Student: Benny Jones **Date:** 7/6/16	**Instructor:** Thara Lowndes **Course:** Math 101 -Summer 2016-Sec. 953 (Choden)	**Assignment:** HW 11: Characterizing Data 1

1. Define and distinguish among mean, median, and mode.

Choose the correct description of the mean below.

- **A.** The mean is the most common value in a data set. It can be strongly affected by outliers.
- **B.** The mean is the most common value in a data set. It is not affected by outliers.
- **C.** The mean is the sum of all the values divided by the number of values. It is not affected by outliers.
- **D.** The mean is the middle value in a data set. It can be strongly affected by outliers.
- **E.** The mean is the middle value in a data set. It is not affected by outliers.
- **F.** The mean is the sum of all the values divided by the number of values. It can be strongly affected by outliers.

Choose the correct description of the median below.

- **A.** The median is the sum of all the values divided by the number of values. It is not affected by outliers.
- **B.** The median is the most common value in a data set. It can be strongly affected by outliers.
- **C.** The median is the most common value in a data set. It is not affected by outliers.
- **D.** The median is the sum of all the values divided by the number of values. It can be strongly affected by outliers.
- **E.** The median is the middle value in a data set. It can be strongly affected by outliers.
- **F.** The median is the middle value in a data set. It is not affected by outliers.

Choose the correct description of the mode below.

- **A.** The mode is the sum of all the values divided by the number of values. It can be strongly affected by outliers.
- **B.** The mode is the middle value in a data set. It can be strongly affected by outliers.
- **C.** The mode is the most common value in a data set. It can be strongly affected by outliers.
- **D.** The mode is the middle value in a data set. It is not affected by outliers.
- **E.** The mode is the most common value in a data set. It is not affected by outliers.
- **F.** The mode is the sum of all the values divided by the number of values. It is not affected by outliers.

2. Decide whether each of the following statements makes sense (or is clearly true) or does not make sense (or is clearly false). Explain your reasoning.

I made a distribution of 15 apartment rents in my neighborhood. One apartment had a much higher rent than all of the others, and this outlier caused the mean rent to be higher than the median rent.

Choose the correct answer below.

- **A.** The statement does not make sense because an outlier with a large value increases the median, but does not affect the mean.
- **B.** The statement makes sense because an outlier with a large value increases the mean, but does not affect the median.
- **C.** The statement does not make sense because an outlier with a large value increases the mean and the median by the same amount.
- **D.** The statement does not make sense because an outlier with a large value decreases the mean, but does not affect the median.

3. Compute the mean, median, and mode of the following data set.

5.74 5.75 5.79 5.78 5.67 5.62

The mean is _____. (Round to two decimal places as needed.)

The median is _____. (Round to two decimal places as needed.)

What is the mode? Select the correct choice below, and, if necessary, fill in the answer box in your choice.

- **A.** There is more than one mode. The values are _____. (Use a comma to separate answers as needed.)
- **B.** There is one mode with a value of _____.
- **C.** There is no mode.

4. Compute the mean, median, and mode of the following data set.

195 190 179 213 205 200 197

The mean is _____. (Round to two decimal places as needed.)

The median is _____. (Round to two decimal places as needed.)

What is the mode? Select the correct choice below, and, if necessary, fill in the answer box in your choice.

- **A.** There is more than one mode. The values are _____. (Use a comma to separate answers as needed.)
- **B.** There is one mode with a value of _____.
- **C.** There is no mode.

5. The following table gives the total area in square miles (land and water) of seven states. Complete parts (a) and (b).

State	Area
1	52,300
2	615,800
3	114,700
4	53,900
5	159,400
6	104,300
7	6,500

a. Find the mean area and median area for these states.

The mean is _____ square miles.
(Round to the nearest integer as needed.)

The median is _____ square miles.
(Round to the nearest integer as needed.)

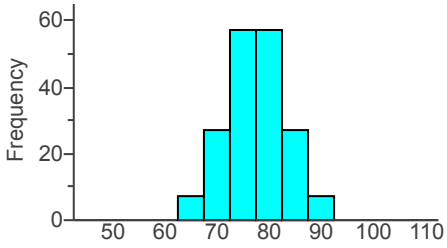
b. Which state is an outlier on the low end? If you eliminate this state, what are the new mean and median areas for this data set?

State _____ is an outlier on the low end.

The new mean is _____ square miles.
(Round to the nearest integer as needed.)

The new median is _____ square miles.
(Round to the nearest integer as needed.)

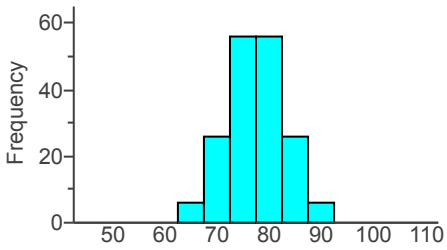
6. The histogram to the right shows the times between eruptions of a geyser for a sample of 300 eruptions (which means 299 times between eruptions). Over the histogram, draw a smooth curve that captures its general features. Then classify the distribution according to its number of peaks and its symmetry or skewness.



Chose the correct answer below.

- **A.** one peak, symmetric, low variation
- **B.** three peaks, symmetric, wide variation
- **C.** one peak, right-skewed, moderate variation
- **D.** two peaks, left-skewed, wide variation

7. The histogram to the right shows the times between eruptions of a geyser for a sample of 300 eruptions (which means 299 times between eruptions). Over the histogram, draw a smooth curve that captures its general features. Then classify the distribution according to its number of peaks and its symmetry or skewness.



Chose the correct answer below.

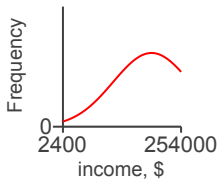
- **A.** one peak, symmetric, low variation
- **B.** three peaks, symmetric, wide variation
- **C.** one peak, right-skewed, moderate variation
- **D.** two peaks, left-skewed, wide variation

8. Suppose you study family income in a random sample of 500 families. Your results can be summarized as the mean family income was $45,000, the median family income was $31,000, the highest and lowest incomes were $254,000 and $2,400, respectively.

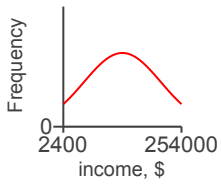
a. Draw a rough sketch of the income distribution, with clearly labeled axes. Choose the correct answer below.

A.

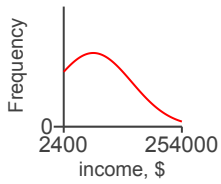
B.



C.



D.



Describe the distribution as symmetric, left-skewed, or right-skewed. Choose the correct answer below.

- right-skewed
- symmetric
- left-skewed

b. How many families in the sample earned less than $31,000? Explain how you know. Choose the correct answer below.

- **A.** 250 families, because the median is the middle value in the sorted data set.
- **B.** 375 families, because the mode is the most common value in a data set.
- **C.** 125 families, because the mean is the average value of income.

c. Based on the given data, can you determine how many families earned more than $45,000? Why or why not? Choose the correct answer below.

- **A.** No, because the number of families that earned more than $45,000 depends on the distribution.
- **B.** Yes, because the mean is the middle value in the sorted data set.

9. What are the quartiles of a distribution? How do we find them?

What are the quartiles of a distribution?

- **A.** The quartiles consist of the lowest value, the median, and the highest value of the data distribution.
- **B.** The quartiles are values that divide the data distribution into quarters.
- **C.** The quartiles consist of the mean, the median, and the standard deviation of the data distribution.
- **D.** When the data distribution is divided equally into four sets, each set of values is called a quartile.

How do we find them?

- **A.** The lower quartile is the median of the data values in the lower half of a data set. The middle quartile is the overall median. The upper quartile is the median of the data values in the upper half of data set.
- **B.** The lower quartile is the mean of the data values in the lower half of a data set. The middle quartile is the overall mean. The upper quartile is the mean of the data values in the upper half of data set.
- **C.** The lower quartile is the mean of the data values in the first quarter of a data set. The middle quartile is the overall mean. The upper quartile is the mean of the data values in the fourth quarter of data set.
- **D.** The lower quartile is the median of the data values in the first quarter of a data set. The middle quartile is the overall median. The upper quartile is the median of the data values in the fourth quarter of data set.

10. Decide whether the following statement makes sense (or is clearly true) or does not make sense (or is clearly false). Explain your reasoning.

The standard deviation for the heights of a group of 5-year-old children is smaller than the standard deviation for the heights of a group of children who range in age from 3 to 15.

Choose the correct answer below.

- **A.** The statement does not make sense because the range of data for the heights of a group of 5-year-old children is larger than the range of data for the heights of a group of children who range in age from 3 to 15.
- **B.** The statement makes sense because the range of data for the heights of a group of 5-year-old children is larger than the range of data for the heights of a group of children who range in age from 3 to 15.
- **C.** The statement does not make sense because the range of data for the heights of a group of 5-year-old children is smaller than the range of data for the heights of a group of children who range in age from 3 to 15.
- **D.** The statement makes sense because the range of data for the heights of a group of 5-year-old children is smaller than the range of data for the heights of a group of children who range in age from 3 to 15.

11. The following data sets show the ages of the first seven presidents and seven recent presidents of a certain country at the time of inauguration. Complete parts (a) through (e).
[1] Click here to view the table.

a. Find the mean, median, and range for each of the two data sets.

Find the mean, median, and range for the first seven presidents.

mean = _____
median = _____
range = _____
(Round to the nearest tenth as needed.)

Find the mean, median, and range for the seven recent presidents.

mean = _____
median = _____
range = _____
(Round to the nearest tenth as needed.)

b. Give the five-number summary for each of the two data sets. Give the five-number summary for the first seven presidents.

low = _____
lower quartile = _____
median = _____
upper quartile = _____
high = _____

Give the five-number summary for the seven recent presidents.

low = _____
lower quartile = _____
median = _____
upper quartile = _____
high = _____

c. Find the standard deviation for each of the two data sets.

Find the standard deviation for the first seven presidents.

s = _____ (Round to the nearest tenth as needed.)

Find the standard deviation for the seven recent presidents.

s = _____ (Round to the nearest tenth as needed.)

d. Based on all your results, compare and discuss the two data sets in terms of their center and variation.

- ○ **A.** The average age of the first seven presidents and the recent seven presidents are approximately the same, and variations in ages of the recent seven presidents and the first seven presidents are approximately the same.
- ○ **B.** The average age of the first seven presidents and the recent seven presidents are approximately the same, but variations in ages of the recent seven presidents are higher.
- ○ **C.** The average age of the first seven presidents and the recent seven presidents are approximately the same, but variations in ages of the recent seven presidents are lower.

1: Data Table

First 7

58 62 58 58 59 58 62

Last 7

52 57 48 65 60 42 50

12. The table to the right gives the cost of living index for six cities in Region 1 and six cities in Region 2 (using an index where 100 represents the average cost of living for all participating cities). Answer parts a through d below.

Region 1		Region 2	
A	98.3	U	155.1
B	109.7	V	111.4
C	132.9	W	144.9
D	111.9	X	181.6
E	211.9	Y	156.8
F	139.3	Z	121.5

a. Find the mean, median, and range for each of the two data sets.

Find the mean, median, and range for Region 1.

Mean = _____

Median = _____

Range = _____

(Type an integer or decimal rounded to two decimal places as needed.)

Find the mean, median, and range for Region 2.

Mean = _____

Median = _____

Range = _____

(Type an integer or decimal rounded to two decimal places as needed.)

b. Give the five-number summary for each of the two data sets.

Give the five number summary for Region 1.

Low Value = _____ Lower Quartile = _____

Median = _____ Upper Quartile = _____

High Value = _____

(Type an integer or decimal rounded to two decimal places as needed.)

Give the five-number summary for Region 2.

Low Value = _____ Lower Quartile = _____

Median = _____ Upper Quartile = _____

High Value = _____

(Type an integer or decimal rounded to two decimal places as needed.)

c. Find the standard deviation for region 1.

The standard deviation for Region 1 is _____.

(Type an integer or decimal rounded to two decimal places as needed.)