CONCORDIA COLLEGE
MOORHEAD, MINNESOTA

- Name: _____
- Date: _____
- Section: ○ Mon/Wed ○ Tue/Thu

BUSN 320: Business Statistics

Problem Set #1: Suggested Solutions

Fall 2026

INSTRUCTIONS:

- These practice problems are intended to help you prepare for Quiz #1.
- You may use your textbook, class notes, calculator, and other course resources while completing these problems.
- The problem set consists of three parts: Definitions, True / False, and Short Answers.
- The practice problems are broader than the quiz and do not necessarily represent the exact questions that will appear on Quiz #1.

Problem 1. Definitions

The following terms are fair game for definition questions on Quiz #1. You should be able to define each term and distinguish it from closely related concepts.

Chapter 1

- Population
- Parameter
- Descriptive Statistics
- Cross-Sectional Data
- Categorical Variable
- Sample
- Statistic
- Inferential Statistics
- Time-Series Data
- Quantitative Variable

Chapter 2

- Frequency
- Percent Frequency
- Frequency Distribution
- Bar Chart
- Relative Frequency
- Cumulative Frequency
- Histogram
- Scatter Diagram

Chapter 3

- Mean
- Mode
- Variance
- Percentile
- Interquartile Range (IQR)
- z-score
- Correlation Coefficient
- Median
- Range
- Standard Deviation
- Quartile
- Five-Number Summary
- Covariance
- Outlier

Problem 2. True / False

For each statement, indicate whether it is TRUE or FALSE. A correct TRUE response will receive full credit. For a FALSE statement, you must briefly explain why the statement is false to receive full credit.

2.A. A numerical measure calculated from a sample is called a parameter.

- FALSE
  A numerical measure calculated from a sample is called a statistic. A parameter is a numerical characteristic of a population.

2.B. A quantitative variable must always take whole-number values.

- FALSE
  A quantitative variable takes numerical values, but these values do not have to be whole numbers. For example, variables such as height, weight, or time may take non-whole-number values.

2.C. Data collected from several businesses at the same point in time are an example of time-series data.

- FALSE
  Data collected from several businesses at approximately the same point in time are cross-sectional data. Time-series data consist of observations collected over multiple time periods.

2.D. The relative frequencies in a frequency distribution should sum to 1.

- TRUE

2.E. A bar chart and a histogram are interchangeable because both use bars to display frequencies.

- FALSE
  A bar chart is generally used to display categorical data, while a histogram is used to display quantitative data grouped into intervals. Although both use bars, they represent different types of data.

2.F. A histogram with a long tail extending toward larger values is right-skewed.

- TRUE

2.G. An extreme observation will generally have a greater effect on the mean than on the median.

- TRUE

Problem 2. True / False (continued)

For each statement, indicate whether it is TRUE or FALSE. A correct TRUE response will receive full credit. For a FALSE statement, you must briefly explain why the statement is false to receive full credit.

2.H. If every observation in a data set is identical, the standard deviation is zero.

- TRUE

2.I. A larger standard deviation indicates that observations tend to be more dispersed around the mean.

- TRUE

2.J. Variance and standard deviation are expressed in the same units as the original data.

- FALSE
  Standard deviation is expressed in the same units as the original data, but variance is expressed in squared units. This is because variance is based on squared deviations from the mean.

2.K. The interquartile range measures the distance between the first and third quartiles.

- TRUE

2.L. An observation with a z-score of zero must also have a value of zero.

- FALSE
  A z-score of zero means that the observation is zero standard deviations from the mean. Therefore, the value of the observation is equal to the mean, which does not necessarily equal zero.

2.M. A positive covariance indicates that two variables tend to move in opposite directions.

- FALSE
  A positive covariance indicates that two variables tend to move in the same direction. Large values of one variable tend to occur with large values of the other variable.

2.N. A correlation coefficient of $r = 0$ proves that there is no relationship of any kind between two variables.

- FALSE
  A correlation coefficient of $r = 0$ indicates that there is no linear association between the two variables. The variables may still have a nonlinear relationship.

Problem 3. Short Answers

3.A. A grocery chain wants to estimate the average amount customers spend per visit. During one week, the company randomly selects 250 transactions from its stores. The average amount spent among these transactions is \$40.10.

- Identify the population of interest.
  - All customer transactions at the grocery chain.
- Identify the sample.
  - The 250 randomly selected transactions.
- Identify the parameter of interest.
  - The population mean amount spent per customer visit.
- Identify the statistic reported in the problem.
  - The sample mean amount spent, $\bar{x} = \$40.10$.

3.B. The following data represent the number of items purchased by 20 customers during a visit to a store.

3 5 2 4 3 6 2 3 4 5

3 2 7 4 3 5 2 4 6 3

- Construct a frequency distribution containing the frequency, relative frequency, percent frequency, and cumulative percent frequency for the number of items purchased.

Items	Frequency	Relative Frequency	Percent Frequency	Cumulative % Freq.
2	4	0.20	20%	20%
3	6	0.30	30%	50%
4	4	0.20	20%	70%
5	3	0.15	15%	85%
6	2	0.10	10%	95%
7	1	0.05	5%	100%
Total	20	1.00	100%	—

- What percentage of customers purchased four or fewer items?
  - The cumulative percent frequency for 4 items gives the percentage of customers who purchased four or fewer items. Therefore,

$$20\% + 30\% + 20\% = \boxed{70\%}.$$

- How many customers purchased at least five items?
  - Customers who purchased at least five items purchased 5, 6, or 7 items. Therefore,

$$3 + 2 + 1 = \boxed{6}.$$

Problem 3. Short Answers (continued)

3.C. The following data represent weekly sales, in thousands of dollars, for a small business.

12 13 13 14 15 17 42

- Calculate the mean, median, mode, and range.
  - Mean: $\frac{12+13+13+14+15+17+42}{7} = \frac{126}{7} = \boxed{18}$
  - Median: ~~12~~, ~~13~~, ~~13~~, 14, ~~15~~, ~~17~~, ~~42~~ $\Rightarrow$ $\boxed{14}$
  - Mode: $\boxed{13}$
  - Range: $42 - 12 = \boxed{30}$
- Explain why the mean and median differ substantially.
  - The value 42 is substantially larger than the other observations. This extreme observation pulls the mean upward, while the median is less affected by extreme observations.
- Which measure of center better describes a typical week in this case? Briefly explain.
  - The median better describes a typical week because it is less affected by the extreme value of 42. Most weekly sales are much closer to the median of \$14,000 than to the mean of \$18,000.

3.D. A retailer records the number of customers entering one of its stores during a sample of five one-hour periods. The observed numbers of customers are shown below.

18 20 20 22 20

Using the sample above, calculate each of the following numerical measures. Show your work where appropriate.

- Sample Variance:

$$s^2 = \frac{\sum_{i=1}^{5}(x_i - \bar{x})^2}{5-1} = \frac{(18-20)^2 + (20-20)^2 + (20-20)^2 + (22-20)^2 + (20-20)^2}{4}$$
$$= \frac{4+0+0+4+0}{4} = \boxed{2}$$

- Sample Standard Deviation:

$$s = \sqrt{s^2} = \sqrt{2} \simeq \boxed{1.41}$$

Problem 3. Short Answers (continued)

3.E. The following data represent the number of customer complaints received by a business over 11 weeks.

2 3 4 4 5 6 7 8 8 9 18

- Find the five-number summary.
  - Minimum $= 2$
  - $Q_1 = 4$
  - $Q_2 = 6$
  - $Q_3 = 8$
  - Maximum $= 18$
- Calculate the interquartile range (IQR).

$$IQR = Q_3 - Q_1 = 8 - 4 = \boxed{4}$$

- Use the $1.5 \times IQR$ rule to determine whether the data set contains any potential outliers.
  - Lower Bound: $Q_1 - 1.5 \times IQR = 4 - 6 = -2$
  - Upper Bound: $Q_3 + 1.5 \times IQR = 8 + 6 = 14$
  - Observations below -2 or above 14 are outliers. Therefore, $\boxed{18}$ is a potential outlier.

3.F. Two pairs of variables have the following correlation coefficients:

$$r_{X,Y} = 0.82 \qquad r_{W,Z} = -0.35$$

- Describe the direction and strength of the linear relationship for each pair.
  - X and Y have a strong positive linear relationship.
  - W and Z have a weak negative linear relationship.
- Which pair exhibits the stronger linear relationship? Briefly explain.
  - X and Y exhibit the stronger linear relationship because $|0.82| > |-0.35|$. The strength of a linear relationship is determined by how close the correlation coefficient is to -1 or 1, regardless of its sign.
- Without calculating the covariance, what sign would you expect the covariance to have for each pair? Briefly explain.
  - The covariance between X and Y should be positive, while the covariance between W and Z should be negative. Covariance and correlation have the same sign, so a positive correlation implies positive covariance and a negative correlation implies negative covariance.

Problem 3. Short Answers (continued)

3.G. A company reports that the average amount customers spend per visit is \$60, with a standard deviation of \$12. One customer spends \$84.

- Calculate the z-score for this customer's spending.

$$z = \frac{x - \bar{x}}{s} = \frac{84 - 60}{12} = \frac{24}{12} = \boxed{2}$$

- Interpret the z-score in the context of the problem.
  - The customer's spending of \$84 is 2 standard deviations above the average amount spent per visit.