E4.2-Peilin Guo, Jia Huang, Hao Wang, Tianyong Dong

Peilin Guo, Jia Huang, Hao Wang, Tianyong Dong

2017/3/26

A description of the problem

Empirical exercise 4.2 concerns the relationship between course evaluations and the professor's beauty, using the data on course evaluations, course characteristics, and professor characteristics for 463 courses at the University of Texas at Austin. Specifically, we want to see how course evaluations (`Course_Eval`) are related to the professor's beauty (`Beauty`).

Before running all R codes, we may first load the package `AER` by running `library(AER)`.

```
library(AER)
```

```
## Loading required package: car
## Loading required package: lmtest
## Loading required package: zoo
##
## Attaching package: 'zoo'
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
## Loading required package: sandwich
## Loading required package: survival
```

Answers to the questions

Question (a)

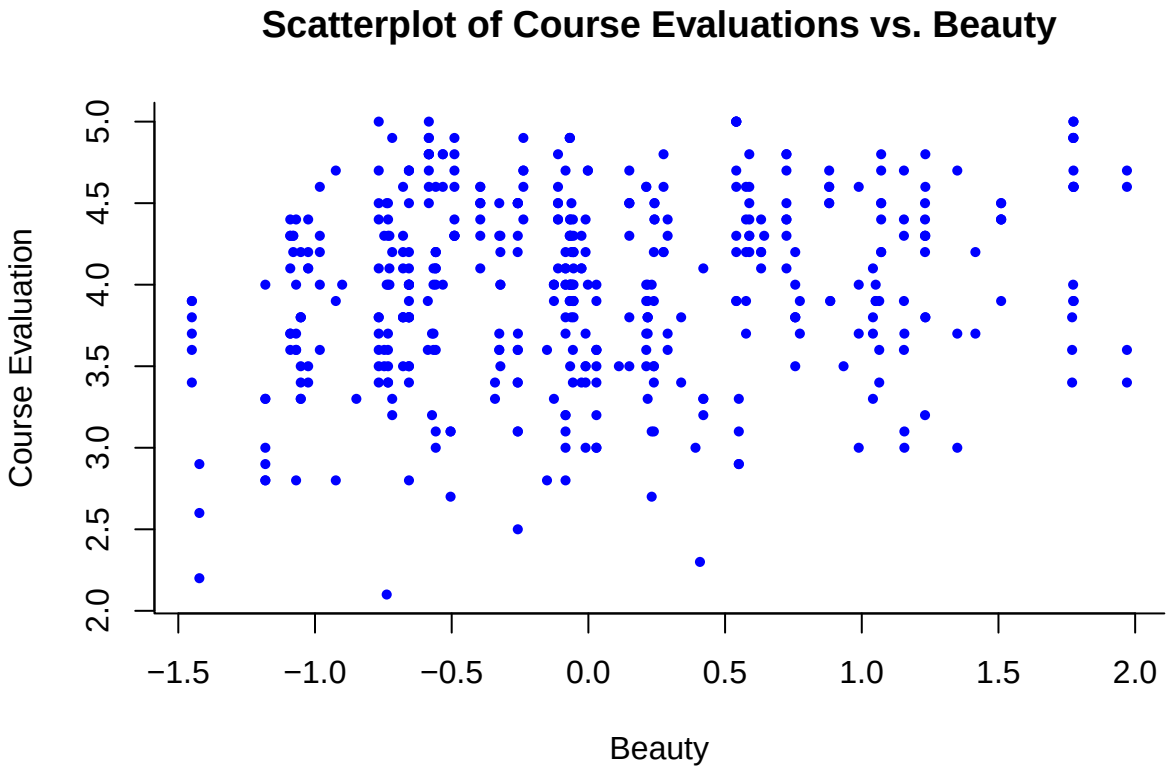
Construct a scatterplot of average course evaluations (`course_eval`) on the professor's beauty (`beauty`). Analyze the relationship between the variables.

Read the data

```
library(foreign)
TRdata <- read.dta("/Users/guopeilin/Desktop/TeachingRatings.dta")
attach(TRdata)
names(TRdata)
```

Create the scatterplot

```r
plot(beauty, course_eval, col="blue", pch=16, cex=0.7, bty="l",
     main="Scatterplot of Course Evaluations vs. Beauty",
     xlab="Beauty", ylab="Course Evaluation")
```



It appears to be a weak positive relationship between Beauty and Course Evaluation.

Question (b)

Run a regression of average course evaluations (Course_Eval) on the professor's beauty (Beauty). Give the estimated intercept and the estimated slope. Explain why the estimated intercept is equal to the sample mean of Course_Eval.

Estimate the regression line of `course_eval` on `beauty`

```r
mod1<-lm(course_eval~beauty)
summary(mod1)
```

```
##
## Call:
## lm(formula = course_eval ~ beauty)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.80015 -0.36304  0.07254  0.40207  1.10373
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.99827    0.02535 157.727  < 2e-16 ***
```

```
## beauty       0.13300    0.03218   4.133 4.25e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5455 on 461 degrees of freedom
## Multiple R-squared:  0.03574,    Adjusted R-squared:  0.03364
## F-statistic: 17.08 on 1 and 461 DF,  p-value: 4.247e-05
```

```
mean(course_eval)
```

```
## [1] 3.998272
```

```
mean(beauty)
```

```
## [1] 4.754221e-08
```

The estimated intercept is 4, and the estimated slope is 0.13. The sample mean of beauty is extremly close to 0. We can use the regression line to calculate the sample mean of Course_Eval: 4+0.13*0.Hence, the sample mean of course eval is equal to the intercept.

$$\widehat{Course - Eval} = 4.00 + 0.13 \times Beauty$$

Queation(c)

Predict Professor Stock's and Professor Watson's course evaluations given that Professor Watson's beauty is 0while Professor Stock's value of Beauty is 1 ### Calculate the predicted course evaluations by the OLS function

```
point<-data.frame(beauty=rbind(mean(beauty),mean(beauty)+sd(beauty)))
predict(lm(course_eval ~ beauty), point)
```

```
##        1        2
## 3.998272 4.103163
```

Professor Stock's and Professor Watson's course evaluations will be 4, 4

Question(d)

Comment on the size of the regression's slope. Comment on the estimated effect of Beauty on Course_Eval and explain.

Compare the increase in Course_Eval when one standard of deviation increse in Beauty with the standard deviation in Course_Eval

```
sd(course_eval)
```

```
## [1] 0.5548656
```

```
sd(beauty)*0.133
```

```
## [1] 0.1048901
```

Because 0 is much smaller than 1, the effect is small.

Question(e)

Does Beauty explain a large fraction of the variance in evaluations across courses? Explain. Because the Multiple R-squared: 0.03574 is much smaller than 1, Beauty doesn't explain a large fraction of the variance in evaluations across courses.