project on 'Harmonised protection of census data in the ESS', which delivered good practices and implementation guidelines for the harmonised disclosure protection of 1 km² grid data.

(11) The measures provided for in this Regulation are in accordance with the opinion of the European Statistical System Committee,

HAS ADOPTED THIS REGULATION:

Article 1

Subject matter

A temporary direct statistical action is hereby established in order to develop, produce and disseminate selected topics of the 2021 population and housing census geocoded to a 1 km² grid ('1 km² grid data').

To this end, a unique harmonised and constant geospatial reference grid for Europe is determined consisting of cells with an area of 1 km². And the specific topics and their breakdowns are established as well as the detailed programme and metadata for the dissemination of 2021 population and housing census data geocoded to the 1 km² reference grid.

Article 2

Definitions

For the purposes of this Regulation, the definitions in Article 2 of Regulation (EC) No 763/2008 shall apply.

The following definitions shall also apply:

(1) 'grid', 'grid cell' and 'grid point' mean grid, grid cell and grid point as defined in Annex II (2.1) of Regulation (EU) No 1089/2010;

(2) 'total population' means all persons of a grid cell whose usual residence is located in that grid cell;

(3) 'data item' means a single measurement contained in the table defined in Annex II of this Regulation;

(4) 'data value' means the information provided by a data item. A data value can be either a 'numerical value' or a 'special value';

([1]) Regulation (EC) No 763/2008 of the European Parliament and of the Council of 9 July 2008 on population and housing censuses (OJ L 218, 13.8.2008, p. 14).

([2]) Commission Regulation (EU) 2017/712 of 20 April 2017 establishing the reference year and the programme of the statistical data and metadata for population and housing censuses provided for by Regulation (EC) No 763/2008 of the European Parliament and of the Council (OJ L 105, 21.4.2017, p. 1).

([3]) Commission Implementing Regulation (EU) 2017/543 of 22 March 2017 laying down rules for the application of Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses as regards the technical specifications of the topics and of their breakdowns (OJ L 78, 23.3.2017, p. 13).

([4]) Commission Implementing Regulation (EU) 2017/881 of 23 May 2017 implementing Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses, as regards the modalities and structure of the quality reports and the technical format for data transmission, and amending Regulation (EU) No 1151/2010 (OJ L 135, 24.5.2017, p. 6).

(5) 'numerical value' means an integer number equal to or greater than '0' that provides statistical information on the observation of that data item;

(6) 'validated data' means data verified by the Member States according to agreed validation rules;

(7) 'observed value' means a numerical value that represents observed or imputed information to the best of knowledge based on all available 2021 census information, in particular before the application of any statistical disclosure control measures;

(8) 'confidential value' means a numerical value which, in order to protect the statistical confidentiality of the data, must not be divulged in accordance with the Member States' protective measures against disclosure of statistical data;

(9) 'special value' means a symbol that is transmitted in a data item instead of a numerical value;

(10) 'flag' means a code that can accompany a particular data item to describe a specific characteristic of its data value.

Article 3

Technical specifications of the 1 km² reference grid

1. In accordance with Section 1.5 of Annex IV of Regulation (EU) No 1089/2010, the statistical 1 km² reference grid for pan-European usage shall be the Equal Area Grid 'Grid_ETRS89-LAEA1000'. The spatial extent of the reference grid in the coordinate system specified for this grid in Section 2.2.1 of Annex II of the same Regulation shall be limited to easting values between 900 000 and 7 400 000 metres and northing values between 900 000 and 5 500 000 metres for the purpose of this Regulation.

2. In accordance with Section 1.4.1.1 of Annex IV of the same Regulation, each individual grid cell of the 1 km² reference grid shall be identified by a unique grid cell code, which is composed of the characters 'CRS3035RES1000mN'. This is followed by the northing value in metres of the grid point in the lower-left corner of the grid cell, followed by the character 'E', followed by the easting value in metres of the grid point in the lower-left corner of the grid cell.

3. The country code of the transmitting Member State as defined in the Interinstitutional Style Guide published by the Publications Office of the European Union, followed by the character '_', shall be prepended to the cell code of each grid cell transmitted by that Member State.

Article 4

Technical specifications of the 1 km² grid data topics and their breakdowns

The technical specifications of the topics set out in the Annex to Regulation (EU) 2017/543 for the 2021 census data shall apply. The breakdowns of the topics for the purpose of this Regulation are specified in Annex I of this Regulation.

Article 5

Programme of the 1 km² grid data

1. The programme of the 1 km² grid data to be transmitted by each Member State to the Commission (Eurostat) for the reference year 2021 is specified in Annex II.

2. Member States shall replace any confidential value with the special value 'not available'.

Article 6

Output harmonisation

1. In order to facilitate Union-wide comparability, the output data values to be disseminated shall be harmonised. To this end, preference shall be given to numerical values over special values as far as possible.

2. In order to ensure sufficiently accurate and reliable information on the spatial distribution of the total population, Member States shall respect the following requirements:

(a) data items on total population shall not be reported as confidential;

(b) data items on total population with an observed value other than '0' shall be marked with the flag 'populated'; and

(c) data items on total population with an observed value '0' shall not be marked with the flag 'populated'.

Article 7

Metadata

Member States shall provide the Commission (Eurostat) with metadata for the 1 km² grid data in accordance with Annex III.

Article 8

Reference date

The reference date of the 1 km² grid data transmitted by each Member State shall be identical to the reference date reported by that Member State in accordance with Article 3 of Regulation (EU) 2017/712.

Article 9

Data and metadata transmission date

1. Member States shall provide the Commission (Eurostat) with validated and aggregated data and with metadata on total population by 31 December 2022.

2. Without prejudice to paragraph 1, Member States shall provide the Commission (Eurostat) with validated and aggregated data and with metadata by 31 March 2024.

Article 10

Technical format for data and metadata transmission

The technical format to be used for data and metadata transmission shall be the SDMX format as implemented through the Census Hub. Member States shall transmit the required data and metadata in accordance with the data structure definitions and related technical specifications provided by the Commission (Eurostat). They shall store the required data and metadata until 31 December 2034 for any later transmission requested by the Commission (Eurostat).

Article 11

Quality requirements

1. Member States shall ensure the quality of the data transmitted.

2. For the purposes of this Regulation, the quality criteria referred to in Article 12(1) of Regulation (EC) No 223/2009 shall apply to the data to be transmitted.

3. At the request of the Commission (Eurostat), Member States shall provide it with additional information necessary to evaluate the quality of the statistical information.

Article 12

Dissemination

1. The Commission (Eurostat) shall disseminate the 1 km² grid datasets referred to in Article 5 as well as the associated metadata referred to in Article 7.

2. For the purposes of this Regulation, the programme of the 1 km^2 grid data and metadata to be transmitted by Member States and disseminated by Eurostat corresponds to the data that Member States disseminate at national level in accordance with Directive 2007/2/EC and its implementing Regulations (EC) No 1205/2008, (EC) No 976/2009 and (EU) No 1089/2010.

Article 13

Entry into force

This Regulation shall enter into force on the twentieth day following that of its publication in the *Official Journal of the European Union*.

This Regulation shall be binding in its entirety and directly applicable in all Member States.

Done at Brussels, 21 November 2018.

For the Commission
The President
Jean-Claude JUNCKER

ANNEX I

Technical specifications of the census topic breakdowns referred to in Article 4

The technical specifications of the breakdowns for the purpose of this Regulation of the census topics specified in the Annex of Regulation (EU) 2017/543 shall be presented as follows:

— Each topic selected for dissemination on the 1 km² reference grid is quoted with its heading from the Annex of Regulation (EU) 2017/543.

— The technical specifications laid down in the Annex of Regulation (EU) 2017/543 referring to that topic in general shall apply.

— Thereafter, the breakdown for that topic is specified.

— All breakdowns are designed to break down any total or subtotal referring to persons.

Topic: Place of usual residence

The breakdown categories of this topic on which a Member State shall report are all cells of the 1 km² reference grid specified in Article 3(1) whose area includes a part of the territory of that Member State, complemented by a single virtual grid cell per Member State to account for persons that are not allocated.

Geographical area based on the 1 km² reference grid [1]		GEO.G.
x.	All grid cells belonging partly or entirely to the territory of the Member State	x.
y.	One virtual grid cell per Member State	y.

[1] The codes 'x.' are grid cell identification codes as specified in Article 3. The code 'y.' shall consist of the character string 'unallocated' prepended by the country code of the reporting Member State as specified in Article 3(3).

If the place of usual residence of a person is unknown within the territory of the reporting Member State that is covered by the reference grid, additional scientifically-based, well-documented, and publicly available statistical estimation methods may be used to allocate this person to a specific grid cell. Persons who are not allocated to any cell of the reference grid shall be allocated to the virtual grid cell GEO.G.y. of that Member State.

Topic: Sex

The breakdown **SEX**. specified in the Annex of Regulation (EU) 2017/543 for this topic shall apply for the purpose of this Regulation.

Topic: Age

The following breakdown categories shall be reported:

Age		AGE.G.
1.	Under 15 years	1.
2.	15 to 64 years	2.
3.	65 years and over	3.

As specified in the Annex of Regulation (EU) 2017/543, the age reached in completed years of age at the reference date shall be reported.

Topic: Current activity status (number of employed persons)

The following breakdown category included in the breakdown **CAS.L.** specified in Regulation (EU) 2017/543 shall be reported:

Current activity status		CAS.L.
1.	Employed persons	1.1.

The specification of 'employed' persons in the Annex of Regulation (EU) 2017/543 shall apply for this category.

Topic: Country/place of birth

The following breakdown categories included in the breakdown **POB.L.** specified in Regulation (EU) 2017/543 shall be reported:

	Country/place of birth	POB.L.
1.	Place of birth in reporting country	1.
2.	Place of birth in other EU Member State	2.1.
3.	Place of birth elsewhere	2.2.

Topic: Place of usual residence one year prior to the census

The following breakdown categories included in the breakdown **ROY.** specified in Regulation (EU) 2017/543 shall be reported:

	Place of usual residence one year prior to the census	ROY.
1.	Usual residence unchanged	1.
2.	Move within the reporting country	2.1.
3.	Move from outside the reporting country	2.2.

A move within the same grid cell shall be reported either as a 'Move within the reporting country' (ROY.2.1.) or as a 'Move from outside the reporting country' (ROY.2.2.) as appropriate.

ANNEX II

Programme of the statistical census data geocoded to the 1 km² reference grid referred to in Article 5

The programme of the 1 km² grid data to be transmitted for the reference year 2021 shall consist of one two-dimensional table that cross-tabulates the set of grid cells **GEO.G.** defined in Annex I against the following selection of categories from the census topic breakdowns specified in Annex I:

Census topic categories to be broken down on the 1 km² reference grid		STAT.G.
0.	SEX.0.: Total population	0.
1.	SEX.1.: Male	1.
2.	SEX.2.: Female	2.
3.	AGE.G.1.: Under 15 years	3.
4.	AGE.G.2.: 15 to 64 years	4.
5.	AGE.G.3.: 65 years and over	5.
6.	CAS.L.1.1.: Employed persons [1]	6.
7.	POB.L.1.: Place of birth in reporting country	7.
8.	POB.L.2.1.: Place of birth in other EU Member State	8.
9.	POB.L.2.2.: Place of birth elsewhere	9.
10.	ROY.1.: Place of usual residence one year prior to the census unchanged	10.
11.	ROY.2.1.: Place of usual residence one year prior to the census: move within reporting country	11.
12.	ROY.2.2.: Place of usual residence one year prior to the census: move from outside of the reporting country	12.

[1] Data on the category 'employed persons' shall be transmitted as far as possible, subject to availability in the transmitting Member State.

ANNEX III

Required metadata for the 1 km² grid data referred to in Article 7

Metadata on data items

1. Where applicable, Member States shall add the following flags to a data item:
   (a) 'provisional';
   (b) 'populated';
   (c) 'revised';
   (d) 'see information attached';
   (e) 'confidential' ([1]).
2. Only data values on 'total population' which are reported under Article 9(1) and which are not considered final data by the Member State at the time of reporting shall be accompanied by the flag 'provisional'.
3. The flag 'populated' shall be applicable only to 'total population' data items under the provisions specified in Article 6(2).
4. For each data value accompanied by at least one of the flags 'revised' or 'see information attached' an explanatory text shall be provided.
5. Each data item whose confidential value has been replaced by the special value 'not available' shall be marked with the flag 'confidential'.

Metadata on the topics

In addition to the metadata on the topics transmitted to the Commission (Eurostat) under Article 6 of Regulation (EU) 2017/712, Member States shall provide metadata on each topic included in Annex I informing about the data sources and methodology used to obtain the data values for that topic on the 1 km² reference grid. In particular, the metadata shall contain:

— information on the reliability and accuracy of the reported data values;

— a description of any methodology used to estimate the data values on the 1 km² reference grid, including reliability and accuracy of the resulting data values;

— a description of any methodology used to allocate persons to specific grid cells under the topic 'place of usual residence', including information on the characteristics of persons under the category GEO.G.y.

Reference metadata

The metadata information and structure laid down in the Annex of Regulation (EU) 2017/881 shall be supplemented for the purpose of this Regulation by the following items specifically referring to the 1 km² grid data:

— Item 3.3. 'Processing and evaluation' shall be supplemented by the additional sub-item 3.3.3. 'Additional information on generic (not topic-related) methodology applied in order to produce the 1 km² grid dataset'.

— Item 3.4. 'Dissemination' shall be supplemented by specific information about statistical disclosure control measures related to the 1 km² grid dataset. Member States shall provide the Commission (Eurostat) with information about the measures related to the harmonised protection of 1 km² grid data, in particular if they used the ESS good practices and implementation guidelines for the harmonised protection of 1 km² grid data.

— Item 4.2. 'Timeliness and punctuality' shall be supplemented by specific calendar date(s) of the transmission and possible revisions of the 1 km² grid data and metadata.

— Item 4. 'Assessment of data quality' shall be supplemented by the additional sub-item 4.7 'Geographic information – data quality' covering geographic quality principles, in particular territorial coverage and comparability, positional accuracy, as well as temporal coherence and completeness of the geographic data used for geocoding.

([1]) This flag is not applicable to data items on total population, as laid down in Article 6(2)(a).

Getting in touch with the EU

In person

All over the European Union there are hundreds of Europe Direct Information Centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by e-mail

Europe Direct is a service that answers your questions about the European Union. You can contact this service

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696 or
- by electronic mail via: https://europa.eu/european-union/contact_en

Finding information about the EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU Publications

You can download or order free and priced EU publications at: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

Open data from the EU

The EU Open Data Portal (http://data.europa.eu/euodp/en) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.

PDF: KS-GQ-18-010-EN-N
Print: KS-GQ-18-010-EN-C

EU legislation on the 2021 population and housing censuses

EXPLANATORY NOTES

Censuses provide a precise and geographically detailed account of population characteristics. This makes them a rich source for analysis, providing key inputs for policy, administration and research. At the EU level, national results are of greater value if they can be compared between Member States, so the EU has been taking continuous steps to harmonise outputs.

The 2021 round also comes at a time of fundamental transition: administrative data sources have become the backbone of the next census in most Member States. This brings efficiency gains, leading to more powerful census systems and addressing fast-evolving user expectations, where the regional detail is increasingly important. A dedicated law to publish key census topics on an EU-wide 1 km^2 grid is a major innovation here. It will allow for much more flexible analysis, even at cross-border level, tailored to research or policy needs.

This publication explains the EU legislation for the 2021 census round, with a focus on the new grid data.

For more information
https://ec.europa.eu/eurostat/

Publications Office

Print: ISBN 978-92-79-96919-5 doi:10.2785/139261
PDF: ISBN 978-92-79-96918-8 doi:10.2785/513433