

STATISTIKENS FRAMSTÄLLNING

Doktorander och examina i högskoleutbildning på forskarnivå

Ämnesområde

Utbildning och forskning

Statistikområde

Högskoleväsende

Produktkod

UF0204

Referenstid

Kalenderåret 2022 för doktorandnybörjare, examina och examinerade

Andra halvåret (1 juli-31 december) 2022 för doktorander

Kontaktuppgifter

Statistikansvarig myndighet	Universitetskanslersämbetet
Kontaktinformation	Analysavdelningen
E-post	analys@uka.se
Telefon	08-563 085 00

Statistikproducent	Statistiska centralbyrån (SCB)
Kontaktinformation	Avdelningen för social statistik och analys
E-post	forskarutbildning@scb.se
Telefon	010-479 50 00 (Statistikservice)

Innehåll

1	Statistikens sammanhang	3
2	Undersökningsdesign	3
2.1	Målstorheter	3
2.2	Ramförfarande	3
2.3	Förfaranden för urval och uteslutning	4
2.4	Insamlingsförfarande.....	4
2.4.1	Datainsamling	4
2.4.2	Mätning.....	4
2.4.3	Bortfallsuppföljning.....	5
2.5	Bearbetningar.....	5
2.6	Granskning.....	5
2.6.1	Granskning under direktinsamlingen	5
2.6.2	Granskning av mikrodata och insamlade statistikvärden.....	5
2.6.3	Granskning av makrodata	6
2.6.4	Granskning av redovisning	6
2.7	Skattningsförfarande	7
2.7.1	Principer och antaganden	7
2.7.2	Skattningsförfarande för målstorheter.....	7
2.7.3	Skattningsförfarande för tillförlitlighet.....	7
2.7.4	Röjandekontroll	7
3	Genomförande	7
3.1	Kvantitativ information.....	7
3.2	Avvikelser från undersökningsdesignen	7

1 Statistikens sammanhang

Statistiken över doktorander och examina i högskoleutbildning på forskarnivå belyser förändringar av antalet doktorandnybörjare, doktorander och examina efter olika redovisningsgrupper såsom kön, forskningsämnesområde och lärosäte. Doktorandernas studiefinansiering, genomströmning och internationella mobilitet illustreras också.

Statistiken är en del av högskolestatistiken, vars syfte är att bidra till en allsidig statistisk belysning av högskolan och dess utveckling över tid i Sverige. Högskolestatistiken produceras av SCB på uppdrag av Universitetskanslersämbetet (UKÄ).

Läs om statistikens kvalitet i kvalitetsdeklarationen som finns på undersökningens hemsida¹, under rubriken *Dokumentation*.

2 Undersökningsdesign

2.1 Målstorheter

Statistiken för 2022 redovisar målstorheter som:

- Antalet aktiva doktorander. Dessa redovisas exempelvis efter kön, lärosäte och forskningsämnesområde.
- Antal doktorandnybörjare. Doktorandnybörjare redovisas exempelvis efter kön, lärosäte och forskningsämnesområde.
- Antalet examina och examinerade. Doktors- och licentiatexamina. Examina redovisas exempelvis efter kön, lärosäte och forskningsämnesområde.

2.2 Ramförfarande

Ramen utgörs av de register över studerande på forskarnivå som förs av de 36 lärosäten som 2022 hade examenstillstånd för högskoleutbildning på forskarnivå². Rampopulationen utgörs av de antagna till utbildning på forskarnivå, de aktiva doktoranderna och de utfärdade examina på forskarnivå som finns med i lärosätenas studieadministrativa register.

Observationsobjekten och uppgiftskällorna skiljer sig åt då observationsobjekten är de individer som finns med i lärosätenas studieadministrativa register, och uppgiftskällorna är administratörer vid lärosätena.

Kontakt med uppgiftskällorna sker via e-post.

¹ <https://www.scb.se/UF0204>

² <https://www.uka.se/integrationer/examenstillstand>

2.3 Förfaranden för urval och uteslutning

Inget urval och ingen uteslutning förekommer i denna undersökning.

2.4 Insamlingsförfarande

2.4.1 Datainsamling

Uppgifterna har huvudsakligen samlats in direkt från lärosätenas studieadministrativa system, Ladok, via ett för ändamålet anpassat uttag, vilket medför en liten uppgiftslämnarbörda. Uppgifter från Handelshögskolan i Stockholm och Världssjöfartsuniversitetet samlas in via SCB:s insamlingsverktyg SIV. De statliga lärosätenas rapportering regleras i förordningen (1993:1153) om redovisning av studier m.m. vid universitet och högskolor. Enskilda utbildningsanordnare som har fått tillstånd att utfärda examina är skyldiga att medverka i uppföljningar och utvärderingar av utbildningen enligt lagen (1993:792) om tillstånd att utfärda vissa examina.

Undersökningen har två insamlingsomgångar. Statistik avseende första halvåret 2022 samlades in i augusti 2022. Statistik avseende andra halvåret 2022 samlades in i februari 2023.

Instruktioner och anvisningar till uppgiftslämnare finns på undersökningens uppgiftslämnarsida³.

2.4.2 Mätning

Uppgiftslämnarna sänder in uppgifter till SCB, antingen som XML-filer via S-FTP, eller som Excel-filer via SCB:s insamlingsverktyg SIV. Instruktioner och definitioner finns tillgängliga på undersökningens uppgiftslämnarsida⁴, under rubriken *Instruktioner*.

Observationsvariablerna överensstämmer med målvariablerna. De huvudsakliga målvariablerna återfinns under rubriken 1.2.2 Variabler i Kvalitetsdeklarationen för Doktorander och examina i högskoleutbildning på forskarnivå på SCB:s webbplats⁵.

Mätning av antalet doktorandnybörjare och examinerade görs per kalenderår. Mätning av antalet doktorander görs avseende hösten. Anledningen är att antalet doktorander som är aktiva överskattas om de mäts per år. Under ett år finns både ett inflöde (doktorandnybörjare) och ett utflöde (examinerade) av doktorander. Alla doktorandnybörjare och examinerade är inte aktiva under hela året. Studiefinansiering redovisas emellertid per kalenderår även för doktorander.

³ <http://www.scb.se/doktorander>

⁴ <http://www.scb.se/doktorander>

⁵ <https://www.scb.se/UF0204>

2.4.3 Bortfallsuppföljning

Inget bortfall förekommer i denna undersökning.

2.5 Bearbetningar

För att identifiera utländska doktorander används uppgifter från Registret över totalbefolkningen (RTB) och STATIV som är SCB:s longitudinella databas för integrationsstudier.

2.6 Granskning

2.6.1 Granskning under direktinsamlingen

För de uppgifter som lämnas av lärosäten som är anslutna till Ladok sker följande kontroller under insamlingsfasen:

- Filnamnen ska ha rätt namnsättning, avse rätt period och antalet filer ska vara tre från varje lärosäte.
- Inga dubbletter av filerna har skickats in.
- Antalet doktorander och examina ska vara i samma storleksordning som under motsvarande period året innan.
- Uppenbara dubbletter, d.v.s. doktorander och examina som redan finns i högskoleregistret, och uppgifter som är äldre än 1998 makuleras maskinellt vid laddning av indata.

För de uppgifter som lämnas via SIV granskas svaren via kontroller som sker direkt i Excelblanketten och via en fellista som skapas i samband med att filer sänds in till SCB.

- I Excelblanketterna kontrolleras i huvudsak om värden saknas för variabler och att ifyllda värden finns i värdemängden. Kontrollerna resulterar i felmeddelanden och uppgiftslämnaren uppmanas att rätta uppgifterna.
- Eventuella kvarstående fel hamnar på en fellista som genereras när filen skickas in. Uppgiftslämnaren måste då utföra rättningen för att filen ska kunna skickas till SCB.

2.6.2 Granskning av mikrodata och insamlade statistikvärden

Mikrodatagranskningen kan delas in i tre huvudsakliga områden: granskning av personnummer, granskning av katalogposter (värdemängder) och granskning av individposter.

Granskning av personnummer görs i en applikation som skapats särskilt för detta ändamål på SCB. Granskningen av personnummer görs av två skäl. Det första skälet är att så långt det är möjligt ersätta Ladoks interrimspersonnummer med personnummer i folkbokföringen. Det andra skälet är att se till att en student som rapporterats på olika nivåer och perioder inte ingår i Universitets- och högskoleregistret med fler än ett personnummer (t.ex. ett riktigt och ett felaktigt).

Granskning av katalogposter (värde mängder) görs via SQL-skript. Granskningen går ut på att identifiera nya och korrigerade värdekoder och värden för t.ex. lärosäten, forskarutbildningsämnen, finansieringstyper och examina. Vidare granskas att vissa kopplingar mellan olika katalogposter finns, t.ex. att ett nytt forskarutbildningsämne kan knytas till en organisationsenhet på lärosätet.

Granskning av individposter görs maskinellt via en applikation. Det finns 100 maskinella granskningsvillkor. Misstänkta fel hamnar på fellistor, en för varje granskningsvillkor. Det som granskas är:

- Överensstämmelse mellan uppgifter från antagning, aktivitet/studiefinansiering och examen och mellan uppgifter som lämnats vid olika tidpunkter. Ofta handlar det om att avgöra om en ny uppgift om forskarutbildningsämne avser en ny antagning, ett byte eller ett korrigerat fel.
- Rimlighetskontroller avseende aktivitetsprocent. Aktiviteten ska inte överstiga 100 procent. Den får inte heller vara för låg eller för hög i förhållande till tidpunkten för antagning, avbrott, uppehåll eller examen.
- Dubletter uppstår oftast när lärosätet skickar in en ny uppgift för en doktorand som i något avseende skiljer sig från den som redan finns i registret. Det kan vara ett nytt forskarutbildningsämne, en ny aktivitetsprocent eller en ny finansieringstyp. Rena dubletter makuleras maskinellt vid inläsning av indata.

SCB har behörighet att i samband med granskningen kontrollera uppgifter i det studieadministrativa systemet Ladok.

2.6.3 Granskning av makrodata

Data laddas i flera steg allteftersom produktionen fortgår. Inför laddning av observationstabellen där rättning sker, måltabellen där granskade uppgifter lagras och det slutliga observationsregistret som används för statistikuttag görs en avstämning av att all relevant data ingår. Avstämningen görs genom att, efter varje laddning, räkna antalet antagna, aktiva, finansieringsposter och examina per lärosäte.

2.6.4 Granskning av redovisning

Granskning av tabellvärden görs med SQL-skript eller i Excel. Tabeller som innehåller tidsserier granskas mot föregående års publicering. Det senaste året granskas med hjälp av kontrollskript mot det slutliga observationsregistret.

Utöver granskning av värden så görs följande kontroller:

- Alla tabeller har tagits fram
- Referensåren i rubriker och tabellhuvuden är de rätta

- Stavningen är korrekt
- Text eller totalsummor som förekommer i flera tabeller överensstämmer med varandra
- En översättning till engelska har gjorts som överensstämmer med ursprungstexten. Korrekta begrepp används i översättningen
- Värdena är högerställda och har tusentalsavgränsare
- Inga formler förekommer i tabellceller
- Texterna i informationsflikarna är aktuella med avseende på omvärldsförändringar
- SoS-loggan finns där den ska vara med

2.7 Skattningsförfarande

2.7.1 Principer och antaganden

Vid skattningsförfarandet antas de insamlade uppgifterna vara heltäckande för målpopulationen, vilket möjliggör totalräkning av statistiken.

2.7.2 Skattningsförfarande för målstorheter

Målstorheterna skattas via summeringar, medianer och andelar.

2.7.3 Skattningsförfarande för tillförlitlighet

Något osäkerhetsmått som komplement till punktskattningarna av målstorheterna beräknas inte.

2.7.4 Röjandekontroll

Inga skyddsmetoder har tillämpats på materialet. Personuppgifterna i materialet har inte bedömts som känsliga, vilket innebär att ingen individ skulle kunna lida men av att personuppgifterna röjs.

3 Genomförande

3.1 Kvantitativ information

Ramen utgörs av de administrativa register som förs över studerande vid de 36 lärosäten som har examenstillstånd för högskoleutbildning på forskarnivå⁶. År 2022 var det 2 920 doktorandnybörjare (preliminärt) och 2 750 doktorsexamina i högskoleutbildning på forskarnivå. Antalet doktorander andra kalenderhalvåret var 17 450 (preliminärt).

3.2 Avvikelser från undersökningsdesignen

Ingen avvikelse från undersökningsdesignen har gjorts.

⁶ <https://www.uka.se/integrationer/examenstillstand>