

Informasjon og oversikt for R-Seminarene

Hei og velkommen til R-seminar i STV1020!

Her kommer litt informasjon om R-seminarene. R-seminarene løper over seks ganger, og etterfølges av en digital prøve 5. mai. Den må bestås for å kunne gå opp til eksamen. Dersom du ikke består ved første forsøk, får du et nytt forsøk 19. mai. Prøven bygger på det vi gjennomgår i seminarene, og er ment som en test for om dere har fått med dere grunnleggende R-kunnskaper. På grunn av koronasituasjonen, vil det ikke være mulig å bruke UiO-maskinene under seminarene. Ta derfor med egen PC ved fysisk seminar.

Før første seminar må dere ha lastet ned R (programmeringsspråket) og RStudio (programmet vi bruker for å skrive programmeringsspråket), og sett at det fungerer.

Før seminaret forventes det at:

- Dere har lest nedlastingsguiden til R og introduksjonen til R (finnes på GitHub-profilen min)
- Dere har lastet ned R og RStudio, se nedlastingsguiden (send meg epost på sarakg@student.sv.uio.no dersom du får problemer med nedlastningen)

Under seminargangen oppfordres det til at:

- Dere møter opp
- Dere jobber med oppgaver

Ved første øyekast kan R virke litt skummelt, men så lenge dere følger med i seminarene og legger ned innsats underveis, burde det gå helt fint!

Boka R for Data Science er nyttig å ta en titt i, også underveis. Den finnes i gratisversjon på nett: <https://r4ds.had.co.nz/>

Oversikt over R-seminargangene

1. **Introduksjon til R:** Introdusere R og RStudio, bli kjent med mappestruktur og filformat, bruk av R som kalkulator, opprette objekter, introduksjon til klasser, bruk av hjelpefiler
2. **Introduksjon til R:** Repetisjon av mappestruktur, laste inn datasett (ulike typer), pakker i R, utforsking av data, introduksjon til plotting
3. **Plotting:** Laste inn data, omkoding av variabler, subsetting i R, mer om plotting
4. **Bivariat analyse:** Laste inn data, missing (NA), statistiske mål, univariat og bivariat analyse
5. **Bivariat regresjon:** Laste inn data, bivariat regresjonsanalyse, printe resultater i Stargazer, plote og tolke resultater
6. **Multipel regresjon:** Laste inn data, multipel regresjonsanalyse, printe resultater, plote og tolke resultater

Kort oversikt over nyttige begreper

- Funksjoner: En funksjon er en ferdig kodesnutt som er laget for å utføre spesifikke oppgaver.
- Pakker: Funksjoner kan pakkes sammen og deles som egne tilleggsmoduler. Disse kaller vi for pakker. Man finner funksjoner i pakker.
- Objekter: Objekter er virtuelle "bokser" hvor vi kan legge informasjon til senere bruk.
- Vektorer: En vektor er en serie med tall eller bokstaver/ord som er lagret i et objekt.
- Argumenter: Er tilleggsinformasjon vi kan legge til i funksjoner, som for eksempel 'color' dersom vi vil at histogrammet vårt skal ha en spesiell farge.