Practice Problems Section 1G

1. Answer the following questions:

    a) Describe a skewed right shape?
    b) Describe a skewed left shape?
    c) Define the median average and explain how it is calculated.
    d) Define the first quartile (Q_1) and explain how it is calculated.
    e) Define the third quartile (Q_3) and explain how it is calculated.
    f) Define the interquartile range (IQR) and explain how it is calculated.

2. Answer the following questions:

    a) If a data set is <u>not</u> normally distributed, what measure of average should we use?

    b) If a data set is <u>not</u> normally distributed, what measure of typical spread should we use?

    c) If a data set is <u>not</u> normally distributed, what are the two statistics that typical values are in between?

    d) If a data set is <u>not</u> normally distributed, approximately what percentage is typical?

    e) If a data set is <u>not</u> normally distributed, how can we use a box plot to find high outliers in the data set?

    f) If a data set is <u>not</u> normally distributed, how can we use a box plot to find low outliers in the data set?

(#3-7) Directions: Analyze the following data sets. Go to www.matt-teachout.org, click on the "Statistics" tab, and then the "Data Sets" tab. Open the "Bear" data, the "Health" data, and the "Car" data. Go to www.lock5stat.com and copy and open StatKey. Under the "Descriptive Statistics and Graphs" menu, click on "One Quantitative Variable". Click on "Edit Data" and copy and paste the indicated data set. Use the graphs and summary statistics to answer the following questions.

3. Bear ages (months)

    a) What is the data measuring and what are the units?
    b) How many numbers are in the data set?
    c) What is the shape of the data set?
    d) What is the minimum value?
    e) What is the maximum value?
    f) What is the average (center)? *(Give the number and the name of the statistic used.)*
    g) How much typical spread does the data set have?
       (Give the number and the name of the statistic used.)
    h) Find two numbers that typical values fall in between.
    i) List all high outliers in this data set. If there are no high outliers, put "none".
    j) List all low outliers in this data set. If there are no high outliers, put "none".

4. Bear Weights (pounds)

    a) What is the data measuring and what are the units?
    b) How many numbers are in the data set?
    c) What is the shape of the data set?
    d) What is the minimum value?
    e) What is the maximum value?
    f) What is the average (center)? *(Give the number and the name of the statistic used.)*
    g) How much typical spread does the data set have?
       (Give the number and the name of the statistic used.)
    h) Find two numbers that typical values fall in between.

i) List all high outliers in this data set. If there are no high outliers, put "none".

j) List all low outliers in this data set. If there are no high outliers, put "none".

5. Women's Systolic Blood Pressure in millimeters of mercury (mm of Hg)

    a) What is the data measuring and what are the units?
    b) How many numbers are in the data set?
    c) What is the shape of the data set?
    d) What is the minimum value?
    e) What is the maximum value?
    f) What is the average (center)? *(Give the number and the name of the statistic used.)*
    g) How much typical spread does the data set have?
       (Give the number and the name of the statistic used.)
    h) Find two numbers that typical values fall in between.
    i) List all high outliers in this data set. If there are no high outliers, put "none".
    j) List all low outliers in this data set. If there are no high outliers, put "none".

6. Men's Diastolic Blood Pressure (mm of Hg)

    a) What is the data measuring and what are the units?
    b) How many numbers are in the data set?
    c) What is the shape of the data set?
    d) What is the minimum value?
    e) What is the maximum value?
    f) What is the average (center)? *(Give the number and the name of the statistic used.)*
    g) How much typical spread does the data set have?
       (Give the number and the name of the statistic used.)
    h) Find two numbers that typical values fall in between.
    i) List all high outliers in this data set. If there are no high outliers, put "none".
   j) List all low outliers in this data set. If there are no high outliers, put "none".

7. Women's Cholesterol in milligrams per deciliter (mg/dL)

    a) What is the data measuring and what are the units?
    b) How many numbers are in the data set?
    c) What is the shape of the data set?
    d) What is the minimum value?
    e) What is the maximum value?
    f) What is the average (center)? *(Give the number and the name of the statistic used.)*
    g) How much typical spread does the data set have?
       (Give the number and the name of the statistic used.)
    h) Find two numbers that typical values fall in between.
    i) List all high outliers in this data set. If there are no high outliers, put "none".
    j) List all low outliers in this data set. If there are no high outliers, put "none".

(#8-12) The following graphs and summary statistics were created from the "Car" data at www.matt-teachout.org and Statcato. Use the Statcato graphs and summary statistics to answer the following questions.

8. Weight of various cars in tons.

Histogram

Box Plot
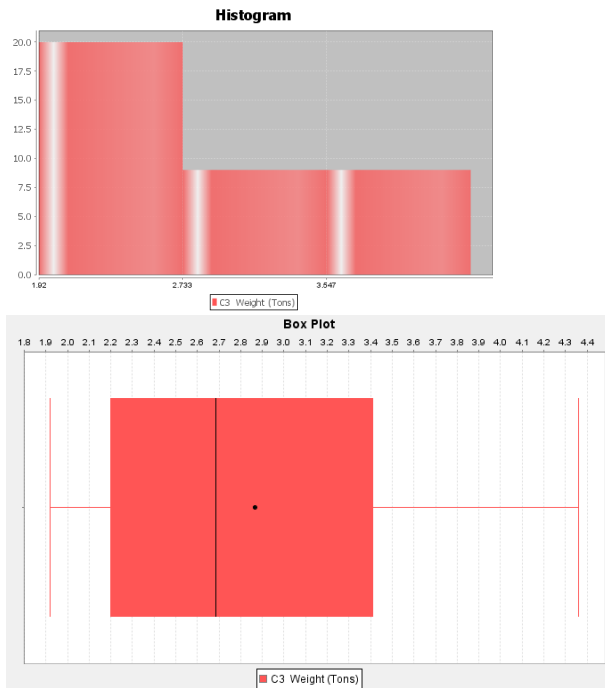
Descriptive Statistics

Variable	Mean	Standard Deviation
Weight (Tons)	2.864	0.706

Variable	Q1	Median	Q3	IQR
Weight (Tons)	2.198	2.685	3.46	1.262

Variable	Min	Max	Range
Weight (Tons)	1.92	4.36	2.440

Variable	N total
Weight (Tons)	38

a) What is the data measuring and what are the units?

b) How many numbers are in the data set?

c) What is the shape of the data set?

d) What is the minimum value?

e) What is the maximum value?

f) What is the average (center)? *(Give the number and the name of the statistic used.)*

g) How much typical spread does the data set have?
   (Give the number and the name of the statistic used.)

h) Find two numbers that typical values fall in between.

i) List all high outliers in this data set. If there are no high outliers, put "none".

j) List all low outliers in this data set. If there are no high outliers, put "none".

9. Gas mileage of various cars in miles per gallon (mpg).
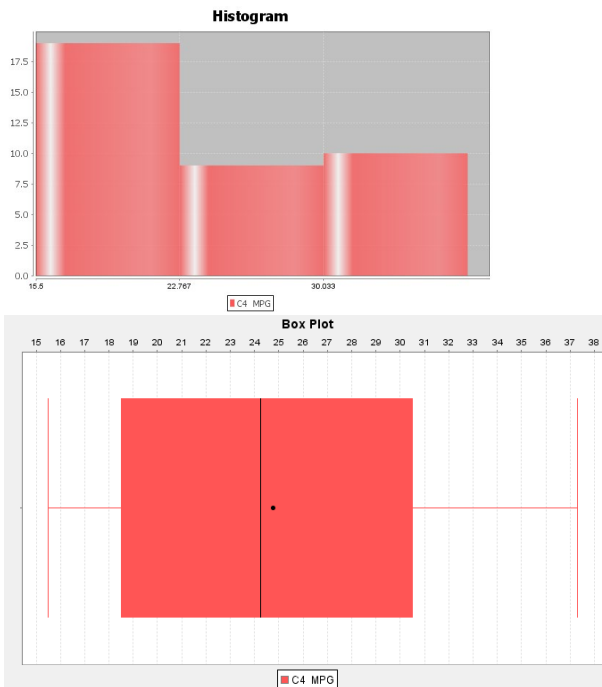
Descriptive Statistics

Variable	Mean	Standard Deviation
MPG	24.761	6.547

Variable	Q1	Median	Q3	IQR
MPG	18.425	24.25	30.6	12.175

Variable	Min	Max	Range
MPG	15.5	37.3	21.800

Variable	N total
MPG	38

a) What is the data measuring and what are the units?
b) How many numbers are in the data set?
c) What is the shape of the data set?
d) What is the minimum value?
e) What is the maximum value?
f) What is the average (center)? *(Give the number and the name of the statistic used.)*
g) How much typical spread does the data set have?
   (Give the number and the name of the statistic used.)
h) Find two numbers that typical values fall in between.
i) List all high outliers in this data set. If there are no high outliers, put "none".
j) List all low outliers in this data set. If there are no high outliers, put "none".

10. The drive ratio of various cars.

Histogram

■ C5 DriveRatio

Box Plot
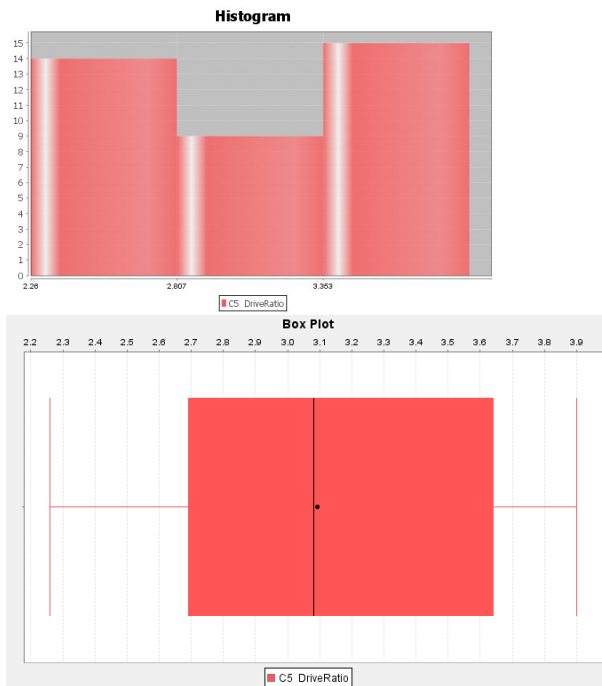
■ C5 DriveRatio

Descriptive Statistics

Variable	Mean	Standard Deviation
DriveRatio	3.093	0.518

Variable	Q1	Median	Q3	IQR
DriveRatio	2.69	3.08	3.655	0.965

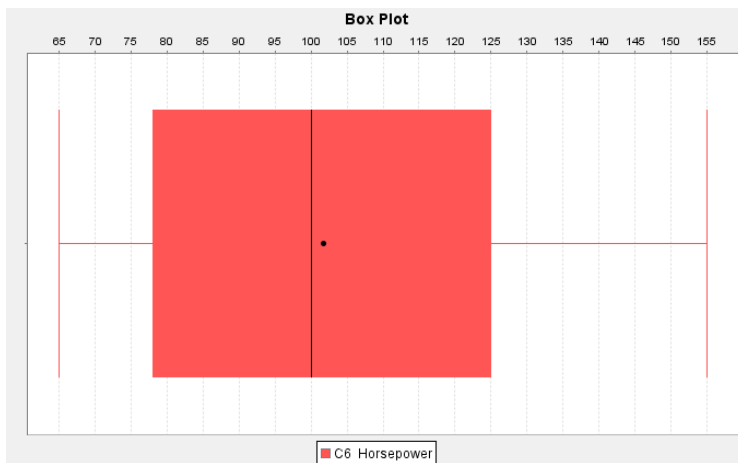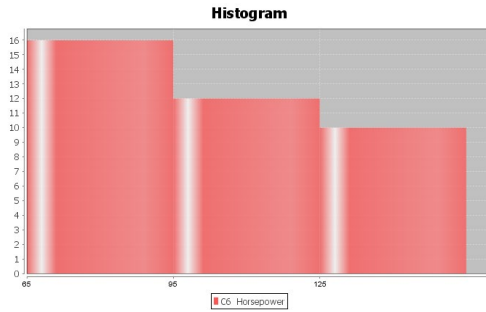| Variable	Min	Max	Range
DriveRatio	2.26	3.9	1.640

Variable	N total
DriveRatio	38

a) What is the data measuring and what are the units?
b) How many numbers are in the data set?
c) What is the shape of the data set?
d) What is the minimum value?
e) What is the maximum value?
f) What is the average (center)? *(Give the number and the name of the statistic used.)*

g) How much typical spread does the data set have?
   (Give the number and the name of the statistic used.)
h) Find two numbers that typical values fall in between.
i) List all high outliers in this data set. If there are no high outliers, put "none".
j) List all low outliers in this data set. If there are no high outliers, put "none".

11. The horsepower of various cars.

Histogram

Box Plot

C6 Horsepower

Descriptive Statistics

Variable	Mean	Standard Deviation
Horsepower	101.737	26.445

Variable	Q1	Median	Q3	IQR
Horsepower	77.25	100.0	125.0	47.75

Variable	Min	Max	Range
Horsepower	65.0	155.0	90.0

Variable	N total
Horsepower	38

a) What is the data measuring and what are the units?
b) How many numbers are in the data set?
c) What is the shape of the data set?
d) What is the minimum value?
e) What is the maximum value?
f) What is the average (center)? *(Give the number and the name of the statistic used.)*
g) How much typical spread does the data set have?
   (Give the number and the name of the statistic used.)
h) Find two numbers that typical values fall in between.
i) List all high outliers in this data set. If there are no high outliers, put "none".
j) List all low outliers in this data set. If there are no high outliers, put "none".

12. The measure of displacement for various cars.

Histogram

Box Plot
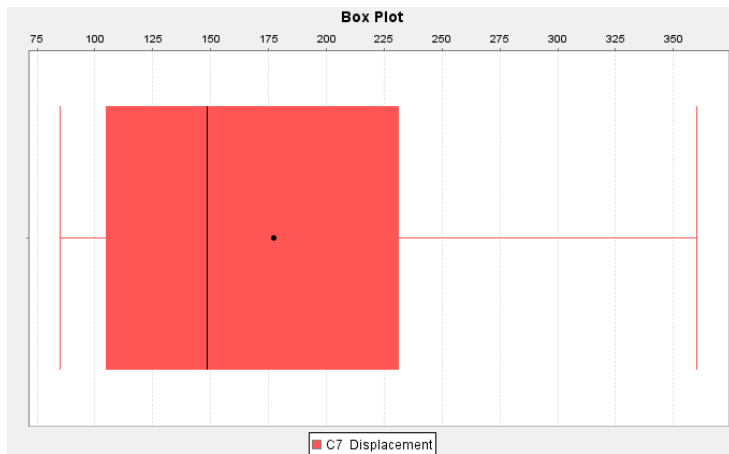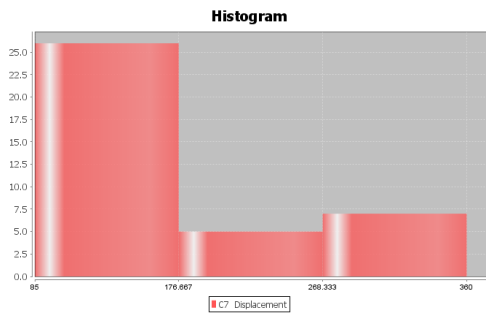
Descriptive Statistics

Variable	Mean	Standard Deviation
Displacement	177.289	88.877

Variable	Q1	Median	Q3	IQR
Displacement	103.25	148.5	237.75	134.5

Variable	Min	Max	Range
Displacement	85.0	360.0	275.0

Variable	N total
Displacement	38

a) What is the data measuring and what are the units?
b) How many numbers are in the data set?
c) What is the shape of the data set?
d) What is the minimum value?
e) What is the maximum value?
f) What is the average (center)? *(Give the number and the name of the statistic used.)*
g) How much typical spread does the data set have?
   (Give the number and the name of the statistic used.)
h) Find two numbers that typical values fall in between.
i) List all high outliers in this data set. If there are no high outliers, put "none".
j) List all low outliers in this data set. If there are no high outliers, put "none".

13. Classify each of the following statistics as a measure of center, spread or position.

   a) Q1
   b) Mean
   c) Variance
   d) Standard Deviation
   e) Minimum
   f) Q3
   g) Mode
   h) IQR
   i) Median
   j) Range
   k) Maximum
   l) Midrange

14. Define each of the following statistics and describe when that statistic should be used.

      a) Q1
      b) Mean
      c) Variance
      d) Standard Deviation
      e) Minimum
      f) Q3
      g) Mode
      h) IQR
      i) Median
      j) Range
      k) Maximum
      l) Midrange