Empirical Exercises 4.2

Zheng Tian

Contents

This file include answers and R codes for completing Empirical Exercise 4.2 in Introduction to Econometrics (3rd edition) by Stock and Watson.

1 Reading the Data

The first step is to read the data file into R. The data files for this problem are `TeachingRatings.dta` and `TeachingRatings.xls`, accompanied by a descriptive file `TeachingRatings_Description.pdf`.

- Read the STATA file

  ```
  library(foreign)
  teachingdata <- read.dta("TeachingRatings.dta")
  ```

- Upon reading the data, we can take a glimpse on the data.
  - Use `head` or `tail` to look at the first or last few observations

    ```
    head(teachingdata)
    ```

2 Summary Statistics

We get the summary statistics of the variables used in the analysis, which is `course_eval` and `beauty`

```
df <- teachingdata[c("course_eval", "beauty")]
sumdf <- summary(df); sumdf
```

```
 course_eval        beauty
Min.   :2.100   Min.   :-1.45049
1st Qu.:3.600   1st Qu.:-0.65627
Median :4.000   Median :-0.06801
Mean   :3.998   Mean   : 0.00000
3rd Qu.:4.400   3rd Qu.: 0.54560
Max.   :5.000   Max.   : 1.97002
```

First, try to use `xtable` to generate the table in html

```
library(xtable)
print(xtable(sumdf), type = "html")
```

We can create a table that looks professional using the following code.

```
library(stargazer)
stargazer(df, type = "html",
  title = "Summary Statistics", label = "tab:sum-stats")
```