

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 **ministers.csv**. Enhetene i datasettet er norske ministre fra 1945 - 2013. Du kan enten laste inn data direkte fra linken:
<http://folk.uio.no/martigso/leeep/ministers.csv>
eller laste ned fra Fronter og inn i R fra lokalt på din data.
2. Lag et boxplot med **parti** på x-aksen og **varighet** på y-aksen. Hvilket parti har høyest varighet på sin median-minister?
3. Variablene **start_aar** og **fodselsaar** viser året ministeren ble gitt ministerposten og året ministeren ble født. Bruk disse variablene til å lage en ny variabel – **alder** – som indikerer alder ved overtakelse av ministerposten. Kontroller at variabelen ble riktig opprettet.
4. Vis hvordan du finner hvor mange ministre som hadde verdien 1 på variabelen **maa_gaa**. Oppgi svaret i en kommentar.
5. Lag et plot av fordelingen på variabelen **varighet**. Basert på plotet, er det grunn til å argumentere for at variabelen er normalfordelt?
6. Lag en korrelasjonsmatrise med variablene **alder**, **varighet**, og **maa_gaa**. Signifikans-test og tolk én av korrelasjonene substansielt.
7. Plot et spredningsdiagram med **maa_gaa** på x-aksen og **varighet** på y-aksen. Tegn en bivariat regresjonslinje mellom y og x. Tolk sammenhengen substansielt.
8. Estimér en OLS-modell med **varighet** som avhengig variabel, og **maa_gaa**, **kjonn** (som kategorisk) og **alder** 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 oppgaven over. Kontroller om restleddene er normalfordelt, og kommenter konsekvensene for modellen basert på det du finner.

Variabelforklaringer:

id Unik id for individuelle ministre

parti Partitilhørighet

start_aar Årstall for start på ministerposten

kjonn Ministerens kjønn

fodselsaar Årstall ministeren ble født

varighet Hvor lenge ministeren satt i ministerposten

maa_gaa Antall ganger noen krevde ministerens avgang