R Basics Practice Problem Solutions

|# Dominic Owusu Abeyie #|

03-12-2024

Data Science for Research Assistants - UChicago

ISLR Chapter 2 Q8

- **(a)** Set **wd** and load college data

```
setwd("~/Dropbox/Github/R/R Module 1 Basics)
college <- read.csv(file="College.csv", header = TRUE)
```

ISLR Chapter 2 Q8

- (b,c) Give data frame college rownames and delete first column afterward.

```
rownames(college) <- college[,1]
college <- college[,-1]
#View(college)
```

ISLR Chapter 2 Q8

- Summary of variables: There are 19 variables. This produces summary of 2 variables in order to fit the page. #The rest can produced by subsetting the data.

```
summary(college[,1:2])
```

```
##    Private                 Apps
##  Length:777         Min.   :   81
##  Class :character   1st Qu.:  776
##  Mode  :character   Median : 1558
##                     Mean   : 3002
##                     3rd Qu.: 3624
##                     Max.   :48094
```

ISLR Chapter 2 Q8

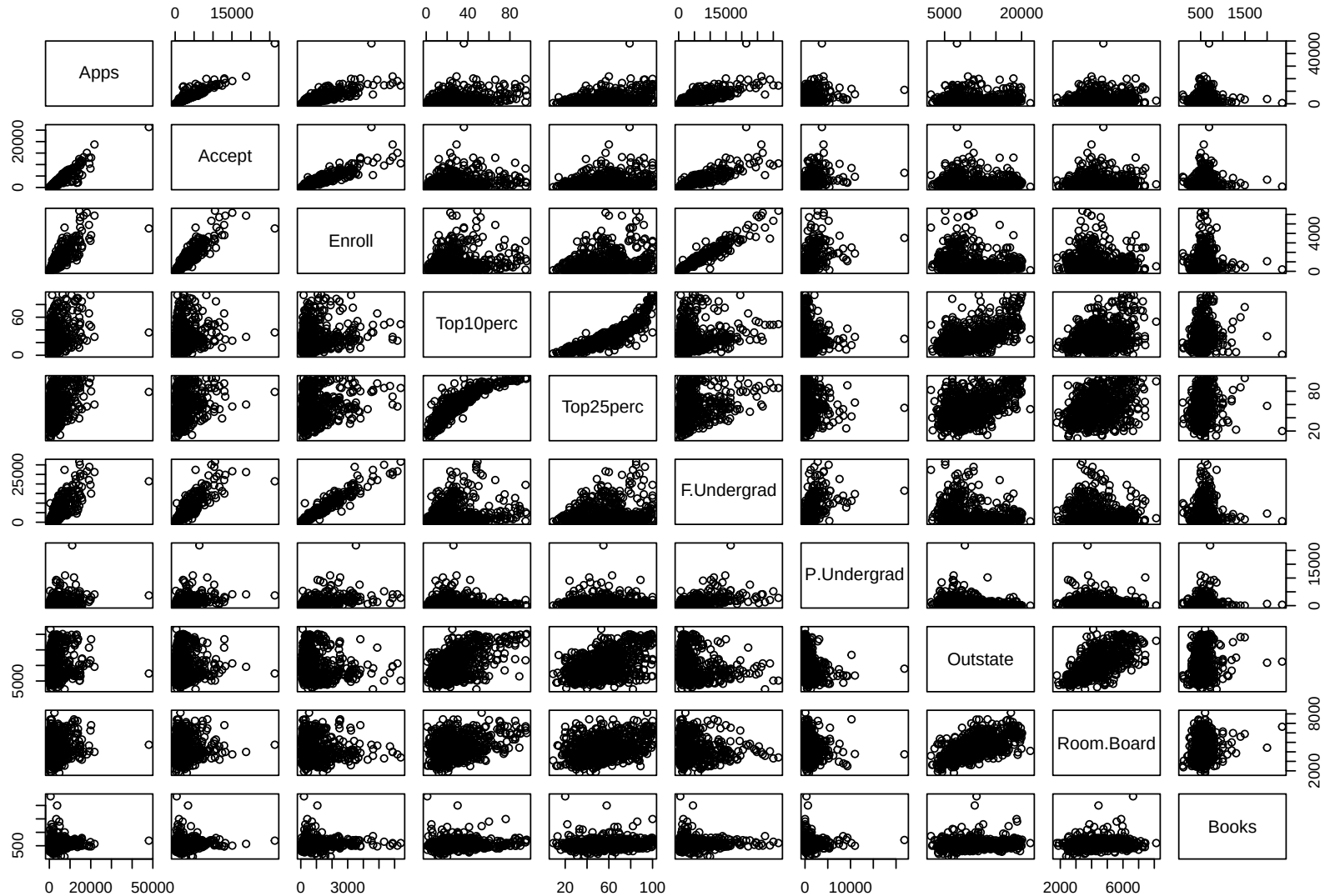
- Summary of variables: *continue*. . .

```
summary(college[,3:5])
```

```
##      Accept          Enroll       Top10perc
##  Min.   :   72   Min.   :  35   Min.   : 1.00
##  1st Qu.:  604   1st Qu.: 242   1st Qu.:15.00
##  Median : 1110   Median : 434   Median :23.00
##  Mean   : 2019   Mean   : 780   Mean   :27.56
##  3rd Qu.: 2424   3rd Qu.: 902   3rd Qu.:35.00
##  Max.   :26330   Max.   :6392   Max.   :96.00
```

ISLR Chapter 2 Q8

```
#Visualize scatter plot of first 10 var.
pairs(college[, 2:11])
```



ISLR Chapter 2 Q8

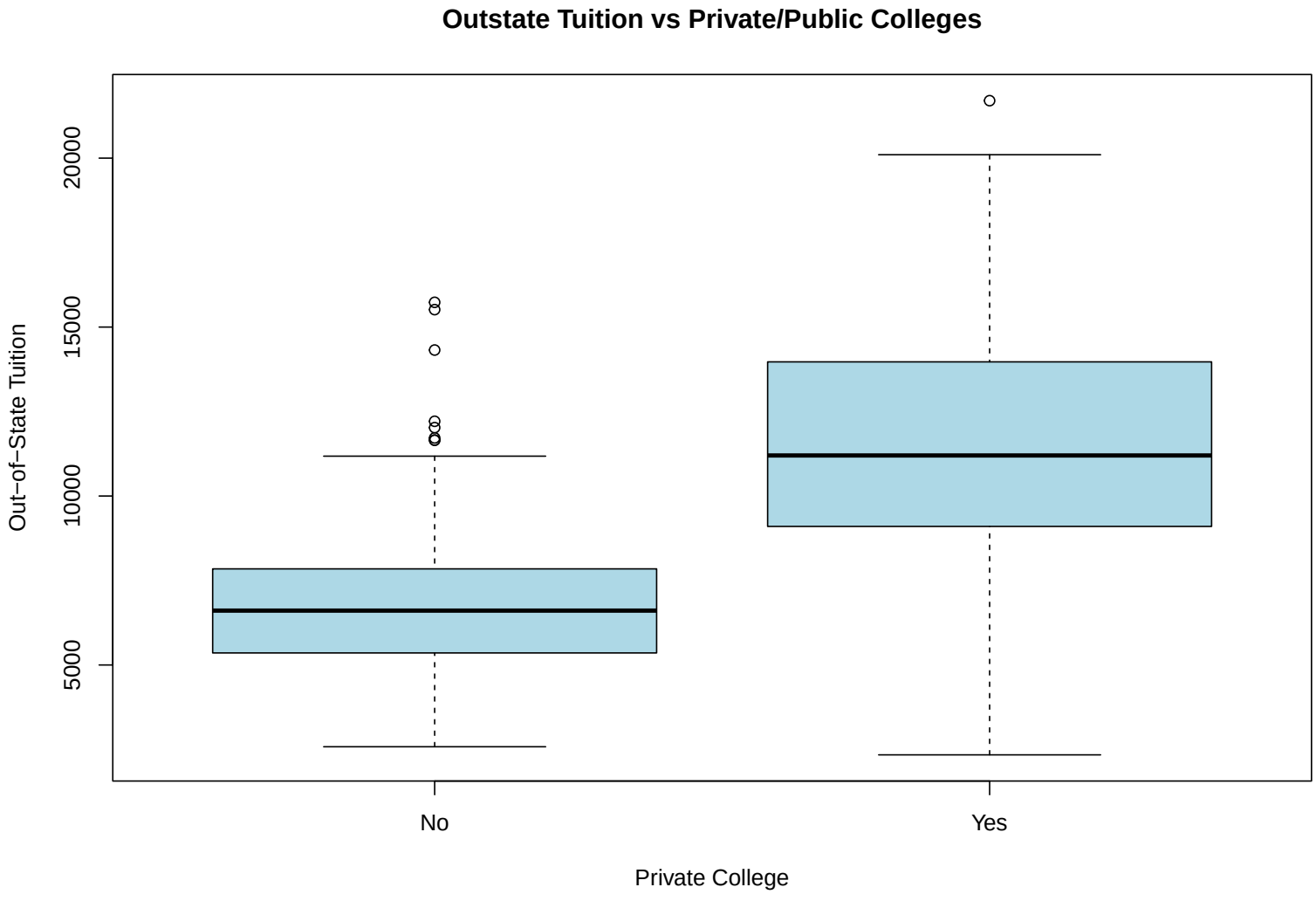
Boxplot Code

```
college$Private <- as.factor(college$Private)

plot(college$Private, college$Outstate,
     main = "Outstate Tuition vs Private/Public Colleges",
     xlab = "Private College",
     ylab = "Out-of-State Tuition",
     col = "lightblue")
```

ISLR Chapter 2 Q8

- Visualize boxplot



ISLR Chapter 2 Q8

- Create a new qualitative variable ("Yes" or "No"), called Elite

```
Elite <- rep("No", nrow(college))
Elite[college$Top10perc > 50] <- "Yes"
Elite <- as.factor(Elite)
college <- data.frame(college, Elite)
```

```
# Check how many elite colleges exist
summary(Elite)
```

```
##  No Yes
## 699  78
```

ISLR Chapter 2 Q8

- Code of boxplots for Outstate vs Elite

```
plot(college$Elite, college$Outstate,
     main = "Outstate Tuition for Elite vs
     Non-Elite Colleges",
     xlab = "Elite Status",
     ylab = "Out-of-State Tuition",
     col = "lightblue")
```

ISLR Chapter 2 Q8

- Boxplots for Outstate vs Elite