

KVALITETSDEKLARATION

Gymnasieungdomars studieintresse

Ämnesområde

Utbildning och forskning

Statistikområde

Befolkningens utbildning

Produktkod

UF0513

Referenstid

Höstterminen 2021

Statistikens kvalitet	3
1 Relevans	3
1.1 Ändamål och informationsbehov	3
1.1.1 Statistikens ändamål	3
1.1.2 Statistikanvändares informationsbehov	3
1.2 Statistikens innehåll	3
1.2.1 Objekt och population	3
1.2.2 Variabler	4
1.2.3 Statistiska mått	5
1.2.4 Redovisningsgrupper	5
1.2.5 Referenstider	6
2 Tillförlitlighet	6
2.1 Tillförlitlighet totalt	6
2.2 Osäkerhetskällor	7
2.2.1 Urval	7
2.2.2 Ramtäckning	8
2.2.3 Mätning	9
2.2.4 Bortfall	10
2.2.5 Bearbetning	12
2.2.6 Modellantaganden	13
2.3 Preliminär statistik jämförd med slutlig	13
3 Aktualitet och punktlighet	13
3.1 Framställningstid	13
3.2 Frekvens	13
3.3 Punktlighet	13
4 Tillgänglighet och tydlighet	13
4.1 Tillgång till statistiken	13
4.2 Möjlighet till ytterligare statistik	13
4.3 Presentation	13
4.4 Dokumentation	13
5 Jämförbarhet och sam användbarhet	13
5.1 Jämförbarhet över tid	13
5.2 Jämförbarhet mellan grupper	16
5.3 Sam användbarhet i övrigt	16
5.4 Numerisk överensstämmelse	16
Allmänna uppgifter	16
A Klassificeringen Sveriges officiella statistik	16
B Sekretess och personuppgiftsbehandling	16
C Bevarande och gallring	17
D Uppgiftsskyldighet	17
E EU-reglering och internationell rapportering	17
F Historik	17
G Kontaktuppgifter	17

Statistikens kvalitet

1 Relevans

1.1 Ändamål och informationsbehov

1.1.1 Statistikens ändamål

Gymnasieungdomars studieintresse syftar till att belysa intresset för högskolestudier hos elever i gymnasieskolans årskurs 3 och orsakerna till att man planerar eller inte planerar att läsa vidare. För elever med studieplaner undersöks även vad de önskar studera, omfattningen och längd av studier samt om de har ett specifikt lärosäte i åtanke och i så fall varför. Under hösten 2021 undersöktes även hur covid-19 pandemin har påverkat elevernas studier och motivation.

Statistiken används huvudsakligen som underlag för information om samt planering av utbildningsprogram och kurser.

1.1.2 Statistikanvändares informationsbehov

Det finns ingen motsvarande statistik över gymnasieelevers attityder till fortsatta studier där det går att undersöka hur dessa förändras över tid. Resultatet presenteras uppdelat på flera olika redovisningsgrupper, vilket möjliggör jämförelser av attityder och planer mellan exempelvis kvinnor och män samt elever på olika gymnasieprogram.

Huvudanvändarna av statistiken är främst universitet och högskolor, Universitetskanslersämbetet och Utbildningsdepartementet. Innehållet i undersökningen har beslutats i samråd med "Användarrådet för utbildningsstatistik".

Eftersom statistiken publiceras innan eleverna faktiskt sökt till universitet/högskola kan resultatet användas som underlag för information om samt planering av utbildningsprogram och kurser.

Då uppgifterna från enkätundersökningarna sparas finns också möjligheten att följa upp elevernas planer i efterhand. Detta blir möjligt genom Universitetskanslersämbetets register över studerande på universitet/högskola. På så vis kan registeruppgifter om elevernas faktiska studier kopplas samman med deras attityder och planer i gymnasieskolan.

1.2 Statistikens innehåll

Målstorheterna avser attityder till högskolestudier hos elever i gymnasieskolans årskurs 3 i Sverige hösten 2021.

1.2.1 Objekt och population

Intressepopulationen består av elever i gymnasieskolans årskurs 3 i Sverige läsåret 2021/22. I målpopulationen ingår endast elever på nationella program. Elever på introduktionsprogram ingår således inte i målpopulationen.

Målobjekt utgörs av en elev på nationella program i gymnasieskolans årskurs 3 läsåret 2021/22. Observationsobjekt utgörs av en elev som enligt uppgift från registret över elever i gymnasieskolan gick i årskurs 2 på nationella program läsåret 2020/21. Målobjekt skiljer sig därmed från

observationsobjekt. Elever som exempelvis tagit studieuppehåll kan ha saknats i årskurs 2 men återkommit till årskurs 3. Elever som fanns med i årskurs 2 kan också ha hoppat av till årskurs 3. Hur täckningsproblematiken hanteras beskrivs i avsnitt 2.2.2.

1.2.2 Variabler

Statistikens viktigaste målvariabler är:

Planer på högskole- eller universitetsstudier inom tre år

Undersökningspersonens egen bedömning av om hon/han planerar att börja studera på universitet/högskola inom de närmaste tre åren.

Orsaker till avsaknad av planer på högskole- eller universitetsstudier inom tre år

Undersökningspersonens egen bedömning av varför hon/han inte planerar att börja läsa på universitet/högskola inom de närmaste tre åren.

Planerat studieupplägg

Undersökningspersonens egen bedömning av om hur hon/han önskar lägga upp studierna i form av fristående kurser eller program.

Planerad typ av program

Undersökningspersonens egen bedömning om vilken typ av program hon/han helst skulle vilja läsa.

Intresse för att läsa hela utbildningen utomlands

Undersökningspersonens egen bedömning av om hon/han är intresserad av att genomföra hela utbildningen utomlands.

Intresse för att läsa en del av utbildningen utomlands

Undersökningspersonens egen bedömning av om hon/han är intresserad av att genomföra en del av studierna utomlands.

Planerad utbildningsinriktning på högskole- eller universitetsstudier

Undersökningspersonens egen bedömning av vilket ämnesområde hon/han planerar att börja studera.

Intresse för att läsa på heltid eller deltid

Undersökningspersonens egen bedömning av om hon/han är intresserad av att genomföra studierna på heltid eller på deltid.

Intresse för att arbeta samtidigt som studierna

Undersökningspersonens egen bedömning av om hon/han är intresserad av att arbeta samtidigt som hon/han studerar.

Intresse för att söka studiemedel från CSN

Undersökningspersonens egen bedömning av om hon/han är intresserad av att söka studiemedel från CSN när hon/han studerar på universitet/högskola.

Orsaker till planer på högskole- eller universitetsstudier inom tre år

Undersökningspersonens egen bedömning av varför hon/han planerar att börja läsa på universitet/högskola inom de närmaste tre åren.

Planerad längd på studierna

Undersökningspersonens egen bedömning av hur länge hon/han planerar att läsa på universitet/högskola.

Information om högskolestudier

Undersökningspersonens egen bedömning av hur mycket information hon/han fått om universitets-/högskolestudier från olika informationskällor.

Distansundervisning i gymnasieskolan

Undersökningspersonens egen bedömning om hur distansundervisning har fungerat.

Distansundervisning ur olika aspekter under coronapandemin i gymnasieskolan

Undersökningspersonens egen bedömning om hur distansundervisning har fungerat utifrån olika aspekter.

Observationsvariablerna utgörs av svaren på frågorna i frågeformuläret och anses stämma väl överens med mål- och intressevariabler.

1.2.3 Statistiska mått

De statistiska mått som används är antal (personer i populationen eller inom en viss redovisningsgrupp) och andel (personer med en viss egenskap i procent av populationen/den aktuella redovisningsgruppen). Statistiken redovisas med 95-procentiga konfidensintervall.

1.2.4 Redovisningsgrupper

De statistiska målstorheterna skattas för olika redovisningsgrupper – delgrupper av populationen – med hjälp av indelningsvariabler. Dessa är:

Kön: Uppgiften om kön hämtas från *Registret över totalbefolkningen (RTB)*.

Född i Sverige/utrikes född: Grupperingen görs utifrån uppgift om födelseland, vilket hämtas från *Registret över totalbefolkningen (RTB)*. Utifrån den variabeln skapas grupperna födda i Sverige och utrikes födda.

Föräldrars utbildningsnivå. Uppgift om vårdnadshavares utbildningsnivå hämtas från *Registret över befolkningens utbildning (Utbildningsregistret)* med aktualitet 2020-12-31. Uppgifterna ligger sedan till grund för följande indelning:

- Båda föräldrar har eftergymnasial utbildning
- En förälder har eftergymnasial utbildning
- Ingen förälder har eftergymnasial utbildning
- Saknas i Utbildningsregistret/ Uppgift saknas

Region (riksområde)

Grupperingen görs utifrån uppgift om det län elevens skola ligger belägen i, vilket hämtas från registret över elever i gymnasieskolan. Grupperna är Stockholms län, Östra Mellansverige, Småland med öarna, Sydsverige, Västsverige, Norra Mellansverige, Mellersta Norrland samt Övre Norrland.

Gymnasieprogram

Grupperingen görs utifrån uppgift om gymnasieprogram från registret över elever i gymnasieskolan. Grupperna är Naturvetenskapsprogrammet, Samhällsvetenskapsprogrammet, Teknikprogrammet och Övriga program.

I övriga program ingår: Barn- och fritidsprogrammet, Bygg- och anläggningsprogrammet, Ekonomiprogrammet, El- och energiprogrammet, Estetiska programmet, Fordons- och transportprogrammet, Handels- och

administrationsprogrammet, Hantverksprogrammet, Hotell- och turismprogrammet, Humanistiska programmet, Industritekniska programmet, International Baccalaureate, Naturbruksprogrammet, Restaurang- och livsmedelsprogrammet, Riksrekryterande utbildningar med egna examensmål, VVS- och fastighetsprogrammet och Vård- och omsorgsprogrammet.

Typ av program

Uppgift om gymnasieprogram kommer från registret över elever i gymnasieskolan och grupperas här i högskoleförberedande program, yrkesprogram.

I högskoleförberedande program ingår: Estetiska programmet, Ekonomi-programmet, Humanistiska programmet, International Baccalaureate, Naturvetenskapsprogrammet, Samhällsvetenskapsprogrammet och Teknikprogrammet.

I yrkesprogram ingår: Barn- och fritidsprogrammet, Bygg- och anläggningsprogrammet, El- och energiprogrammet, Fordons- och transportprogrammet, Handels- och administrationsprogrammet, Hantverksprogrammet, Hotell- och turismprogrammet, Industritekniska programmet, Naturbruksprogrammet, Restaurang- och livsmedelsprogrammet, Riksrekryterande utbildningar med egna examensmål, VVS- och fastighetsprogrammet och Vård- och omsorgsprogrammet

1.2.5 Referenstider

Insamlingen skedde under perioden oktober 2021– januari 2022 och avser höstterminen 2021. I tabeller och diagram redovisas uppgifterna såsom avseende hela läsåret 2021/22.

2 Tillförlitlighet

2.1 Tillförlitlighet totalt

Den totala tillförlitligheten hos de redovisade statistikvärdena påverkas av ett flertal osäkerhetskällor. I avsnitt 2.2 Osäkerhetskällor görs en genomgång av dessa källor och deras konsekvenser.

De statistikvärden som tas fram utgår ifrån undersökningens design som är ett stratifierat obundet slumpmässigt urval. Förfarandet ger upphov till slumpmässig osäkerhet då ett specifikt urval av populationen undersöks. Bortfall av individer i urvalet kan ge upphov till både slumpmässiga och systematiska fel. Hjälpinformation används för att ta fram kalibrerade uppräkningsvikter som kan kompensera främst för systematiska fel som uppstår som ett resultat av bortfall.

Den osäkerhet som orsakas av att skattningarna baseras på ett urval uppskattas med hjälp av 95-procentiga konfidensintervall. Det har dock inte genomförts några metodstudier för att kunna redovisa mått över storleken på andra osäkerhetskällor än urvalsfelet. I avseende att reducera täckningsfel och bortfallsfel, har kalibrerade vikter beräknats.

2.2 Osäkerhetskällor

De osäkerhetskällor som har störst betydelse för Gymnasieungdomars studieintresse bedöms vara urval samt bortfall.

2.2.1 Urval

Rampopulationen stratifierades efter kön, utrikes född/född i Sverige, typ av gymnasieprogram och riksområde. Från rampopulationen drogs ett stratifierat urval om 5 836 individer. Vid bestämningen av urvalsstorleken togs hänsyn till det förväntade bortfallet.

I varje stratum drogs ett obundet slumpmässigt urval. Urvalet fördelades så att felmarginalen skulle bli lika stor i alla stratum, baserat på ett antagande om skattningarnas ungefärliga storlek.

För att skatta målstorheterna används ett estimationsförfarande 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 estimationsförfarandet innebär.

I kalibreringen används följande registervariabler (hjälpvariabler):

Kön, riksområde, gymnasieprogram, vårdnadshavarens utbildningsnivå, födelseland, slutbetyg i årskurs 9. Särskilt vårdnadshavarens utbildningsnivå och slutbetyg i årskurs 9 har visat sig vara starka variabler.

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

I synnerhet är det punkterna (i) och (iii) som varit vägledande. Om punkt (i) är uppfylld kan man förvänta sig att användandet av kalibrerade vikter leder till en reduktion av bortfallsskevhets. Är punkt (ii) uppfylld bör såväl variansen (samplingvariens + bortfallsvariens) som bortfallsskevhets reduceras för estimatorer baserade på de (viktiga) undersökningsvariablerna. Slutligen, om punkt (iii) är uppfylld bör variansen (samplingvariens + bortfallsvariens) minska för estimatorer av storheter definierade på redovisningsgruppsnivå.

För att reducera för den täckningsskevhet som uppstår till följd av den täckningsproblematik som finns, har de summor som används för att fastställa de kalibrerade vikterna hämtats från det uppdaterade registret för årskurs 3. Detta förfarande bygger på ett underförstått, icke kontrollerbart, modellantagande om hur undersökningsvariablerna är relaterade till kalibreringsvariablerna.

De kalibrerade vikterna bestäms så att när skattningar görs för hela populationen motsvarar skattningarna summorna hos respektive registervariabel.

Osäkerhetsintervall beräknas enligt formeln för 95-procentiga konfidensintervall:

$\text{punktestimat} \pm 1,96 * (\text{skattning av estimatorns medelfel})$

där såväl punkt- som medelfelsestimaten beräknas med hjälp av ett av SCB egenutvecklat estimationsprogram, CLAN97.

Under förutsättning att stickprovsstorleken är stor samt, att

- (i) den estimator som används är approximativt väntevärdesriktig, dvs. icke-skev
- (ii) estimatorns varians endast beror av samplingvariansen och bortfallsvariansen
- (iii) den hjälpinformation som används i kalibreringen är starkt korrelerad med de okända svarssannolikheterna

gäller att intervallet har en konfidensgrad på ca 95 procent, dvs. under förutsättningarna är sannolikheten ca 95 procent att man med det använda urvals- och estimationsförfarandet erhåller ett intervall som täcker in den sökta storheten.

Naturligtvis är inte samtliga ovanstående förutsättningar perfekt uppfyllda i praktiken. Även om hjälpvariablerna är klokt valda kommer estimatorn att föra med sig en viss skevhet, delvis till följd av kvarstående bortfalls- och täckningsfel men också till följd av andra fel av systematisk karaktär. Vidare är det troligt att estimatorns varians underskattas. Utöver samplingvariansen och bortfallsvariansen tillkommer även ett variansbidrag från andra osäkerhetskällor av slumpmässig natur, men vanligtvis kan endast en del av detta variansbidrag "fångas upp" av den varians estimator som används. Tyvärr är det inte möjligt att avgöra i vilken utsträckning som avvikelser finns från förutsättningarna ovan och vilka effekter dessa avvikelser får för tolkningen av de osäkerhetsintervall som presenteras i termer av empirisk konfidensgrad. En rekommendation är dock att man bör vara försiktig med att tolka de presenterade intervallen som regelrätta konfidensintervall.

2.2.2 Ramtäckning

Undersökningens rampopulation består av personer som gick årskurs 2 på gymnasieskolan läsåret 2020/21. Den hämtades från Skolverkets elevregister 2020/21 och bestod vid urvalsdragningen av 105 166 personer som även fanns med i Registret över totalbefolkningen (RTB). Att rampopulationen består av elever i årskurs 2 läsåret 2020/21 beror på att inget register avseende elever i årskurs 3 läsåret 2021/22 fanns färdigt vid tiden för urvalsdragningen.

I och med att rampopulationen skiljer sig från målpopulationen förekommer täckningsbrister i denna undersökning. Övertäckningen i statistiken, dvs. enheter i rampopulationen som inte tillhör målpopulationen, utgörs av elever som avbrutit eller gjort uppehåll i sina studier under eller efter årskurs 2. Undertäckningen i statistiken, dvs. enheter i målpopulationen som inte finns med i rampopulationen, utgörs av elever som inte fanns i gymnasieskolans årskurs 2 läsåret 2020/21, men som var studerande i gymnasieskolans årskurs 3 ett år senare. Det kan exempelvis handla om elever som haft studieuppehåll.

Vid tidpunkten för viktberäkningar fanns ett register bestående av personer som gick årskurs 3 på gymnasieskolan läsåret 2021/22 tillgängligt. För att reducera för täckningsproblematiken, har de summor som används för att fastställa de kalibrerade vikterna hämtats från det uppdaterade registret för årskurs 3. Det går därmed att identifiera och ta hänsyn till övertäckningen med hjälp av det nya registret. Även undertäckningen identifieras med årskurs 3-registret.

I de fall en individ inte var inskriven i gymnasieskolans årskurs 3 läsåret 2021/22 och heller inte besvarat enkäten behandlas denne också som övertäckning.

I de fall en individ som inte var inskriven i gymnasieskolans årskurs 3 läsåret 2021/22 ändå besvarat enkäten räknas denne istället som svarande och därmed som del av populationen.

En matchning mellan årskurs 2-registret och årskurs 3-registret visar följande:

Runt 4,0 procent av de elever som fanns i rampopulationen (årskurs 2-registret) fanns inte med i årskurs 3-registret. Dessa elever har till största del definierats som övertäckning. Undantaget är de som svarat på enkäten (43 elever). Av urvalet på 5 836 elever, var det 241 elever som inte fanns med i årskurs 3-registret. Av dessa svarade 43 elever på enkäten och dessa fick kvarstå i urvalet. Övriga 198 elever i urvalet som saknades i årskurs 3, definierades som övertäckning.

Runt 2,6 procent av de elever som ingick i årskurs 3-registret fanns inte med i rampopulationen (årskurs 2-registret). Dessa elever har definierats som undertäckning.

Genom att kalibrera mot hjälpsummor från årskurs 3-registret i estimationen görs ett försök att justera för de täckningsbrister som uppstår i och med att urvalet dras ifrån årskurs 2-registret. Kvarstående täckningsbrister och deras effekt på andelsskattningarna är svåra att bedöma då ingen särskild studie gjorts av täckningsfel.

2.2.3 Mätning

Undersökningen har genomförts som en post- och webbenkät. I enkäten finns både frågor med fasta och öppna svarsalternativ.

Den första undersökningen av liknande karaktär gjordes 1993. Erfarenheter från hur frågorna fungerat i de tidigare undersökningarna har tagits till vara och justeringar/förbättringar har gjorts vid behov. Inför årets undersökning har frågeformuleringar och svarsalternativ i vissa frågor justerats. Större förändringar presenteras i avsnitt 5.1 *Jämförbarhet över tid*.

Ingen mätfelsstudie har gjorts, därför är det svårt att göra en bedömning av effekterna på resultaten. Det finns dock inga indikationer på att det finns speciellt stora problem i någon grupp av populationen.

Det finns ingen fråga som har fått en nämnvärt stor andel svarande för alternativen "annat" eller "övrigt". Detta alternativ finns med för att inte tvinga in uppgiftslämnarna på svarsalternativ. Skulle någon fråga ha en hög andel svar för dessa alternativ skulle det kunna innebära problem med att tolka resultaten.

I fråga 8 och 10 där eleven tillfrågas om vilket ämnesområde de vill läsa finns information från det öppna svarsalternativet *Annat*. I insamlingen 2021/22 har ingen tolkning av de fritext svar som finns i det öppna svarsalternativet *Annat*. Det är inte självklart att den tolkning av det öppna svaret motsvarar den bedömning individen själv angett som ämnesområde. Till exempel finns det öppna svar där individen tolkat Polisutbildning, YH och övriga utbildningar som högskoleutbildningar. Detta är utbildningen som egentligen inte räknas som högskoleutbildning.

2.2.4 Bortfall

Svarsbortfall, vanligen förkortat till bortfall, uppstår när värden på en eller flera observationsvariabler (t.ex. registervariabler eller enkätfrågor) i en undersökning inte kan inhämtas.

- Saknas alla värden för en individ är det fråga om objektsbortfall.
- Saknas enbart vissa värden, handlar det om partiellt bortfall.

Följande vägda bortfallsmått har använts vid beräkning av objektsbortfallets storlek:

$$\text{Bortfallsandel} = 1 - \frac{\sum_S d_k}{\sum_S d_k + \sum_B d_k + u \sum_O d_k},$$

där d_k är designvikten, dvs. en vikt som tar hänsyn till urvalsdesignen och u är andelen med okänd status som tillhör målpopulationen. Här har u satts lika med 1, vilket innebär att bortfallsmåttet har beräknats med antagandet att alla som inte har svarat i bortfallet tillhör målpopulationen. Undantaget är de 346 elever som ingick i urvalet, men som inte återfanns i årskurs 3-registret. Dessa elever har exkluderats ur beräkningen av objektsbortfallets storlek.

Summeringarna i detta bortfallsmått har gjorts över följande delmängder:

- S: Element tillhörande målpopulationen för vilka fullständiga eller partiella svar erhållits.
- B: Element tillhörande målpopulationen som utgör objektsbortfall.
- O: Element vars målpopulationsstatus är okänd (i detta fall de som inte svarat).

Objektsbortfallet var 67,7 procent. Objektsbortfallet storlek varierade mellan olika grupper i urvalet. I tabellen nedan presenteras objektsbortfallet efter några viktiga redovisningsgrupper.

Tabell 1 Objektsbortfall i olika grupper, procent

Redovisningsgrupp	Designviktad bortfallsandel
Kön	
Kvinna	60,3

Man	74,7
Gymnasieprogram	
Naturvetenskapligt program	53,7
Samhällsvetenskapligt program	64,8
Teknikprogrammet	61,2
Övriga program	73,3
Riksområde	
Stockholms län	67,8
Östra Mellansverige	66,6
Småland med öarna	66,4
Sydsverige	69,7
Västsverige	64,9
Norra Mellansverige	70,6
Mellersta Norrland	69,6
Övre Norrland	71,4
Totalt	67,7

Partiellt bortfall kan bero på att frågan är svår att förstå, är känslig, att uppgiftslämnaren glömmer att besvara frågan, att lämpligt alternativ saknas eller att instruktionerna vid hoppfrågor misstolkas. Till partiellt bortfall räknas även dubbelmarkeringar och svar som inte kan tydas. I denna undersökning varierar det partiella bortfallet mellan 0–3 procent för de flesta frågor.

För att kompensera för objektsbortfallet har en *kalibreringsestimator* använts (se avsnittet om urval och estimation).

Om bortfallet skiljer sig från de svarande med avseende på undersökningsvariablerna så kan skattningarna som grundar sig på enbart de svarande bli skeva. Om exempelvis de svarande till högre andel vill påbörja högskolestudier än de som inte svarat finns risk för överskattning av andelen som planerar att påbörja högskolestudier.

Vikter har beräknats med hjälp av kalibrering. Något försök att mäta kvarstående skevheter efter kalibrering gjordes inte för läsåret 2019/20. För undersökningsomgången avseende läsåret 2013/14 gjordes dock ett sådant försök. Slutsatserna från detta arbete sammanfattas nedan:

Kalibrerade skattningar utifrån hjälpinformation har jämförts med skattningar där ingen hjälpinformation utöver urvalsinformation används. Skattningar har gjorts för några viktiga undersökningsvariabler samt för en registervariabel (hushållsinkomst).

För de studerade undersökningsvariablerna är effekten av kalibrering liten. Den har försumbar betydelse för resultaten.

För registervariabeln hullshållsinkomst totalt så uppstår en liten överskattning av medianinkomsten om kalibrering inte används. När kalibrering används så elimineras denna överskattning nästan helt. För vissa redovisningsgrupper finns dock skevheter orsakade av bortfallet. Skevheter bedöms dock inte som allvarliga. De kalibrerade skattningarna ligger i samtliga studerade redovisningsgrupper relativt nära den kända medianinkomsten i populationen.

Det är okänt hur de som inte svarat på undersökningen skulle ha svarat och vilka konsekvenser det får för den statistik som baseras på de svarande. Den analys som genomförts tyder dock på att de skevheter som bortfallet orsakar inte är allvarliga. Det är inga stora skillnader mellan skattningar från hela urvalet och skattningar från de svarande för de variabler som studerats.

En mer ingående dokumentation av bortfallsanalysen för läsåret 2013/14 finns hos SCB. Se vidare 5.5 *Upplysningstjänster* för kontaktuppgifter.

2.2.5 Bearbetning

Registrering av data sker genom att inkomna enkäter läses in maskinellt (skanning). Skanningprogrammet kontrollerar så att valida värden registreras vid skanningen. De värden som faller ut som fel i dessa kontrolleras rättas upp manuellt med stöd av skanningprogrammet, s.k. verifiering.

Eftersom de flesta variabler i undersökningen är kategoriska med slutna svarsalternativ bör inte skanningen ge upphov till några stora fel vid registrering av data. Givetvis uppstår dock fel, därför görs granskning av skannade data.

Efter verifieringen görs ytterligare kontroller av registrerade data. Logiska kontroller görs så att svar på olika frågor inte är orimliga i förhållande till varandra. I de fall där endast ett svarsalternativ är tillåtet kontrolleras så att inte dubbelmarkeringar finns. Rättning av dubbelmarkeringar för kategoriska variabler sker efter en viss prioritetsordning. I vissa fall sker rättningen utifrån svaren på andra frågor i enkäten. Eftersom antalet upprättningar i denna undersökning har varit litet bedöms effekterna på de slutliga skattningarna som relativt små.

Partiellt bortfall (saknade värden) redovisas oftast i en egen kategori när statistiken publiceras. Ingen imputering av saknade värden görs.

Punktskattningar och konfidensintervall har beräknats med hjälp av ett av SCB egenutvecklat estimationsprogram, ETOS. Skattningarna har förts över till Excel där ytterligare bearbetningar har gjorts för att skapa de diagram som publicerats i temarapporten, samt de tabeller som publicerats separat.

I några fall i rapporten redovisas resultat baserade på den skattade skillnaden mellan två grupper. Samtliga är fall där skillnaden till 95 procent är statistiskt säkerställd (95 procents konfidensgrad).

Det är svårt att bedöma effekten av de fel som kan ha uppstått vid bearbetningen. Ingen studie har gjorts för att försöka bedöma storleken på denna felkälla.

2.2.6 Modellantaganden

Denna källa bidrar inte till osäkerhet i statistiken

2.3 Preliminär statistik jämförd med slutlig

Endast slutgiltig statistik redovisas.

3 Aktualitet och punktlighet

3.1 Framställningstid

Resultaten presenteras i mars året efter referensterminen dvs. ca 2-3 månader efter läsårets slut.

3.2 Frekvens

Undersökningen genomförs och publiceras vartannat år.

3.3 Punktlighet

Statistiken följer publiceringsplanen för Sveriges officiella statistik, www.scb.se/hitta-statistik/publiceringskalendern/.

4 Tillgänglighet och tydlighet

4.1 Tillgång till statistiken

Statistiken publiceras i temarapporten *Gymnasieungdomars studieintresse 2020/21*. Resultaten redovisas också i färdiga tabeller på SCB:s webbplats. Alla resultat redovisas på www.scb.se/UF0513 och sammanfattas i statistiknyheter.

4.2 Möjlighet till ytterligare statistik

För att tillgodose behovet av att kunna följa upp elevernas eventuella högskolestudier lagras primärdata vid SCB. SCB utför på beställning specialbearbetningar av primärmaterial. Forskare, utredare m.fl. kan efter särskild prövning få tillgång till aidentifierade mikrodata.

4.3 Presentation

Statistiken redovisas i form av text, tabeller och diagram.

4.4 Dokumentation

Framställningen av statistikregistret och statistiken beskrivs i *Dokumentation av statistikens framställning (StaF)*. Statistikens kvalitet beskrivs i innevarande dokument, *Kvalitetsdeklaration*. Detaljerad information om mikrodata finns beskrivet i *Dokumentation av mikrodata (MetaPlus)*. Samtliga dokumentationer finns på SCB:s webbplats, www.scb.se/UF0513.

5 Jämförbarhet och sammanvändbarhet

5.1 Jämförbarhet över tid

Sedan den första undersökningen 1992/93 har gymnasieskolan som skolform genomgått förändringar. Även undersökningen har ändrats över tid, både vad gäller referenstid och hur olika frågor ställs. Dessa förändringar bör tas hänsyn till vid analys av materialet. De största förändringarna beskrivs nedan.

Förändringar inom gymnasieskolan

Under åren har gymnasieutbildningarna förändrats, vilket även påverkar jämförbarheten över tid. I och med de olika gymnasiereformer som genomförts har programstruktur och innehåll förändrats. Från och med läsåret 2011/12 kom reformen "Gy11", vilken förde med sig förändringar inom bl.a. ämnesplaner, gymnasieprogram och behörighetskrav till gymnasieskolan. De elever som sökte gymnasieutbildning till läsåret 2011/12 omfattades av den nya reformen. Det innebär att majoriteten av de elever som fanns i årskurs 3 läsåret 2013/14 därmed ingick i den nya gymnasieskolan. Av detta skäl görs från och med 2013/14 en brytning i de tidsserier som publiceras för *program* och jämförelser över tid rekommenderas därmed inte.

Förändring i redovisningsgrupp "Riksområde (region)"

Hur redovisningsgruppen riksområde (region) tagits fram har varierat mellan olika undersökningsomgångar. Från och med undersökningen 2013/14 används skolans län som grund för variabeln, men i vissa tidigare undersökningar har istället elevens hemkommun varit utgångspunkt. Om hemkommun använts istället för skolans län i undersökningen 2013/14 hade det resulterat i att 4 procent av urvalet hade kopplats till ett annat riksområde än det område de nu kopplades till. Skillnader mellan åren beroende på detta bör således vara förhållandevis små, men jämförelser över tid på regionnivå bör ändå göras med detta i åtanke.

Förändring i urvalsstorlek

Inför undersökningen år 2011/12 ökades urvalet från 5 303 till 21 977 personer. För undersökningen därefter, 2013/14, minskades urvalet till 5 014 personer och var ungefär detsamma under 2015/16, 2017/18, 2019/20 och 2021/22. Under 2021/22 gjordes en mindre justering av urvalet. Antalet elever i urvalet som läste naturvetenskapsprogrammet, samhällsvetenskapsprogrammet eller teknikprogrammet minskade medan i övriga program ökade urvalet något. Detta kan påverka jämförbarheten mellan undersökningarna men ingen studie av detta har genomförts.

Förändring i variabel "Planer på högskole- eller universitetsstudier om tre år"

Till undersökningen av läsåret 2003/04 gjordes förändringar i den viktigaste frågan om intresset för högskolestudier. Bland annat ändrades ett av svarsalternativen. Nedan följer en fullständig redovisning av förändringarna i frågan om intresset för högskolestudier.

Introduktionsprogram

I undersökningen avseende läsåret 2019/20 har elever på introduktionsprogram uteslutits från målpopulationen. Anledningen till att de har uteslutits är att kvaliteten på skattningarna för denna grupp var för dålig 2017/18, främst på grund av få svarande elever. Det är heller inte helt relevant att fråga elever på introduktionsprogram om studieplaner på eftergymnasial nivå, då de ännu inte påbörjat studier på ett nationellt program i gymnasieskolan.

Blankett läsåret 2003/04 och senare

1	Har du planer på att börja läsa på universitet eller högskola inom de närmaste tre åren?	<input type="checkbox"/>	1 Ja	→ Gå till fråga 3
		<input type="checkbox"/>	2 Nej	→ Gå till fråga 2
		<input type="checkbox"/>	3 Vet inte	→ Tack för ditt svar! Skicka in blanketten till SCB.

Blankett före läsåret 2003/04

1	Har du planer på att börja läsa på universitet eller högskola inom de närmaste tre åren?	<input type="checkbox"/>	1 Ja	→ Gå till fråga 3
		<input type="checkbox"/>	2 Jag har inte bestämt mig än	→ Gå till fråga 3
		<input type="checkbox"/>	3 Nej	→ Gå till fråga 2

Dessa förändringar fick till följd att andelen som var intresserade att läsa på högskolan inom tre år ökade kraftigt (en kontrollgrupp som fick den gamla blanketten användes i undersökningen). Av de elever som erhöll den nya blanketten (läsåret 2003/04 och senare) planerade **60 procent** att läsa vidare på högskolan inom de närmaste tre åren. Motsvarande andel som erhöll den gamla blanketten (läsåret 2002/03 och tidigare) var endast **50 procent**. Skillnaden är statistiskt säkerställd.

Detta medför att inga jämförelser kan göras mellan undersökningar gjorda innan läsåret 2003/04 och efter 2003/04 i fråga om gymnasieelevernas studieintresse. Det visade sig att nästintill all ökning i studieintresset (2003/04 var studieintresset 60 procent och 2002/03 var det 49 procent) kunde hänföras till förändringen i frågekonstruktionen. Däremot kan jämförelser från undersökningen 2003/04 och senare göras.

Till undersökningen läsåret 2011/12 gjordes en mindre ändring i samma fråga. Förändringen medförde främst att de som inte visste om de hade planer på att börja läsa på universitet eller högskola inom de närmaste tre åren ändå fick svara på frågorna om hur mycket information respondenten fått om högskolestudier från ett antal informationskällor.

1	Har du planer på att börja läsa på universitet eller högskola inom de närmaste tre åren?	<input type="checkbox"/> Ja → Gå till fråga 4
		<input type="checkbox"/> Nej → Gå till fråga 2
		<input type="checkbox"/> Vet inte → Gå till fråga 22

Förändring i variabler om utbildningsinriktning på högskole- eller universitetsstudier.

Läsåret 2015/16 förändrades formuleringen av frågorna, liksom grupperingen av ämnesområdena. Svartalernativet *Jag har inte bestämt mig* lades också till i den första av de två frågorna. I den andra frågan lades ett hopp till för dem som angett *Jag har inget andrahandsval*. På grund av omgrupperingen av ämnesområdena bör jämförelser med tidigare undersökningsomgångar göras med stor försiktighet.

5.2 Jämförbarhet mellan grupper

I den publicerade statistiken ingår ett flertal olika redovisningsgrupper och jämförelser kan göras mellan dessa, bl.a. mellan kvinnor och män samt inrikes och utrikes födda.

5.3 Samanvändbarhet i övrigt

Sammanvändbarhet med annan utbildningsstatistik är möjlig. Genom att utnyttja SCB:s registerstatistik om övergången från gymnasieskolan till högskolan ges exempelvis möjlighet att få en uppfattning om andelen gymnasieelever som fullföljer/inte fullföljer sina studieplaner.

Undersökningen använder till viss del samma redovisningsgrupper och indelningar som annan av SCB:s statistik. Vid sammanvändning bör dock hänsyn tas till om statistiken är grundad på urvalsdata eller inte.

5.4 Numerisk överensstämmelse

Det finns inga brister i numerisk överensstämmelse då inga avrundningar görs.

Allmänna uppgifter

A Klassificeringen Sveriges officiella statistik

För statistik som ingår i Sveriges officiella statistik (SOS) gäller särskilda regler för kvalitet och tillgänglighet, se lagen (2001:99) och förordningen (2001:100) om den officiella statistiken samt Statistiska centralbyråns föreskrifter (SCB-FS 2016:17) om kvalitet för den officiella statistiken.

B Sekretess och personuppgiftsbehandling

I myndigheternas särskilda verksamhet för framställning av statistik gäller sekretess enligt 24 kap. 8 § offentlighets- och sekretesslagen (2009:400).

För att skydda enskilda personers eller företags sekretessreglerade uppgifter säkerställs att de inte kan röjas direkt eller indirekt i den statistik som offentliggörs.

Vid behandling av personuppgifter, dvs. information som direkt eller indirekt kan hänföras till en person som är i livet, gäller lagen (2001:99) och förordningen (2001:100) om den officiella statistiken samt personuppgiftslagen (1998:204).

Registret för *Gymnasieungdomars studieintresse* är ett personregister och enligt personuppgiftslagen ska registerutdrag lämnas på begäran.

C Bevarande och gallring

Enligt lagen om den officiella statistiken (SFS 2001: 99) ska personuppgifter gallras hos de statistikansvariga myndigheterna när uppgifterna inte längre behövs för sitt ändamål. Med stöd av förordningen (2001:100) om den officiella statistiken 12 § har Riksarkivet föreskrivit i RA-MS 2007:64 undantag från gallring av uppgifter i undersökningen "Gymnasieungdomars studieintresse".

D Uppgiftsskyldighet

Uppgiftsskyldighet till denna statistik föreligger inte.

E EU-reglering och internationell rapportering

Undersökningen är inte berörd av någon EU-reglering och ingen internationell rapportering sker.

F Historik

SCB genomförde den första undersökningen om gymnasieungdomars studieintresse under våren 1993 bland de elever som gick i gymnasieskolans avgångsklasser läsåret 1992/93. Nästa undersökning gjordes på hösten samma år och då bland elever i gymnasieskolans avgångsklasser läsåret 1993/94. Syftet med att ändra referensperiod från vårtermin till hösttermin var att kunna presentera statistiken tidigare. Mellan läsåren 1993/94 och 2003/04 genomfördes undersökningen varje år. Därefter har den endast genomförts vartannat år.

G Kontaktuppgifter

Statistikansvarig myndighet	Statistiska centralbyrån
Kontaktinformation	Jessica Baldgrim
E-post	Jessica.Baldgrim@scb.se
Telefon	010-479 66 44