

Kursbeskrivning: Statistiska Databaser och Register

AN 7,5 högskolepoäng

Engelsk titel: Statistical Data-bases and Registers, AL, 7,5 ECTS-credits

Avancerad nivå och mastersnivå

Kursen kan också ingå som moment i vissa kurser på avancerad nivå

Beskrivning av kursen

Kursen består av ett moment:

1. Statistiska Databaser och Register, 7,5 högskolepoäng

Idag är databaser en integrerad del i statistikproduktion. Mycket statistik tas fram genom att sambearbeta uppgifter från olika källor. Uppgifterna måste då ha lagrats på ett sådant sätt att sambearbetningen underlättas. Statistik redovisas också i form av databaser. Data kan t ex ha lagrats så att användaren kan konstruera sina egna tabeller eller studera helt egna frågeställningar.

Kursen tar bl a upp hur databaser och register byggs upp, användning av SQL, datamining, hjälpinformation, samkörning, samordning och kvalitetsaspekter på statistiska databaser. Mikrosimulering.

Kursens innehåll ger kunskaper som är av stor nytta vid framtagning av statistik som bygger på data från olika källor. Den som vill att insamlade uppgifter skall kunna användas i framtiden av andra har också stor nytta av kursen.

Kursens lärandemål

Efter att ha genomgått kursen förväntas studenten

- känna till och redogöra för några vanliga svenska databaser och deras egenskaper samt
- kunna bedöma relevans och kvalitet hos olika ramar rent generellt
- känna till och redogöra för principerna för mikrosimulering och genomföra enklare mikrosimuleringar
- använda några av våra vanligare program för databashantering såsom SQL behärska något enkelt program för datamining
- genomföra urvalsdragningar från större databaser t ex pps och stratifierade urval samt sedan kunna utnyttja hjälpinformationen i registren till effektiva skattningar.

Undervisning

Undervisningen består av 14 undervisningstillfällen. Schema framgår av institutionens hemsida. För att få gott resultat på kursen är det viktigt att närvara vid samtliga undervisningstillfällen.

Kursen motsvarar 7,5 högskolepoäng vilket i sin tur motsvarar 10 halvtidsveckor eller 200 arbetstimmar.

Kursen kan sägas bestå av tre delar – statistiska aspekter på registerstatistik och konstruktion (fyra tillfällen med Anders och Britt Wallgren), infologiska aspekter på register och databaser (sex tillfällen med främst Bo Sundgren) samt datamining och övrigt (fyra tillfällen med olika lärare, Databaser inom hälsoområdet, inom privata ekonomiska sektorn, datamining, beslutsträdstekniker). I arbetsinsats kommer delarna att motsvara ungefär 80, 80 resp. 40 arbetstimmar.

Kunskapskontroll och examination och betygskriterier

Momentet examineras genom individuella inlämningsuppgifter, muntliga prov och skriftliga prov. Uppgifter och prov poängbedöms och kan tillsammans ge maximalt 100 poäng. De tre delarna motsvarar 40, 40 resp. 20 poäng. Förutom poäng åsätts prov och inlämningsuppgifter även betyget ej godkänd eller godkänd. För godkänt resultat på momentet (betyg A-E) krävs dels att samtliga uppgifter åsatts betyget godkänd, dels att totalpoängen överstiger 50 poäng. Allt skall vara uppnått vid samma "kurstillfälle" dvs. under våren 2008. Betyget på momentet baseras på en sammanräkning av poängen som alltså maximalt kan vara 100 poäng.

Betygssättning sker efter en sjugradig betygsskala

- A. Utmärkt – motsvarar 90 – 100 poäng
- B. Mycket bra – motsvarar 80 – 89 poäng
- C. Bra – motsvarar 70 – 79 poäng
- D. Tillfredsställande – motsvarar 60 – 69 poäng
- E. Tillräckligt – motsvarar 50 – 59 poäng
- Fx Otillräckligt – motsvarar 30 – 49 poäng eller
mer än 50 poäng utan att alla uppgifter och prov godkänts
- F. Helt Otillräckligt – motsvarar mindre än 29 poäng

För att få godkänt på hela kursen krävs lägst betyget E på momentet. Betyget på kursen blir detsamma som på momentet.

Första delen (ca 3 ECTS-poäng) examineras med muntlig tentamen på kursboken. Kursboken bör tas med på tentamen. Andra delen examineras preliminärt med en inlämningsuppgift (få fram uppgifter ur verkliga databaser, ca 1 ECTS-poäng) och en skriftlig tentamen med teorifrågor och en konstruktion av lämplig datamodell (2 ECTS-poäng). Sista delen (Ca 1.5 ECTS-poäng) tenteras med en hemtentamen.

Studerande som godkänts på prov får ej genomgå förnyat prov för högre betyg. Normalt förfaller inlämningsuppgifter och skrivningsresultat efter vårterminen 2008. Efter detta skall student som ännu ej godkänts på hela momentet göra aktuell termins samtliga uppgifter och prov.

Kurslitteratur:

- Wallgren, A. och Wallgren, B. (senaste upplagan) *Register-based Statistics – Administrative Data for Statistical Purposes*. Chichester, Wiley.
- Register-based statistics in the Nordic Countries – Review of best practices with focus on population and social statistics (2007), Geneva, United Nations (finns att ladda ner på sidan

www.unece.org/stats/publications)

- Kompletterande material som utdelas i samband med undervisningen. (Detta är en viktig del av kurslitteraturen)

Lärare:

Kursansvarig: Daniel Thorburn, 08 – 16 29 56, Daniel.thorburn@stat.su.se

Del 1, Anders och Britt Wallgren

Del 2, Bo Sundgren och Alf Fyhrlund

Del 3, Mats Lanner, Sven Cnattingius, Per Weidenman.