HW2_Microeconometrics_Zhesheng Zheng_15220162202530

I use CGSS2012 data to do an analysis about effect of self-confidence on the income. I choose the difference between "Which tier do you think you will be in 10 years?" and "Which tier do you think you are now?" to measure the self-confidence. And I choose the personal income in 2011 to measure the income.

```r
library(readstata13)
```

```
## Warning: package 'readstata13' was built under R version 3.5.3
```

```r
library(lmtest)
```

```
## Warning: package 'lmtest' was built under R version 3.5.3

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```r
data<-read.dta13('cgss2012_14.dta', nonint.factor=T)
```

```
## Warning in read.dta13("cgss2012_14.dta", nonint.factor = T):
##   s2:
##   Missing factor labels - no labels assigned.
##   Set option generate.factors=T to generate labels.

## Warning in read.dta13("cgss2012_14.dta", nonint.factor = T):
##   a01:
##   Missing factor labels - no labels assigned.
##   Set option generate.factors=T to generate labels.

## Warning in read.dta13("cgss2012_14.dta", nonint.factor = T):
##   a02:
##   Missing factor labels - no labels assigned.
##   Set option generate.factors=T to generate labels.

## Warning in read.dta13("cgss2012_14.dta", nonint.factor = T):
##   a03:
##   Missing factor labels - no labels assigned.
##   Set option generate.factors=T to generate labels.

## Warning in read.dta13("cgss2012_14.dta", nonint.factor = T):
##   a04:
##   Missing factor labels - no labels assigned.
##   Set option generate.factors=T to generate labels.

## Warning in read.dta13("cgss2012_14.dta", nonint.factor = T):
##   a1402:
##   Missing factor labels - no labels assigned.
##   Set option generate.factors=T to generate labels.

## Warning in read.dta13("cgss2012_14.dta", nonint.factor = T):
##   a1403:
```

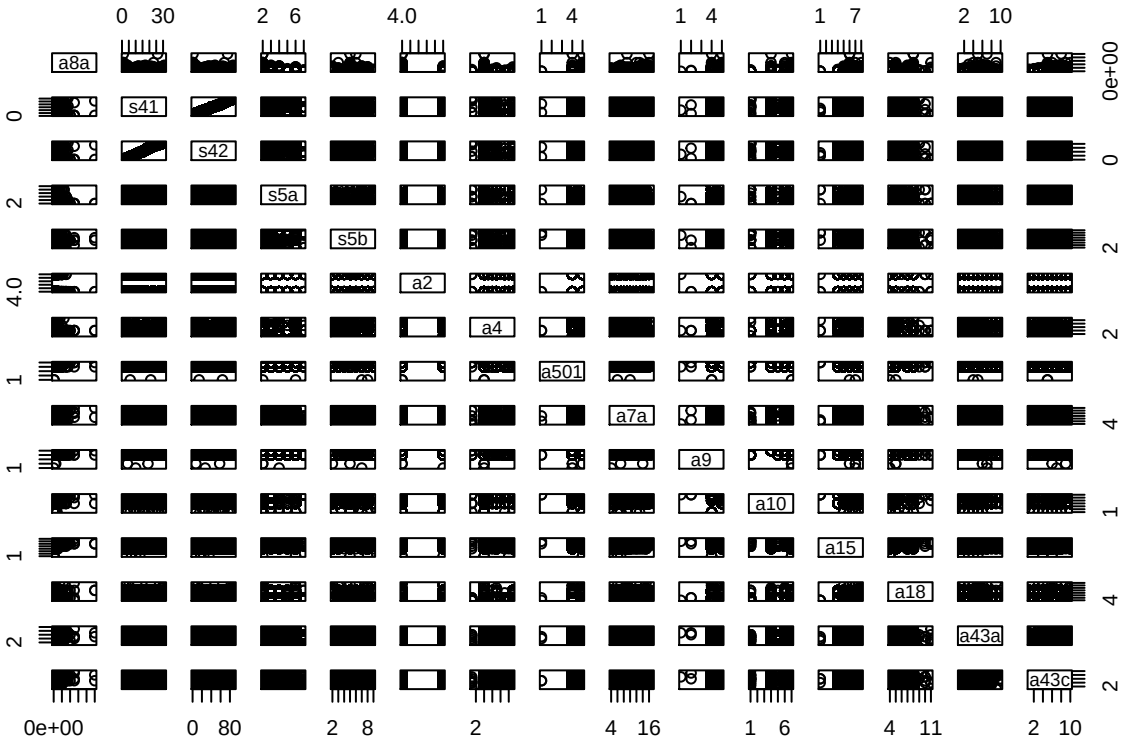
```
##   Duplicated factor levels detected - generating unique labels.

## Warning in read.dta13("cgss2012_14.dta", nonint.factor = T):
##   iscodad:
##   Duplicated factor levels detected - generating unique labels.

## Warning in read.dta13("cgss2012_14.dta", nonint.factor = T):
##   iscodad:
##   Missing factor labels - no labels assigned.
##   Set option generate.factors=T to generate labels.

## Warning in read.dta13("cgss2012_14.dta", nonint.factor = T):
##   iscomom:
##   Duplicated factor levels detected - generating unique labels.
```

```
attach(data)
sgss<-cbind(a8a,s41,s42,s5a,s5b,a2,a4,a501,a7a,a9,a10,a15,a18,a43a,a43c)
detach(data)
sgss[sgss<0]<-NA
sgss[sgss>1000000]<-NA
sgss<-na.omit(sgss)
sgss<-as.data.frame(sgss)
pairs(sgss)
```




```
attach(sgss)
sgss$a43<-sgss$a43c-sgss$a43a
fit<-lm(a8a~s41+s42+s5a+s5b+a2+a4+a501+a7a+a9+a10+a15+a18+sgss$a43)
```

```r
coeftest(fit)
```

```
##
## t test of coefficients:
##
##             Estimate Std. Error  t value  Pr(>|t|)
## (Intercept) 35512.75    9731.36   3.6493 0.0002642 ***
## s41         -1326.06     435.97  -3.0416 0.0023590 **
## s42           349.90     142.84   2.4496 0.0143164 *
## s5a         -2464.14     279.90  -8.8035 < 2.2e-16 ***
## s5b           955.52     197.86   4.8292 1.391e-06 ***
## a2          -8125.64     585.58 -13.8762 < 2.2e-16 ***
## a4           -959.83     190.23  -5.0455 4.599e-07 ***
## a501        -3403.04     831.34  -4.0935 4.282e-05 ***
## a7a          3340.64     119.01  28.0696 < 2.2e-16 ***
## a9            776.58    1264.26   0.6143 0.5390587
## a10           272.62     457.64   0.5957 0.5513796
## a15          1392.17     276.71   5.0312 4.956e-07 ***
## a18          2556.13     298.80   8.5548 < 2.2e-16 ***
## sgss$a43     -146.56     205.18  -0.7143 0.4750801
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The linear regression's result shows that the self-confidence have not effect on the personal income.