Problem 1

First, create the environment and load packages.

```
rm(list = ls())
library(ggplot2)
library(dplyr)
library(data.table)
library(patchwork)

theme_set(theme_bw())
dir = paste0("/Users/tombearpark/Documents/princeton/1st_year",
             "/ECO517/honore/assignments/pr2/")
```

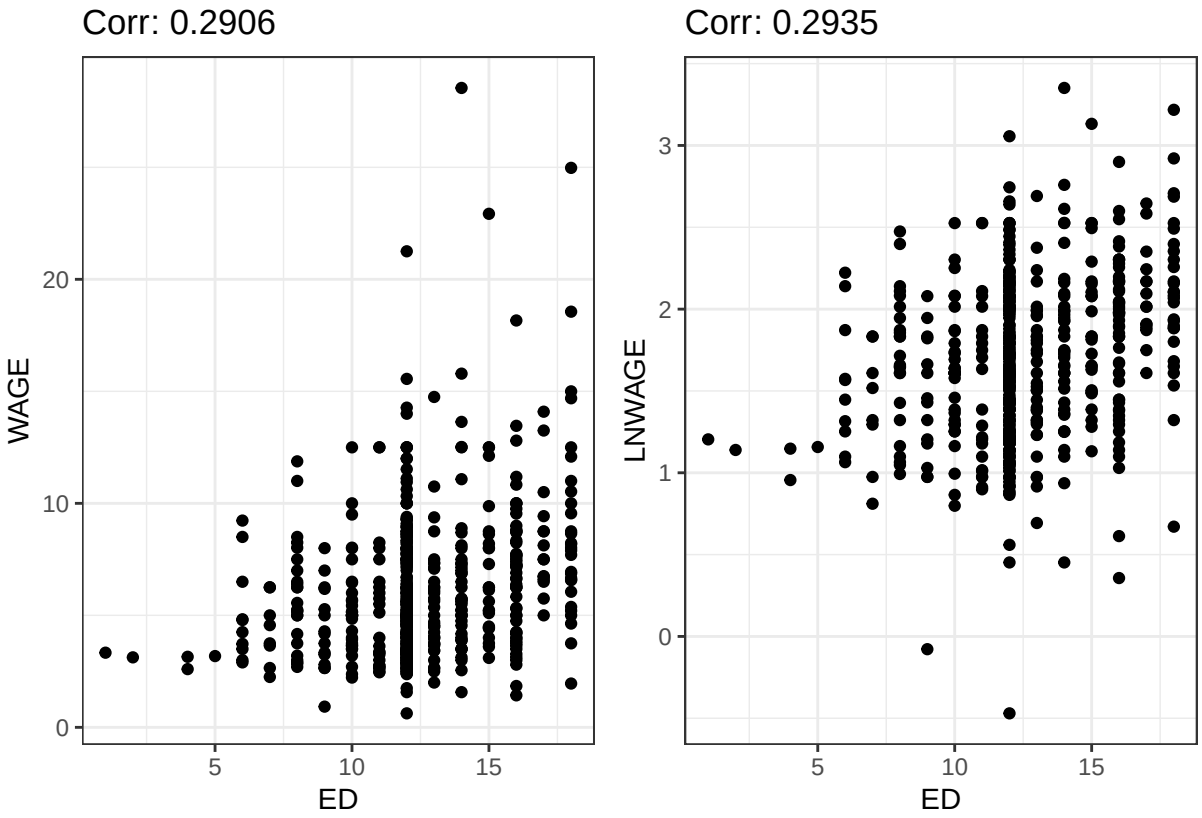
Now, lets define some functions that load in the datasets and plot visualisations

```
get_data = function(dir, name){
  df = fread(paste0(dir, name)) %>%
    as.data.frame() %>%
    rename(LNWAGE = V11,
           ED = V1) %>%
    mutate(WAGE = exp(LNWAGE))
  return(df)
}

# Consider a visualisation
plot_correlations = function(df){

  corr = round(cor(df$WAGE, df$ED), 4)
  corr_ln = round(cor(df$LNWAGE, df$ED), 4)

  ggplot(data = df) +
    geom_point(aes(x = ED, y = WAGE)) +
    ggtitle(paste0("Corr: ", corr)) +
  ggplot(data = df) +
    geom_point(aes(x = ED, y = LNWAGE))+
    ggtitle(paste0("Corr: ", corr_ln))
}
```

Problem 1

Nowe we can run the functions to answer question 1

```
df = get_data(dir = dir, name = "Cps78")
plot_correlations(df)
```



Add a new chunk by clicking the *Insert Chunk* button on the toolbar or by pressing *Cmd+Option+I*.

When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Cmd+Shift+K* to preview the HTML file).

The preview shows you a rendered HTML copy of the contents of the editor. Consequently, unlike *Knit*, *Preview* does not run any R code chunks. Instead, the output of the chunk when it was last run in the editor is displayed.