IST 777
Joe Hernandez
7/10/19

Homework #1

1. Personal definition of terms
    a. Mean: A measure of central tendency by taking the average (sum of all values divided by number of values)
    b. Median: A measure of central tendency by observing the middle value (values are sorted when finding middle number)
    c. Mode: A measure of central tendency by observing the most frequent value
    d. Variance: A measure of dispersion where the sum of squared deviations is divided by the number of observations
    e. Standard Deviation: A measure of dispersion much smaller than the variance (square root of variance)
    f. Histogram: A graphical representation of data dispersion and frequency of values
    g. Normal Distribution: Commonly known as the bell-shaped curve showing a symmetrical distribution of data with similar standard deviation values
    h. Poisson Distribution: Characterizes events with low probabilities showing a right skewed distribution
3. Use data() function to get list of data sets included with base r

```
> data()
```

Choose data from list use summary() command
Islands chosen measuring the area of land mass by square miles

```
> summary(islands)
   Min. 1st Qu.  Median    Mean       3rd Qu.    Max.
   12.0   20.5   41.0      1252.7      183.2      16988.0
```

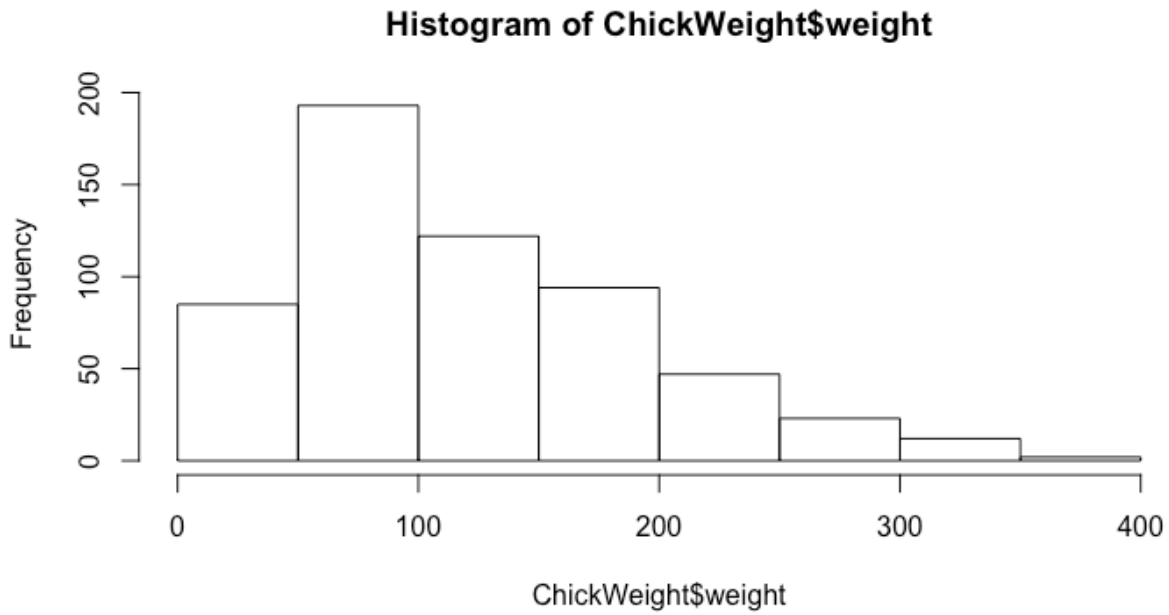
Mean is the average number of the data set
1252.7 is the mean for this dataset where the smallest landmass is 12 and largest is 16988.0
Median is the middle number of the data set
41.0 is the median for this dataset where 7 out of 48 landmasses are above 1000 square miles

4. Use data() function to get list of data sets included with base r

```
> data()
```

Choose data from list use summary() command
Use hist() command to create histogram

```
> hist(ChickWeight$weight)
```



The shape of the histogram shows a right tailed Poisson distribution
This may fit the distribution as there is a low probability for chicks to weigh significantly high