

STATISTIKENS FRAMSTÄLLNING

Undersökningen om vuxnas deltagande i utbildning

Ämnesområde

Utbildning och forskning

Statistikområde

Befolkningens utbildning

Produktkod

UF0538

Referenstid

Aktuella förhållanden augusti 2022 - januari 2023

Genomförda utbildningar 12 månader bakåt från mättillfället som var någon gång mellan augusti 2022 - januari 2023

Kontaktuppgifter

Statistikansvarig myndighet	Statistiska Centralbyrån
Kontaktinformation	Mattias Fritz
E-post	Mattias.fritz@scb.se

Telefon	010 – 479 62 67
----------------	-----------------

Innehåll

1	Statistikens sammanhang	4
2	Undersökningsdesign	4
2.1	Målstorheter	4
2.2	Ramförfarande	4
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	5
2.4.1	Datainsamling	5
2.4.2	Mätning	5
2.4.3	Bortfallsuppföljning	6
2.5	Bearbetningar	7
2.6	Granskning	7
2.6.1	Granskning under direktinsamlingen	8
2.6.2	Granskning av mikrodata och insamlade statistikvärden	8
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	8
2.7.3	Skattningsförfarande för tillförlitlighet	9
2.7.4	Röjandekontroll	11
3	Genomförande	11
3.1	Kvantitativ information	11
3.2	Avvikelser från undersökningsdesignen	11

1 Statistikens sammanhang

Statistiken bygger på en intervjuundersökning riktad mot vuxna personer i åldern 18–69 år. Undersökningen är EU-reglerad och visar bland annat andel vuxna befolkningens deltagande i utbildning efter utbildningsnivå, hushållstyp, födelseland, kön och ålder. Statistiken visar hur den vuxna befolkningens deltagande i utbildning under de senaste 12 månaderna varit samt vilka grupper som deltar mer i utbildning och vilka som deltar mindre. Det handlar också om skälen eller hinder till deltagande i utbildning samt nyttan med utbildning och information om utbildningsmöjligheter. Statistikens kvalitet beskrivs i Kvalitetsdeklarationen som finns på www.scb.se/uf0538.

2 Undersökningsdesign

2.1 Målstorheter

De viktigaste målstorheterna är antal och andel vuxna personer i åldern 18–69 år som deltagit i formell eller icke-formell utbildning under de senaste 12 månaderna från mättillfället. Den enskilt viktigaste målstorheten är andelen deltagande personer i icke-formell utbildning.

2.2 Ramförfarande

Ramen som används för att dra urvalet från är Registret över totalbefolkningen (RTB) där alla folkbokförda personer finns med. Rampopulationen är avgränsad med åldersgränser så att ingen person i rampopulationen var under 18 år eller över 69 år under mätperioden september 2022 – januari 2023. Ramen framställdes från RTB med aktualitet 2022-01-31. Aktualiteten för ramen är viktigt för att minska täckningsfelen.

Observationsobjekten är personerna i ovan nämnda population. Observationsobjekten stämmer överens med uppgiftskällorna. Inför datainsamlingen har uppgifter om adress hämtats från RTB och telefonnummer har hämtats från ITESCO Eniro och Försäkringskassan.

2.3 Förfaranden för urval och uteslutning

2.3.1 Urvalsförfarande

Urvalsdesignen är sannolikhetsurval av typen stratifierat obundet slumpmässigt urval utan återläggning. Urvalet av observationsobjekt dras direkt från urvalsramen som omfattar rampopulationen. Ramen avseende folkbokförda personer i Sverige i åldern 18–69 år delades på två grupper, en för de mellan 25–69 år och en för de mellan 18–24 år,

innan urvalet drogs. Gruppen 25-69 år delades upp i 18 urvalsstrata medan 18-24 år delades upp på fyra urvalsstrata. Stratifieringen gjordes i syfte att uppnå Eurostats precisionskrav i skattningen för den viktigaste målstorheten, andelen deltagande personer i icke-formell utbildning för gruppen 25-64 år och andelen deltagare i formell utbildning för gruppen 18-24 år.

Urvalsstorleken för respektive stratum har bestämts på ett sådant sätt att precisionen förväntas bli densamma för alla stratum.

Planeringsvärden för svarsandelar i stratum hämtade för den senaste undersökningsomgången som genomfördes 2016 och har använts för att optimera allokeringen.

$$m_h = \frac{\tilde{p}_h(1 - \tilde{p}_h)}{V(\hat{p}_h) + \tilde{p}_h(1 - \tilde{p}_h) / N_h}$$

där

m_h = antal svar som krävs för att uppnå precisionskrav, stratum h

$\tilde{p}_h$ = planeringsvärde för andel personer som deltar i icke-formell utbildning, stratum h

$V(\hat{p}_h)$ = önskad precision, stratum h

N_h = antal personer i urvalsramen, stratum h

Därefter beräknas planerad stickprovsstorlek per stratum

$$n_h = m_h / \tilde{a}_h$$

där

$\tilde{a}_h$ = planeringsvärde för svarsandel, stratum h

2.3.2 Uteslutning från insamling (cut-off)

Ingen del av målpopulationen utesluts från direktinsamling

2.4 Insamlingsförfarande

2.4.1 Datainsamling

Statistiken baseras på direktinsamlade uppgifter från personer som blivit intervjuade via telefon eller som själva anger sina svar i ett webbformulär. För att minska uppgiftslämnarbördan har alla variabler som går att samla in från administrativa register hämtats på detta vis. De variabler som inte finns i register har inhämtats via direktinsamling från observationsobjekten.

De administrativa register som använts för datainsamling är

- Registret över totalbefolkningen (RTB)
- Inkomst- och taxeringsregistret (IoT)
- Utbildningsregistret (UREG)
- Registerbaserad arbetsmarknadsstatistik (RAMS)

För direktinsamling startade fältarbetet med att ett informationsbrev skickades ut med information om undersökningen, varför det är viktigt att delta men att det är frivilligt och hur de insamlade uppgifterna används. I brevet fanns inloggningsuppgifter till webbenkät. Om inget webbsvar kommit in inom två veckor började intervjuare att ringa urvalspersonerna. Vid telefonintervju läses frågorna upp av intervjuaren som registrerar svaren och vid insamling via webbenkät är det urvalspersonen själv som registrerar sina svar. Urvalet var uppdelat på tre grupper och direktinsamlingen pågick i sex veckor till vardera gruppen. Under den tiden skickades två påminnelser ut. I slutet av insamlingsperioden skickades vädjanbrev ut till de som inte redan deltagit. De personer som deltog fick ett presentkort utskickat som tack för sitt deltagande.

2.4.2 Mätning

Undersökningen avser primärt icke-formell utbildning och det ställs därför fördjupande frågor om icke-formella utbildningar (dock högst två) per undersökningssperson. För formell utbildning ställs endast fördjupande frågor om den senaste utbildningen och för informella utbildningsaktiviteter ställs inga fördjupande frågor. De fördjupande frågorna handlar om utbildningsaktiviteternas egenskaper.

För registeruppgifter har registerversioner använts med aktualitet så nära referenstiden hösten 2022 som möjligt.

Vid datainsamlingen används SCB:s statistikinsamlingsverktyg (SIV) för att registrera svaren från dem som deltar i undersökningen. I såväl telefonintervjuformuläret som webbenkätformuläret finns inbyggda logiska kontroller för att förebygga mätfel. Ett exempel på kontroll är timmar och kronor. Om uppgiften överstiger eller understiger ett visst värde kommer en uppmaning om att kontrollera rimligheten i svaret. Uppgiftslämnare som inte kan genomföra intervju på svenska erbjuds tolkintervju på ett språk de önskar. Webbenkäten finns inte på annat språk än svenska.

Frågeformulären för webb och telefon är uppbyggda att fånga deltagande i olika utbildningsformer som sedan följs av fördjupningsfrågor om de utbildningsformer urvalspersonen deltagit i. Om en urvalsperson inte deltagit i en viss utbildningsform sker ett hopp förbi följdfrågorna. Det innebär att en urvalsperson som deltagit i mycket utbildning får svara på många frågor medan en

urvalsperson som inte deltagit i någon form av utbildning får svara på ett mycket begränsat antal frågor.

Information om de slutliga observationsvariablerna och statistikens detaljerade innehåll finns på SCB:s webbplats. Där beskrivs alla observationsvariabler och värdemängder m.m. Dokumentationen finns på www.metadata.scb.se under *Vuxnas deltagande i utbildning (AES)*.

2.4.3 Bortfallsuppföljning

SCB arbetar aktivt med att nå alla personer i urvalet. Flera kontaktförsök genomförs vid olika tidpunkter i syfte att få in så stor andel svar som möjligt.

2.5 Bearbetningar

Kodning har skett för variablerna utbildningsaktiviteters inriktning, yrke och näringsgren. Om inte kodning sker vid mättillfället har det kodats i efterhand.

För kodning av yrke har en förprogrammerad lista med 7 000 yrken lagts in i frågeformuläret med sökfunktion som genererar förslag. I de fall yrket inte finns i listan blir det fråga om kodning av det öppna svaret i efterhand.

När det gäller näringsgren har Arbetsgivardeklarationer på individnivå använts för att hämta uppgift om näringsgren. Uppgiften avser juli 2022. Om personen svarat att hen är anställd av bemanningsföretag eller arbetar som konsult har frågor om arbetsstället samlats in för att kunna kodas i efterhand.

För kodning av utbildningsinriktning har kodning skett vid mättillfället för formell utbildning. Det handlade om att klassificera inriktningen till en av 20 huvudgrupper. Vid kodning av inriktning för icke-formell utbildning har hjälpinformation om utbildningens namn och huvudsakliga innehåll använts vid kodning i efterhand. En lista med 13 000 kursnamn som erhållits vid tidigare undersökningsomgångar har använts vid kodning av det öppna svaret i efterhand.

All kodning som skett i efterhand har kodats av kodningsexperter på SCB i ett kodningsverktyg med relevant hjälpinformation. Detta kodningsarbete har inte kvalitetskontrollerats via s.k. kontrollkodning.

En stor del av variablerna i det slutliga observationsregistret är härledda utifrån en eller oftast flera frågor i frågeformuläret. Inga imputeringar av partiellt bortfall eller objektsbortfall har gjorts.

2.6 Granskning

2.6.1 Granskning under direktinsamlingen

En stor mängd inbyggda kontroller fanns i frågeformulären för att undvika ologiska svar eller dubbelmarkeringar där så inte var tillåtet. Om intervjupersonen uppger ett svar som inte verkar rimligt uppmanas intervjuaren att stämma av svaret med intervjupersonen. I webbenkäten uppmanas uppgiftslämnare kontrollera sitt svar. Det var dock möjligt att ignorera sådana felmeddelanden. Därför har samma kontroller gjorts i efterhand.

2.6.2 Granskning av mikrodata och insamlade statistikvärden

Efter en vecka av datainsamlingsperioden levererades den första testfilen. Där kunde materialet granskas så att frågeformuläret fungerade som förväntat och svar fanns där det skulle finnas svar. Insamlade data granskades kontinuerligt under datainsamlingsperioden då leveranser av mikrodata skedde i flera omgångar. Poster för kodning kunde också laddas så att kodningsmaterialet kunde granskas. I veckovisa informationsbrev till intervjuarna kunde information ges för att förbättra insamlingen och minska mätfel.

2.6.3 Granskning av makrodata

Löpande under datainsamlingen granskades centrala skattningar för att bedöma rimligheten i dem. Eftersom antalet variabler är stort var det främst selektiv makrogranskning som genomfördes av de viktigaste variablerna, andel deltagare i formell och icke-formell utbildning. Innan leverans till Eurostat togs skattningar fram och rimlighetsgranskades för en stor mängd indikatorer med olika skärningar på kön, ålder och utbildningsnivå.

2.6.4 Granskning av redovisning

Inför publicering laddades samtliga delar av materialet i SCB:s webbpubliceringsverktyg. Det innebar att det slutgiltiga materialet som ska publiceras kunde granskas precis i den form det publiceras.

2.7 Skattningsförfarande

Undersökningen är en urvalsundersökning med bortfall. Kalibrering används för att kompensera för bortfallet.

2.7.1 Principer och antaganden

Vikter har beräknats så att resultat kan redovisas för hela populationen och inte bara för de svarande. Vikterna kompenserar för objektsbortfallet men inte för det partiella bortfallet. Om vikterna inte används så kan resultaten bli helt missvisande, speciellt som olika urvalspersoner haft olika sannolikhet att komma med i urvalet.

De kalibrerade vikterna har den 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 registrertotal som används för att fastställa de kalibrerade vikterna. Detta innebär bland annat att alla antalsskattningar för hjälpvariablerna som baseras på de kalibrerade vikterna överensstämmer med de storheter som finns i undersökningens rampopulation.

Den estimator som används i undersökningen är en kalibreringsestimator. För en allmän beskrivning av kalibreringsestimatorn, se "Estimation in the presence of Nonresponse and Frame Imperfections" (2001).

Syftet med kalibreringsestimatorn är att reducera urvalsfel och bortfallskevheter med hjälp av extern information, så kallade hjälpvariabler. Kalibreringen (punktskattningar och konfidensintervall) och de tabeller som presenteras i temarapporten har tagits fram med hjälp av SAS och variansprogrammet ETOS.

2.7.2 Skattningsförfarande för målstorheter

I en undersökning uppkommer fel om vi inte lyckas få svar från alla personer och om de avviker från de svarande med avseende på undersökningsvariablerna. Detta fel kallas för bortfallsfel. Kunskap om bortfallsfelet kan i regel bara fås på ett indirekt och approximativt sätt genom att utnyttja registervariabler.

Bortfallsfel kan reduceras genom att använda ett effektivt uppräkningsförfarande (vägning som kompenserar för bortfallet). I följande avsnitt redovisas hur det görs i denna undersökning.

Vägning görs genom att beräkna s.k. kalibrerade uppräkningsvikter. Detta förfarande utnyttjar hjälpinformation i form av registervariabler. I kalibreringen används följande registervariabler (hjälpvariabler):

Gruppen 18–24 år: Kön, åldersgrupp, utbildningsnivå, kommungrupp, födelseland, föräldrarnas utbildningsnivå, sysselsättningsstatus och deltagande i formell utbildning HT 2021

Gruppen 25–69 år: Kön, åldersgrupp, utbildningsnivå, födelseland, sysselsättningsstatus och deltagande i formell utbildning HT 2021, , civilstånd och hushållsinkomst

Var och en av de hjälpvariabler som används i kalibreringen har valts utifrån minst ett av följande tre kriterier:

- (i) de samvarierar med svarsbenägenheten

(ii) samvarierar med (viktiga) undersökningsvariabler

(iii) avgränsar (viktiga) redovisningsgrupper

Vikterna i denna undersökning kan beskrivas med formeln

$$w_k = 1 + (\sum_s \mathbf{x}_k - \sum_r \mathbf{x}_k) (\sum_r \mathbf{x}_k \mathbf{x}_k')^{-1} \mathbf{x}_k$$

där w_k = vikt/uppräkningsstal för objekt k

$$\sum_s \mathbf{x}_k = \text{hjälpotaler för hela urvalet}$$

$$\sum_r \mathbf{x}_k = \text{hjälpotaler för de svarande}$$

Vikterna/uppräkningsstalen multipliceras med objektens variabelvärden för att skapa statistikvärden gällande för populationen. För beräkning av skattningen av totaler används följande formel:

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

där w_k = vikt/uppräkningsstal för objekt k

y_k = variabelvärde för objekt k

Och för beräkning av skattningen av medelvärden används följande formel:

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

summering sker av de svarande (r).

Svarsandelar för olika grupper har beräknats enligt nedanstående formel

$$SA = \frac{T_s}{T_s + (1-a)T_b}$$

där T_s är antal svarande personer, T_b utgörs av antal personer bortfallet och a är en skattning av hur stor andel av bortfallet som inte tillhör målpopulationen. Skattning av a görs enligt

$$a = \frac{T_o}{T_o + T_s}$$

2.7.3 Skattningsförfarande för tillförlitlighet

Konfidensintervall är beräknade enligt en metod som i 95 procent täcker in det sanna värdet och beräknas enligt formeln:
 $\text{punktskattning} + 1,96 * \text{skattning av estimatorns medelfel}$.

Konfidensintervallet är korrekt under förutsättning att bortfallsfel inte finns. I undersökningar med bortfall finns i princip alltid bortfallsfel. Kalibrerade vikter reducerar bortfallsfel om hjälpinformationen är stark, men en del av bortfallsfelet finns kvar.

2.7.4 Röjandekontroll

I publicerade tabeller och diagram redovisas inga resultat för variabler med färre än 30 svar samt för svarsalternativ med färre än 3 svar. I det första fallet sekretessprickas svar för samtliga svarsalternativ. I det andra fallet sekretessprickas enbart svarsalternativet med färre än 3 svar.

3 Genomförande

3.1 Kvantitativ information

Urvalet i AES 2022 bestod av 7 400 individer i åldern 25–69 år och 3 300 individer i åldern 18–24 år. Antalet svarande uppgick till ungefär 4 600, vilket ger en ovägd svarsandel på 43 procent.

3.2 Avvikelser från undersökningsdesignen

Inga avvikelser har gjorts.