R Lab 2

Vinayak Mathur

1/25/22

Problem 1:

The energy consumption for 90 gas-heated homes during a winter heating season is given in the file "Energy_Usage.csv" (see Data module on canvas). The variable reported is BTU/number of heating degree day.

a)

Use R codes to find the five point summary, mean and standard deviation of the data. (4 points)

```
energy <- read.csv("Energy_Usage.csv", header = TRUE)
summary(energy)
```

```
##       BTU
##  Min.   : 2.970
##  1st Qu.: 7.947
##  Median : 9.835
##  Mean   :10.038
##  3rd Qu.:12.045
##  Max.   :18.260
```

```
sd(energy$BTU)
```

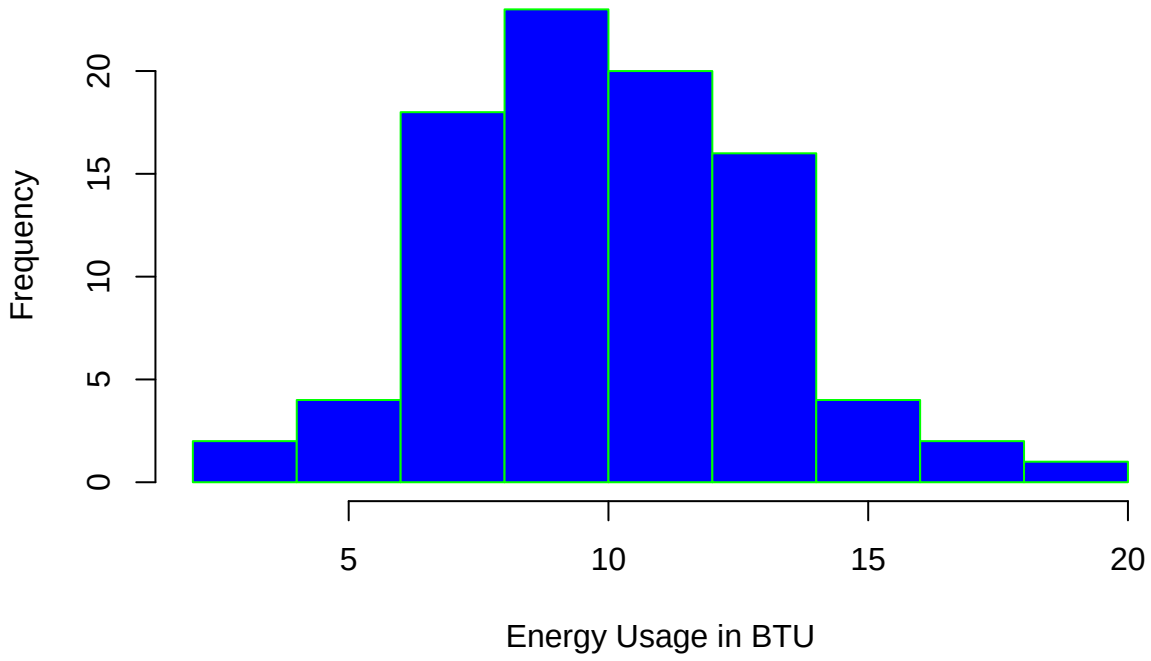
```
## [1] 2.86799
```

Ans: Standard Deviation = 2.86799

b)

Construct a histogram and a stem-and-leaf diagram of energy usage. Use appropriate scale for the stem-and-leaf diagram. Then use SOCS method to describe the data. (6 points)

```
hist(energy$BTU,
     main="Histogram for Energy usage of 90 homes",
     xlab="Energy Usage in BTU",
     border="green",
     col="blue")
```




```
stem(energy$BTU, scale = 2)
```

```
##
##   The decimal point is at the |
##
##    2 |
##    3 | 0
##    4 | 0
##    5 | 269
##    6 | 04678899
##    7 | 2223667799
##    8 | 03345566778
##    9 | 134456688888
##   10 | 002333444556
##   11 | 011234677
##   12 | 223367799
##   13 | 144456
##   14 | 024
##   15 | 12
##   16 | 19
##   17 |
##   18 | 3
```

```
summary(energy$BTU)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.970   7.947   9.835  10.038  12.045  18.260
```

Ans: Shape -> Shape of the distribution is symmetric with a single peak. Outliers -> The data does not have any outliers. Min. 1st Qu. Median Mean 3rd Qu. Max. 2.970 7.947 9.835 10.038 12.045 18.260