Assignment: Week 4 - Scores

Name: Ramani, Aarti

Date: 2023-01-04

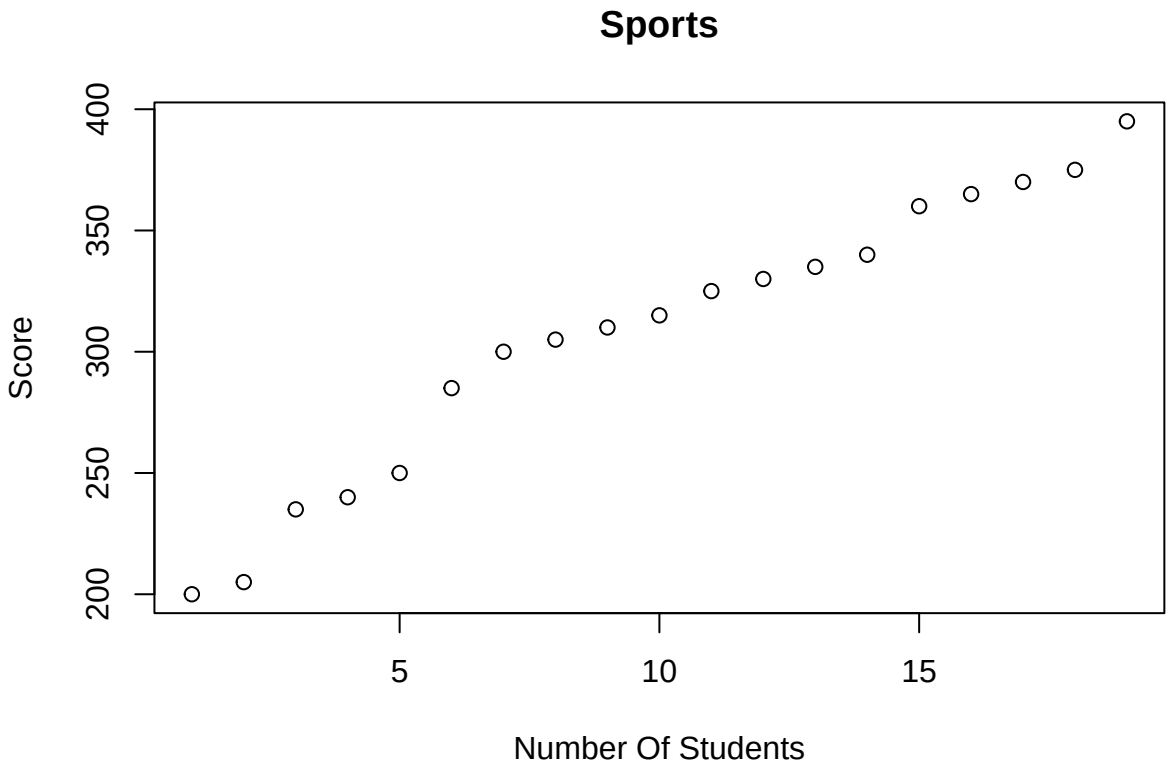
```
setwd("C:/Masters/GitHub/Winter2022/Ramani-DSC520")
scores_df <- read.csv(file="C:/Masters/GitHub/Winter2022/Ramani-DSC520/data/scores.csv", header = TRUE)

#1. What are the observational units in this study?
colnames(scores_df)
```

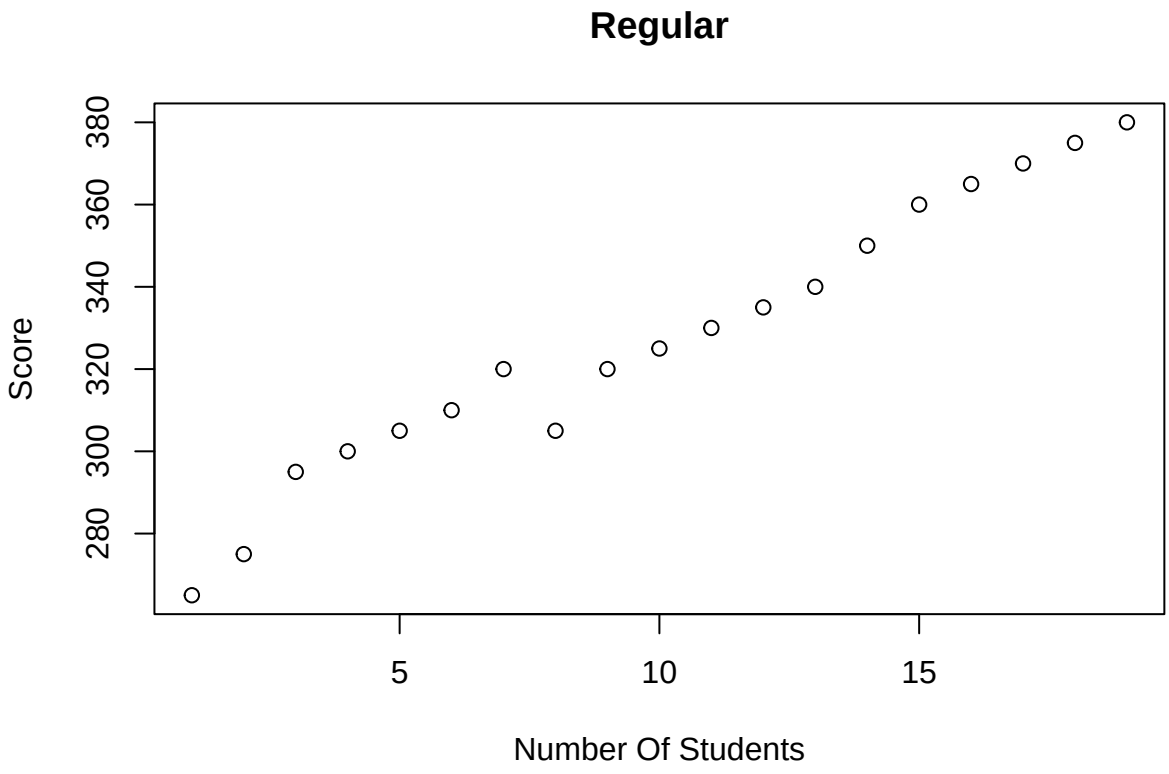
```
## [1] "Count"   "Score"   "Section"
```

```
# Answer: Observation Units are the course grades and total points earned in the course


#2. Identify the variables mentioned in the narrative paragraph and determine
#   which are categorical and quantitative?
str(scores_df)
```

```
## 'data.frame':    38 obs. of  3 variables:
##  $ Count  : int  10 10 20 10 10 10 10 30 10 10 ...
##  $ Score  : int  200 205 235 240 250 265 275 285 295 300 ...
##  $ Section: chr  "Sports" "Sports" "Sports" "Sports" ...
```

```
#Answer - The Categorical variables is "Section" (Sports and Regular)
# The Quantitative variable is the "Score"

#3. Create one variable to hold a subset of your data set that contains only
#   the Regular Section and one variable for the Sports Section.
sports <- subset(scores_df, scores_df$Section=="Sports")
regular <- subset(scores_df, scores_df$Section=="Regular")

#4. Use the Plot function to plot each Sections scores and the number of
#   students achieving that score. Use additional Plot Arguments to label
#   the graph and give each axis an appropriate label.
sports_ls <- sports[,2]
regular_ls <- regular[,2]
par(mfrow=c(2,1))
plot(sports_ls, xlab="Number Of Students", ylab="Score", main="Sports")
plot(regular_ls, xlab="Number Of Students", ylab="Score", main="Regular")
```


Sports
Score
200
250
300
350
400
5
10
15
Number Of Students


Regular
Score
280
300
320
340
360
380
5
10
15
Number Of Students

```
# a. Comparing and contrasting the point distributions between the two section,
#    looking at both tendency and consistency: Can you say that one section
#     tended to score more points than the other? Justify and explain your answer.
#Answer - Looking at the plot, sports section students scored more than the regular section students.

# b. Did every student in one section score more points than every student in
#    the other section? If not, explain what a statistical tendency means in this context.
# Answer - Sports section scored more than the regular section

# c. What could be one additional variable that was not mentioned in the
#    narrative that could be influencing the point distributions between the two sections?
#Answer - Variable that was not mentioned in the narrative is Counts.
```