

Et praktisk eksempel

En ting er å lese om R, men en annen ting er å gjøre det i praksis. La oss jobbe gjennom et konkret, enkelt eksempel og kommentert grundig hva vi gjorde slik at den som kommer etter oss (våre kolleger eller oss sjøl i framtida) forstår hva vi gjør.

Oppgavebeskrivelse

Vi vil laste inn to SPSS-datasett. Det ene ligger i en undermappe (relativt til skriptet) kalt *data*. Det andre ligger godt hjem langt, langt vekk et sted på *C*-disken min. Vi vil laste inn begge to og se hvordan det de ser ut.

```
# For å laste inn SPSS-filer trenger vi en pakke som gjør det. Haven er bra.
# Dersom du ikke har haven fra før må den installeres. En pakke trenger bare
# installeres én gang. Siden jeg har pakka fra før har
# jeg kommentert vekk neste kode for. Skal du installere koden, fjern
# emneknaggen.

# install.packages("haven")

# Så må vi laste inn pakka for å kunne ta dens funksjoner i bruk. Dette må vi
# gjøre hver gang vi starter en ny sesjon.

library(haven)

# Vi laster inn den første SPSS-fila. Vi gir den navnet atferd. Siden den
# ligger på mappa data må vi spesifisere dette når vi oppgir hvor den ligger og
# hva den heter. Husk også filendelsen. Noen operativsystem er glad i skjule
# fil-endelsen, men den er en viktig del av alle filers navn.
atferd <- read_sav(file = "data/behavior.sav")

# Den andre fila ligger langt vekk på C. For å gjøre det litt enklere for meg
# sjøl vil jeg lagre filstien (og navnet) i en vektor. Jeg gjør ofte dette om
# filstien er lang, for å ikke gjøre import-funksjonen så lang.
# Husk dette med skråstreker: Vi må enten bruke \\ eller /, ikke en enkelt \
filsti <- "C:\\Program Files\\IBM\\SPSS\\Statistics\\26\\Samples\\English\\accidents.sav"

# Så laster vi inn den andre fila.
ulykker <- read_sav(file = filsti)

# Dette er altså det samme som å skrive
```

```
# ulykker <- read_sav(file = "C:\\Program Files\\IBM\\SPSS\\Statistics\\26\\Samples\\Engli
# Siden vi definerte filsti tidligere.
```

```
# Nå har vi fått de to datasetta våre. Hvordan ser de ut?
```

```
summary(atferd)
```

ROWID	Run	Talk	Kiss	Write
Min. : 1.0	Min. :1.060	Min. :0.420	Min. :0.270	Min. :0.710
1st Qu.: 4.5	1st Qu.:4.300	1st Qu.:0.560	1st Qu.:2.625	1st Qu.:2.810
Median : 8.0	Median :6.440	Median :0.920	Median :4.080	Median :4.440
Mean : 8.0	Mean :5.539	Mean :1.523	Mean :3.996	Mean :4.166
3rd Qu.:11.5	3rd Qu.:7.050	3rd Qu.:1.675	3rd Qu.:5.460	3rd Qu.:5.790
Max. :15.0	Max. :7.620	Max. :5.710	Max. :7.920	Max. :6.420

Eat	Sleep	Mumble	Read
Min. :0.560	Min. :0.150	Min. :1.330	Min. :0.420
1st Qu.:1.135	1st Qu.:4.145	1st Qu.:3.685	1st Qu.:1.780
Median :1.810	Median :6.020	Median :3.960	Median :4.290
Mean :3.148	Mean :5.311	Mean :4.431	Mean :3.905
3rd Qu.:4.470	3rd Qu.:6.970	3rd Qu.:5.430	3rd Qu.:5.365
Max. :7.620	Max. :8.250	Max. :7.690	Max. :7.270

Fight	Belch	Argue	Jump
Min. :4.750	Min. :2.190	Min. :1.480	Min. :1.580
1st Qu.:6.780	1st Qu.:4.500	1st Qu.:4.070	1st Qu.:4.500
Median :7.330	Median :6.420	Median :4.830	Median :5.460
Mean :7.035	Mean :5.731	Mean :5.027	Mean :5.265
3rd Qu.:7.585	3rd Qu.:6.810	3rd Qu.:6.085	3rd Qu.:6.795
Max. :8.380	Max. :7.790	Max. :7.290	Max. :7.520

Cry	Laugh	Shout
Min. :1.000	Min. :0.770	Min. :1.060
1st Qu.:4.450	1st Qu.:1.030	1st Qu.:4.495
Median :5.520	Median :1.600	Median :5.480
Mean :4.999	Mean :1.993	Mean :5.317
3rd Qu.:5.895	3rd Qu.:2.500	3rd Qu.:7.050
Max. :7.630	Max. :6.400	Max. :7.670

```
summary(ulykker)
```

agecat	gender	accid	pop
Min. :1.00	Min. :0.0	Min. :54123	Min. :187791

1st Qu.:1.25	1st Qu.:0.0	1st Qu.:57334	1st Qu.:196416
Median :2.00	Median :0.5	Median :60967	Median :199633
Mean :2.00	Mean :0.5	Mean :60801	Mean :199035
3rd Qu.:2.75	3rd Qu.:1.0	3rd Qu.:64610	3rd Qu.:202586
Max. :3.00	Max. :1.0	Max. :66804	Max. :208239

Datasetta er forøvrig [lånt fra SPSS](#). Du trenger ikke bruke disse datasetta i din egen gjennomgang, bare finn to andre. Putt dem gjerne på forskjellige plasser for å øve på å laste dem inn fra ulik lokasjon. Hvis du putter filstien i en vektor slik jeg gjorde i det ene eksemplet, husk at du enten må putte begge filstiene i ulike objekter, eller at du må laste inn den *første* fila før du overskriver filstia med den *andre* filas filsti.

```
# Dette funker
filsti_atferd <- "data/behavior.sav"
filsti_ulykker <- "C:\\Program Files\\IBM\\SPSS\\Statistics\\26\\Samples\\English\\accidents.sav"

atferd <- read_sav(file = filsti_atferd)
atferd <- read_sav(file = filsti_ulykker)

# Dette funker også, men du må huske på å aldri endre rekkefølgen.
filsti <- "data/behavior.sav"
atferd <- read_sav(file = filsti)

filsti <- "C:\\Program Files\\IBM\\SPSS\\Statistics\\26\\Samples\\English\\accidents.sav"
ulykker <- read_sav(file = filsti)

# Dette vil laste inn det samme datasettet to ganger, og dem ulike navn.
# Begge vil være ulykker, hvis filsti blei definert sist.
filsti <- "data/behavior.sav"
filsti <- "C:\\Program Files\\IBM\\SPSS\\Statistics\\26\\Samples\\English\\accidents.sav"

atferd <- read_sav(file = filsti)
ulykker <- read_sav(file = filsti)
```