Introduction to Econometrics with R

M. Arnold, M. Schmelzer, A. Gerber

20 Mai 2016

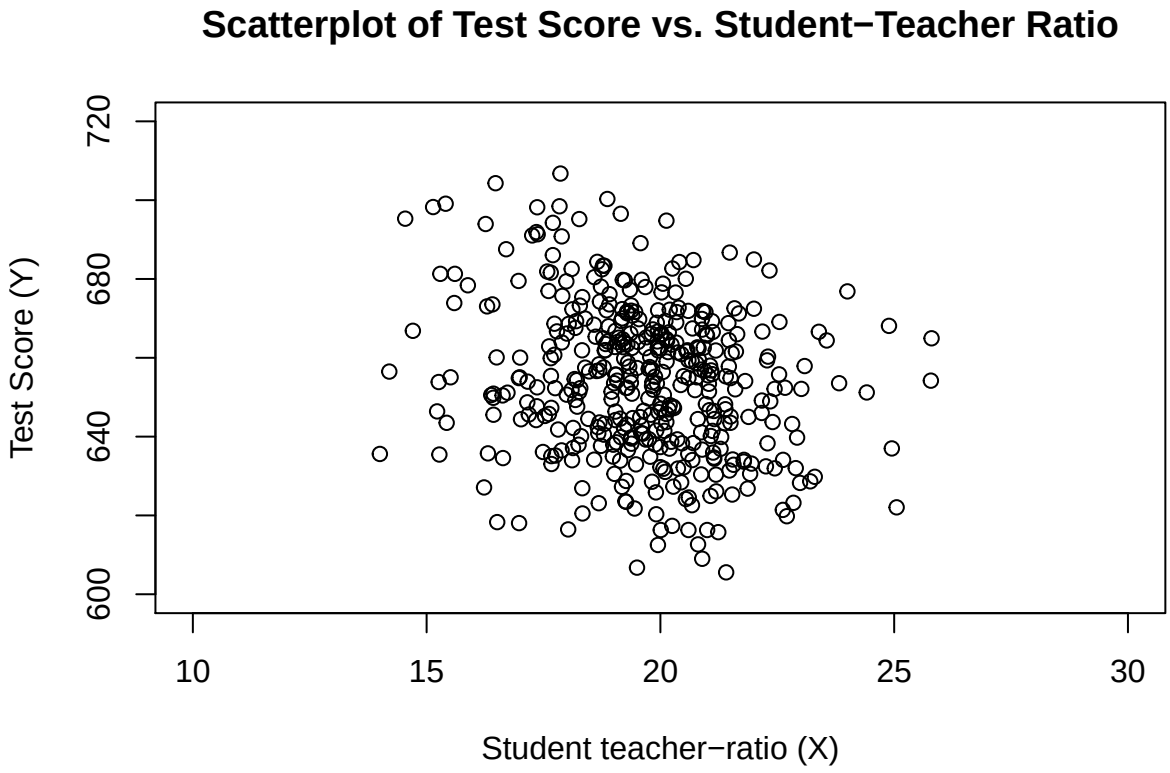
Chapture 4

```
library(AER)                                                    # contains the dataset
data(CASchools)

CASchools$tsratio  <- CASchools$students/CASchools$teachers     # teacher-student-ratio
CASchools$score    <- (CASchools$read + CASchools$math)/2       # average test-score

avg_tsratio        <- mean(CASchools$tsratio)
avg_score          <- mean(CASchools$score)
sd_tsratio         <- sd(CASchools$tsratio)
sd_score           <- sd(CASchools$score)
quantiles          <- c(0.10, 0.25, 0.4, 0.5, 0.6, 0.75, 0.9)
qstratio           <- quantile(CASchools$tsratio, quantiles)
qmath              <- quantile(CASchools$score, quantiles)
qmath
```

```
 10%      25%      40%      50%      60%      75%      90%
```

630.3950 640.0500 649.0700 654.4500 659.4000 666.6625 678.8600

```
plot(score ~ tsratio,
     data = CASchools,
     main = "Scatterplot of Test Score vs. Student-Teacher Ratio",
     xlab = "Student teacher-ratio (X)",
     ylab = "Test Score (Y)",
     xlim = c(10,30),
     ylim = c(600, 720))
```




```
attach(CASchools)
beta_1 <- sum((tsratio - mean(tsratio))*(score - mean(score))) / sum((tsratio - mean(tsratio))^2)
beta_0 <- mean(score) - beta_1 * mean(tsratio)
```

```
library(xtable)
plot(score ~ tsratio,
     data = CASchools,
     main = "Scatterplot of Test Score vs. Student-Teacher Ratio",
     xlab = "Student teacher-ratio (X)",
     ylab = "Test Score (Y)",
     xlim = c(10,30),
     ylim = c(600, 720))

linear_model <- lm(score ~ tsratio, data = CASchools)
linear_model
```

```
##
## Call:
## lm(formula = score ~ tsratio, data = CASchools)
##
## Coefficients:
## (Intercept)      tsratio
##      698.93        -2.28
```

```r
abline(linear_model)
```