Seminaropppgave 1

Torkel

2022-10-05

# Oppgave 1

Det er 15 varibler og 1816 observasjoner.

Variablene heter height, weight, male, earn, earnk, ethnicity, education, mother\_education, father\_education, walk, exercise, smokenow, tense, angry, age

Kan sees gjennom ***names()***

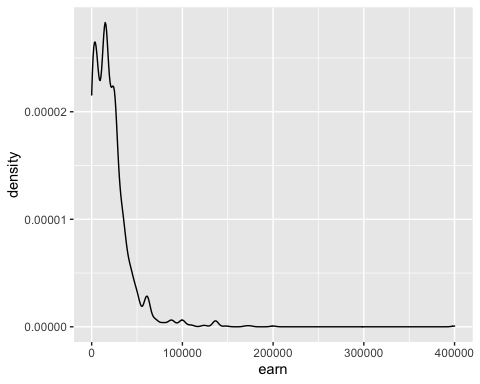
names(earnings)

## [1] "height" "weight" "male" "earn"   
## [5] "earnk" "ethnicity" "education" "mother\_education"  
## [9] "father\_education" "walk" "exercise" "smokenow"   
## [13] "tense" "angry" "age"

# Oppgave 2

Lager et densityplot med ***ggplot*** og ***geom\_density***

options(scipen=999)  
earn <- earnings$earn  
earn.density <- ggplot(data = earnings) + geom\_density(aes(x=earn) )  
show(earn.density)



# Oppgave 2.2

Har vurdert at det ikke er nødvendig

# Oppgave 3

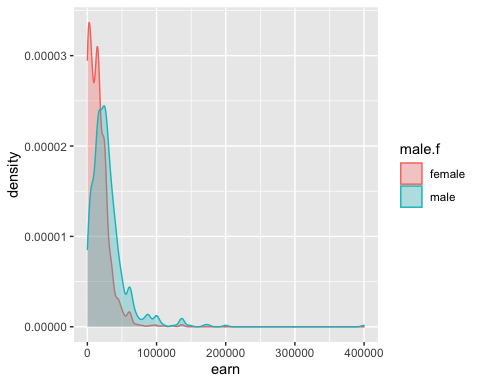
Lager en ny faktorvariabel ***male.f*** fra variabelen ***male*** i ***earnings*** hvor 1 = male og 0 er female

earnings$male.f <- factor(earnings$male, labels = c("female", "male"))

# Oppgave 3b

Her er det nye plottet, basert på de to kjønnene. Man ser at det er flere kvinner med lav inntekt, og flere menn med høy inntekt.

earn.density2 <- ggplot(earnings, aes(earn, color=male.f, fill=male.f)) + geom\_density(alpha = 0.3)  
show(earn.density2)



# Oppage 3c

Plottet viser hvor mange som har hva slags lønn, og sammenlikner menn og kvinner, med en strek for medianlønnen i datasettet. Får ikke til å lage median gitt kjønn :(

earn.density2 + geom\_vline(xintercept = median(earn), col = "red", lwd = 0.3)

