

# STATISTIKENS FRAMSTÄLLNING

## Befolkningens utbildning

### Ämnesområde

Utbildning och forskning

### Statistikområde

Befolkningens utbildning

### Produktkod

UF0506

### Referenstid

31 december 2021 (befolkningens storlek)

Vårterminen 2021 (avklarade utbildningar)

### Kontaktuppgifter

<b>Statistikansvarig myndighet</b>	Statistiska centralbyrån
<b>Kontaktinformation</b>	<b>Tomas Westling</b> , Sektionen för utbildningsstatistik <b>Michael Karlsson</b> , Sektionen för individregister
<b>E-post</b>	Fornamn.efternamn@scb.se
<b>Telefon</b>	010 - 479 40 00

## Innehåll

1	Statistikens sammanhang.....	3
2	Undersökningsdesign .....	3
2.1	Målstorheter .....	3
2.2	Ramförfarande .....	3
2.3	Förfaranden för urval och uteslutning .....	3
2.4	Insamlingsförfarande.....	4
2.4.1	Datainsamling .....	4
2.4.2	Mätning .....	4
2.4.3	Bortfallsuppföljning.....	5
2.5	Bearbetningar.....	5
2.6	Granskning.....	7
2.6.1	Granskning under direktinsamlingen .....	7
2.6.2	Granskning av mikrodata och insamlade statistikvärden.....	7
2.6.3	Granskning av makrodata .....	8
2.6.4	Granskning av redovisning .....	8
2.7	Skattningsförfarande .....	8
2.7.1	Principer och antaganden .....	8
2.7.2	Skattningsförfarande för målstorheter.....	9
2.7.3	Skattningsförfarande för tillförlitlighet.....	9
2.7.4	Röjandekontroll .....	9
3	Genomförande .....	9
3.1	Kvantitativ information.....	9
3.2	Avvikelser från undersökningsdesignen .....	9

## 1 Statistikens sammanhang

Statistiken över befolkningens utbildning mäter utbildningsnivå och utbildningsinriktning. Statistiken bygger på individregistret Befolkningens utbildning - i dagligt tal kallat Utbildningsregistret (UREG). Registrets första version avsåg förhållandena per 1985-12-31. Därefter genomförs årliga uppdateringar. Före 1985 togs detaljerad och totalräknad statistik över utbildningsnivå fram i samband med folk- och bostadsräkningarna 1930 respektive 1970.

Till statistiken över befolkningens utbildning är också knutet vård och utveckling av Svensk utbildningsnomenklatur SUN, som är den officiella svenska klassificeringen av utbildningsnivå- och utbildningsinriktning.

I detta dokument beskrivs översiktligt upplägg och genomförande av statistiken om befolkningens utbildning. Dokumentet kompletterar *kvalitetsdeklarationen* för Befolkningens utbildning som finns tillgänglig på [www.scb.se/UF0506](http://www.scb.se/UF0506) under rubriken *Dokumentation*.

Under första kvartalet av 2020 började covid-19 spridas i Sverige och i världen. Det har inte påverkat framställningen av Utbildningsregistret.

## 2 Undersökningsdesign

### 2.1 Målstorheter

De viktigaste målstorheterna i undersökningen är antalet och andelen personer inom olika utbildningsnivåer och inriktningar till och med vårterminen respektive år. Statistiken redovisas efter exempelvis kön och ålder. De vanligaste åldersavgränsningarna vid redovisning av utbildningsnivå och utbildningsinriktning är 25-64 år respektive 16 år och äldre.

### 2.2 Ramförfarande

Ramen för Utbildningsregistret är Registret över totalbefolkningen (RTB). De objekt för vilka variabelvärden observeras är poster i RTB där varje post motsvarar en person.

Undersökningens observationsobjekt är individer. Rampopulation för Utbildningsregistret är samtliga folkbokförda personer enligt Registret över totalbefolkningen (RTB).

### 2.3 Förfaranden för urval och uteslutning

Inget urval och ingen uteslutning förekommer i statistiken över befolkningens utbildning.

## 2.4 Insamlingsförfarande

### 2.4.1 Datainsamling

Statistiken över befolkningens utbildning bygger på uppgifter från SCB:s utbildningsregister (UREG). Utbildningsregistret baseras till största del på registeruppgifter och en kompletterande årlig enkätundersökning. Både interna och externa register används, de huvudsakliga källregistren är dock interna register på SCB.

Några av de interna registren är

- Universitets- och högskoleregistret, examinerade
- Universitets- och högskoleregistret, avklarade poäng
- Universitets- och högskoleregistret, Registret över forskarutbildade
- Register över avgångna från gymnasieskolan
- Avgångna från grundskolan
- Folk- och bostadsräkningen 1990 (FoB90)

Några av de externa registren är

- Arbetsförmedlingen (AF), Utbildningsbakgrund
- Socialstyrelsens register över hälso- och sjukvårdspersonal

För gruppen nyinvandrade utrikes födda (20-59 år) som under året (oktober-september före aktuellt årsskifte) blivit folkbokförda och saknar uppgift om utbildning i Utbildningsregistret genomförs en årlig enkätundersökning (Utbildning i annat land än Sverige) där uppgifter om personernas högsta utbildning insamlas.

Uppgifter om i populationen ingående personers aktuella folkbokföringsförhållanden samt födelseland och medborgarskap erhålls via RTB.

### 2.4.2 Mätning

För information om vilka mätinstrument som användes vid den primära datainsamlingen till respektive register som används i Utbildningsregistret hänvisas till dokumentation för respektive register. Utbildningsregistret kompletteras med några "egna" insamlingar med enkäter där det inte finns några tillgängliga uppgifter via register. Dels görs en enkät, "Utbildning i annat land än Sverige", till nyinvandrade utrikes födda och dels görs en insamling av examinerade från vissa eftergymnasiala utbildningar som inte rapporteras via reguljära högskolestatistiken eller yrkeshögskolestatistiken.

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](http://www.metadata.scb.se) under *Befolkningens utbildning*.

Det går även i Kvalitetsdeklarationen för Befolkningens utbildning hitta information om observationsvariablerna.

### **2.4.3 Bortfallsuppföljning**

Eftersom statistiken över befolkningens utbildning är registerbaserad, förekommer inget objektsbortfall i egentlig mening. I Utbildningsregistret förekommer dock att personer dels helt saknar uppgift om utbildning, dels saknar uppgift om en viss avklarad utbildning (men har uppgift avseende annan utbildning). Personer som aldrig har gått i skola redovisas också som bortfall/uppgift saknas. Ingen möjlighet finns dock att kontrollera vilka som saknar *uppgift om utbildning* och vilka som *helt saknar utbildning*. I arbetet med att framställa Utbildningsregistret undersöks kontinuerligt möjligheterna att samla in uppgifter om avklarade utbildningar utomlands via nya källor.

## **2.5 Bearbetningar**

För varje individ som ingår i Utbildningsregistret aktuellt år läggs samtliga kända utbildningar från respektive källregister in i Utbildningsregistrets ackumulerade s.k. "bruttofil". Uppgifterna om utbildning från källregistren kodas enligt Svensk utbildningsnomenklatur (SUN 2020). Visst material kodas direkt vid källan av respektive "dataleverantör" såsom exempelvis data från Universitets- och högskoleregistret. De flesta källors utbildningsdata kodas av UREG-personal, såsom exempelvis registret över Gymnasieavgångna och Universitets- och högskolerådets och Skolverkets evalveringsdata.

### **Svensk utbildningsnomenklatur (SUN 2020)**

Svensk utbildningsnomenklatur (SUN 2020) är en standard för klassificering av enskilda utbildningar samtidigt som den ger ett system för sammanförande till grupper av utbildningar.

De senaste 20 årens förändringar i utbildningssystemet ställde krav på en uppdaterad SUN. Våren 2019 fastställde SCB den nya versionen av SUN, benämnd SUN 2020 (tidigare SUN 2000), som innehåller en uppdaterad inriktningsstruktur medan nivåstrukturen är oförändrad. I Utbildningsregistret implementerades SUN 2020 för läsåret 2018/19 med publicering våren 2020. Utöver en uppdaterad struktur för inriktningar har innehållsbeskrivningar för såväl nivåer som inriktningar förbättrats och förtydligats med fler exempel på utbildningar som ingår eller inte ingår i respektive kod.

Standarden är anpassad till den internationella nomenklaturen ISCED97. För ytterligare information hänvisas till *SUN 2020 Svensk utbildningsnomenklatur*. Publikationen finns att hämta i PDF-format på SCB:s webbplats, [www.scb.se/sun](http://www.scb.se/sun).

För de viktigaste förändringarna i den redovisade statistiken vid övergången från SUN 2000 till SUN 2020, se *Kvalitetsdeklarationen*, avsnitt 5.1 *Jämförbarhet över tid*.

### **Val av högsta utbildning**

Från den ackumulerade s.k. "bruttofilen" över varje persons samtliga kända utbildningar väljs den högsta utbildningen enligt på förhand bestämda prioriteringsregler.

Prioriteringsreglerna bygger på antaganden om kvaliteten i olika källregister. Detta kan leda till bearbetningsfel om den riktiga högsta utbildningen finns i ett källregister som är lågt prioriterat. Vid klassificeringen av utbildningsnivå och utbildningsinriktning enligt SUN 2020 kan det förekomma bearbetningsfel när Utbildningsregistret har använt uppgifter som översatts från gamla SUN till SUN 2000 och som sedan översatts till SUN 2020.

En utbildning för varje individ väljs som dennes f.n. högsta utbildning enligt följande:

I första hand väljs utbildningen med högst utbildningsnivå enligt SUN 2020.

I händelse av att flera utbildningar har samma nivå, används i tur och ordning max-värdet av följande variabler för att välja ut den högsta utbildningen:

högsta(=senaste) examensår/termin,

högsta (=bästa enligt intern kodsättning) källa,

högsta(=senaste) år för inläggning.

Vissa källor behandlas p.g.a. sin lägre kvalitet annorlunda vid inläggning i "genomgångna utbildningar". Dessa källor används endast för personer som helt saknar poster från källor som bedöms vara av bättre kvalitet.

För varje person som är folkbokförd i Sverige respektive årsskifte väljs den f.n. högsta utbildningen.

Högsta utbildning avser vårterminens slut för respektive årsskifte. För vissa personer saknas uppgift om utbildning helt.

Ingen imputering görs för personer som saknar uppgift om högsta utbildning.

## 2.6 Granskning

Granskning av statistiken över befolkningens utbildning sker i flera led. Dels sker en granskning vid framställning av respektive källregister, dels sker en granskning vid framställningen av Utbildningsregistret och till sist granskas de tabeller som framställs över befolkningens utbildning i samband med publicering.

### 2.6.1 Granskning under direktinsamlingen

Vid datainsamlingen från administrativa källor utanför SCB, som exempelvis Arbetsförmedlingen, används en mall för databeställningar där population, referenstid och variabler specificeras. När datamaterialet kommer till SCB kontrolleras att levererat material motsvarar det som angetts i beställningen.

SCB:s enkätenhet samlar in det frågeformulär som årligen går ut till nyinvandrade utrikesfödda och som utgör en indatakälla till Utbildningsregistret. De inkomna frågeformulären granskas i flera olika skeden av bearbetningen. Inflödet följs upp och jämförelser görs med tidigare år med avseende på svarsfrekvenser och ett flertal processvariabler.

Universitets- och högskoleregistret som är en källa till Utbildningsregistret, har sin insamling via LADOK. Granskning vid insamlingen sker inom Universitets- och högskoleregistret.

SCB samlar på uppdrag av Skolverket in uppgifter om förskola, fritidshem, skola, vuxenutbildning och annan pedagogisk verksamhet. En del av Skolverksregistren är också viktiga indatakällor för UREG. Granskningen görs i framställningen av varje skolregister.

### 2.6.2 Granskning av mikrodata och insamlade statistikvärden

Dubblettkontroller görs i varje indatakälla, både på personnummer och på den kombination av variabler som ligger till grund för val av högsta utbildning, den så kallade nyckeln. De variabler som ingår i nyckeln är för varje person utbildningsnivå, utbildningsinriktning, examensår, examenstermin, källkod och källår.

I första hand väljs utbildningen med högst utbildningsnivå enligt SUN 2020. Om flera utbildningar har samma nivå, används i tur och ordning maxvärdet av resterande variabler för att välja ut den högsta utbildningen. I händelse av att det fortfarande finns dubletter på nyckeln läggs även utbildningsinriktning till i maxvalet. Då väljs raden med högsta inriktningen.

Rimlighetskontroller görs i varje enskild källa under arbetets gång. Förutom dubblettkontroller görs kontroller av valida värden i utbildningsnivå och utbildningsinriktning. Rimligheten kontrolleras också med avseende på om de utbildningsnivåer som förekommer i

källan är valida för just den källan. Liknande kontroller görs också för utbildningsinriktning. Rimlighetskontroller görs också för att upptäcka fel i kombinationer av utbildningsnivå- och inriktning.

Examensår med tidigare årtal eller samma årtal som när personen är född kontrolleras och sätts till okänt värde; \*\*\*\*. Personernas examenstermin sätts i dessa fall också till okänd.

I varje källa kontrolleras populationens storlek jämfört med tidigare år. Ogiltiga personnummer med avseende på bokstäver eller blanksteg kontrolleras.

Personer med samordningsnummer kontrolleras mot RTB och i de fall det finns en koppling till ett riktigt personnummer används detta i UREG.

I registerproduktionen körs också en procedur som tar fram personernas senaste personnummer. Uppdateringar av personnummer görs utifrån resultatet av proceduren. Dubbletthantering kan behövas efter uppdateringen.

### **2.6.3 Granskning av makrodata**

Utöver de makrokontroller som görs i registerproduktionen av bl.a. frekvenser och jämförelser med tidigare årgångar tas ett antal Exceltabeller fram. Exceltabellerna innehåller aggregat där olika kombinationer av variabler granskas mot varandra.

Jämförelser görs också med statistik som är publicerad på de externa källornas hemsidor.

### **2.6.4 Granskning av redovisning**

Alla tabeller i statistiken över befolkningens utbildning granskas, läggs ut på intern server, och granskas igen innan de publiceras. Granskning sker genom jämförelser mot föregående års motsvarande tabeller. Publicering av statistiken över befolkningens utbildning sker både i Statistikdatabasen och som fasta Exceltabeller. Kontroller görs så att redovisningen i de båda publiceringsformerna är konsistenta. Vidare görs kontroller mot publicerade tabeller med parallella produktionsskript.

## **2.7 Skattningsförfarande**

### **2.7.1 Principer och antaganden**

Statistiken över befolkningens utbildning bygger på uppgifter från Utbildningsregistret som i sin tur bygger på antal interna och externa källregister. Utgångspunkten för skattningsförfarandet är att datamaterialet är näst intill heltäckande vilket ska möjliggöra totalräkning av statistiken. I kvalitetsdeklaration för Befolkningens utbildning finns statistikens osäkerhetskällor beskrivna. I de fall

uppgift om en individs utbildningar helt saknas redovisas detta som "Uppgift saknas" i de tabeller som publiceras. Det berör framför allt utrikes födda personer som endast har utländska utbildningar.

#### **2.7.2 Skattningsförfarande för målstorheter**

De statistiska målstorheterna i statistiken över befolkningens utbildning består av antal och andel personer med en viss utbildningsnivå och utbildningsinriktning. Dessa skattas genom summering av det totala antalet aktuella registerposter i aktuell version av Utbildningsregistret.

#### **2.7.3 Skattningsförfarande för tillförlitlighet**

Ingen skattning över tillförlitligheten genomförs i statistiken över befolkningens utbildning.

#### **2.7.4 Röjandekontroll**

En bedömning har gjorts att inga skyddsmetoder behövs. Undantaget är en tabell med redovisning av utbildningsnivå efter födelseland. Där särredovisas endast födelseländer med fler än 500 personer folkbokförda i Sverige.

### **3 Genomförande**

#### **3.1 Kvantitativ information**

Vid redovisning av statistik över befolkningens utbildning ingår personer 16 år och äldre. Storleken på den populationen i årgång 2021 av Utbildningsregistret är 8 491 000 personer. Av dessa saknar 2,6 procent helt och hållet uppgift om utbildning i registret. Andelen för utrikes födda är 9,2 procent, och 0,6 procent för inrikes födda.

En vanlig delpopulation som redovisas är personer i åldern 25-64 år. Storleken på den populationen är 5 325 000 personer. Av dessa saknar 2,6 procent helt och hållet uppgift om utbildning i registret. Andelen för utrikes födda är 8,4 procent, och 0,5 procent för inrikes födda.

#### **3.2 Avvikelser från undersökningsdesignen**

Inga avvikelser har gjorts.