Project 1

Jared Barber

1/9/2022

Exercise 1.2

a) Calculate the sample mean and median for the sample values

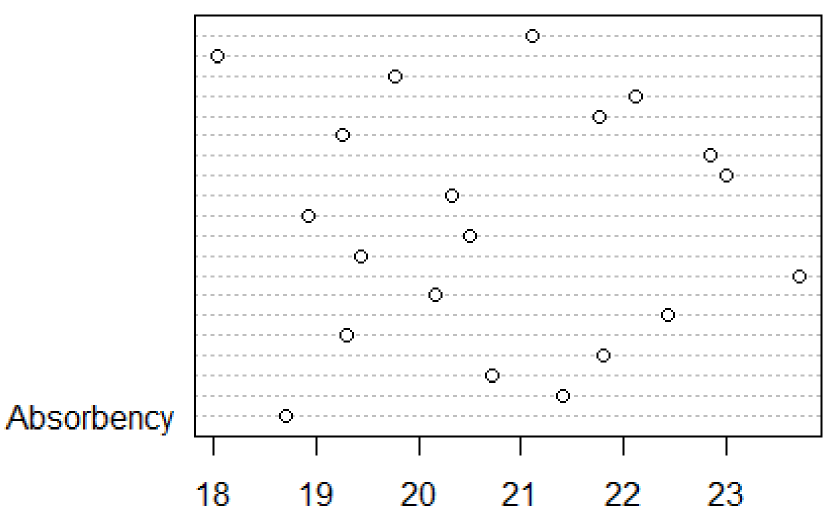
```
mean(Ex01.02$absorbency)

## [1] 20.7675

median(Ex01.02$absorbency)

## [1] 20.61
```

b) Compute the 10% trimmed mean

```
mean(Ex01.02$absorbency, trim=0.10)

## [1] 20.74312
```

c) Do a dot plot of the absorbency data

```
dotchart(Ex01.02$absorbency, labels="Absorbency")
```



d) Using only the mean, median, and trimmed mean, do you have evidence of outliers in the data?

No, based only on these measures, there is no evidence of outliers in the data. If there were outliers, there should be a significant difference between the mean and median or the mean and trimmed mean, which is not the case.

Exercise 1.8

Compute the variance and standard deviation for the water absorbency data from Exercise 1.2

```
var(Ex01.02$absorbency)

## [1] 2.532914

sd(Ex01.02$absorbency)

## [1] 1.591513
```

Exercise 1.14

a) Find the sample mean and median

The sorted data is : 565, 568, 569, 570, 572, 572, 573, 575.

The mean is

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

= (565 + 568 + 569 + 570 + 572 + 572 + 573 + 575)/ 8 = **570.5**

The median is (570 + 572)/2 = **571**

b) Find the sample variance, standard deviation, and range

The variation =

$$\frac{\sum(\bar{x} - x_i)^2}{n - 1}$$

$$\frac{(570.5 - 565)^2 + (570.5 - 568)^2 + (570.5 - 569)^2 + (570.5 - 570)^2 + (570.5 - 572)^2 + (570.5 - 572)^2}{8 - 1}$$

$$$$

$$\frac{70}{7} = 10$$

The standard deviation is the square root of the variance. √10 = **3.162278**

The range is 575-565 = **10**

c) Using the calculated statistics in parts (a) and (b), can you comment on the quality of the tires?

The mean and median indicate that the sample diameters are approximately correct. There is some variation, but given that this variation is small, the quality of the tires is probably acceptable.

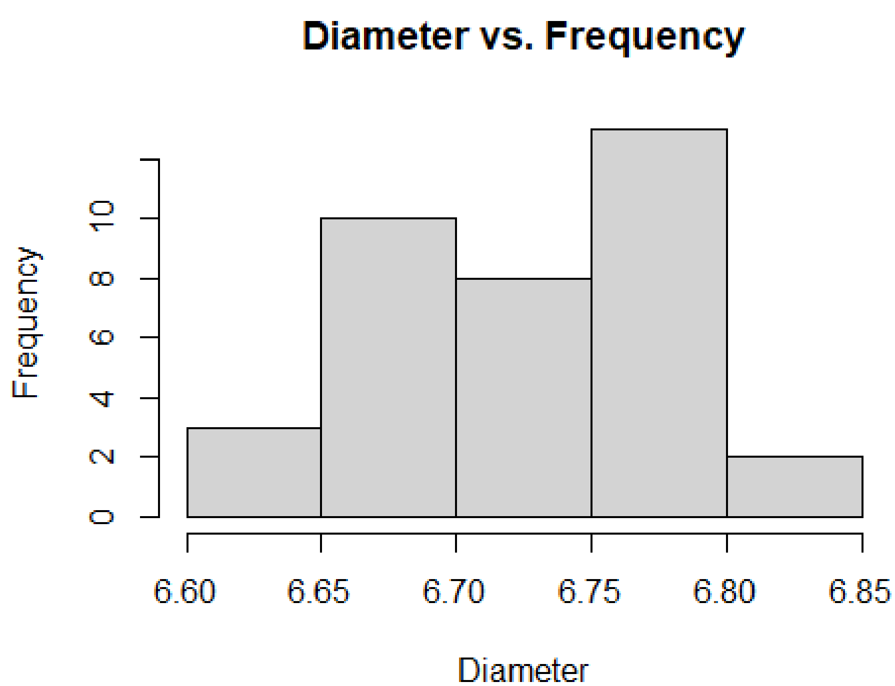
Exercise 1.22

a) Find the mean and standard deviation

```
mean(Ex01.22$diameter)

## [1] 6.726111

sd(Ex01.22$diameter)

## [1] 0.05357386
```

b) Draw a frequency histogram

```
hist(Ex01.22$diameter, main='Diameter vs. Frequency', xlab='Diameter')
```



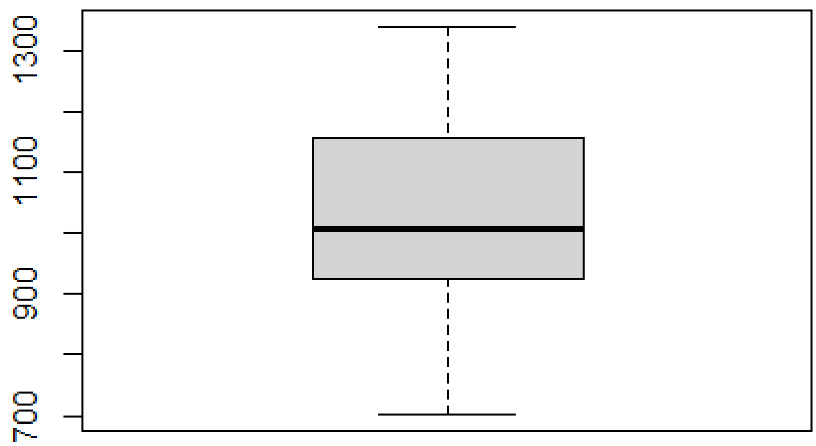
c) Comment on whether or not there is any clear indication the data came from a bell-shaped population.

Based on the shape of the graph and the standard deviation, there is some evidence that the sample came from a population with a bell-shaped distribution.Most of the data points fall in the middle of the spectrum, and fewer fall on the edges.

Exercise 1.30

Construct a boxplot for the data

```
boxplot(Ex01.30$Lifetime)
```



To make computations easier, the data can be sorted

```
Ex01.30$Lifetime[order(Ex01.30$Lifetime)]

##  [1]  702  765  785  811  832  855  896  902  905  918  919  920  923  929
936
## [16]  938  948  950  956  958  958  970  972  978 1009 1009 1022 1035 1037
1045
## [31] 1067 1085 1092 1102 1122 1126 1151 1156 1157 1157 1162 1170 1195 1195
1196
## [46] 1217 1237 1311 1333 1340
```

Q1 is the value x at the position where 25% of the values in the sorted data set are less than x. To find this value, take the weighted average of the numbers at the 13th and 14th positions.

0.75(923)+0.25(929) = **924.5**

Q3 is the value y at the position where 75% of values in the sorted data set are less than y. To find this value, take the weighted average of the numbers at the 37th and 38th positions.

0.25(1151)+0.75(1156) = **1154.75**

Using R,

```
quantile(Ex01.30$Lifetime)

##      0%     25%     50%     75%    100%
##  702.00  924.50 1009.00 1154.75 1340.00
```

IQR is Q3-Q1 = 1154.75 - 924.5 = **230.25**

A low-lying outlier is a value that is 1.5*IQR less than Q1, or below 924.5-1.5(230.25) = 579.125. No values in the data are less than or equal to 579.125, so there are no low-lying outliers.

A high outlier is a value that is 1.5*IQR greater than Q3, or above 1154.75 + 1.5(230.25) = 1500.125. No values in the data are greater than or equal to 1500.125, so there are no outliers.