Name:_____ Class:_____ Date:_____

CHAPTER 3 - Measures of Central Tendency and Dispersion

1. What is the purpose of measures of central tendency?
   a. to find the most typical value of a distribution of scores
   b. to find the most surprising value of a distribution of scores
   c. to find the most significant value of a distribution of scores
   d. to find the most important value of a distribution of scores

ANSWER: a

2. Which statement describes the three commonly used measures of central tendency (mode, median, and mean)?
   a. They define "typical" or "average" in different ways and will usually have different values.
   b. They will always fall in the same order: the mean will have the lowest value, the median will always be in the middle, and the mode will have the highest value.
   c. They will always have the same value.
   d. They will always fall in the same order: the mean will have the highest value, followed by the median and the mode.

ANSWER: a

3. How does the mode measure "central tendency"?
   a. in terms of the most average case
   b. in terms of the most important score
   c. in terms of the most central case
   d. in terms of the most common score

ANSWER: d

4. What do measures of dispersion provide?
   a. an indication of the size of the sample
   b. an indication of the adequacy of the selection criteria for the sample
   c. an indication of the typical or most common score
   d. an indication of the variety within the distribution of scores

ANSWER: d

5. Which data set shows the greatest variability?
   a. 0, 5, 9
   b. 9, 10, 11
   c. 31, 33, 35
   d. 50, 60, 170

ANSWER: d

6. Which measure of central tendency would be the MOST appropriate for summarizing data on the gender of university professors?
   a. mean
   b. mode
   c. median
   d. quartile

ANSWER: b

CHAPTER 3 - Measures of Central Tendency and Dispersion

7. What is the definition of the variation ratio (*v*)?
   a. the proportion of cases not in the modal category
   b. the proportion of cases with a score of one
   c. the proportion of cases with a score greater than one
   d. the proportion of cases in the modal category

ANSWER: a

8. Which of the following is a characteristic of a "good" measure of dispersion?
   a. It should be complex to calculate.
   b. It should be complex to interpret.
   c. It should describe the average or typical deviation of the scores.
   d. It should use only the scores in the middle of a distribution.

ANSWER: c

9. Which of the following distributions most clearly illustrates a scenario where the mode fails to represent the central tendency of a distribution?
   a. 115, 125, 132, 160, 185, 202
   b. 80, 82, 82, 82, 83, 84
   c. 315 488, 533, 696, 769, 769
   d. 1, 87, 229, 229, 229, 229, 480, 534

ANSWER: c

10. Which variation ratio (*v*) below shows the greatest degree of homogeneity, if *v* varies from 0.00 to 1.00?
   a. 0.25
   b. 0.50
   c. 0.75
   d. 1.00

ANSWER: a

11. Which of the following is an appropriate measure of dispersion for variables measured at the nominal level?
   a. the range
   b. the mode
   c. the average deviation
   d. the variation ratio

ANSWER: d

12. How does the median define "central tendency"?
   a. in terms of the most likely or probable score
   b. in terms of the central case
   c. in terms of the most common score
   d. in terms of the largest score

ANSWER: b

13. What is the median if the scores of an even number of cases are arranged from high to low?
    a. the same as the mode
    b. the middle score
    c. the average of the two middle scores
    d. the average of the highest and lowest scores

ANSWER: c

14. What is the median score in a ranked list of 25 scores?
    a. the 12^{th} case
    b. the 13^{th} case
    c. the average of the scores of the 12^{th} and 13^{th} cases
    d. the average of all of the scores

ANSWER: b

15. Why can the median NOT be found for variables measured at the nominal level?
    a. The scores cannot be ranked or ordered.
    b. The distribution is skewed.
    c. All of the cases are representative.
    d. There is no central tendency.

ANSWER: a

16. Which measure of central tendency is the MOST appropriate one to use for ordinal level variable?
    a. range
    b. median
    c. mean
    d. mode

ANSWER: b

17. In a classroom of grade three students, more students are right-handed than left-handed. What does a variation ratio of .35 for a "hand preference" nominal variable represent?
    a. less dispersion than if the variation ratio was .30
    b. more dispersion than if the variation ratio was .65
    c. 1 – (35/100) = .65 of the variation possible in hand preference
    d. 1 – (65/100) = .35 of the variation possible in hand preference

ANSWER: d

18. Which of the following is among the problems with using the range (*R*) as a measure of dispersion?
    a. *R* can be used only for nominal level variables.
    b. *R* is very difficult to calculate.
    c. *R* ignores the most extreme scores.
    d. *R* is based on only the most extreme scores.

ANSWER: d

19. Which of the following has the same location as the first quartile in a distribution of scores?
    a. the first percentile
    b. the third percentile
    c. the twenty-fifth percentile
    d. the seventy-fifth percentile

ANSWER: c

20. Which of the following is the second quartile (Q_2) equal to in value?
    a. the median
    b. the mode
    c. the range
    d. the mean

ANSWER: a

21. What can be concluded if a student's score on the test is the same as the third quartile (Q_3)?
    a. that the student distribution of the scores is skewed
    b. that the student scored higher than 75% of the people who took the test
    c. that the student scored higher than 25% of the people who took the test
    d. that the student's score is "typical" because it is the same value as the median

ANSWER: b

22. Which of the following describes the symbol Q_3?
    a. the third quintile
    b. the first quartile
    c. the third quartile
    d. the third percentile

ANSWER: c

23. The following is the monthly cost of four different one-bedroom apartments: $1,200; $1,050, $1,320, $1,250. What is the median cost?
    a. $1,200
    b. $1,250
    c. the average of $1,200 and $1,250
    d. the average of $1,050 and $1,320

ANSWER: c

24. Why is the interquartile range sometimes preferred to the range when describing the dispersion of a distribution?
    a. The interquartile range is more sensitive to outlying values in the distribution.
    b. The range is more sensitive to outlying values in the distribution.
    c. The interquartile range is a more concise summary of the distribution's dispersion.
    d. The interquartile range is a measure of the distribution's dispersion, whereas the range is actually an indication of its central tendency.

ANSWER: b

25. What is one advantage of the interquartile range (*Q*) over the range (*R*)?
    a. *Q* ignores the first and third quartiles.
    b. *Q* can be used for nominal level variables.
    c. *Q* is based on only the middle 50% of the scores.
    d. *Q* includes the most extreme scores.

ANSWER: c

26. For which of the following would interquartile range be the MOST appropriate measure of dispersion?
    a. for nominal level variables
    b. for discrete interval-ratio level variables only
    c. for continuous interval-ratio level variables only
    d. for ordinal level variables

ANSWER: d

27. What is the purpose of the expression $\Sigma(X_i)$?
    a. It directs the researcher to find the middle score.
    b. It directs the researcher to find the average of the scores.
    c. It directs the researcher to find the median.
    d. It directs the researcher to find the sum of the scores.

ANSWER: d

28. On a survey, age was divided into three categories: younger than 18 years, 18 to 21 years, and older than 21 years. The researcher computed the mean age of the respondents by adding up the scores and then dividing by the number of cases. Why was this a mistake?
    a. because the data are interval-ratio and should not be treated as ordinal
    b. because the data are only nominal and should not be treated as ordinal
    c. because the data are continuous and should not be treated as discrete
    d. because the data are ordinal and should not be treated as interval-ratio

ANSWER: d

29. Which measure of central tendency is the MOST appropriate one to use for the variable age (measured in years)?
    a. mode
    b. range
    c. quartile
    d. mean

ANSWER: d

Charity

The following four problems refer to a data set of six volunteers at a charitable organization.

People	Years served in organization	Employment status
Steve	1	Part-time worker
Hasan	5	Full-time worker
Amanda	5	Retired

CHAPTER 3 - Measures of Central Tendency and Dispersion

Xi	2	Student
Jorge	3	Full-time worker
Candace	2	Full-time worker

30. What is mean of years served in the organization?
    a. 2
    b. 2.5
    c. 3
    d. 5

ANSWER: c

31. What is the median of years served in the organization?
    a. 2
    b. 2.5
    c. 3
    d. Both 2 and 5

ANSWER: b

32. What is the mode of employment status?
    a. student
    b. part-time worker
    c. full-time worker
    d. there is no mode for this distribution

ANSWER: c

33. What is the mean of employment status?
    a. student
    b. part-time worker
    c. full-time worker
    d. There is no mean for this distribution.

ANSWER: d

34. Which measure of central tendency is affected by every score in the distribution?
    a. range
    b. mode
    c. median
    d. mean

ANSWER: d

35. How does the standard deviation formula differ for samples and populations?
    a. X_i is used to represent the score of a variable in the population, while *N* is used to represent scores in a sample.
    b. X_i is used to represent the score of a variable in the sample, while *N* is used to represent scores in a population.
    c. μ is used to represent the population mean, while $\overline{X}$ is used to represent a sample's mean.

d. Standard deviations can only be calculated for samples, not for populations.

ANSWER: c

36. What is the "least squares" principle?

a. The mean is the point in a distribution around which the variation of the scores is minimized.

b. The range of an inter-ratio level variable can never exceed the value of its mean.

c. The interquartile range will always contain a variable's median, but not necessarily the variable's mean.

d. Taking the square root of a variable's mean will create a more symmetrical distribution.

ANSWER: a

37. What two pieces of information are universally important for summarizing data?

a. central tendency and dispersion

b. frequency distribution and histogram

c. mean and mode

d. mean and standard deviation

ANSWER: a

38. What does the expression $\sum\left(X_i - \overline{X}\right)$ instruct us to do?

a. sum the scores

b. subtract every score from the mean

c. sum the scores and then subtract the mean

d. subtract the mean from each score and add up the deviations

ANSWER: d

39. The mean square footage of a 20-home Ontario suburban neighbourhood is 1775 square feet. The developers are considering building a 13,500-square-foot mansion on the edge of the neighbourhood. What measure of the neighbourhood's central tendency would be most affected if the mansion was built?

a. the mode

b. the median

c. the mean

d. the variation ratio

ANSWER: c

40. The distribution of student test scores is severely left skewed. What would this mean?

a. that the mean of the distribution is less than the median

b. that the mode of the distribution is less than the median

c. that the median of the distribution is less than the mean

d. that the mode of the distribution is less than the mean

ANSWER: a

41. When does a distribution's mean necessarily equal its median?

a. when the distribution has two modes

b. when the distribution is only slightly skewed

c. when the interquartile range is half the size of the range

d. when the distribution is symmetrical

ANSWER: d

Park

The following two questions refer to the following Table, which shows park space in five neighbourhoods of a fictional city.

City neighbourhood	Square footage of park space
Montclair East	25,471
South Lawndale	30,755
Everton	180
East Lakeside	26,983
Dunsterville	22,870

42. Which of the following describes the distribution of park space?

a. negatively skewed

b. positively skewed

c. symmetric

d. unclear from the information given

ANSWER: a

43. What would be the most appropriate measure of central tendency for the distribution of park space?

a. the mode

b. the median

c. the mean

d. the interquartile range

ANSWER: b

44. What is the term for a distribution that has its peak on the left and a tail extending far to the right?

a. positively skewed

b. negatively skewed

c. normally distributed

d. bimodally distributed

ANSWER: a

45. Which statement describes the distribution if a distribution of test scores has a mean of 70 and a median of 80?

a. It has a positive skew.

b. It has few very high scores.

c. It has a negative skew.

d. It has a mode of 75.

ANSWER: c

46. When will the mean and the median in any distribution have the same value?

a. when the distribution is positively skewed

b. when the distribution is negatively skewed
c. when the distribution has extreme scores
d. when the distribution is symmetrical

ANSWER: d

47. In an election campaign debate about the level of prosperity in Canada, the Prime Minister says: "Average income is $53,000." One of his opponents responds: "The average Canadian earns only $38,000." Both statements are true. Why is it possible that both statements are true?

a. because the opponent is using the mean
b. because the Prime Minister is using the mean
c. because the Prime Minister is citing the median
d. because the opponent is using the mode

ANSWER: b

48. In preparing a report showing mid-semester grades, an instructor notes that the mean, median, and mode are all exactly 76.00. What can the instructor conclude from this observation?

a. that the distribution of grades is unskewed
b. that the distribution has extreme scores
c. that there is a positive skew in the distribution because a few students have very high grades
d. that there is a negative skew in the distribution because a few students have very low grades

ANSWER: a

49. Which is the preferred measure of central tendency for the variable income distribution, which almost always has a positive skew?

a. range
b. median
c. mode
d. mean

ANSWER: b

50. Which statement describes how the standard deviation (*s*) differs from the average deviation (AD)?

a. The standard deviation decreases in value as distributions become more heterogeneous.
b. The standard deviation squares the deviations to eliminate minus values.
c. The standard deviation is based on the median rather than the mean.
d. The standard deviation does not use all the scores.

ANSWER: b

51. Which measure of dispersion is MOST important and MOST commonly used in social science research?

a. the range
b. the median
c. the standard deviation
d. the average deviation

ANSWER: c

52. The mean height of a 22-member university women's hockey team is 175.2 cm. If we subtracted each of the 22 players' heights from the mean value and added these differences together, what would the overall sum of the deviations be?

a. It is impossible to tell from the information given.

b. 0

c. 175.2/2 = 87.6

d. 175.2

ANSWER: b

53. The distribution of film genre preferences in a retirement community of 150 seniors is as follows: Drama—37%; Comedy—32%; Musicals—8%; Action—18%; Documentaries—5%. Why is it impossible to calculate a standard deviation for this variable?

a. The variable's categories are not mutually exclusive.

b. The variable's categories do not sum to 100%.

c. The variable is too skewed to calculate the variance.

d. The variable does not have a mean value.

ANSWER: d

54. The mean test grade for an introductory sociology class was 72. The variance was 9. How would we calculate the standard deviation for the distribution?

a. $\sqrt{9} = 3$

b. (72 - 9)/9 = 7

c. $(72 - \sqrt{9}) = 69$

d. $9^2 = 81$

ANSWER: a

55. Four students have applied to a co-op education program, and only one can be accepted. They have each taken six skills tests and all the students have exactly the same average score. The standard deviation of the test scores are: Student A = 3.12, Student B = 0.27, Student C = 13.45, Student D = 6.45. Which student should be accepted to the program if consistency of performance is among the criteria for acceptance?

a. Student A

b. Student B

c. Student C

d. Student D

ANSWER: b

56. What would be the preferred measure of dispersion for a variable where the income of a sample was measured in dollars per year?

a. the average deviation

b. the standard deviation

c. the quartile deviation

d. the variation ratio

ANSWER: b

57. What occurs when a distribution of scores becomes more variable?

a. The value of the standard deviation becomes unpredictable.
b. The value of the standard deviation decreases.
c. The value of the standard deviation stays the same.
d. The value of the standard deviation increases.

ANSWER: d

58. When the lower limit for the standard deviation is 0, what is the upper limit?
a. It is a function of the number of cases.
b. It is undefined; there is no upper limit.
c. It is the score of the mean.
d. It is the score of the median.

ANSWER: b

59. Which of the following describes a situation where a distribution of scores has a mean of 30 and a range of 0?
a. The quartile deviation Q is 10.
b. There is no dispersion in the distribution.
c. The standard deviation is 1.
d. The variance is 30.

ANSWER: b

60. Which of the following is the lowest valid score for a standard deviation?
a. -3
b. -1
c. 0
d. .5

ANSWER: c

61. What is the difference between variance and standard deviation?
a. Variance is an indicator of dispersion, while standard deviation is an indicator of central tendency.
b. Standard deviations are more affected by outlying values than are variances.
c. Standard deviations are always smaller than variances when variance >1.
d. There are no differences; the terms are synonyms.

ANSWER: c

62. If extreme variability in pulse rate indicates a health problem, which of the following statistics summarizing a patient's six hourly pulse checks would be most alarming to a physician?
a. standard deviation = .25 beats per minute
b. standard deviation = 1.6 beats per minute
c. standard deviation = 11.8 beats per minute
d. standard deviation = 28.0 beats per minute

ANSWER: d

63. What would happen to the value of the standard deviation if the standard deviation for a distribution of 20 scores was calculated and then the 5 highest scores were removed and recalculated?

a. It would increase by 5.
b. It would stay the same.
c. It would increase by 2.
d. It would decrease.

ANSWER: d

64. Researchers wish to compare the dispersion of student grades across eight universities of varying sizes. What strategy would allow them to properly compare grade dispersion across the eight distributions?

a. adding up the deviations for the smallest and the largest university
b. dividing the squared deviations by the number of cases (*n*) for each university
c. calculating the mode for each of the 8 universities
d. calculating the median for each of the 8 universities

ANSWER: b

Mining

For the following two questions, refer to the table below. The table shows means and standard deviations in monthly accident events for two mining companies in 2014.

Jackson Minerals Inc.	H & W Energy Resources
$X?3^{-} = 5.2$	$X?3^{-} = 6.6$
$s = 2.6$	$s = 3.9$

65. Which of the following conclusions can be drawn from these data?

a. Jackson Minerals had more accidents in 2014 and more month-to-month variability in number of accidents.
b. Jackson Minerals had more accidents in 2014, but less month-to-month variability in number of accidents.
c. H & W Energy Resources had more accidents in 2014 and more month-to-month variability in number of accidents.
d. H & W Energy Resources had more accidents in 2014, but less month-to-month variability in number of accidents.

ANSWER: c

66. What would be the variance for Jackson Minerals?

a. $\sqrt{2.6} = 1.61$
b. $5.2 - 2.6 = 2.6$
c. $2.6^2 = 6.76$
d. $5.2 + 2.6 = 7.8$

ANSWER: c

67. What is the mode an appropriate measure of central tendency for?

a. nominal data
b. ordinal data
c. interval-ratio data
d. all of the above

ANSWER: d

68. What is the difference between central tendency and dispersion in a distribution of data?
· Central tendency gives us the most common (mode), the midpoint (median), or the average (mean) score in a distribution. One could say that these statistics summarize a lot of data to give us some sense of what is "typical."
· Dispersion, on the other hand, conveys variability. It indicates how the data are spread around their central tendency. In other words, do the data tend to cluster around the mode, median, or mean? Or are the data very spread out across a wide range of values?
ANSWER: Answers will vary

69. A sociology professor is wondering if his current students are performing at a level that is better or worse than that of his past students. He has given the same final exam to his current class that he gave to his class 10 years ago. Compute the mode, mean, median, range, and standard deviation for both classes.

Exam scores: current class				
35	45	65	75	87
80	69	71	53	90
99	95	70	82	73
93	67	61	57	74
72	77	71	81	83

Exam scores: class 10 years ago				
56	77	75	76	59
74	51	89	55	79
67	77	69	91	68
90	65	79	69	79
87	86	98	91	95

ANSWER:

Current class:	Class 10 years ago:
Mode = 71	Mode = 79
Median = 73	Median = 77
Mean = 73	Mean = 76.08
Range = 64	Range = 47
s = 14.86	s = 12.70

70. Summarized below is some information about auto theft rates (number of auto thefts per 100,000 population) for a sample of 17 cities in two different years. Express this statistical information in words. What changes occurred in the overall shape of the distribution of this variable? In central tendency? In dispersion?

	1980	2010
Mean	150.32	125.17
Median	117.17	123.01
Standard deviation	12.23	7.01

ANSWER:

In 1980, the mean auto-theft rate was much higher than the median. This indicates a positively skewed distribution of rates among the 17 cities, in which an extreme score(s) or outlier(s) inflates the mean relative to the median. The median therefore provides a more useful, accurate measure of central tendency for auto-theft rates in 1980. Overall, the median auto-theft rate increased slightly between 1980 (117.17) and 2010 (123.01).

By 2010, the mean and median rates are very similar, and thus the distribution is likely symmetric. Moreover, the standard deviation in 1980 is much higher than in 2010. Again, this likely reflects the positively skewed

Name:_____ Class:_____ Date:_____

distribution in 1980 (as with the mean, the standard deviation is inflated by outliers).