Empirical Exercise

HongYuxi

2017/4/4

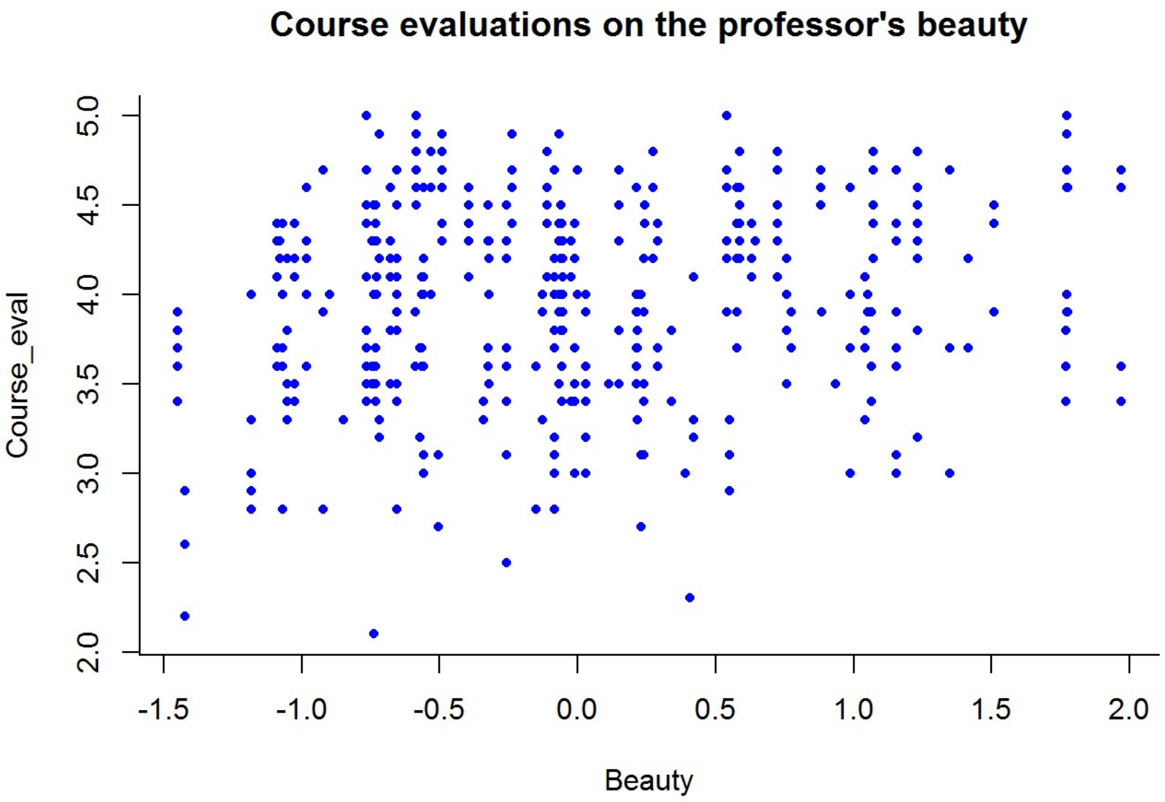
Empirical Exercise Anwers

4.2a

```
library("foreign", lib.loc="C:/Program Files/R/R-3.3.3/library")
classdata <- read.dta("C:\\Users\\hyx\\Documents\\TeachingRatings.dta")
head(classdata[c("age", "beauty")])
```

```
##   age     beauty
## 1  36  0.2899157
## 2  59 -0.7377322
## 3  51 -0.5719836
## 4  40 -0.6779634
## 5  31  1.5097942
## 6  62  0.5885687
```

```
df <- classdata[c("beauty", "course_eval")]
plot(df$beauty, df$course_eval, col = "blue", pch =16, cex = 0.7, bty =
"l", main = "Course evaluations on the professor's beauty", xlab = "Beaut
y", ylab = "Course_eval")
```



```
cor(df$beauty,df$course_eval)
```

```
## [1] 0.1890391
```

4.2b

course_eval=4+0.13beauty

The sampel mean of Beauty is 0;

The estimated intercept = (the mean of Course_Eval) -0.133(the mean ofBeauty).

Thus, the estimated intercept is equal to the sample mean of Course_Eval.

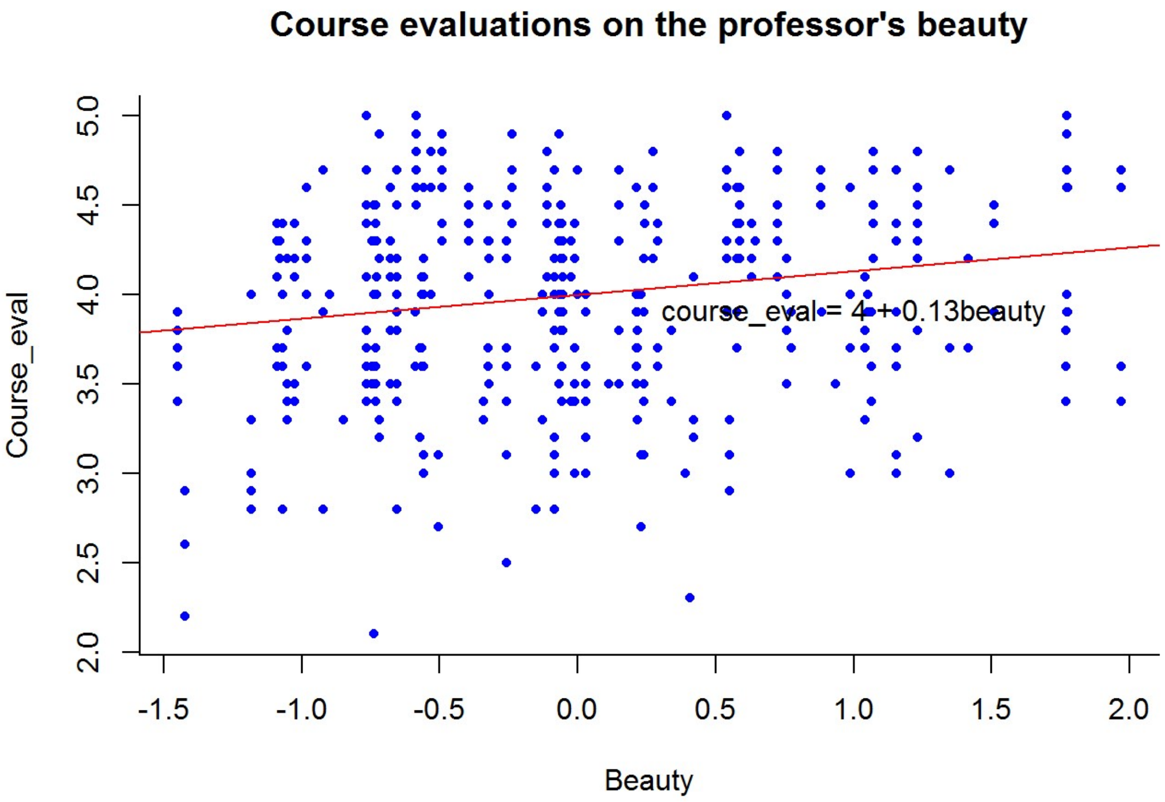
```
mod1 <-lm(course_eval ~ beauty, data = df)
summary(mod1)
```

```
##
## Call:
## lm(formula = course_eval ~ beauty, data = df)
##
## 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
```

```
summary(df)
```

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

```
intercept <- coef(mod1)[1]
slope <- coef(mod1)[2]
texteq <- paste("course_eval = ", round(intercept, 1), " + ", round(slope,
2), "beauty", sep = "")
plot(df$beauty, df$course_eval, col = "blue", pch =16, cex = 0.7, bty =
"l", main = "Course evaluations on the professor's beauty", xlab = "Beaut
y", ylab = "Course_eval")
abline(intercept, slope, col="red")
text(1, 3.9, texteq, cex.lab = 0.5, font.lab = 3)
```



4.2c

```
std<-sapply(df,sd)
print(std)
```

```
##      beauty course_eval
##   0.7886477   0.5548656
```

The standard deviation of Beauty is 0.789.

Thus Professor atson’s predicted course evaluations is 4.00

Professor Stock’s predicted course evaluations is 4.105

4.2d

The effect is “small”.

The standard deviation of course evaluations is 0.55 and the standard deviation of beauty is 0.789.

A one standard deviation increase in beauty is expected to increase course evaluation by 0.133 *0.789 =0.105It’s only around 1/5 of a standard deviation of course evaluations. So,the effect is small.

4.2e

R^2=0.03574, which means the “Beauty” only explains 3.574% of the variance in course evaluations.

Appendix–The problem codes

```
library("foreign", lib.loc="C:/Program Files/R/R-3.3.3/library")
classdata <- read.dta("C:\\Users\\hyx\\Documents\\TeachingRatings.dta")
head(classdata[c("age", "beauty")])
df <- classdata[c("beauty", "course_eval")]
plot(df$beauty,df$course_eval, col = "blue", pch =16, cex = 0.7, bty = "l",
main = "Course evaluations on the professor's beauty", xlab = "Beauty", yla
b = "Course_eval")
cor(df$beauty,df$course_eval)
mod1 <-lm(course_eval ~ beauty, data = df)
summary(mod1)
summary(df)
intercept <- coef(mod1)[1]
slope <- coef(mod1)[2]
texteq <- paste("course_eval = ", round(intercept, 1), " + ", round(slope,
2), "beauty", sep = "")
plot(df$beauty, df$course_eval, col = "blue", pch =16, cex = 0.7, bty =
"l", main = "Course evaluations on the professor's beauty", xlab = "Beaut
y", ylab = "Course_eval")
abline(intercept, slope, col="red")
text(1, 3.9, texteq, cex.lab = 0.5, font.lab = 3)
std<-sapply(df,sd)
print(std)
```