

Quiz repetisjon - 8. april

1. I lavaan så får vi et datasett, *HolzingerSwineford1939* Hvor mange variabler og observasjoner inneholder datasettet?

```
?HolzingerSwineford1939
dim(HolzingerSwineford1939)
HS <- HolzingerSwineford1939# short name
```

2. Hvor mange skoler kommer elevene fra?
3. Lag en aldervariabel *age* ved å kombinere *ageyr* og *agemo* i en formel

```
HS$age <- HS$ageyr+...?
```

3. Hva er median alder?
4. Lag et 90% konfidensintervall for alder

```
t.test(HS$age)
```

5. Utfør en hypotesetest om det er forskjell i alder mellom skolene. Bruk signifikansnivå 1% og konkluder i et lettfattelig språk
6. Lag et histogram for *x9*. Er denne variabelen tilnærmet normalfordelt?
7. Er det noe missing i dette datasettet? Hvilken ID har missing, eventuelt?

```
naniar::vis_miss(HS)
```

7. Lag en summeskår for testene *x1* - *x9*

```
HS$sum <- rowSums(select(HS, paste0("x", 1:9)))
```

Lag en graf med to boksplott som sammenlikner summeskårene for de to kjønnene. Ser det ut til at det er noen forskjell?

8. Utfør en hypotesetest om det noen forskjell i kjønnenes gjennomsnittskår. Bruk signifikansnivå 5% og konkluder lettfattelig.

```
t.test(HS$sum~HS$sex)
```

9. Kjør en multippel regresjon med summeskår som avhengig variabel, og kjønn, age, school og grade som prediktorer. Hva blir forklaringskraften?

```
model <- lm(sum ~ sex+age+school+grade, data=HS)
```

10. Sjekk forutsetningene. Er noen brutt?

```
performance::check_model(model)
```

11. Vi skal nå se på faktorstrukturen til de 9 prøvene x1-x9 Les hva itemene sier i ?Holzinger-Swineford1939. Hvor mange faktorer tror du det er? Gjør en parallelanalyse for antall faktorer for å se om det stemmer.

```
fa.parallel(select(HS, paste0("x", 1:9)))
```

12. Gjør en EFA for å se hvilke items som lader på hvilken av faktorene

```
f <- fa(select(HS, paste0("x", 1:9)),nfactors=?? )  
f$loadings
```

Grupper indikatorene x1-x9 slik at de utgjør hver sin faktor.