

# Seminar 6: oppgaver

Lise Rødland

April 26, 2021

Datasettet dere skal jobbe med i dag inneholder blant annet følgende variabler:

- grspnum – Inntekt i kroner
- eduyrs – Antall år med fullført utdanning
- gndr - 1 = mann, 2 = kvinne
- agea - Alder

## Oppgaver

1. Last inn datasettet ESS9. Dette er en versjon av spørreundersøkelsen European Social Survey, runde 9 (fra 2018) med tall fra Norge. Du finner datasettet i seminar-modulen på Canvas. Har du problemer med å laste inn datasettet så slå opp i dokumentet “Guide for innlastning av datasett” som ligger på Canvas for å finne riktig funksjon.
  2. Skap et oversiktlig datasett det er lett å jobbe med:
    - Subset datasettet/opprett et nytt objekt hvor du kun inkluderer de variablene som skal inngå i regresjonsanalysen din.
    - Skift navn på variablene «agea», «eduyrs» og «grspnum» til noe du synes virker mer intuitivt.
    - Omkod variabelen «gndr» (kjønn) slik at mann = 0 og kvinne = 1.
    - Klarer du å gjøre dette i samme kodesekvens? (Tips: bruk %>% mellom hver funksjon).
  3. Fjern enheter som har missing på en eller flere av de relevante variablene i det nye datasettet. Hvor mange observasjoner sitter du igjen med? Hva skjer dersom du fjerner enheter med missing fra det opprinnelige datasettet?
  4. Multippel regresjon:
    - Estimer en lineær OLS-regresjonsmodell med inntekt som avhengig variabel og utdanning, alder og kjønn som uavhengige variabler.
    - Presenter resultatene i en fin tabell ved hjelp av «stargazer» funksjonen. Legg inn tittel og navn på avhengig og uavhengige variabler.
    - Oppgi og tolk koeffisientene til utdanning, kjønn og alder.
- \*\* BONUSOPPGAVER \*\*** 5. Plott regresjonslinjen til en modell med utdanning som uavhengig variabel og inntekt som avhengig variabel og gjør kort rede for hva plottet viser. 6. Sjekk og beskriv formen til residual-fordelingen, enten med et tetthets-plott eller et histogram.