

```

library(MASS)
library(tidyverse)
library(GGally)
library(labelled)
library(dplyr)
library(rstatix)
library(ivreg)
library(simstudy)
library(ivmodel)
library(ggplot2)
library(car)
library(lmtest)
library(tseries)
library(ggfortify)
library(plotly)
library(haven)
library(broom)

setwd("C:\\Users\\smjofess\\Documents\\Uni Dokumente\\Labour Economics PS2")

# load data
minwage <- read_dta("minwage.dta")

# swap the "state" variable with the state names (not a must, but makes it clearer)
minwage$state <- recode(minwage$state, "0='PA'; 1='NJ'")

# select data that is available in both survey waves
df <- minwage %>%
  filter(sample == 1)

##### a) #####
wages <- df %>%
  group_by(state) %>%
  summarise(wage_st2 = mean(wage_st2),
            wage_st = mean(wage_st),
            dw = mean(dw))

# i)
difference <- wages$wage_st2 - wages$wage_st

# ii)
diff_row <- wages %>%
  filter(state %in% c("NJ", "PA")) %>%
  summarise(wage_st2 = -diff(wage_st2),
            wage_st = -diff(wage_st),
            dw = -diff(dw),
            state = "diff")

wages <- bind_rows(wages, diff_row)
print(wages)

##### b) #####
fte <- df %>%
  group_by(state) %>%
  summarise(fte2 = mean(fte2),
            fte = mean(fte),
            dfte = mean(dfte))

# i)
difference_fte <- fte$fte2 - fte$fte

# ii)
diff_row_fte <- fte %>%
  filter(state %in% c("NJ", "PA")) %>%
  summarise(fte2 = -diff(fte2),
            fte = -diff(fte),
            dfte = -diff(dfte),
            state = "diff")

```

```

fte <- bind_rows(fte, diff_row_fte)
print(fte)

##### c) #####
df$state <- recode(df$state, "'PA'=0; 'NJ'=1")

# ii)
## for wages
reg_wages <- lm(dw ~ state, data = df)
summary(reg_wages)

## for employment
reg_emp <- lm(dfte ~ state, data = df)
summary(reg_emp)

# iii)
## for wages
reg_wages_c <- lm(dw ~ state + co_owned + as.factor(chain), data = df)
summary(reg_wages_c)

## for employment
reg_emp_c <- lm(dfte ~ state + co_owned + as.factor(chain), data = df)
summary(reg_emp_c)

##### d) #####
# only consider restaurants in NJ
nj <- df %>%
  filter(state == 1)

# ii)
## construct variable low_wage, which is 1 if the wage is below 5, and 0 otherwise
nj$low_wage <- ifelse(nj$wage_st < 5, 1, 0)

# for wages
reg_wages_d <- lm(dw ~ low_wage, data = nj)
summary(reg_wages_d)

# for employment
reg_emp_d <- lm(dfte ~ low_wage, data = nj)
summary(reg_emp_d)

##### e) #####
# only consider restaurants in PA
pa <- df %>%
  filter(state == 0)

pa$low_wage <- ifelse(pa$wage_st < 5, 1, 0)

# i)
## for wages
reg_wages_e <- lm(dw ~ low_wage, data = pa)
summary(reg_wages_e)

## for employment
reg_emp_e <- lm(dfte ~ low_wage, data = pa)
summary(reg_emp_e)

# ii)
df$NJ <- df$state
df$low_wage <- ifelse(df$wage_st < 5, 1, 0)

reg_f <- lm(dw ~ NJ + low_wage + NJ * low_wage, data = df)
summary(reg_f)

```

```
reg_g <- lm(dfte ~ NJ + low_wage + NJ * low_wage, data = df)
summary(reg_g)
```