Statistics & Business Decision Making

Homework assignment: Measures of central tendency (**first**)

1. What is the mode of a data set?
   The mode is the value that occurs most frequently in the set.

2. What is the mean of a data set?
   The mean is defined as the sum of the values in the set divided by the number of values in the set.

3. What is the median of a data set?
   The median is defined as the middle value when the numbers are arranged in increasing or decreasing order.

4. The mean, median and mode of a distribution are measures of

   (a) the number of observations.
   (b) variability or spread.
   (c) the center. (correct)
   (d) both (b) and (c)
   (e) none of these.

5. Fifty percent of the observations will be at or above the

   (a) maximum   (b) mean   (c) median (d) third quartile   (e) first quartile

6. Find the mean of these numbers { 3 -7 5 11 -2}

Mean = (3-7+5+11-2)/5 = 10/5 = 2

7. When there is an outlier, which is the better measure of central tendency?

(a)   the mean   (b) the median (correct)

8. The science test grades are posted. The class did very well. All students taking the test scored over 75. Unfortunately, 4 students were absent for the test and the computer listed their scores as 0 until the test is taken. Assuming that no score repeated more times than the 0's, what measure of central tendency would most likely give the best representation of this data?
The mode = 0, but the mode does not fairly represent the center. Since the distribution is not symmetric, the median is a better measure of center than the mean.

9.   Here are the numbers of hours that each of a random sample of college students from a large class studied for their most recent statistics test:

$$2 \quad 4 \quad 22 \quad 6 \quad 1 \quad 4 \quad 1 \quad 5 \quad 7 \quad 4$$

(a)   Find the mode of this data set. 4

(b)   Find the median of this data set. Show your method.
Rearranging: 1 1 2 4 <u>4 4</u> 5 6 7 22
Median = (4+4)/2 = 4

(c)   Explain what the median tells you about the amount of time students studied for the test.
Fifty percent of the students studied less than or equal to 4 hours.
Fifty percent of the students studied greater than or equal to 4 hours.

(d)   Calculate the mean of this data set. Explain why the mean is larger than the median.
Mean = 56/10 = 5.6
The outlier (22) tends to increase the mean.

(e)   It was later discovered that the student who wrote 22 hours actually studied only 12 hours. Which, if any, of the following would change as a result: mean, median, quartiles, a boxplot, sample size of the data? Describe how these would change.
The mean would decrease. The maximum on the box plot would decrease. Everything else would remain the same.

(f)   Suppose that 10 were added to each value in the original distribution, calculate the new mean and median? How do these compare to the mean and median of the original distribution.
New mean = 156/10 = 15.6. The new mean and new median are 10 greater than the old mean and old median.

10. Which of these is *not* true of the mean $\bar{x}$ of the lengths in inches of a sample of brook trout?

       (a) $\bar{x}$ must take a value greater than 0.
       (b) $\bar{x}$ is measured in inches.
       (c) $\bar{x}$ would not change if we measured these trout in centimeters instead of inches.
       (d) Both (b) and (c).
       (e) Both (a) and (c).
1 cm = 0.39 inches. If we multiply the lengths of all the fish by (0.39 inches/1cm), then the mean length in centimeters would be multiplied by 0.39 inches/1cm.

11.   The weekly salaries of six employees at McDonalds are $140, $220, $90, $180, $140, $200. For these six salaries, find: (a) the mean (b) the median (c) the mode.
Mean = $970/6 = $161.67   REARRANGING: 90   140   140   180   200   220

Median = (140+180)/2 = $160

Mode = $140

2

12. Use the frequency table below to find these measures of central tendency.

Number of pets	
Number	Frequency
1	3
2	6
3	4
4	2

REARRANGING: 1 1 1 2 2 2 2 2 2 3 3 3 3 4 4

Mean = $(35/15)$ = <u>2.3</u>

Median= <u>2</u>

Mode = <u>2</u>

13. A student has made the following grades on his first four math tests:
80, 86, 92, and 95.

What grade must he make on the next test to raise his average to a 90?
Mean = $(80+86+92+95+x)/5 = 90$

➔ $80+86+92+95+x = 450$ → $x = 97$

14.   Given the set of data {8 4 5 5 6 2 }, demonstrate that adding a constant (e.g., 5) to every score increases all measures of central tendency by that amount.

Original data {2 4 5 5 6 8}   New data after adding 5 to each score: {7 9 10 10 11 13}

Mean = $\sum x$ /n = 30/6 = 5   Mean = 60/6 = 10

Median = 5   Median = 10

Mode = 5   Mode = 10

15.   Given the same set of data {8 4 5 5 6 2 }, show that multiplying each score by a constant multiplies all measures of central tendency by that constant.

Original data {2 4 5 5 6 8} New data after multiply each score by 5: {10 20 25 25 30 40}

Mean = $\sum x$ /n = 30/6 = 5   Mean = 150/6 = 25

Median = 5   Median = 25

Mode = 5   Mode = 25

A distribution is **symmetric** if the right and left sides of the graph are approximately mirror images of each other. A distribution is **skewed to the right** if the right side of the graph extends much further out than the left side. It is **skewed to the left** if the left side of the graph extends much further out than the right side.
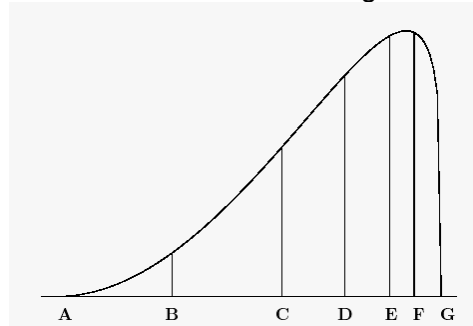
16. The figure above is the density curve of a distribution. This distribution is
   (a) Roughly symmetric.
   (b) Skewed to the left. (correct)
   (c) Skewed to the right.
   (d) Positively correlated.
   (e) Negatively correlated.

17. For a distribution above, which of the following is true?
   (a) mean < median < mode (correct)
   (b) median < mean < mode
   (c) mode < mean < median
   (d) mode < median < mode

18. Sketch a histogram for a distribution that is skewed to the right (positively skewed).