

# Statistik og Dataanalyse for Softwareteknologi

## ECTS

5

## Forudsætninger

Det forventes at den studerende har matematik svarende til adgangskravet for uddannelsen samt har bestået MSE1 eller tilsvarende kursus. Det anbefales yderligere at tage et databasekursus (fx TS-DBS1) sideløbende eller tidligere have haft om databaser.

## Formål

Formålet med kurset er at give studerende en praktisk introduktion til sandsynlighedsteori, statistik og dataanalyse, med et stærkt fokus på softwareteknologiske anvendelser. De studerende skal opnå de nødvendige værktøjer til at indsamle, behandle, analysere og visualisere data hentet fra databaser, samt lære hvordan disse analyser og deres resultater effektivt kan klargøres og integreres i softwareapplikationer og webbaserede løsninger. Kurset sigter mod at gøre de studerende i stand til kritisk at vurdere data og analytiske resultater i en softwarekontekst.

## Viden

Efter kurset har den studerende grundlæggende viden om:

- Grundlæggende sandsynlighedsteori og stokastiske variable, herunder middelværdi og varians
- De vigtigste diskrete og kontinuerte sandsynlighedsfordelinger (fx binomial-, Poisson- og normalfordelingen) og deres anvendelser
- Deskriptiv statistik og datavisualiseringsteknikker
- Principperne bag estimation (punkttestimation og konfidensintervaller) samt hypotesetest (nul- og alternativ hypotese, p-værdi, signifikansniveau, type I og II fejl)
- Simpel lineær regression og residualanalyse
- Hvordan statistiske metoder kan anvendes på data hentet fra databaser

## Færdigheder

Efter kurset kan den studerende:

- Hente data fra databaser og udføre relevant databehandling og analyse
- Anvende relevant software (fx Python med Pandal, NumPy, SciPy, Matplotlib, scikit-learn) til statistiske beregninger og dataanalyse
- Beregne og fortolke deskriptiv statistik og visualisere data ved hjælp af histogrammer, scatterplots og boxplots
- Opstille, udføre og fortolke statistiske hypotesetest (fx t-test, chi-i-anden test) og konfidensintervaller
- Udføre og fortolke simple lineære regressioner og residualanalyse
- Bearbejde og strukturere resultaterne af statistiske analyser, så de er egnede til integration med softwareapplikationer, herunder klargøring til visning i fx webbaserede frontends (f.eks. via API'er eller filformater).
- Programmatisk hente, filtrere og manipulere data fra en database

## Kompetencer

Efter gennemført kursus har den studerende kompetencer til at:

- Selvstændigt designe og implementere en dataanalyse-pipeline, der involverer datahentning fra en database, statistisk analyse og klargøring af resultater til frontend-visning
- Kritisk vurdere og vælge passende statistiske metoder til analyse af data relateret til typiske softwareteknologiske problemstillinger
- Kommunikere resultaterne af statistiske analyser klart og effektivt til både tekniske og ikke-tekniske interessenter, herunder forståelse af metodernes antagelser, begrænsninger og usikkerhed.

## Undervisningsform og aktiviteter

Undervisningen består af en variation mellem asynkrone videoforelæsninger, online diskussioner af emnerne samt øvelsestimer, hvor de studerende arbejder med opgaver.

## **Eksamen**

Eksamensforudsætning:

Ingen

Eksamensform:

Eksamen er skriftlig (4 timer) og gennemføres ved fysisk fremmøde på Campus Horsens.

Eksamen består af to dele:

- Del 1: En Flowlock-eksamen i WISEflow, 3 timer.
- Del 2: En WISEflow-eksamen uden FLOWlock, 1 time.

Del 2 vil først være tilgængelig, når Del 1 er afsluttet og afleveret.

Den endelige karakter er baseret på en samlet vurdering af del 1 og del 2.

Ekstern bedømmelse.

Tilladte værktøjer:

I Del 1 må studerende medbringe alle egne noter, bøger og andet skriftligt eller trykt materiale. Studerende må også tilgå PDF-filer gemt lokalt på deres bærbare computer og anvende en personlig lommeregner. Internetadgang er ikke tilladt, og eksamen afvikles i Flowlock-tilstand.

I Del 2 er alle hjælpemidler tilladt, herunder brug af computer til beregninger. Dog er internetadgang strengt forbudt, og brug af generative AI-værktøjer som CoPilot, ChatGPT, Bing, Gemini eller lignende tjenester er ikke tilladt.

Enhver form for kommunikation under eksamen er strengt forbudt og vil medføre bortvisning af alle involverede parter fra eksamen.

Reeksamen:

Reeksamen kan være mundtlig.

## **Karakterbeskrivelse**

Karakter efter 7-trinsskalaen.

## **Kursusansvarlig**

Richard Brooks (rib)

## **Gældende fra**

01-02-2026

## **Kurstype**

Obligatorisk for Softwareingeniør Trainee, 2. semester