Name _____ Date _____ Block _____ **READING GUIDE**

Chapter 1: Exploring Data

Key Vocabulary:

- individuals
- variable
- categorical variable
- quantitative variable
- two way table
- marginal distributions
- conditional distribution
- association
- distribution
- range
- spread
- frequency
- outlier
- center
- shape
- skewed left
- skewed right
- symmetric
- dot plot
- histogram
- stemplot
- split stems
- back-to-back stemplot
- time plot
- mean
- Σ
- $\overline{x}$
- nonresistant
- median
- resistant
- quartiles
- Q_1, Q_3
- IQR
- five-number summary
- minimum
- maximum
- boxplot
- modified boxplot
- standard deviation
- variance

INTRO Analyzing Categorical Data

1. How is statistics defined?
2. Define data analysis?
3. Define individual.
4. Define variable.
5. What is a categorical variable?
6. What is a quantitative variable?
7. Define distribution.
8. How should data be explored?
9. Drawing conclusions that go beyond the given data is referred to as _____.
10. What are the two primary ways to produce data?

1.1 Displaying Distributions with Graphs

1. What is the difference between a frequency table and a relative frequency table?

2. What type of data are *pie charts* and *bar graphs* used for??

3. Pie Charts can only be used when?

4. How is a two-way table setup?

5. Which is more informative when comparing group counts or percents? Why?

6. Explain the four step process to organizing a statistical problem.

7. What do you need to be cautious of when variables seem to have a strong association?

1.2 Describing Distributions with Numbers

8. How do you make a dot plot?

9. When examining a distribution, you can describe the overall pattern by its

   S_____ O_____ C_____ S_____

10. If a distribution is *symmetric*, what does its dot plot look like?

11. If a distribution is *skewed right*, what does its dot plot look like?

12. If a distribution is *skewed left*, what does its dot plot look like?

13. What is the difference between unimodal, bimodal, and multimodal data?

14. How do you make a *stemplot*?

15. When is it advantageous to split stems on a stemplot?

16. When is a *back to back stemplot* useful?

17. How is the *stemplot* of a distribution related to its histogram?

18. What is a *histogram*?

19. When is it better to use a *histogram* rather than a *stemplot* or *dotplot*?

20. What is meant by *frequency* in a histogram?

21. What is the difference between a *bar-graph* and a *histogram*?

22. Define *outlier*.

1.3 Describing Quantitative Data with Numbers

1. In statistics, what are the most common measures of center?

2. Explain how to calculate the *mean,* $\bar{x}$.

3. Explain how to calculate the *median, M*.

4. Explain why the median is *resistant* to extreme observations, but the mean is *nonresistant*.

5. In a symmetric distribution where are the mean and median in relation to each other? What about in a distribution that is skewed?

6. What is the difference between *"average"* value and *"typical*" value?

7. Explain how to calculate Q_1 and Q_3 and *IQR*.

8. When does an observation become an *outlier*?

9. What is the *five-number summary*?

10. How much of the data falls between each quartile?

11. How much of the data falls between Q1 and Q3?

12. Describe a *boxplot*.

13. What does *standard deviation* measure?

14. What is the relationship between *variance* and *standard deviation*?

15. When does *standard deviation* equal zero?

16. What are the units for the standard deviation of a distribution?

17. Is *standard deviation* resistant or nonresistant to extreme observations? Explain.

18. Use a five number summary when…

19. Use $\bar{x}$ and s when…