***Business Analytics, 2e, GE* (Evans)**

**Chapter 4 Descriptive Statistical Measures**

1) Which of the following is an example of a subset?

A) The number of individuals who have a Ford car

**B) The number of IT employees out of all employees working in an office of Google**

C) The number of members in the Democratic Party

D) The population of Canada

Answer: B

Diff: 1

Blooms: Understand

Topic: Populations and Samples

LO1: Explain the difference between a population and a sample.

LO2: Explain the fundamentals of sampling methods, experiment designs, and sampling distributions

2) The purpose of sampling is to \_\_\_\_\_\_\_\_.

A) enumerate all the values in the population

B) measure all items of interest for a particular interest or investigation

**C) obtain sufficient information to draw a valid inference about a population**

D) calculate all variables and observations within a population

Answer: C

Diff: 1

Blooms: Remember

Topic: Populations and Samples

LO1: Explain the difference between a population and a sample.

LO2: Explain the fundamentals of sampling methods, experiment designs, and sampling distributions

3) In statistical notation, the elements of a data set are typically labeled as \_\_\_\_\_\_\_\_.

A) summation operators

B) letters in capitals

C) Greek letters

**D) subscripted variables**

Answer: D

Diff: 1

Blooms: Remember

Topic: Populations and Samples

LO1: Understand statistical notation.

LO2: Compare and contrast methods of summarizing and describing data

4) According to statistical notations, what does Σ stand for?

**A) to act as a summation operator**

B) to represent sample statistics

C) to represent population measures

D) to represent the number of items in a population

Answer: A

Diff: 1

Blooms: Remember

Topic: Populations and Samples

LO1: Understand statistical notation.

LO2: Compare and contrast methods of summarizing and describing data

5) Which of the following measures of location is calculated using the formula, where *n* is the number of observations?

A) midrange

**B) sample mean**

C) mode

D) median

Answer: B

Diff: 1

Blooms: Remember

Topic: Measures of Location

LO1: List different measures of location.

LO2: Compare and contrast methods of summarizing and describing data

6) Which of the following is true from the equation: , where is the mean of the sample?

**A) The sum of the deviations above the mean are the same as the sum of the deviations below the mean**

B) Half the data lie above the mean of the values

C) The specific set of values does not have any outliers affecting the mean

D) Half the data lie below the mean of the values

Answer: A

Diff: 2

Blooms: Remember

Topic: Measures of Location

LO1: Compute the mean, median, mode, and midrange of a set of data.

LO2: Compare and contrast methods of summarizing and describing data

Use the table below to answer the following question(s).

Below is a table showing the costs per order of items bought by a computer hardware store.

|  |  |  |
| --- | --- | --- |
| **Supplier** | **Item Description** | **Cost per Order** |
| Aloham Technologies | Monitors | $250 |
| Aloham Technologies | Hard Disks | $220 |
| Aloham Technologies | Power Supply | $ 88 |
| Aloham Technologies | Graphics Card | $300 |
| Aloham Technologies | Processor | $325 |
| Aloham Technologies | Speakers | $ 88 |
| MindRootHardwares | Flashdrive | $ 9 |
| MindRootHardwares | Graphics Card | $299 |
| MindRootHardwares | CD-ROM | $ 88 |
| MindRootHardwares | Bluetooth Device | $ 80 |

7) Calculate the mean cost per order.

A) $220.54

**B) $174.70**

C) $159

D) $88

Answer: B

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Location

LO1: Compute the mean, median, mode, and midrange of a set of data.

LO2: Compare and contrast methods of summarizing and describing data

8) Calculate the mode for the cost per order data.

A) $174.17

B) $325

**C) $88**

D) $206.5

Answer: C

Diff: 1

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Location

LO1: Compute the mean, median, mode, and midrange of a set of data.

LO2: Compare and contrast methods of summarizing and describing data

9) Calculate the median for the cost per order data.

A) $325

**B) $154**

C) $174.17

D) $88

Answer: B

Diff: 1

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Location

LO1: Compute the mean, median, mode, and midrange of a set of data.

LO2: Identify different business uses for statistics and the major statistical tools businesses use

10) One of the measures of location is calculated as 88. Which of the following measures of location corresponds to this value?

A) mean

**B) mode**

C) median

D) midrange

Answer: B

Diff: 1

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Location

LO1: List different measures of location.

LO2: Identify different business uses for statistics and the major statistical tools businesses use

11) The measure of location that specifies the middle value when the data are arranged from least to greatest is the \_\_\_\_\_\_\_\_.

A) outlier

B) mean

**C) median**

D) mode

Answer: C

Diff: 1

Blooms: Remember

Topic: Measures of Location

LO1: List different measures of location.

LO2: Compare and contrast methods of summarizing and describing data

12) Which of the following is true for a median?

A) A median is only meaningful for interval or ordinal data and not for ratio data.

B) Medians can be calculated no matter how the data is arranged.

C) Medians are affected by outliers.

**D) For an even number of observations, the median is the mean of the two middle numbers.**

Answer: D

Diff: 1

Blooms: Remember

Topic: Measures of Location

LO1: List different measures of location.

LO2: Compare and contrast methods of summarizing and describing data

13) Which of the following is a difference between a mean and a median?

A) A mean divides the data half above it and half below it; a median does not.

**B) A median is not affected by outliers; a mean is affected by outliers.**

C) A mean is an observation that occurs most frequently; a median is the average of all observations.

D) A median is not meaningful for ratio data; a mean is meaningful to ratio data.

Answer: B

Diff: 1

Blooms: Remember

Topic: Measures of Location

LO1: List different measures of location.

LO2: Compare and contrast methods of summarizing and describing data

14) The \_\_\_\_\_\_\_\_ is the observation that occurs most frequently.

**A) mode**

B) mean

C) outlier

D) median

Answer: A

Diff: 1

Blooms: Remember

Topic: Measures of Location

LO1: List different measures of location.

LO2: Compare and contrast methods of summarizing and describing data

15) Which of the following is true for a mode?

A) Modes are the mid values of data arranged from least to greatest.

B) Modes are affected by outliers.

C) Modes cannot be used in data having repetitious values.

**D) The highest bar in a histogram is its mode.**

Answer: D

Diff: 1

Blooms: Remember

Topic: Measures of Location

LO1: List different measures of location.

LO2: Compare and contrast methods of summarizing and describing data

16) The \_\_\_\_\_\_\_\_ is the average of the greatest and least values in the data set.

A) mean

B) median

**C) midrange**

D) mode

Answer: C

Diff: 1

Blooms: Remember

Topic: Measures of Location

LO1: List different measures of location.

LO2: Compare and contrast methods of summarizing and describing data

17) Which of the following types of conditions is most likely to render a midrange value useless?

A) having repetitious values in the data set

B) having the data arranged from least to greatest in value

C) having a small sample size

**D) having extreme values in a data**

Answer: D

Diff: 1

Blooms: Remember

Topic: Measures of Location

LO1: List different measures of location.

LO2: Compare and contrast methods of summarizing and describing data

18) Which of the following is a similarity between a midrange and a mean?

A) Both measures are calculated using all the values in a data set.

**B) Both measures are affected by outliers.**

C) Both measures divide the data into two equal halves.

D) Both measures can only be used for small sample sizes.

Answer: B

Diff: 1

Blooms: Understand

Topic: Measures of Location

LO1: List different measures of location.

LO2: Compare and contrast methods of summarizing and describing data

19) Which of the following is an example of a measure of dispersion?

A) median

B) mode

**C) variance**

D) midrange

Answer: C

Diff: 1

Blooms: Remember

Topic: Measures of Dispersion

LO1: List different measures of dispersion.

LO2: Compare and contrast methods of summarizing and describing data

20) Which of the following describes dispersion in statistics?

**A) the degree of variation in the data**

B) the central value in the data

C) the frequency of values in the data

D) the measure of outliers in the data

Answer: A

Diff: 1

Blooms: Remember

Topic: Measures of Dispersion

LO1: List different measures of dispersion.

LO2: Compare and contrast methods of summarizing and describing data

21) The difference between the first and third quartiles is referred to as the \_\_\_\_\_\_\_\_.

A) standard deviation

B) variance

**C) interquartile range**

D) midrange

Answer: C

Diff: 1

Blooms: Remember

Topic: Measures of Dispersion

LO1: List different measures of dispersion.

LO2: Compare and contrast methods of summarizing and describing data

22) Which of the following is true of midspread?

A) It is an example of a measure of location.

B) It is calculated by finding the difference between the highest and lowest values in the data set.

C) It is affected by extreme values.

**D) It is calculated using only the middle 50% of the data.**

Answer: D

Diff: 1

Blooms: Understand

Topic: Measures of Dispersion

LO1: List different measures of dispersion.

LO2: Compare and contrast methods of summarizing and describing data

23) Which of the following describes variance?

A) It is the difference between the maximum value and the minimum value in the data set.

B) It is the difference between the first and third quartiles of a data set.

C) It is the average of the squared deviations of the observations from the mean.

D) It is the average of the greatest and least values in the data set.

Answer: C

Diff: 1

Blooms: Remember

Topic: Measures of Dispersion

LO1: List different measures of dispersion.

LO2: Compare and contrast methods of summarizing and describing data

24) Which of the following is true of variance?

**A) The formula to calculate variance of a population is not the same as the formula to calculate variance of a sample.**

B) Its value is inversely proportional to the degree to which the data is spread from the mean.

C) It only requires the middle 50% of data to be calculated.

D) It is the square root of standard deviation.

Answer: A

Diff: 1

Blooms: Understand

Topic: Measures of Dispersion

LO1: List different measures of dispersion.

LO2: Identify different business uses for statistics and the major statistical tools businesses use

25) Which of the following describes standard deviation?

A) It is the average of the greatest and least values in the data set.

**B) It is the square root of the variance.**

C) It is the difference between the first and third quartiles of a data set.

D) It is the average of the squared deviations of the observations from the mean.

Answer: B

Diff: 1

Blooms: Remember

Topic: Measures of Dispersion

LO1: List different measures of dispersion.

LO2: Compare and contrast methods of summarizing and describing data

**26) \_\_\_\_\_\_\_\_ states that for any set of data, the proportion of values that lie within k standard deviations (*k > 1*) of the mean is at least *1 - 1/k2*.**

A) Prime number theorem

B) Bertrand's postulate

C) Oppermann's conjecture

**D) Chebyshev's theorem**

Answer: D

Diff: 1

Blooms: Remember

Topic: Measures of Dispersion

LO1: Explain Chebyshev's theorem.

LO2: Compare and contrast methods of summarizing and describing data

27) Using Chebyshev's theorem for standard deviation, calculate the percentage of data that lie within five standard deviations of the mean.

A) 89%

B) 75%

**C) 96%**

D) 50%

Answer: C

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Dispersion

LO1: Explain Chebyshev's theorem.

LO2: Compare and contrast methods of summarizing and describing data

28) In the equation *Cp = (upper specification - lower specification)/total variation*, what does *Cp* denote?

A) capacity variation index

B) capability pattern

C) process capability index

D) capability push

Answer: C

Diff: 1

Blooms: Remember

Topic: Measures of Dispersion

LO1: List different measures of dispersion.

LO2: Compare and contrast methods of summarizing and describing data

29) Which of the following is the z-score equation for the *i*th observation?

A) *Zt =*

B) *Zi =*

C) *Z1=*

D) Z*i=*

Answer: B

Diff: 1

Blooms: Remember

Topic: Measures of Dispersion

LO1: Compute a standardized value (z-score) for observations in a data set.

LO2: Compare and contrast methods of summarizing and describing data

30) In the z-score formula, which of the following is true if the value in the numerator is a negative value?

**A) that the *xi* value lies to the left of the mean**

B) that the mean is of lesser value than the *xi* value

C) that the mean is of negative value

D) that the numerator value cannot be divided by the standard deviation

Answer: A

Diff: 2

Blooms: Understand

Topic: Measures of Dispersion

LO1: Compute a standardized value (z-score) for observations in a data set.

LO2: Compare and contrast methods of summarizing and describing data

31) A z-score of 1.0 means that \_\_\_\_\_\_\_\_.

A) the observation is -1.0 standard deviation to the right of the mean

B) the observation is -1.0 standard deviation to the left of the mean

C) the observation has no deviation from the mean

D) the observation is 1.0 standard deviation to the right of the mean

Answer: D

Diff: 2

Blooms: Understand

Topic: Measures of Dispersion

LO1: Compute a standardized value (z-score) for observations in a data set.

LO2: Identify different business uses for statistics and the major statistical tools businesses use

32) Which of the following is the equation for calculating the coefficient of variation (CV)?

**A) CV = standard deviation/mean**

B) CV = standard deviation - z-score/mean (total variation)

C) CV = value of observation's distance from mean/standard deviation

D) CV = mean/(standard deviation)2

Answer: A

Diff: 2

Blooms: Understand

Topic: Measures of Dispersion

LO1: Define and compute the coefficient of variation.

LO2: Compare and contrast methods of summarizing and describing data

33) How is the return to risk described in financial statistics?

A) as the relative measure of the distance an observation is from the mean

**B) as the reciprocal of coefficient of variation**

C) as the square root of variance

D) as the ratio of excess returns to its standard deviation

Answer: B

Diff: 1

Blooms: Understand

Topic: Measures of Dispersion

LO1: Define and compute the coefficient of variation.

LO2: Identify different business uses for statistics and the major statistical tools businesses use

34) In finance, the \_\_\_\_\_\_\_\_ is the ratio of a fund's excess returns (annualized total returns minus Treasury bill returns) to its standard deviation.

A) field ratio

B) Sortino ratio

C) Calmar ratio

**D) Sharpe ratio**

Answer: D

Diff: 1

Blooms: Remember

Topic: Measures of Dispersion

LO1: Define and compute the coefficient of variation.

LO2: Identify different business uses for statistics and the major statistical tools businesses use

35) Calculate the coefficient of variation from the following data:

z-score = 1.32; standard deviation = 0.173; mean = 4.7; total variation = 0.6

A) 27.16

B) 156.66

**C) 0.04**

D) -0.5

Answer: C

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Dispersion

LO1: Define and compute the coefficient of variation.

LO2: Compare and contrast methods of summarizing and describing data

36) Which of the following values of the coefficients of variation of stocks represents the least risky stock?

A) 1.0

**B) 0.005**

C) 0.5

D) 0.045

Answer: B

Diff: 2

Blooms: Understand

Topic: Measures of Dispersion

LO1: Define and compute the coefficient of variation.

LO2: Identify different business uses for statistics and the major statistical tools businesses use

37) Which of the following describes a positively skewed histogram?

**A) a histogram that tails off toward the right**

B) a histogram that has no fluctuation in mass

C) a histogram where more mass tails off toward the left

D) a histogram where mass is only concentrated in the middle

Answer: A

Diff: 1

Blooms: Remember

Topic: Measures of Shape

LO1: Explain the nature of skewness and kurtosis in a distribution.

LO2: Compare and contrast methods of summarizing and describing data

38) The \_\_\_\_\_\_\_\_ measures the degree of asymmetry of observations around the mean.

A) coefficient of variation

B) return to risk factor

**C) coefficient of skewness**

D) coefficient of kurtosis

Answer: C

Diff: 1

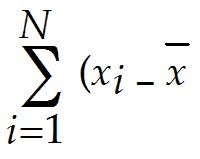
Blooms: Remember

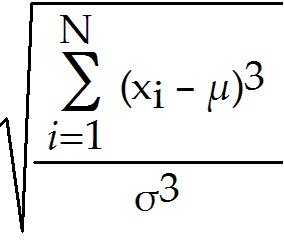
Topic: Measures of Shape

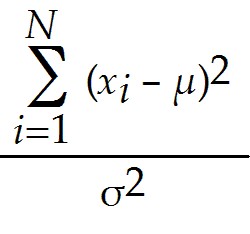
LO1: Explain the nature of skewness and kurtosis in a distribution.

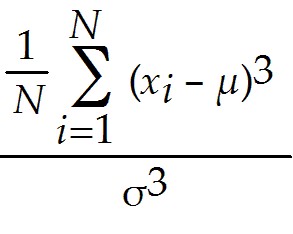
LO2: Compare and contrast methods of summarizing and describing data

39) The coefficient of skewness is computed as \_\_\_\_\_\_\_\_.

A) *CS =* )3

B) *CS* = 

C) CS = 

**D) CS = **

Answer: D

Diff: 1

Blooms: Remember

Topic: Measures of Shape

LO1: Interpret the coefficients of skewness and kurtosis.

LO2: Compare and contrast methods of summarizing and describing data

40) Which of the following is true of the coefficient of skewness (CS)?

A) If the CS is positive, the distribution of values tails off to the left.

B) If the CS value is between 0.5 and 1, the skewness is considered to be moderate.

C) If the CS is closer to zero, the degree of skewness is considered to be high.

D) If the CS value is between -0.5 and -1, the skewness is considered negligible.

Answer: B

Diff: 1

Blooms: Remember

Topic: Measures of Shape

LO1: Interpret the coefficients of skewness and kurtosis.

LO2: Compare and contrast methods of summarizing and describing data

41) In statistics, \_\_\_\_\_\_\_\_ refers to the peakedness or flatness of a histogram.

A) Sharpe ratio

B) entropy rate

C) Markov chain

**D) kurtosis**

Answer: D

Diff: 1

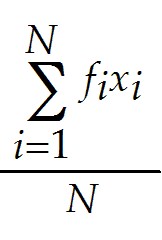
Blooms: Remember

Topic: Measures of Shape

LO1: Explain the nature of skewness and kurtosis in a distribution.

LO2: Compare and contrast methods of summarizing and describing data

42) Which of the following cases require the mean of a sample to be calculated using the formula

= , where N represents the sample size?

A) the sample has no mode value

B) there is direct access to the raw data

**C) sample data are grouped in a frequency distribution**

D) the coefficient of skewness is 1 or less than -1

Answer: C

Diff: 1

Blooms: Understand

Topic: Descriptive Statistics for Grouped Data

LO1: Calculate the mean, variance, and standard deviation for grouped data.

LO2: Compare and contrast methods of summarizing and describing data

43) The \_\_\_\_\_\_\_\_ is a formal statistical measure for categorical data, such as defects or errors in quality control applications or consumer preferences in market research.

A) variance

**B) proportion**

C) mean

D) skewness

Answer: B

Diff: 1

Blooms: Remember

Topic: Descriptive Statistics for Categorical Data: The Proportion

LO1: Calculate a proportion.

LO2: Identify different business uses for statistics and the major statistical tools businesses use

44) \_\_\_\_\_\_\_\_ is a measure of the linear association between two variables, X and Y.

A) Kurtosis

B) Proportion

C) Skewness

**D) Covariance**

Answer: D

Diff: 1

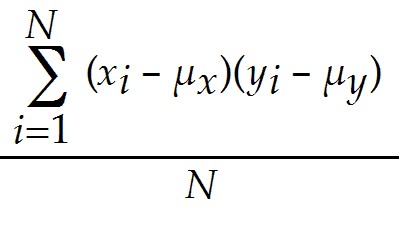
Blooms: Remember

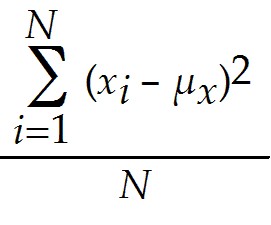
Topic: Measures of Association

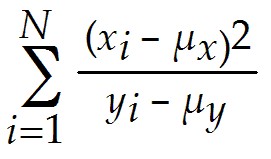
LO1: Explain the importance of understanding relationships between two variables. Explain the difference between covariance and correlation.

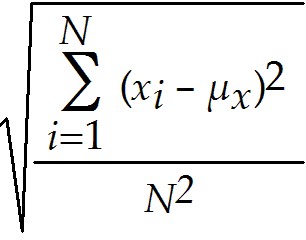
LO2: Compare and contrast methods of summarizing and describing data

45) Which of the following formula is used to compute covariance between two variables X and Y?

**A) cov (X,Y) = **

B) cov (X,Y) = 

C) cov (X,Y) = 

D) cov (X,Y) = 

Answer: A

Diff: 1

Blooms: Remember

Topic: Measures of Association

LO1: Explain the importance of understanding relationships between two variables. Explain the difference between covariance and correlation.

LO2: Compare and contrast methods of summarizing and describing data

46) Which of the following is true of covariance, between two variables, when one of the deviations from the mean is positive and the other is negative?

A) the degree of linear association is high between the two variables

B) there is no covariance between the two variables

C) the covariance will be negative

D) the covariance will be positive

Answer: C

Diff: 2

Blooms: Understand

Topic: Measures of Association

LO1: Explain the importance of understanding relationships between two variables. Explain the difference between covariance and correlation.

LO2: Compare and contrast methods of summarizing and describing data

47) \_\_\_\_\_\_\_\_ is a measure of the linear relationship between two variables, X and Y, which does not depend on the units of measurement.

A) Kurtosis

B) Proportion

C) Skewness

**D) Correlation**

Answer: D

Diff: 1

Blooms: Remember

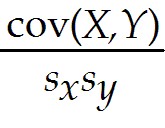
Topic: Measures of Association

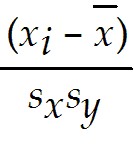
LO1: Explain the importance of understanding relationships between two variables. Explain the difference between covariance and correlation.

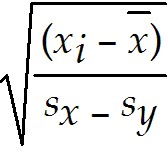
LO2: Compare and contrast methods of summarizing and describing data

48) Which of the following is the equation used for computing the sample correlation coefficient?

A) *rxy=* cov(*X,Y*)(*sxsy)*

**B) *rxy=* **

C) *rxy=* 

D) *rxy=* 

Answer: B

Diff: 1

Blooms: Remember

Topic: Measures of Association

LO1: Calculate measures of covariance and correlation.

LO2: Compare and contrast methods of summarizing and describing data

49) For two variables, a positive correlation coefficient indicates \_\_\_\_\_\_\_\_.

**A) a linear relationship exists for which one variable increases as the other also increases**

B) a linear relationship exists for one variable that increases while the other decreases

C) that the two variables have no linear relationship with each other

D) a nonlinear relationship with no linear correlation between the two variables

Answer: A

Diff: 2

Blooms: Understand

Topic: Measures of Association

LO1: Calculate measures of covariance and correlation.

LO2: Compare and contrast methods of summarizing and describing data

Use the spreadsheet below to answer the following question(s).

Below is the table showing rate of shoes sold per day and the highest-priced shoe sold that day for a one-week period. The rate of shoes sold per day (X) and the price of the shoes (Y).

|  |  |
| --- | --- |
| **Shoes sold in % (X)** | **Price (Y)** |
| 15 | $48 |
| 27 | $55 |
| 10 | $40 |
| 13 | $88 |
| 25 | $53 |
| 9 | $47 |
| 18 | $40 |
|  |  |

**50) Use Excel to calculate the mean for X.**

**A) 16.71**

B) 7.04

C) 2.67

D) 0.02

Answer: A

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Location

LO1: Compute the mean, median, mode, and midrange of a set of data.

LO2: Use a modern software tool to perform statistical calculations.

51) Use Excel to calculate the variance of X.

A) 271.33

B) 16.47

**C) 49.57**

D) 7.04

Answer: C

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Dispersion

LO1: Compute the range, interquartile range, variance, and standard deviation of a set of data.

LO2: Use a modern software tool to perform statistical calculations.

52) Use Excel to calculate the standard deviation for Y.

A) 49.57

**B) 16.47**

C) 7.04

D) 2.6

Answer: B

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Dispersion

LO1: Compute the range, interquartile range, variance, and standard deviation of a set of data.

LO2: Use a modern software tool to perform statistical calculations.

53) Use Excel to calculate the covariance between variables X and Y.

**A) 2.67**

B) 0.02

C) 16.71

D) 7.04

Answer: A

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Association

LO1: Calculate measures of covariance and correlation.

LO2: Use a modern software tool to perform statistical calculations.

54) Use Excel to calculate the z-score of X observation 27.

A) -0.24

B) 0

C) 0.18

**D) 1.46**

Answer: D

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Dispersion

LO1: Compute a standardized value (z-score) for observations in a data set.

LO2: Use a modern software tool to perform statistical calculations.

55) Use Excel to calculate the coefficient of variation of X.

**A) 0.42**

B) 0.31

C) 7.04

D) 16.47

Answer: A

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Dispersion

LO1: Define and compute the coefficient of variation.

LO2: Use a modern software tool to perform statistical calculations.

56) Use Excel to calculate the correlation between variables X and Y.

A) 16.47

**B) 0.02**

C) 16.71

D) 7.04

Answer: B

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

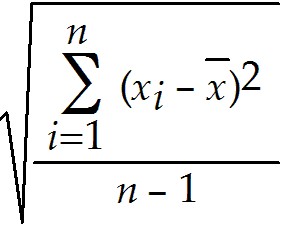
Topic: Measures of Association

LO1: Calculate measures of covariance and correlation.

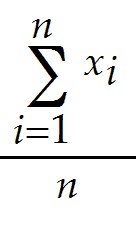
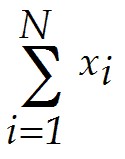
LO2: Use a modern software tool to perform statistical calculations.

57) Calculate the standard deviation for cost per order in the following table.

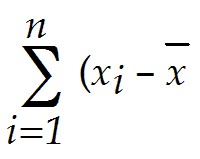
|  |  |
| --- | --- |
| **Observation** | **Cost per order** |
| X1 | $ 1345 |
| X2 | $ 4364 |
| X3 | $ 1768 |
| X4 | $ 2278 |
| X5 | $ 4332 |
| X6 | $ 2984 |
| X7 | $ 3326 |

Answer: The formula for standard deviation, *S* =  where n, is the number of observations. (*xi* - )2 is the square of the difference between each observation minus the mean.

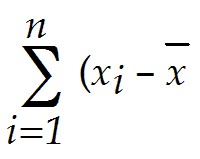
In order to calculate the standard deviation, we must first calculate the mean  of all the cost per order.

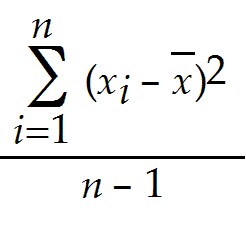
= , where, is the sum of observations, *n,* the number of observations.

Therefore,  = 20397 / 7 = 2913.86.

Now, we have to calculate)2, which is calculated by adding all the squares of each observation minus the mean.

|  |  |  |  |
| --- | --- | --- | --- |
| **Observation** | **Cost per order**  **in $** | **(xi - mean)** | **(xi - mean)2** |
| X1 | 1345 | -1568.86 | 2461312.73 |
| X2 | 4364 | 1450.14 | 2102914.31 |
| X3 | 1768 | -1145.86 | 1312988.59 |
| X4 | 2278 | -635.86 | 404314.31 |
| X5 | 4332 | 1418.14 | 2011129.16 |
| X6 | 2984 | 70.14 | 4920.02 |
| X7 | 3326 | 412.14 | 169861.73 |
|  |  |  |  |
| **Sum** | 20397 |  | 8467440.86 |
| **Count** | 7 |  |  |
| **Mean** | 2913.86 |  |  |
| **Variance** | 1411240.14 |  |  |
| **Standard**  **Deviation** | 1187.96 |  |  |

Therefore, )2= 8467440.86

From this variance can be calculated using  = 8467440.86 / 6 = 1411240.14

Standard Deviation is the square root of variance, therefore, standard deviation = 

= 1187.96.

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Dispersion

LO1: Compute the range, interquartile range, variance, and standard deviation of a set of data.

LO2: Compare and contrast methods of summarizing and describing data

58) Below is the data collected from a manufacturing process for a part whose dimensions are specified as 4.00 ± 0.1 centimeters. Calculate the process capability index within three standard deviations of the mean.

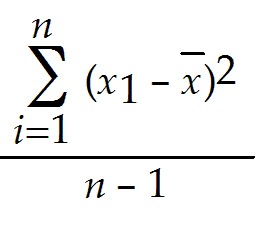
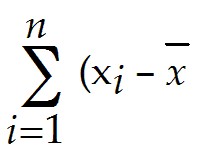
|  |
| --- |
| Manufacturing Measurements |
| 4.78 |
| 3.99 |
| 5.02 |
| 4.23 |
| 4.57 |
| 4.44 |
| 5.18 |

Answer: The process capability index, denoted as Cp = upper specification - lower specification / total variation.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Manufacturing Measurements** | **(Xi - mean {X})** | **(Xi - mean {X})2** |  |  |
| 4.78 | 0.18 | 0.032 | **Mean** | 4.6 |
| 3.99 | -0.61 | 0.37 | **Standard Deviation** | 0.424 |
| 5.02 | 0.42 | 0.18 | **Variance** | 0.18 |
| 4.23 | -0.37 | 0.14 | **Mean -3\*Stdev** | 3.33 |
| 4.57 | -0.03 | 0.00099 | **Mean +3\*Stdev** | 5.87 |
| 4.44 | -0.16 | 0.03 | **Total Variation** | 2.54 |
| 5.18 | 0.58 | 0.33 |  |  |
|  |  |  | **Lower Specification** | 3.9 |
| **Sum** |  | 1.08 | **Upper Specification** | 4.1 |
|  |  |  | **Specification Range** | 0.2 |
|  |  |  |  |  |
|  |  |  | **Cp** | .08 |

In order to calculate the process capability index, the mean of the measurements should be first calculated using the formula: sum of all measurements / number of measurements.

The mean is found to be: 4.6.

From the mean, we calculate variation using the formula ; where *n* is the number of observations. )2is the sum of all the squares of each observation minus the mean.

The variation is found to be: 0.18.

The standard deviation is the square root of variance, therefor, standard deviation =  = 0.42.

Now *Mean - 3\*Stdev*, and *Mean+3\*Stdev* are calculated.

*Mean - 3\*Stdev* = 4.6 × 3(0.44) = 3.33, and,

*Mean + 3\*Stdev* = 4.6 + 3(0.44) = 5.87.

Therefore *total variation = (Mean + 3\*Stdev) - (Mean - 3\*Stdev)* = 5.87 - 3.33 = 2.54.

The process capability index, Cp, is then calculated as 0.2 / 2.54 = 0.08.

Therefore, Cp = 0.08.

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Dispersion

LO1: Compute the range, interquartile range, variance, and standard deviation of a set of data.

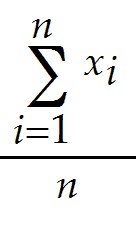
LO2: Identify different business uses for statistics and the major statistical tools businesses use

59) The table below shows the sales per day at a grocery store over a period of a week. Calculate the z-score for sales on Day 4.

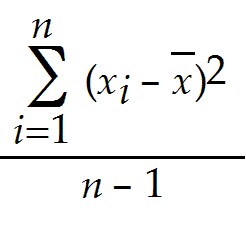
|  |  |
| --- | --- |
| **Day** | **Sales per day in $ (Xi)** |
| 1 | 5000 |
| 2 | 5304 |
| 3 | 6430 |
| 4 | 7889 |
| 5 | 5333 |
| 6 | 6109 |
| 7 | 6734 |

Answer: Z-score is calculated using the formula,*z4 =*, where *s* is the standard deviation, *xi* are observations,  is the mean of the observation.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Day** | **Sales per day in $ (Xi)** | **(Xi - mean**  **{X})** | **(Xi - mean**  **{X})^2** | **z-score** |
| 1 | 5000 | -1114.14 | 1241314.31 | -1.10 |
| 2 | 5304 | -810.14 | 656331.45 | -0.80 |
| 3 | 6430 | 315.86 | 99765.73 | 0.31 |
| 4 | 7889 | 1774.86 | 3150117.88 | 1.75 |
| 5 | 5333 | -781.14 | 610184.16 | -0.77 |
| 6 | 6109 | -5.14 | 26.44898 | -0.005 |
| 7 | 6734 | 619.86 | 384222.88 | 0.61 |
| **Sum** | 42799 |  | 6141962.86 |  |
| **Mean** | 6114.14 |  |  |  |
| **Variance** | 1023660.48 |  |  |  |
| **Standard Deviation** | 1011.76 |  |  |  |

First, the mean has to be calculated using the formula = , n is the number of observations, and*x* are number of observations.

The mean is calculated to be: 42799/7 = 6114.14.

From the mean we calculate the variance using the formula s2 =, where (*xi -* )2is the sum of the squares of observations minus the mean.

Therefor variance is calculated as 6141962.86/6 = 1023660.48.

Standard deviation is the square root of variance, therefor, standard deviation = = 1011.76.

Now, the z-score for the 4th day can be calculated as 1774.86/1011.76 = 1.75.

Therefore, z4 = 1.75.

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

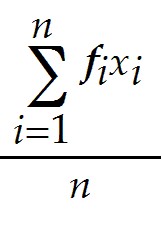
Topic: Measures of Dispersion

LO1: Compute a standardized value (z-score) for observations in a data set.

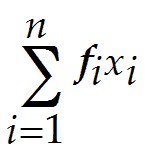
LO2: Compare and contrast methods of summarizing and describing data

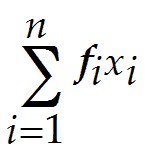
60) The table below shows the daily frequency of female customers at a particular ice cream shop during discount sale periods. The data is collected for 5 days. Calculate the mean for the data.

|  |  |
| --- | --- |
| **Female Customer Frequency** |  |
| **Days (x)** | **Frequency (f)** |
| 1 | 0 |
| 2 | 5 |
| 3 | 8 |
| 4 | 11 |
| 5 | 18 |

Answer: For grouped data, mean, =  , where *n* is the number of observations, and *fi*x*i* is individual frequencies multiplied by the corresponding number of days.

|  |  |  |
| --- | --- | --- |
| **Female Customer Frequency** |  |  |
|  |  |  |
| **Days (x)** | **Frequency (f)** | **Frequency\*Days** |
| 1 | 0 | 0 |
| 2 | 5 | 10 |
| 3 | 8 | 24 |
| 4 | 11 | 44 |
| 5 | 18 | 90 |
| **Sum** | 42 | 168 |
|  |  |  |
|  | **Mean** | 4 |

First we multiply individual frequencies by their corresponding number of day. Then we add the multiplied values to obtain .

Here, =168.

*n* is calculated to be 42.

Therefore the mean,  = 168/42 = 4.

 = 4.

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

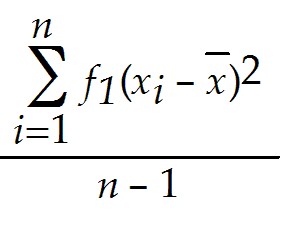
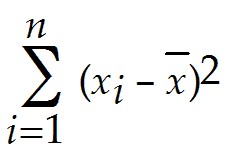
Topic: Descriptive Statistics for Grouped Data

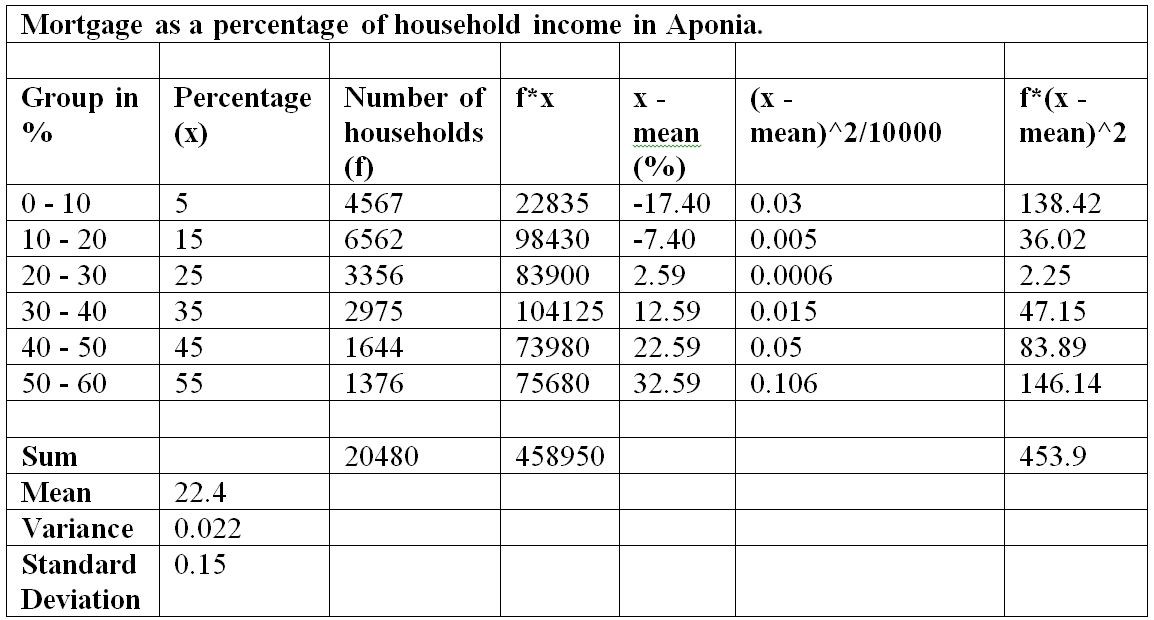
LO1: Calculate the mean, variance, and standard deviation for grouped data.

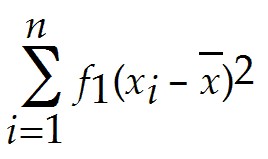
LO2: Compare and contrast methods of summarizing and describing data

61) The following table lists the amounts paid in mortgages as a percentage of the household income in the town Aponia. Calculate the variance and standard deviation of the data, if mean is given as 22.4%. The percentage (x) values are estimated as midpoints of each percentage range.

|  |  |  |
| --- | --- | --- |
| **Mortgage as a percentage of household income in Aponia** |  |  |
| **Group in %** | **Percentage (x)** | **Number of households (f)** |
| 0 - 10 | 5 | 4567 |
| 10 - 20 | 15 | 6562 |
| 20 - 30 | 25 | 3356 |
| 30 - 40 | 35 | 2975 |
| 40 - 50 | 45 | 1644 |
| 50 - 60 | 55 | 1376 |

Answer: For grouped data such as this, variance, s2 =, where n is the sum of frequencies, and,  is the sum of all the squares of each observation minus the mean.



Since mean is given,  is calculated to be 453.9.

Note that the square of (x - mean) is divided by 10000 to simplify the calculations.

Sum of frequencies is calculated to be 20480.

Substituting these values in the variance formula, we get, S2 = 453.9/20479 = 0.022.

Therefore, the variance value is 0.02.

Standard deviation is the square root of variance, therefore, standard deviation, *s* = = 0.15.

Therefore, the standard deviation value is 0.15.

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

Topic: Descriptive Statistics for Grouped Data

LO1: Calculate the mean, variance, and standard deviation for grouped data.

LO2: Compare and contrast methods of summarizing and describing data

**62) A sample is a subset of a population.**

**Answer: TRUE**

Diff: 1

Blooms: Remember

Topic: Populations and Samples

LO1: Explain the difference between a population and a sample.

LO2: Explain the fundamentals of sampling methods, experiment designs, and sampling distributions

***63) One of the properties of the mean is that the sum of the deviations of each observation from the mean is zero.***

***Answer: TRUE***

Diff: 1

Blooms: Remember

Topic: Measures of Location

LO1: Compute the mean, median, mode, and midrange of a set of data.

LO2: Compare and contrast methods of summarizing and describing data

64) For an odd number of observations, the median is the mean of the two middle numbers.

Answer: FALSE

Diff: 1

Blooms: Remember

Topic: Measures of Location

LO1: Compute the mean, median, mode, and midrange of a set of data.

LO2: Compare and contrast methods of summarizing and describing data

65) The larger the variance, the more the data are spread out from the mean.

Answer: TRUE

Diff: 1

Blooms: Remember

Topic: Measures of Dispersion

LO1: Compute the range, interquartile range, variance, and standard deviation of a set of data.

LO2: Compare and contrast methods of summarizing and describing data

**66) According to the empirical rules of standard deviation in statistics, approximately 68% of the observations will fall within three standard deviations of the mean.**

**Answer: FALSE**

Diff: 2

Blooms: Apply

AACSB: Analytic Skills

Topic: Measures of Dispersion

LO1: State the Empirical Rules and apply them to practical data.

LO2: Compare and contrast methods of summarizing and describing data

**67) Two different sets of data can never have the same z-scores.**

**Answer: FALSE**

Diff: 1

Blooms: Understand

Topic: Measures of Dispersion

LO1: Compute a standardized value (z-score) for observations in a data set.

LO2: Compare and contrast methods of summarizing and describing data

**68) In stock trading, the higher the coefficient of variation for the stock, the smaller the relative risk is for the return provided.**

**Answer: FALSE**

Diff: 1

Blooms: Remember

Topic: Measure of Dispersion

LO1: Define and compute the coefficient of variation.

LO2: Identify different business uses for statistics and the major statistical tools businesses use

**69) Skewness describes the lack of symmetry of data.**

**Answer: TRUE**

Diff: 1

Blooms: Remember

Topic: Measure of Shape

LO1: Explain the nature of skewness and kurtosis in a distribution.

LO2: Identify different business uses for statistics and the major statistical tools businesses use

**70) The value of a coefficient of skewness falling between 0.5 and -0.5 indicates relative symmetry.**

**Answer: TRUE**

Diff: 1

Blooms: Remember

Topic: Measure of Shape

LO1: Interpret the coefficients of skewness and kurtosis.

**71) If the distribution of observations are such that mode < median < mean, the histogram would be said to be negatively skewed.**

**Answer: FALSE**

Diff: 1

Blooms: Understand

Topic: Measure of Shape

LO1: Explain the nature of skewness and kurtosis in a distribution.

LO2: Compare and contrast methods of summarizing and describing data

72) Two variables can only have a statistical relationship if there is a cause-and-effect factor between them.

Answer: FALSE

Diff: 1

Blooms: Understand

Topic: Measures of Association

LO1: Explain the importance of understanding relationships between two variables. Explain the difference between covariance and correlation.

LO2: Compare and contrast methods of summarizing and describing data

**73) A correlation of 0 indicates that the two variables have no linear relationship with each**

**other.**

**Answer: TRUE**

Diff: 1

Blooms: Remember

Topic: Measures of Association

LO1: Calculate measures of covariance and correlation.

LO2: Compare and contrast methods of summarizing and describing data

74) According to the interquartile range, extreme outliers are more than 3\*IQR away from the left of Q1 or to the right of Q3.

Answer: TRUE

Diff: 1

Blooms: Remember

Topic: Outliers

LO1: Compute the range, interquartile range, variance, and standard deviation of a set of data.

LO2: Compare and contrast methods of summarizing and describing data

**75) State the three empirical rules in standard deviation.**

**Answer: The three empirical rules in standard deviation are:**

**1. Approximately 68% of the observations will fall within one standard deviation of the mean, or between  - s and + *s.***

**2. Approximately 95% of the observations will fall within two standard deviations of the mean, or within + 2s.**

**3. Approximately 99.7% of the observations will fall within three standard deviations of the mean, or within + 3s.**

Diff: 1

Blooms: Remember

Topic: Measures of Dispersion

LO1: State the Empirical Rules and apply them to practical data.

LO2: Compare and contrast methods of summarizing and describing data

76) Give an account of Excel's *Descriptive Statistics* Tool.

Answer: Excel provides a useful tool for basic data analysis, *Descriptive Statistics.* Click on *Data Analysis* in the *Analysis* group under the *Data* tab in the Excel menu bar. Select *Descriptive Statistics* from the list of tools. You need to enter only the range of the data, which must be in a *single row or column.* If the data are in multiple columns, the tool treats each row or column as a separate data set, depending on which you specify. This means that if you have a single data set arranged in a matrix format, you would have to stack the data in a single column before applying the *Descriptive Statistics* tool. Check the box *Labels in First Row* if labels are included in the input range. You may choose to save the results in the current worksheet or in a new one. For basic summary statistics, check the box *Summary statistics;* you need not check any others. One important point to note about the use of the tools in the *Analysis Toolpak* versus Excel functions is that while Excel functions dynamically change as the data in the spreadsheet are changed, the results of the *Analysis Toolpak* tools do not.

Diff: 1

Blooms: Remember

Topic: Excel Descriptive Statistics Tool

LO1: Use the Excel Descriptive Statistics tool to summarize data.

LO2: Identify different business uses for statistics and the major statistical tools businesses use

77) What are the different interpretations for correlation coefficient values?

Answer: A correlation of 0 indicates that the two variables have no linear relationship to each other. Thus, when one variable changes, we cannot reasonably predict what the other variable might do. A positive correlation coefficient indicates a linear relationship for which one variable increases as the other also increases. A negative correlation coefficient indicates a linear relationship for which one variable that increases while the other decreases.

Diff: 1

Blooms: Remember

Topic: Measures of Association

LO1: Calculate measures of covariance and correlation.

LO2: Compare and contrast methods of summarizing and describing data

78) How can outliers be identified in a data set?

Answer: Outliers can make a significant difference in the results we obtain from statistical analyses. The first thing to do is to check the data for possible errors, such as a misplaced decimal point or an incorrect transcription to a computer file. Histograms can help to identify possible outliers visually. The empirical rule and z-scores can also be used to identify an outlier as one that is more than three standard deviations from the mean. Outliers can also be identified based on the interquartile range. "Mild" outliers are often defined as being between 1.5\*IQR and 3\*IQR to the left of Q1 or to the right of Q3, and "extreme" outliers, as more than 3\*IQR away from these quartiles.

Diff: 1

Blooms: Remember

Topic: Outliers

LO1: Identify outliers in data.

LO2: Compare and contrast methods of summarizing and describing data

79) What is statistical thinking? How is it important to making business decisions?

Answer: Statistical thinking is a philosophy of learning and action for improvement that is based on the principles that (1) all work occurs in a system of interconnected processes, (2) variation exists in all processes, and (3) better performance results from understanding and reducing variation.

Understanding business processes provides the context for determining the effects of variation and the proper type of action to be taken. Any process contains many sources of variation. Statistical analysis can provide better insight into the facts and nature of relationships among the many factors that may have contributed to an event and enable managers to make better decisions.

Diff: 2

Blooms: Understand

AACSB: Analytic Skills

Topic: Statistical Thinking in Business Decisions

LO1: State the principles of statistical thinking.

LO2: Identify different business uses for statistics and the major statistical tools businesses use