

KURSPLAN

Statistisk analys och datavisualisering, Forskarnivå

Statistical Analysis and Data Visualization, Post-graduate level

6 högskolepoäng/ECTS

Kurskod: IT0945F

Kursplanen gäller från: 2021-07-01

Datum för fastställande: 2021-05-03

Version: 1

Ämne för utbildning på forskarnivå: Informationsteknologi

Utbildningsnivå: Forskarnivå

1 Kursens benämning, omfattning och nivå

Kursen ges av Högskolan i Skövde och benämns Statistisk analys och datavisualisering, Forskarnivå och omfattar 6 högskolepoäng. Kursen tillhör forskarnivå.

2 Mål

Efter avslutad kurs ska doktoranden kunna:

- beskriva och skilja mellan olika metoder inom deskriptiv (beskrivande) och statistisk inferens,
- demonstrera användningen av multivariata visualiseringsmetoder för explorativ (undersökande) dataanalys,
- formulera statistiska hypoteser och testa dessa med hjälp av statistisk programvara,
- utvärdera och motivera användningen av specifika statistik- och visualiseringsmetoder för vissa dataanalysuppgifter,
- analysera och bedöma statistiska analysmetoder som använts i ett urval av vetenskapliga artiklar från olika forskningsområden, och
- argumentera för hur experimentell design och statistisk rapportering kan påverka debatten inom vetenskapssamhället och samhället i övrigt.

3 Innehåll

Kursen behandlar viktiga statistiska metoder inom deskriptiv (beskrivande) och statistisk inferens, samt explorativ (undersökande) analys genom multivariat datavisualisering. Kursen inleds med en introduktion

till statistiska begrepp, vanliga sannolikhetsmodeller och deras tillämpningar. Klassiska univariata och bivariata visualiseringsmetoder beskrivs, följt av multivariata visualiseringsmetoder baserade på 'machine learning', som inkluderar klustring, dimensionsreducerande projektioner och tvådimensionell avbildning.

Kursens avsnitt om statistisk inferens tar upp det frekventistiska angreppssättet, men ger även en kort översikt av Bayesianska metoder. Grunden för statistisk hypotesprövning och dess grund i den vetenskapliga metoden diskuteras. I detta avsnitt förklaras vanligt förekommande experimentell design och samlingsmetoder. Ett stort urval av hypotestester, inklusive anpassningsgrads-, centralmåts- och spridningsmåts- tester behandlas, för scenarier med ett, två eller flera stickprov. Statistiska regressionsmetoder beskrivs också i statistisk inferens.

Kursen avslutas med en kritisk analys av statistiska analysmetoder som använts i vetenskapliga artiklar från olika forskningsområden och diskussioner om hur felaktig experimentell design, felaktig tolkning av statistiska tester och oetisk statistisk rapportering kan påverka debatten negativt inom vetenskapssamhället och samhället i övrigt.

Kursen använder statistiska verktygslådor och bibliotek i ett eller flera vetenskapliga datorprogram och programmeringsmiljöer. I de praktiska övningarna används data från verkligheten, varav några väljs för att främja diskussioner om samhällsfrågor såsom klimatförändringar, ekonomisk ojämlikhet, ojämlikhet mel-

lan könen och diskriminering.

4 Undervisningsformer

Undervisningen består av föreläsningar, laborationer, projektarbeten och seminarier.

Undervisningen bedrivs på engelska.

5 Examination

Kursen bedöms med betygen Underkänd eller Godkänd.

Betyg för kursen utfärdas först när samtliga examinationsmoment godkänts. Enskilda examinationsmoment registreras inte.

Poängregistrering av examinationen för kursen sker enligt följande system:

Examinationsmoment	Omfattning	Betyg
Laborationer	2 hp/ECTS	U/G
Seminarium	1 hp/ECTS	U/G
Projekt rapport ¹	3 hp/ECTS	U/G

¹ Bestämmer kursens slutbetyg vilket utfärdas först när samtliga moment godkänts.

För att få godkänt slutbetyg på kursen krävs att samtliga examinationsmoment har bedömts som godkända.

6 Behörighet

Kursens behörighetskrav är grundläggande behörighet för utbildning på forskarnivå, dvs avlagd examen på avancerad nivå eller slutförda kursfordringar om minst 240 hp, varav minst 60 hp på avancerad nivå, eller motsvarande.

För att uppfylla kravet på särskild behörighet krävs att den sökande har fullgjort kursfordringar om minst 60 högskolepoäng, inklusive självständigt uppsatsarbete omfattande minst 15 högskolepoäng på avancerad nivå, inom området informationsteknologi, näraliggande tillämpningsområden eller andra ämnesområden som bedöms som direkt relevanta för avhandlingsarbetet.

Vidare krävs godkänt betyg i gymnasiekursen Engelska 6 eller motsvarande. Motsvarande kunskaper visas normalt genom ett internationellt erkänt språktest, t ex IELTS, TOEFL eller liknande.

7 Ämne för utbildning på forskarnivå

Kursen tillhör forskarutbildningsämnet informationsteknologi vid Högskolan i Skövde.

8 Inrättande av kurs och fastställande av kursplan

Kursen är inrättad av utbildningskommittén för utbildning på forskarnivå inom informationsteknologi 2021-05-03. Denna kursplan är fastställd av utbildningskommittén för utbildning på forskarnivå inom informationsteknologi 2021-05-03. Kursplanen gäller från 2021-07-01.

9 Överlappning av annan kurs

Kursen kan inte ingå i examen med annan kurs, vars innehåll helt eller delvis överensstämmer med innehållet i denna kurs.

10 Övrigt

Ytterligare information lämnas om kursen på Högskolans utbildningssidor på webben inför respektive kurstillfälle.

Nationella och lokala styrdokument för Högskolans verksamhet finns tillgängliga på Högskolans webbplats.

Efter kursens slut sker en utvärdering av kursen där doktorandens erfarenheter och synpunkter inhämtas. Utvärderingens främsta syfte är att bidra till förbättringar och utveckling av kursen. Doktoranden ska informeras om resultatet av kursutvärderingen och eventuella beslut om åtgärder.

11 Kurslitteratur och övriga läromedel

Referenslitteratur

Chen, C., Härdle, W. K., Unwin, A. (2008). *Handbook of Data Visualization*. Berlin Heidelberg: Springer-Verlag. URL: doi:10.1007/978-3-540-33037-0 ISBN 9783662500743.

Efron, B., & Hastie, T. (2016). *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*. (Institute of Mathematical Statistics Monographs). Cambridge, UK: Cambridge University Press. URL: doi:10.1017/CBO9781316576533 .

Field, A. (2017). *Discovering statistics using IBM SPSS Statistics*. (5th ed.) London: SAGE. ISBN 9781526419521.

Nussbaumer Knaflic, C. (2015). *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Hoboken, New Jersey: John Wiley & Sons. ISBN 978-1-119-00225-3.

Peck, R., Olsen, C. & Devore, J. (2019). *Introduction to Statistics and Data Analysis*. (6th ed.) Cengage Learning. ISBN 9781337793612.

Wasserman, L. (2004). *A Concise Course in Statistical Inference. All of Statistics*. New York, NY: Springer. URL: doi.org/10.1007/978-0-387-21736-9 ISBN

9781441923226.

Weiss, N. A. (2016). *Introductory Statistics*. Pearson. ISBN 9780321989178.

Vetenskapliga artiklar enligt lärarnas anvisningar.