ST-314 Data Analysis 3

Ivan Timothy Halim

January 28, 2019

Part 1

a. Average Drying Time of Manufacturer : 32.67 min
Average Drying Time of Competition : 32.8 min

b. The manufacturers claim is truthful in that the average drying time of the manufacturer is 0.1 minutes less than the competition. However, it is also misleading because the observed differences were no more than what might reasonably occur by chance.

c. This is a problem because there might be a sampling bias in the form of convenience and voluntary response bias. The company is only reporting on the data that supports its claims while ignoring the data that rejects them.

Part 2

Categorical Variable

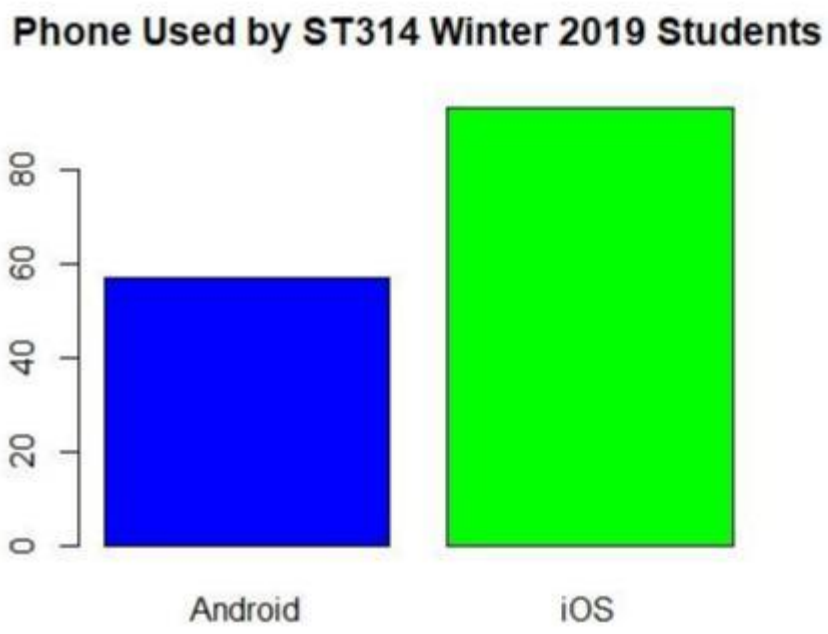
a. I choose the Type of Phone Used by ST314 Winter 2019 Students.

b.

```
st314data = read.csv(file.choose(), header = TRUE)
table(st314data$Phone)
```

```
##
## Android     iOS
##      57      93
```

```
barplot(table(st314data$Phone),
        main = "Phone Used by ST314 Winter 2019
        Students", col = c("blue","green"))
```



c. 38% of students use Android

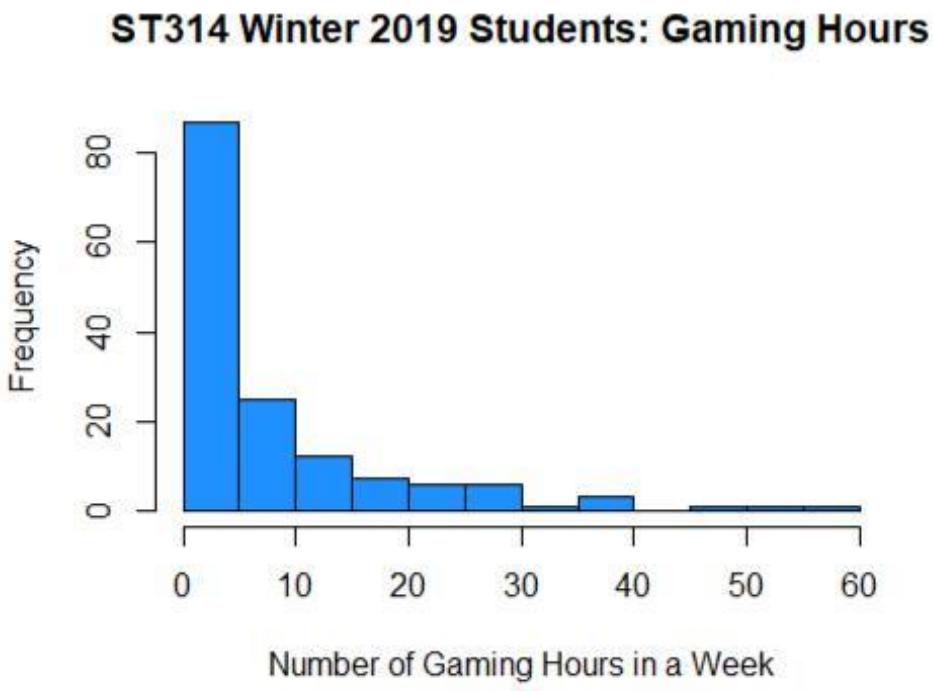
62% of students use iOS

Quantitative Variable

a. I choose the Gaming Hours per Week of ST314 Winter 2019 Students

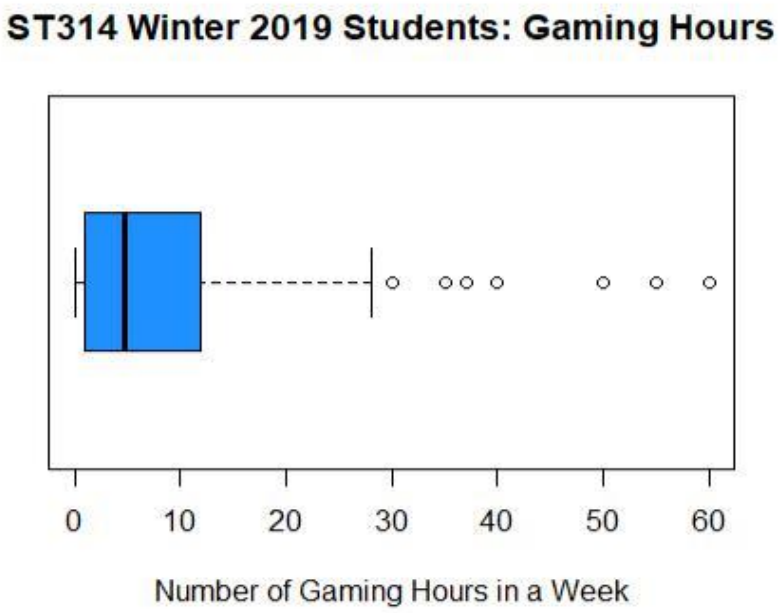
b.

```
hist(st314data$GamingHours,
     main = "ST314 Winter 2019 Students: Gaming Hours",
     xlab = "Number of Gaming Hours in a Week", col =
     "dodgerblue")
```



c.

```
boxplot(st314data$GamingHours,
        main = "ST314 Winter 2019 Students: Gaming
        Hours", xlab = "Number of Gaming Hours in a
        Week", col = "dodgerblue", horizontal = TRUE)
```



d. The histogram is useful for looking at the frequency of each observations. The boxplot is useful for looking at the spread of the data.

e.

```
summary(st314data$GamingHours)
```

```
##    Min.  1st Qu.  Median    Mean  3rd Qu.    Max.
##   0.000    1.250   4.750   8.899   11.500  60.000
```

f. The distribution is unimodal and positively skewed, ranging from 0 hours to 60 hour, with many students hovering between 0 to 5 hours. The average gaming hours in a week is 8.9 hours with a standard deviation of 11.4 hours. The median gaming hours is 4.75 hours. There are several extreme observations ranging from 30 hours to 60 hours.

g. For skewed data, the median is a better representation of a typical value because the outliers might influence the mean of the data.