

Instruksjoner

- Prøven skal besvares med et fungerende R-script som lastes opp i innleveringsmappen på Fronter.
- Scriptet skal inneholde nødvendig kode for å besvare oppgavene samt kommentarer markert med `#` som forklarer fremgangsmåten dere har valgt. Der oppgavene ber dere oppgi bestemte verdier eller tolkninger skal disse også oppgis som kommentarer i scriptet.
- Sørg for at koden er oversiktlig. For å skille oppgavene fra hverandre, anbefales overskrifter av typen:
 `### Oppgave 1 #####` i scriptet
 `# Oppgave 1:`
 eller lignende.
- Lykke til!

Oppgaver

1. Last inn datasettet **tillit.csv**. Enhetene i datasettet er respondenter til en spørreundersøkelse. Du kan enten laste inn data direkte fra linken:
<http://folk.uio.no/martigso/oof/tillit.csv>
eller laste ned fra Fronter og inn i R fra lokalt på din data.
2. Lag et boxplot med **stemt** på x-aksen og **tillit_politikere** på y-aksen. Hvilken av kategoriene på **stemt** har høyest median?
3. Kod om de som **ikke** hadde stemmerett på variabelen **stemt** til NA. Kontroller at det ble riktig.
4. Vis hvordan du finner antall missingverdier på variabelen **tillit_politikere**. Kommenter hvor mange det er.
5. Lag et plot av fordelingen på variabelen **tillit_politikere**. Basert på plotet, er det grunn til å argumentere for at variabelen er normalfordelt?
6. Lag en korrelasjonsmatrise med variablene **tillit_politikere**, **lykke**, og **inntekt_husholdning**. Bruk listwise deletion av missingverdier. Signifikanstest og tolk én av korrelasjonene.
7. Plot et spredningsdiagram med **lykke** på x-aksen og **tillit_politikere** på y-aksen. Tegn en bivariat regresjonslinje mellom y og x. Tolk sammenhengen substansielt.
8. Estimér en OLS-modell med **tillit_politikere** som avhengig variabel, og **lykke**, **kjonn** (som kategorisk) og **inntekt_husholdning** som uavhengige variabler. Er sammenhengen fra oppgave 7 signifikant forskjellig fra 0?
9. Lag en variabel med restleddene til enhetene i datasettet basert på OLS-modellen i oppgave 7. Kontroller om restleddene er normalfordelt, og kommenter konsekvensene for modellen basert på det du finner.

Variabelforklaringer:

tillit_politikere Tillit til politikere, 1=Lav 10=Høy.

stemt Stemte respondenten ved forrige valg? ()

lykke Hvor lykkelig er respondenten? 1=Lav 10=Høy

kjonn Kjønn

inntekt_husholdning Husholdningsinntekt, 1=Lav (1. desil) 10=Høy (10. desil)