

STATISTIKENS FRAMSTÄLLNING

Högutbildade utrikes födda

Ämnesområde

Utbildning och forskning

Statistikområde

Befolkningens utbildning

Produktkod

UF0529

Referenstid

2018

Kontaktuppgifter

Statistikansvarig myndighet	Statistiska centralbyrån
Kontaktinformation	Tomas Westling
E-post	fornamn.efternamn@scb.se
Telefon	010-479 61 78

Innehåll

1	Statistikens sammanhang.....	3
2	Undersökningsdesign	3
2.1	Målstorheter	3
2.2	Ramförfarande	3
2.3	Förfaranden för urval och uteslutning	4
2.3.1	Urvalsförfarande.....	4
2.3.2	Uteslutning från insamling (cut-off)	4
2.4	Insamlingsförfarande.....	4
2.4.1	Datainsamlingsmetoder	4
2.4.2	Mätning.....	5
2.4.3	Bortfallsuppföljning	5
2.5	Bearbetningar.....	5
2.6	Granskning.....	6
2.6.1	Granskning under insamlingen	6
2.6.2	Granskning av mikrodata	7
2.6.3	Granskning av makrodata	7
2.6.4	Granskning av redovisning	7
2.7	Skattningsförfarande	7
2.7.1	Principer och antaganden	7
2.7.2	Skattningsförfarande för målstorheter.....	9
2.7.3	Skattningsförfarande för tillförlitlighet.....	9
2.7.4	Röjandekontroll	10
3	Genomförande	10
3.1	Kvantitativ information.....	10
3.2	Avvikelser från undersökningsdesignen	10

1 Statistikens sammanhang

Statistiken beskriver situationen på arbetsmarknaden för högutbildade utrikes födda i åldern 25–64 år som invandrade till Sverige under perioden 2006–2017. Med högutbildad avses en person som hade minst en treårig eftergymnasial utbildning enligt den senaste versionen av SCB:s Utbildningsregister. I undersökningen ingår även inrikes födda personer i åldern 25–64 år med minst en treårig eftergymnasial utbildning som en jämförelsegrupp.

Statistiken bygger på en enkätundersökning som genomfördes under hösten 2018.

I detta dokument beskrivs översiktligt design och genomförande av statistiken om Högutbildade utrikes födda. Dokumentet kompletterar *kvalitetsdeklarationen* för Högutbildade utrikes födda som finns tillgänglig på www.scb.se/UF0529 under rubriken *Dokumentation*.

2 Undersökningsdesign

2.1 Målstorheter

De statistiska målstorheterna i undersökningen är totaler och andelar som belyser situationen på arbetsmarknaden under en viss mätvecka hösten 2018. De viktigaste målstorheterna rör huvudsaklig sysselsättning och arbetets överensstämmelse med utbildningen.

2.2 Ramförfarande

Undersökningens rampopulation konstruerades med hjälp av Registret över befolkningens utbildning (Utbildningsregistret, 2017-12-31) som är ett internt register på SCB. Ramen omfattade 175 865 utrikes födda personer mellan 25 och 64 år som hade minst en treårig eftergymnasial utbildning samt hade invandrat till Sverige mellan 2006 och 2017. I ramen har inte ingått personer med motsvarande 180 eller fler avklarade högskolepoäng (utbildningsnivå 532 enligt svensk utbildningsnomenklatur för klassificering av utbildning, SUN 2000) som högsta utbildning. Det innebär att 3 355 personer exkluderades från urvalsramen.

Populationen stratifierades efter världsdel (Norden; EU28 utom Norden; Övriga Europa utom Norden och EU28; Afrika; Asien; Övriga), utbildningsinriktning (pedagogik; humaniora; samhällsvetenskap; naturvetenskap och teknik; hälso- och sjukvård; övriga utbildningsinriktningar) och ålder (25-34 år; 35-44 år; 45-54 år; 55-64 år). Sammanlagt bildades 144 strata.

Ramen för kontrollgruppen (inrikes födda) omfattade 926 143 personer i åldern 25 till 64 år med minst treårig eftergymnasial utbildning. I ramen har inte ingått personer med motsvarande 180 eller fler avklarade högskolepoäng (utbildningsnivå 532 enligt svensk

utbildningsnomenklatur för klassificering av utbildning, SUN 2000) som högsta utbildning. Det innebar att 148 472 personer exkluderades från ramen.

Kontrollgruppen stratifierades på utbildningsinriktning (pedagogik, humaniora och samhällsvetenskap; naturvetenskap och teknik; hälso- och sjukvård; övriga utbildningsinriktningar), ålder (25-34 år; 35-44 år; 45-54 år; 55-64 år) och kön. Sammanlagt bildades 32 strata.

Från rampopulationen har personer som inte var folkbokförda i Registret över totalbefolkning (RTB) 2018-06-30 exkluderats. De individer som sedan återstår utgör observationsobjekt liksom uppgiftskällor. Kontaktvägen till uppgiftskällorna är folkbokföringsadress enligt RTB.

2.3 Förfaranden för urval och uteslutning

2.3.1 Urvalsförfarande

Undersökningen är en urvalsundersökning. Från ramen drogs ett stratifierat obundet slumpmässigt urval om totalt 13 006 personer, 10 007 utrikes födda och 2 999 inrikes födda personer. Ramen för utrikes födda stratifierades efter födelseland (världsdel), utbildningsinriktning och ålder. Ramen för inrikes födda stratifierades efter utbildningsinriktning, ålder och kön. Vid bestämningen av urvalsstorleken togs hänsyn till det förväntade bortfallet.

Urvalet fördelades så att den förväntade felmarginalen för redovisningsgruppen "Övriga utbildningsinriktningar" tilläts vara något större än övriga redovisningsgrupper.

Urvalsstorleken i ett visst stratum beror till viss del av hur många personer som finns där (N), dels vad som ska mätas (p), och dels hur stor felmarginalen tilläts vara (B). Parametern p är okänd. Som planeringsvärde för p användes 0,5. Parametern B är approximativt felmarginalen kring ett skattat värde p för ett visst stratum.

Urvalsstorleken n fås ur formeln:

$$n = (Npq)/(N-1)D + pq; \text{ där } q = 1-p \text{ och } D=B^2/4$$

2.3.2 Uteslutning från insamling (cut-off)

Ingen del av målpopulationen utesluts från direktinsamling.

2.4 Insamlingsförfarande

2.4.1 Datainsamlingsmetoder

Statistiken baseras på uppgifter insamlade direkt från de individer som ingick i undersökningen. De variabler som används för att skapa

redovisningsgrupper hämtades främst från SCB:s utbildningsregister (UREG).

Baserat på erfarenheter från tidigare undersökningsomgångar genomfördes undersökningen som pappers- och webbformulär och utskick skedde till personernas folkbokföringsadress. Det första utskicket skedde den 18 september 2018 och inkluderade en pappersenkät samt ett brev med information om undersökningen och inloggningsuppgifter till webbformuläret. Ett tack- och påminnelsekort skickades ut den 1 oktober och även detta inkluderade inloggningsuppgifter. Den 15 oktober skickades ett påminnelsebrev ut, vilket inkluderade både pappersenkät och inloggningsuppgifter. Datainsamlingen avslutades den 26 november 2018.

2.4.2 Mätning

Utrikes födda kunde välja mellan att besvara frågeformuläret på svenska eller engelska. De fick även informationsbrev och samtliga påminnelser på både svenska och engelska.

För mer information om innehåll i enkäten, se avsnitt *Bilagor* i temarapporten "Arbetsmarknaden 2018 för högutbildade utrikes födda" på SCB:s webbplats, www.scb.se/UF0529.

Information om de slutliga observationsvariablerna och statistikens detaljerade innehåll finns på SCB:s webbplats. Där beskrivs alla variabler och värdemängder m.m. Dokumentationen finns på www.metadata.scb.se under *Högutbildade utrikes födda*.

Det går även att i kvalitetsdeklarationen hitta information om observationsvariablerna.

2.4.3 Bortfallsuppföljning

Objektsbortfallet utgörs av personer som har valt att inte besvara enkäten eller som inte kunnat nå exempelvis på grund av att de saknat folkbokföringsadress. Påminnelsebrev skickades ut vid två tillfällen. Båda påminnelserna var bestämda på förhand och räknas som fasta åtgärder.

2.5 Bearbetningar

Den kodning som görs avser variabler som används för att skapa redovisningsgrupper. Dessa variabler hämtas från SCB:s utbildningsregister.

Redovisningsgruppen *Utbildningsinriktning* skapas från uppgift om inriktning på personens högsta utbildning enligt Svensk utbildningsnomenklatur (SUN).

Uppgifterna ligger sedan till grund för följande indelning:

- Pedagogik, humaniora och samhällsvetenskap
- Naturvetenskap och teknik
- Hälsa- och sjukvård
- Övriga utbildningsinriktningar

Redovisningsgruppen *Invandringsår* skapas från uppgift om senaste datum då personen blev folkbokförd hos Skatteverket och grupperas enligt följande:

- 2006–2009
- 2010–2013
- 2014–2017

Födelseregion skapas utifrån uppgift om födelseland och grupperas enligt följande:

- Norden
- EU28 (utom Norden)
- Övriga Europa (utom Norden och EU28)
- Afrika
- Asien
- Övriga födelseregioner

Redovisningsgruppen *Invandringsskäl* hämtas direkt från en av enkätfrågorna.

I de fall då uppgifter saknas på vissa av frågorna för de svarande i undersökningen sker ingen imputering. Partiellt bortfall redovisas som *Uppgift saknas* i tabellerna.

För varje svarande elev har en vikt beräknats. Vikterna har beräknats utifrån urvalsdesignen samt antaganden om objektsbortfall och täckningsfel. Läs mer i avsnitt 2.7.2 *Skattningsförfarande för målstorheter*.

2.6 Granskning

De insamlade uppgifterna granskas under insamlingen, efter avslutad insamling samt i aggregerad form.

2.6.1 Granskning under insamlingen

Registrering av data från pappersblanketter skedde genom att inkomna frågeformulär lästes in maskinellt (skanning). Skanningsprogrammet kontrollerade så att korrekta värden registrerades vid skanningen, till exempel om svar var otydligt markerade. De värden som föll ut som fel vid dessa kontroller rättades upp manuellt med stöd av skanningsprogrammet, genom så kallad verifiering. Vid förekomst av dubbelmarkeringar, det vill säga

om fler svarsalternativ fanns angivna för en fråga som bara fick ha ett svar, gjordes en markering i datafilen för senare manuell kontroll.

2.6.2 Granskning av mikrodata

När alla inkomna svar var registrerade och samlade i en datafil kontrollerades materialet ytterligare. Rättning av dubbelmarkeringar skedde nu efter en viss prioritetsordning. I vissa fall gjordes rättningen utifrån svaren på andra frågor i enkäten. Ytterligare kontroller som gjordes var bland annat att endast valida värden förekom i materialet samt att så kallade hoppinstruktioner hade efterföljts och att rätt antal svar för varje fråga hade lämnats.

2.6.3 Granskning av makrodata

En granskning på makronivå görs av de variabler som ingår i den publicerade statistiken. Jämförelser görs med föregående undersökningsomgångs statistik och eventuella förändringar över tid undersöks mer ingående.

2.6.4 Granskning av redovisning

Alla tabeller och diagram läggs ut på intern server, och granskas igen innan de publiceras. Det kontrolleras att antals- och andelsuppgifter i temarapporten, tabellerna på webbplatsen och i statistiknyheten stämmer överens med grunddata. Rubriker och förklaringar till tabeller och diagram granskas så att de är korrekta. En läsgrupp har lämnat synpunkter på innehållet i temarapporten.

2.7 Skattningsförfarande

2.7.1 Principer och antaganden

För att skatta målstorheterna används ett förfarande som baseras på s.k. kalibrerade uppräkningsvikter. Detta förfarande, som utnyttjar hjälpinformation i form av registervariabler, används främst för att kompensera för den skevhet som kan förmodas uppstå till följd av de bortfalls- och täckningsproblem som finns, men även för att reducera den slumpmässiga osäkerhet, den varians, som skattningsförfarandet innebär.

De hjälpvariabler som används i skattningsförfarandet har valts utifrån kriterier om hur de samvarierar med svarsbenägenhet och viktiga undersökningsvariabler.

I tablan nedan beskrivs hjälpvariablerna och de kategorier som används.

Hjälpvariabel	Kategorier
Yrke	Yrke som kräver högskoleutb.
	Yrke som ej kräver högskoleutb.
	Yrke okänt
Utbildningsinriktning	Pedagogik
	Humaniora
	Samhällsvetenskap
	Naturvetenskap och teknik
	Hälso- och sjukvård
	Övriga utbildningsinriktningar
Inkomst	Mindre än 100 000
	100 000 -299 999
	300 000 eller mer
Grund för bosättning	Uppgift saknas
	Studier
	Arbete
	Anhöriga
	Flyktingskäl
	Övriga skäl
Källa för uppgift om högsta utbildning	Svensk utbildning
	Utländsk utbildning
	Övriga
Födelseland	Norden
	EU28 (utom Norden)
	Övriga Europa
	Afrika
	Asien
	Övriga
Kön	Man
	Kvinna
Invandringsår	2006-2009
	2010-2013
	2014-2017
Ålder	25-34
	35-44
	45-54
	55-64

Hjälpvariablerna *Grund för bosättning*, *Källa för uppgift om högsta utbildning*, *Födelseland* och *Invandringsår* är endast aktuella för utrikes födda personer. För inrikes födda har vissa utbildningsinriktningar grupperats ytterligare.

Utifrån de valda hjälpvariablerna beräknas därefter kalibrerade vikter. Vikterna har egenskapen att de vid skattning av populationstotalen, för var och en av de hjälpvariabler som används för att fastställa de kalibrerade vikterna, resulterar i en skattning som överensstämmer exakt med den registertotal som används för att fastställa de kalibrerade vikterna. För utrikes födda har vikterna justerats i avseende att kompensera för den övertäckning som urvalsramen innehåller. Detta innebär att de skattade populationstotalerna är lägre än motsvarande registertotaler. Ett antagande görs att övertäckningen i urvalsramen är lika stor som den övertäckning som observerats under insamlingen. Ingen justering för undertäckning har gjorts för utrikes födda.

Vikterna kan beskrivas med formeln:

$$w_k = d_k \times v_k$$

där

w_k = den totala vikten för objekt k

d_k = designvikt

v_k = kalibreringsvikt

Designvikten är den del av vikten som beror på urvalsdesignen. Kalibreringsvikten justerar för bortfall och täckningsbrister.

2.7.2 Skattningsförfarande för målstorheter

För beräkning av skattningen av totaler används följande formel:

$$\hat{y} = \sum_r w_k y_k$$

där

y_k = variabelvärde för objekt k

summering sker av de svarande (r)

För beräkning av skattningen av medelvärden används följande formel:

$$\hat{\bar{y}} = \frac{\sum_r w_k y_k}{\sum_r w_k}$$

2.7.3 Skattningsförfarande för tillförlitlighet

Konfidensintervallen är beräknade enligt en metod som med 95 procents säkerhet täcker in det sanna värdet och beräknas enligt formeln

punktskattning $\pm 1,96$ * (skattning av estimatorns medelfel)

där såväl punkt- som medelfelsskattning beräknas med hjälp av ett av SCB egenutvecklat program, ETOS2.

2.7.4 Röjandekontroll

Ingen röjandekontroll har gjorts utifrån bedömningen att inga skyddsmetoder behövs. Undersökningen är en urvalsundersökning och statistiken som publiceras är inte på en så pass detaljerad nivå att det bedömts föreligga risk för röjande.

Vid publicering redovisas dock inte skattningar på enskilda frågor med färre än 30 svarande och inga enskilda cellvärden med färre tre svarande redovisas. Alltför osäkra skattningar redovisas inte heller.

3 Genomförande

3.1 Kvantitativ information

Undersökningens rampopulationer bestod vid urvalsdragningen av 175 865 utrikes födda personer och 926 143 inrikes födda personer. Från rampopulationerna drogs två stratifierade obundna slumpmässiga urval. Bland utrikes födda drogs ett urval av 10 007 personer och bland inrikes födda drogs ett urval av 2 999 personer.

Svarsfrekvensen var 38 procent bland utrikes födda och 59 procent bland inrikes födda.

3.2 Avvikelser från undersökningsdesignen

Inga avvikelser har gjorts.