Econometrics_Problem_Set_1.R

Bernardo

Sun Sep 08 23:24:44 2019

```
library(readstata13)
```

```
## Warning: package 'readstata13' was built under R version 3.4.4
```

```
library(plyr)
```

```
## Warning: package 'plyr' was built under R version 3.4.4
```

```
## create dataset
cpsdata <- read.dta13("cps09mar.dta")

## compute hourly wages
cpsdata$hourly_wage <- cpsdata$earnings / cpsdata$hours / cpsdata$week
head(cpsdata)
```

```
##   age female hisp education earnings hours week union uncov region race
## 1  52      0    0        12   146000    45   52     0     0      1    1
## 2  38      0    0        18    50000    45   52     0     0      1    1
## 3  38      0    0        14    32000    40   51     0     0      1    1
## 4  41      1    0        13    47000    40   52     0     0      1    1
## 5  42      0    0        13   161525    50   52     1     0      1    1
## 6  66      1    0        13    33000    40   52     0     0      1    1
##   marital hourly_wage
## 1       1    62.39316
## 2       1    21.36752
## 3       1    15.68627
## 4       1    22.59615
## 5       1    62.12500
## 6       5    15.86538
```

```
## convert annual earnings to log annual earnings
logwage <- cpsdata
logwage$earnings <- log(cpsdata$earnings)

## compute mean of log wages conditional on region
logwage_region <- by(data = logwage$earnings, logwage$region, mean)
logwage_region
```

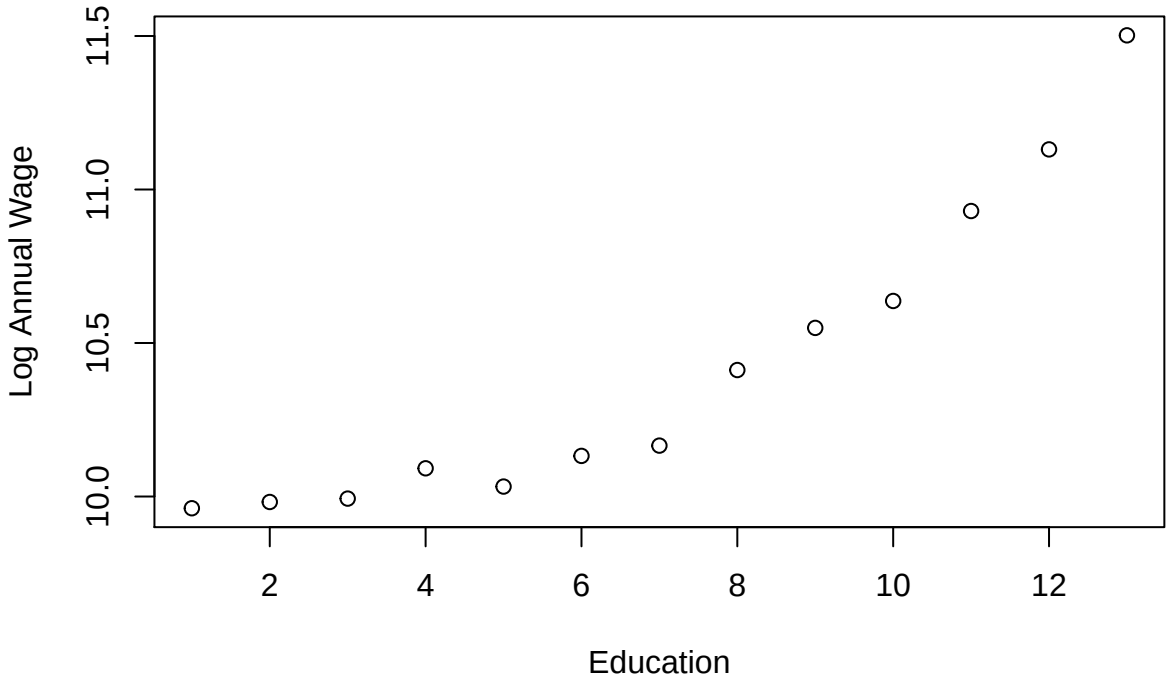
```
## logwage$region: 1
## [1] 10.77313
## ------------------------------------------------------------
## logwage$region: 2
## [1] 10.61838
## ------------------------------------------------------------
## logwage$region: 3
## [1] 10.60997
## ------------------------------------------------------------
## logwage$region: 4
## [1] 10.68738
```

```r
## compute mean of log wages conditional on education
logwage_educ <- by(logwage$earnings, logwage$education, mean)
logwage_educ
```

```
## logwage$education: 0
## [1] 9.961781
## ------------------------------------------------------------
## logwage$education: 4
## [1] 9.982151
## ------------------------------------------------------------
## logwage$education: 6
## [1] 9.993212
## ------------------------------------------------------------
## logwage$education: 8
## [1] 10.09185
## ------------------------------------------------------------
## logwage$education: 9
## [1] 10.03237
## ------------------------------------------------------------
## logwage$education: 10
## [1] 10.13228
## ------------------------------------------------------------
## logwage$education: 11
## [1] 10.16576
## ------------------------------------------------------------
## logwage$education: 12
## [1] 10.41195
## ------------------------------------------------------------
## logwage$education: 13
## [1] 10.54923
## ------------------------------------------------------------
## logwage$education: 14
## [1] 10.63689
## ------------------------------------------------------------
## logwage$education: 16
## [1] 10.9299
## ------------------------------------------------------------
## logwage$education: 18
## [1] 11.13059
## ------------------------------------------------------------
## logwage$education: 20
## [1] 11.5021
```

```r
## create dataframe from logwage_educ
logwage_df <- cbind(logwage_educ)

## plot log annual earnings against education
plot(logwage_df, xlab = "Education", ylab = "Log Annual Wage")
```


Log Annual Wage
10.0
10.5
11.0
11.5
2
4
6
8
10
12
Education