

Teknisk rapport

En beskrivning av genomförande och metoder

2021-10-26

Beteckning: Uppföljning av digitaliseringsstrategin - Förskolepersonal



Innehåll

1	Sammanfattning.....	3
2	Genomförande och metoder.....	4
2.1	Population och urval.....	4
2.2	Blankettfrågor och registeruppgifter	4
2.3	Datainsamling.....	5
2.3.1	Kontroller.....	6
2.4	Bortfall	7
2.5	Viktberäkning och estimation.....	8
2.5.1	Viktberäkning.....	8
2.5.2	Estimation.....	8
2.6	Statistikens tillförlitlighet	9
2.6.1	Tillförlitlighet totalt.....	9
2.6.2	Osäkerhetskällor	9
2.6.3	Urval.....	10
2.6.4	Ramtäckning	10
2.6.5	Mätning.....	10
2.6.6	Bortfall	10
2.6.7	Bearbetning.....	11
2.6.8	Modellantagande	11
2.7	Beskrivning av tabeller och datafil.....	11
2.7.1	Tabeller	11
2.7.2	Jämförbarhet och användbarhet	12
2.7.3	Sekretess och utlämnande	13
2.7.4	Övrigt	13

3	Bilagor	14
----------	----------------------	-----------

1 Sammanfattning

Den här rapporten beskriver hur SCB har genomfört uppdraget Uppföljning av digitaliseringsstrategin inom skolväsendet och vilka metoder som har använts.

Om ni har frågor är ni välkommen att höra av er till de ansvariga på SCB: Tom Berglind, telefon 010-479 66 09 eller e-post tom.berglind@scb.se alt. Jonathan Baquedano, telefon 010 - 479 69 48 eller e-post jonathan.baquedano@scb.se.

Uppdragets syfte: Skolverket ska på uppdrag av regeringen följa upp målen i digitaliseringsstrategin och vilket genomslag denna strategi får i verksamheterna. Den första uppföljningen av digitaliseringsstrategin genomfördes 2018. I enlighet med ett tidigare uppdrag har Skolverket vart tredje år från 2008 och fram till 2015 följt upp kompetens, tillgång och användning avseende digital teknik och digitala verktyg i förskola, skola och kommunal vuxenutbildning.

Uppdragets parter: Enheten Utbildning och arbete vid SCB genomförde under perioden februari 2021 – maj 2021 en enkätundersökning på uppdrag av Skolverket.

Kontaktpersoner: Undersökningsledare på SCB var Tom Berglind och Sara Westling var metodstatistiker. Skolverkets kontaktperson gentemot SCB var Magnus Eriksson.

Leverans: Preliminära resultat i form av tabeller levererades vecka 26 enligt ny överenskommelse med Skolverket. Mikrodata och uppdaterade tabeller levereras vecka 39.

Urval: Urvalet bestod av 1 000 personer varav 40 personer utgjorde övertäckning. Det innebär att urvalet, där övertäckningen är borträknad, var 960 personer. Det var totalt 430 personer som besvarade frågeblanketten, vilket var 44,8 procent av urvalet (övertäckning borträknad). Den designvägda svarsandelen, vilken tar hänsyn till om olika personer haft olika stor sannolikhet att komma med i urvalet, var 42,8 procent. Denna svarsandel kan ses som en skattning av den svarsandel som en totalundersökning skulle ha resulterat i.

2 Genomförande och metoder

2.1 Population och urval

Populationen, det vill säga de objekt som man vill kunna dra slutsatser om, utgjordes av personal inom förskola. En ram bestående av 113 302 personer hämtades från registret över personal i förskola, med uppgifter avseende 2020-10-15.

För att kunna dra ett urval från populationen skapades en urvalsram som avgränsade och identifierade objekten i populationen. Urvalsramen i undersökningen skapades med hjälp av data från registret över personal i förskola, uppgifter avseende 2020-10-15. Urvalet samordnades med undersökningen Utvärdering av läroplan för att minimera uppgiftslämnarbördan.

Urvalsramen stratifierades efter huvudman, föräldrarnas utbildningsnivå och personalens utbildningsnivå. Där personalens utbildning delades upp i examen eller ej examen (med examen räknas lärar- eller förskollärarexamen). Totalt bildades 8 strata. Från urvalsramen drogs ett stratifierat obundet slumpmässigt urval om 1 000 personer (se Tabell 1 nedan).

Ett stratifierat obundet slumpmässigt urval innebär att alla objekt inom ett stratum har samma sannolikhet att komma med i urvalet.

Tabell 1 Urvalsram och urval fördelat på stratum

Stratum-nummer	Huvudman	Utbildning	Personalutbildning	Population	Urval
111	Kommunal	Hög	Examen	17 585	125
101	Kommunal	Låg	Examen	20 950	125
110	Kommunal	Hög	Ej examen	20 981	125
100	Kommunal	Låg	Ej examen	27 452	125
511	Enskild	Hög	Examen	5 100	120
501	Enskild	Låg	Examen	1 692	180
510	Enskild	Hög	Ej examen	11 957	100
500	Enskild	Låg	Ej examen	5 585	100
Summa				113 302	1 000

Urvalsstorleken bestämdes i samråd mellan SCB och Skolverket

2.2 Blankettfrågor och registeruppgifter

Skolverket utformade frågorna i frågeblanketten i samarbete med SCB.

Grundläggande för bra kvalitet i en undersökning är kvaliteten på de data som samlas in. För att minska risken för mätfel har blanketten därför genomgått ett mättekniskt test i form av en mätteknisk standardgranskning. Det innebär att mättekniker på SCB gjort en djupgående granskning av frågeblankettens disposition, frågor och svarsalternativ, för att identifiera problem, analysera konsekvenser och ta fram utvecklade förslag på lösningar. Granskningen finns dokumenterad i en skriftlig rapport.

Blanketten bestod av 22 numrerade frågor, flera av dem hade delfrågor vilket genererade totalt 44 frågor.

Referensperiod för frågorna avser mättillfället, det vill säga tidpunkten då respondenten svarar på frågorna.

Förutom de variabler som samlades in via frågeblanketten hämtades följande variabler:

Enhetens storlek:

- Antal barn på förskoleenheten
- Andelen barn med föräldrar med eftergymnasial utbildning på förskoleenheten

Uppgifter om huvudmannen:

- Huvudmannatyp
- Antal barn per huvudman
- Antal förskoleenheter per huvudman
- SKR:s kommungruppsindelning

Individuppgifter:

- Personalens utbildning
- Utbildningskategori

2.3 Datainsamling

Innan utskick gjordes en kontroll av personerna i urvalet mot de senaste folkbokföringsuppgifterna, för att ta fram aktuella adressuppgifter. Vid kontrollen och under insamlingsfasen framkom det att 40 personer inte längre tillhörde populationen utan utgjorde känd övertäckning, det vill säga den övertäckning som identifierats.

Tolv personer svarade att de har en pågående anställning i annan förskola än det som stod i brevet. Av dessa tolv personer kunde åtta personer knytas till den förskola de angivit, övriga fyra personer behåller den ursprungliga förskolan som angetts i brevet.

Totalt tre utskick gjordes. Samtliga utskick gjordes med post. I ett informationsbrev ombads urvalspersonerna att besvara frågorna och skicka svaret till SCB. I informationsbrevet kunde urvalspersonerna läsa om undersökningens bakgrund, syfte och att undersökningen genomfördes i samarbete mellan Skolverket och SCB. Informationsbrevet informerade också om att uppgifter hämtades från register och att en avidentifierad datafil levereras till Skolverket. Brevet informerade även om dataskyddslagstiftningen samt offentlighets- och sekretesslagen och att det var frivilligt att medverka i undersökningen.

Det första utskicket i form av ett informationsbrev med inloggningsuppgifter och pappersblankett genomfördes den 24 februari 2021. Ett påminnelsebrev skickades den 10 mars och en påminnelse med frågeblankett skickades den 24 mars.

Insamlingen avslutades 3 maj.

Vid samtliga utskick fick uppgiftslämnaren information om att frågorna även kunde besvaras på webben samt inloggningsuppgifter.

Totalt svarade 430 personer på frågeblanketten, vilket var 44,8 procent av urvalet efter att den kända övertäckningen, det vill säga den övertäckning som identifierats, tagits bort. Den designvägda svarsandelen, vilken tar hänsyn till att olika personer haft olika stor sannolikhet att komma med i urvalet var 42,8 procent.

Tabell 2 Beskrivning av inflödet. Antal och andel svar

	Antal	Andel
Efter första utskicket	221	23,0
Efter andra utskicket	119	12,4
Efter tredje utskicket	90	9,4
Totalt antal svar	430	44,8
Bortfall	530	55,2
Urval exkl. övertäckning	960	100,0

Det var 191 personer som valde att besvara webblanketten, vilket var 44,4 procent av samtliga svarande.

De frågeblanketter som besvarats på papper registrerades med hjälp av skanning.

2.3.1 Kontroller

Kontroller har genomförts bland annat för att säkerställa att endast valida värden förekommer i materialet.

De inkomna pappersblanketterna granskades i flera olika skeden av bearbetningen.

Handskrivna tecken kontrollerades efter skanningen och tecken som tolkats felaktigt korrigerades.

Vid frågor där svaret skulle anges i form av skrivna siffror, gjordes en rimlighetsbedömning. Orimliga värden togs ut för kontroll i syfte att säkerställa att rätt värde registrerats.

Under datainsamlingen granskades otillåtna värden löpande, till exempel om det fanns frågor där uppgiftslämnare markerat flera svar trots att frågan skulle besvaras med ett alternativ. Hur otillåtna värden hanterats beskrivs i avsnittet 3.7.1.1 Datafil.

Datamaterialet granskades före leverans med hjälp av frekvenstabeller för att upptäcka och korrigera eventuella kvarstående fel.

SCB kan inte garantera att den utvalda personen själv besvarat frågeblanketten. Någon kontroll av att rätt person har besvarat frågeblanketten har inte gjorts.

2.4 Bortfall

Bortfallet består dels av objektsbortfall, som innebär att frågeblanketten inte är besvarad alls, dels av partiellt bortfall som innebär att vissa frågor i blanketten inte är besvarade. Om de urvalspersoner som utgör bortfall skiljer sig från de svarande med avseende på undersökningsvariablerna, kan det systematiska felet (bias) i form av bortfallsskevheter öka för skattningarna som grundar sig på enbart de svarande. För att reducera bortfallsskevheter har vikter beräknats med hjälp av kalibreringsestimation (se avsnitt 2.5 och vidare information i bilaga 3).

Objektsbortfall kan bland annat bero på att uppgiftslämnaren inte är villig att delta i undersökningen, att uppgiftslämnaren inte går att nå eller att uppgiftslämnaren är förhindrad att medverka. Objektsbortfallet i denna undersökning redovisas i Tabell 3 nedan.

Tabell 3 Beskrivning av objektsbortfall

	Antal
Ej avhörda	517
Avböjd medverkan	5
Ej anträffad	8
Totalt	530

Med "Ej avhörda" menas att ingen uppgift om varför frågeblanketten inte är besvarad har lämnats. Med "Avböjd medverkan" menas att SCB meddelats att uppgiftslämnaren inte vill medverka i undersökningen. "Ej anträffad" innehåller personer som saknar adress i Registret över totalbefolkningen (RTB), har hemlig adress, är tillfälligt bortresta samt postreturer.

Partiellt bortfall kan bero på att en fråga är svår att förstå, är känslig, att uppgiftslämnaren missar att besvara frågan eller att instruktionerna vid hoppfrågor misstolkas.

Det partiella bortfallet varierar mellan 0,1 och 1,1 procent för de flesta av frågorna. Det högsta partiella bortfallet, 23,8 procent, finns i fråga 19f.

2.5 Viktberäkning och estimation

För varje svarande person (kallas objekt i fortsättningen) har en vikt beräknats. Syftet med detta är att kunna redovisa resultat för hela populationen och inte bara för de svarande. Vikten kallas därför även för uppräkningsstal.

Vikterna/uppräkningsstalen multipliceras med objektens variabelvärden för att skapa statistikvärden gällande för populationen. Om vikterna inte används vid beräkning av skattningar så kan resultaten bli missvisande. Vikterna kompenserar för objektbortfallet, men inte för det partiella bortfallet.

Vikterna har beräknats utifrån urvalsdesignen samt antaganden om objektbortfall och ramtäckning. Beräkningen gjordes i SAS med hjälp av ett av SCB utvecklat makro (ETOS).

2.5.1 Viktberäkning

I en urvalsundersökning är skattningarna behäftade med *urvalsosäkerhet* beroende på att endast en delmängd av populationen studeras. Bortfall gör att antalet svar minskar ytterligare vilket ökar osäkerheten. Om de objekt som utgör bortfall dessutom avviker från de svarande med avseende på undersökningsvariablerna kan även så kallad *bortfallskevheter* uppstå.

Både urvalsosäkerhet och bortfallsskevheter kan reduceras genom att använda ett effektivt uppräkningsförfarande, så kallad *kalibrering*. För detaljer se bilaga 3.

2.5.2 Estimation

Utifrån undersökningens design fås en designvikt. Den kan användas för att ta fram skattningar när bortfall inte förekommer. De vikter som tas fram med kalibreringsestimation justerar designvikterna utifrån registervariablerna och ges av

$$w_k = d_k \cdot v_k$$

där w_k är vikt/uppräkningsstal för objekt k , d_k är designvikt för objekt k och v_k ger justeringen för objekt k utifrån registervariabler.

Vikterna bygger på antagandet att ramen återspeglar populationen väl och därmed att över- och undertäckningen är försumbar.

För en teknisk beskrivning av kalibreringen se bilaga 3. För en mer utförlig beskrivning av kalibreringsestimatorens se Lundström och Särndal (2001): *Estimation in the Presence of Nonresponse and Frame Imperfections*, Statistics Sweden.

Skattningar av totaler ges av

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

där w_k är vikt/uppräkningsstal för objekt k och y_k är variabelvärde för objekt k . Notera att summering sker över de svarande r .

Skattningar av medelvärden ges av

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

där w_k är vikt/uppräkningsstal för objekt k och y_k är variabelvärde för objekt k . Summering sker över de svarande r .

2.6 Statistisk tillförlitlighet

2.6.1 Tillförlitlighet totalt

Statistiken är behäftad med osäkerhet. Vid bedömning av hur olika osäkerhetskällor påverkar statistiken från en undersökning skiljer man på slumpmässiga och systematiska avvikelser. Slumpmässiga avvikelser orsakar enligt statistisk teori beräkningsbar osäkerhet hos de skattade resultaten. Systematiska avvikelser påverkar resultaten i en viss riktning. Med total osäkerhet avses den sammanlagda osäkerheten = slumpmässiga + systematiska avvikelser. Det är svårt att ange hur tillförlitligheten påverkas av eventuella systematiska avvikelser (skevhet) då det ofta kräver resurskrävande utvärderingsinsatser. En sådan utvärdering har inte gjorts för detta uppdrag.

2.6.2 Osäkerhetskällor

Bortfall bedöms vara den största källan till osäkerhet när statistik ska tas fram på aggregerad nivå. Om statistik ska tas fram på mer detaljerad nivå, exempelvis på stratumnivå, så är även urval en betydande källa till osäkerhet. Täckningsfel och bearbetningsfel bedöms vara litet till försumbart i sammanhanget.

Vid framställning av statistik bör bortfallets fördelning beaktas. Det var totalt 430 personer som besvarade frågeblanketten, vilket var 44,8 procent av urvalet (övertäckning borträknad). Den designvägda svarsandelen, vilken tar hänsyn till om olika personer haft olika stor sannolikhet att komma med i urvalet, var 42,8 procent. Denna

svarsandel kan ses som en skattning av den svarsandel som en totalundersökning skulle ha resulterat i.

2.6.3 Urval

Denna kvalitetskomponent avser osäkerhet som uppkommer på grund av att endast ett urval av populationen undersöks. Urvalsosäkerheten är således den avvikelse mellan ett skattat värde och det faktiska värdet som beror på att man inte undersöker alla objekt i populationen. Urvalsosäkerheten minskar med en ökad urvalsstorlek. Urvalsosäkerheten bör beaktas när man drar slutsatser från undersökningen.

2.6.4 Ramtäckning

Under- och övertäckning innebär att ram- och målpopulation inte helt stämmer överens. Undertäckning innebär att vissa objekt som ingår i målpopulationen saknas i ramen. Övertäckning innebär att objekt som inte ingår i målpopulationen ändå finns i ramen. Ett sätt att minska täckningsbrister är att använda bra och uppdaterade register för att ta fram ramar.

Registret över personal i förskola samlas in en gång per år, den 15 oktober. Denna undersökning genomfördes under perioden februari – maj 2021, vilket medför att vissa förändringar kan ha skett sedan oktober 2020. 40 objekt som hade valts ut visade sig inte tillhöra populationen och utgör därmed övertäckning. Sammanfattningsvis är SCB:s bedömning att bristerna i ramen, både med avseende på under- och övertäckning, inte snedvrider statistiken i någon nämnvärd omfattning.

2.6.5 Mätning

Ett fel som kan uppstå vid mätning är att lämnade uppgifter skiljer sig från faktiska uppgifter. Felet kallas mätfel och kan uppkomma då uppgiftslämnaren inte minns de faktiska uppgifterna, missförstår frågan eller medvetet svarar felaktigt.

2.6.6 Bortfall

Skevheter i skattningar på grund av bortfall inträffar om objekten i bortfallet och de svarande skiljer sig åt avseende frågorna/variablerna i undersökningen.

Eventuella bortfallsfel kan reduceras genom att på ett effektivt sätt justera undersökningens vikter. En beskrivning av förfarandet finns i den bortfallsanalys som gjorts i samband med valet av hjälpinformation till kalibreringen (se bilaga 3). I bortfallsanalysen redovisas även svarsandelar för viktiga delgrupper.

Någon utvärdering av kvarstående skevheter efter kalibrering har inte gjorts, men med tanke på att bortfallet är stort i vissa grupper finns det risk för att skevheter kvarstår även efter kalibrering.

Förutom eventuell skevhet innebär bortfallet att den slumpmässiga osäkerheten i statistiken ökar.

2.6.7 Bearbetning

Vid den manuella och maskinella bearbetningen av data kan osäkerhet uppstå. Exempel på bearbetningsbrister är registreringsfel och kodningsfel. Dessa kan förhindras och upptäckas i de kontroller som genomförs vid dataregistreringen. I den här undersökningen bedöms osäkerheten på grund av registrering vara litet eftersom frågeblanketten endast hade fasta svarsalternativ.

2.6.8 Modellantagande

Inga modellantaganden har gjorts mer än de som redovisats ovan.

2.7 Beskrivning av tabeller och datafil

2.7.1 Tabeller

Skattningarna i tabellerna har räknats upp till populationsnivå, vilket innebär att resultatet avser hela populationen och inte endast de svarande. Antalsuppgifterna är således skattningar av antal personer i populationen med den aktuella egenskapen. Att antalet personer är olika för olika frågor beror på hoppinstruktioner då olika delpopulationer avses.

För de följdfrågor som föregås av hoppinstruktioner har selekteringar genomförts vid tabellframställandet. Det betyder att till exempel för fråga 4 har de som svarat "Ja" i fråga 3 inte ingått i tabellframställningen.

Om en följdfråga är besvarad men dess styrfråga inte är besvarad och svaret på följdfrågan gör att ett visst svarsalternativ på styrfrågan är självklart så har hänsyn till detta tagits vid tabellframställningen.

I andelstabellerna är värdena avrundade till närmaste heltal och värden under 0,5 är därmed avrundade till noll.

Tabellerna har sekretessgranskats (prickats) så att inga uppgifter för enskilda objekt röjs. Mycket osäkra skattningar, där konfidensintervallet är större än andelsskattningen, har också prickats.

Skolverket har även beställt särskild redovisning och sammanslagningar av flertalet variabler vilka finns specificerade i tabellbeställningen, bilaga 4.

De tabeller som presenteras i denna rapport har tagits fram enligt de formler som presenterats i avsnitt "Viktberäkning och estimation" med hjälp av SPSS/SAS och variansprogrammet ETOS.

2.7.1.1 Datafil

Svarsdatafilen kompletterades med vikter. De bifogade vikterna är avsedda att användas vid skattning av målstorheter i populationen, exempelvis totaler och medelvärden.

Vid analyser av samband med hjälp av exempelvis regressionsanalys bör det uppmärksammas att det inte är givet hur vikterna ska hanteras. Urvalets design samt bortfall är komplikationer som kan störa analyser av samband.

Om det är partiellt bortfall (uppgift saknas) i en fråga har värdet 88 angetts. Vid hoppfrågor har värdet 0 angetts för dem som inte ska besvara följdfrågorna.

Vid dubbelmarkeringar i enkäterna har rättningar gjorts enligt överenskommelse mellan Skolverket och SCB. För information om hur dubbelmarkeringar hanterats, se bilaga 5.

Svar från öppna frågor (fråga 19 och 21) från webbenkäter finns i klartext i datafilen. När det gäller pappersenkät finns dessa öppna svar som sökbara bildfiler separat (TIFF-format) och är markerade med en "1" i datafilen.

Övriga överväganden som har gjorts i kodningsarbetet redovisas i bilaga 6.

2.7.1.2 Statistiska mått

Materialet lämpar sig att redovisas som totaler eller procentuella andelar för hela målpopulationen eller fördelat på olika redovisningsgrupper.

2.7.1.3 Redovisningsgrupper

Ofta redovisas statistik inte bara för hela populationen utan också för delgrupper (redovisningsgrupper). Redovisningsgrupper i den här undersökningen kan avgränsas med hjälp av registervariabler (exempelvis huvudmannatyp och utbildningsnivå).

2.7.2 Jämförbarhet och användbarhet

2.7.2.1 Jämförbarhet över tid

Detta är femte gången som SCB genomför undersökningen. Första gången var 2008 och därefter 2012, 2015 och 2018.

Undersökningarna 2015, 2018 och 2021 genomfördes på liknande sätt med avseende på urvalsmetod. Tidigare år 2008 och 2012 skapades urvalsramen med hjälp av data från företagsdatabasen (FDB).

En del frågor går att jämföra med tidigare års blankettfrågor. I årets undersökning finns det ett fåtal nya frågor i enkäten jämfört med tidigare år.

2.7.3 Sekretess och utlämnande

SCB har gjort en sekretessprövning för utlämnandet av data på mikronivå.

De register som framställts hos SCB med anledning av detta uppdrag avidentifieras 3 månader efter leverans räknat från uppdragets slutleverans som sker i samband med leverans av denna tekniska rapport den 26 oktober 2021. Därefter är det inte möjligt att göra några rättelser som kräver koppling till aktuella identiteter, om inte annat överenskommes.

2.7.4 Övrigt

Öppna svar

I informationsbrevet informerades urvalspersonerna att fritextsvar levereras avidentifierat till Skolverket. SCB har därför innan leverans tagit bort eller maskerat text för nedan löpnummer och fråga:

203845201 F21_txt

3 Bilagor

Bilaga 1. Informationsbrev

Bilaga 2. Frågeblankett

Bilaga 3. Kalibreringsrapport

Bilaga 4. Tabellbeställning

Bilaga 5. Dubbelmarkeringar

Bilaga 6. Kodning av datafil