

SOK-2008-2022-oppgave4.R

OddVi

2022-10-14

```
# Datasettet inneholder informasjon om generøsiteten av forldrepermisjonsordninger  
# og mødres yrkesdeltakelse i 24 OECD-land i 2021  
# Foreldrepermisjonene varierer mellom land.  
# For å gjøre ulike landers rettigheter sammenlignbare presenteres permisjonsordningers gunstighet  
# ved hjelp av "full rate equivalent":  
# Uker med 100% støtte = Foreldrepermisjonens varighet i uker * utbetalingssats  
# (i prosent av gjennomsnittlig inntekt)
```

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.1.3
```

```
## -- Attaching packages ----- tidyverse 1.3.2 --
```

```
## v ggplot2 3.3.5      v purrr  0.3.4  
## v tibble  3.1.6      v dplyr  1.0.7  
## v tidyr   1.1.4      v stringr 1.4.0  
## v readr   2.1.1      v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag()     masks stats::lag()
```

```
library(ggplot2)  
women<-read.csv2("women.csv")
```

```
women$tot_full_rate<-as.numeric(women$tot_full_rate)  
women$fem_emp_rate_0_2<-as.numeric(women$fem_emp_rate_0_2)  
women$fem_emp_rate_6_14<-as.numeric(women$fem_emp_rate_6_14)
```

```
women %>%  
  ggplot(aes(x=tot_full_rate,y=fem_emp_rate_0_2))+  
  geom_point()+  
  ylim(0, 100)+  
  labs(x = "Uker med 100% støtte", y = "Yrkesdeltakelse blant mødre hvis yngste barn er 0-2 år")+  
  geom_smooth(method=lm, se=FALSE) -> kids_0_2
```

```
women %>%
  ggplot(aes(x=tot_full_rate,y=fem_emp_rate_6_14))+
  geom_point()+
  ylim(0, 100)+
  labs(x = "Uker med 100% støtte", y = "Yrkesdeltakelse blant mødre hvis yngste barn er 6-14 år")+
  geom_smooth(method=lm, se=FALSE) -> kids_6_14

library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'

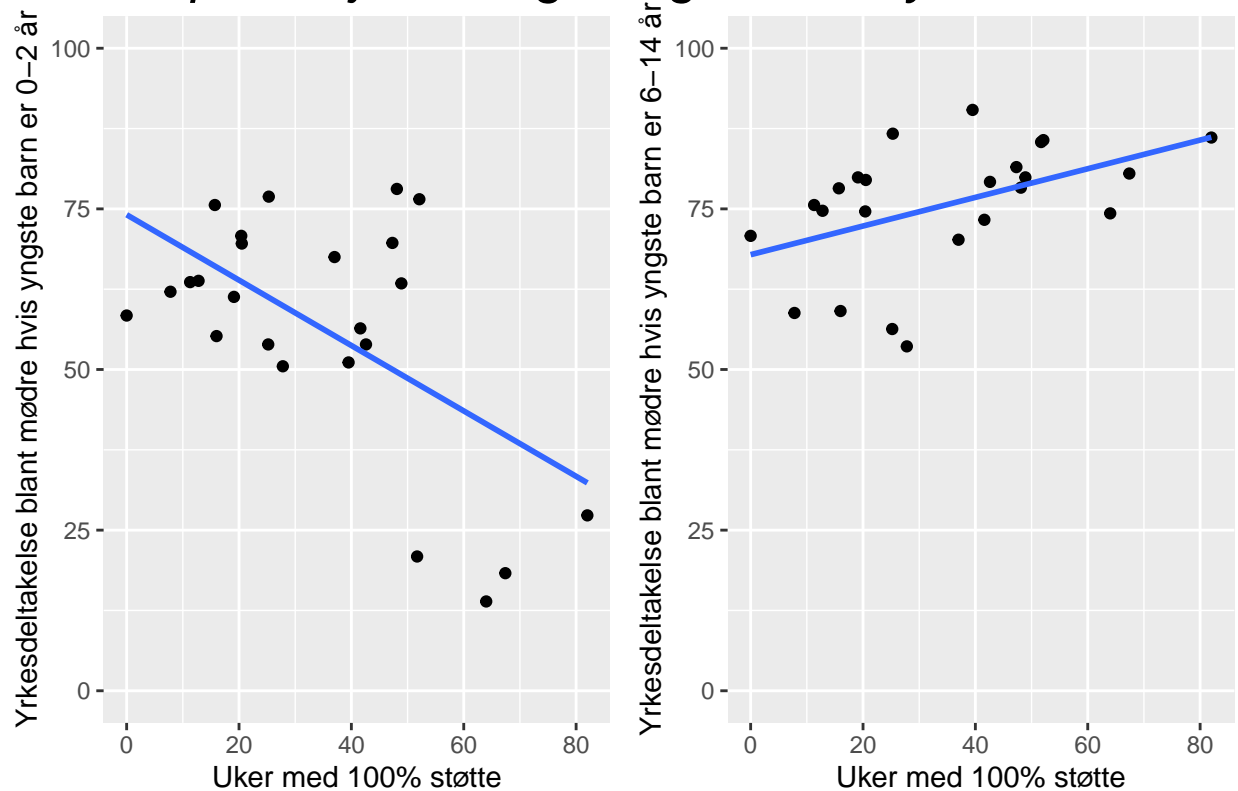
## The following object is masked from 'package:dplyr':
##
## combine
```

```
library(grid)
grid.arrange(kids_0_2,kids_6_14, nrow = 1, top = textGrob("Sammenhengen mellom foreldrepermisjons lengde
```

```
## 'geom_smooth()' using formula 'y ~ x'
```

```
## 'geom_smooth()' using formula 'y ~ x'
```

foreldrepermisjons lengde og mødres yrkesdeltakelse



#Utfordring 4.2.2

Sammenhenger

```
yrkesdeltakelse_under_2 <-lm(fem_emp_rate_0_2 ~ tot_full_rate , data = women)
summary(yrkesdeltakelse_under_2)
```

```
##
## Call:
## lm(formula = fem_emp_rate_0_2 ~ tot_full_rate, data = women)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -27.618  -8.372  -2.976   10.188   28.926
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    74.0856     6.2761   11.804 5.45e-11 ***
## tot_full_rate  -0.5089     0.1572   -3.238 0.00378 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15.7 on 22 degrees of freedom
## Multiple R-squared:  0.3227, Adjusted R-squared:  0.2919
## F-statistic: 10.48 on 1 and 22 DF,  p-value: 0.00378
```

```
yrkesdeltakelse_6_14 <-lm(fem_emp_rate_6_14 ~ tot_full_rate , data = women)
summary(yrkesdeltakelse_6_14)
```

```
##
## Call:
## lm(formula = fem_emp_rate_6_14 ~ tot_full_rate, data = women)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -20.468  -4.363   2.007   6.057  13.724
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   67.87049     3.58134   18.951 4.1e-15 ***
## tot_full_rate  0.22292     0.08968    2.486  0.021 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.956 on 22 degrees of freedom
## Multiple R-squared:  0.2193, Adjusted R-squared:  0.1838
## F-statistic: 6.178 on 1 and 22 DF,  p-value: 0.02102
```

Hva ser vi?

*# For de med barn under 2 år faller yrkesdeltakelsen med økende antall uker med full støtte.
Vi ser av figuren at det er stor spredning i dataene.
En test for lineær sammenheng viser justert R^2 lik 0.29 for denne gruppen som innebærer*

en svak sammenheng.

For de med barn i alderen 6-14 år ser vi en stigende yrkesdeltakende basert på datagrunnlaget.

Vi ser også her at det er stor spredning i data.

En test for lineær sammenheng viser justert R^2 lik 0.18 for denne gruppen som innebærer

en meget svak sammenheng.