

Chapter 1: Exploring Data

Key Vocabulary:

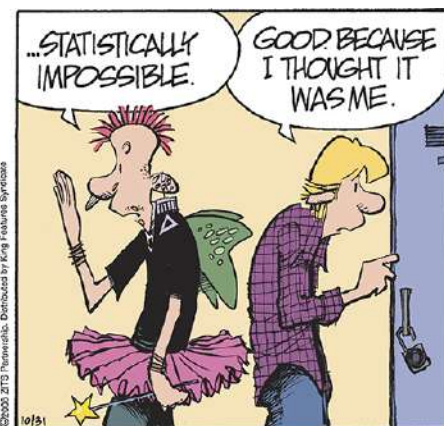
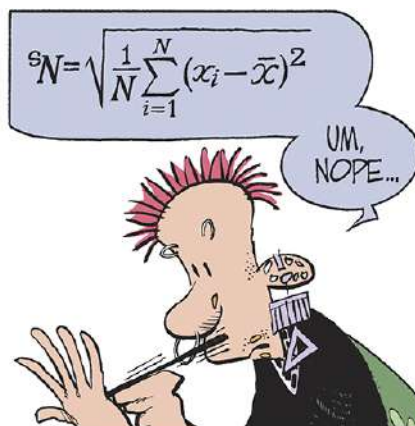
- individual
- variable
- frequency table
- relative frequency table
- distribution
- pie chart
- bar graph
- two-way table
- marginal distributions
- conditional distributions

- side-by-side bar graph
- association
- dotplot
- stemplot
- histogram
- SOCS
- outlier
- symmetric
- Σ



- $\bar{x}$
- spread
- variability
- median
- quartiles
- Q_1, Q_3

- IQR
- five-number summary
- minimum
- maximum
- boxplot
- resistant
- standard deviation
- variance



Data Analysis: Making Sense of Data (pp.2-6)

1. Individuals are...
2. A variable (*it is NOT X*) is ...
3. When you first meet a new data set, ask yourself:
 - Who...
 - What...
 - Why, When, Where and How...
4. Explain the difference between a *categorical* variable and a *quantitative* variable. Give an example of each.
5. Give an example of a categorical variable that has number values.
6. Define *distribution*:

Give an example of a distribution:

6. Answer the two questions for the *Check Your Understanding* on page 5:
7. Define *inference*.

1.1 Analyzing Categorical Data (pp.8-22)

1. A *frequency* table displays...
2. A *relative frequency table* displays...
3. What type of data are *pie charts* and *bar graphs* used for?
4. *Categories* in a bar graph are represented by _____ and the *bar heights* give the category _____.
5. What is a *two-way table*?
6. Define *marginal distribution*.
7. Answer the two questions for the *Check Your Understanding* on page 14.
8. What is a *conditional distribution*? Give an example demonstrating how to calculate one set of conditional distributions in a two-way table.

9. What is the purpose of using a *segmented bar graph*?
10. Answer question one for the *Check Your Understanding* on page 17.
11. Describe the four steps to organizing a statistical problem:
- State...
 - Plan...
 - Do...
 - Conclude...
12. Explain what it meant by an *association* between two variables; Give an example.

~~13. SKIP SIMPSON'S PARADOX.~~

1.2 Analyzing Categorical Data (pp.27-42)

1. What is a *dotplot*? When would you use it? Draw an example.

2. [VERY IMPORTANT CONCEPT!!] When examining a distribution, you can describe the overall pattern by its

S_____ O_____ C_____ S_____

- I call this CUSS and BS. You can use either mnemonic.
- Center, Unusual, Shape, Spread and Be Specific,
- Make sure you understand how to compare distributions in context.

3. Describe Shape

a) If a distribution is symmetric, what does it look like?

b) If a distribution is *skewed to the right*, what does it look like?

c) If a distribution is *skewed to the left*, what does it look like?

d) Describe and illustrate the following distributions:

i) Unimodal

ii) Bimodal

iii) Multimodal

4. Answer questions in *Check Your Understanding* on page 31.
5. How are a *stemplot* and a *histogram* similar?
6. When is it beneficial to *split the stems* on a stemplot?
7. When is it best to use a *back-to-back stemplot*?
8. Answer questions in *Check Your Understanding* on page 34.

9. List the three steps involved in making a histogram.

10. Why is it advantageous to use a relative frequency histogram instead of a frequency histogram?

11. Do *Technology Corner* (page 38) problem and sketch your graphs.

12. Answer *Check Your Understanding* questions on pages 39.

13. Answer *Check Your Understanding* questions on pages 41.

1.3 Analyzing Categorical Data (pp.50-67)

Describe Center

1. Explain how to calculate by hand **and** find on the calculator the *mean*, $\bar{x}$ (Xbar).
2. What is the meaning of Σ (sigma)?
3. Explain the difference between $\bar{x}$ and μ (mu).
4. **Define resistant measure.**
5. Explain why the mean is not a resistant measure of center.
6. What is the *median* of a distribution? Explain how to calculate by hand **and** find on the calculator.

7. Explain why the median is a resistant measure of center?

Describe Center (continued)

8. How does the shape of the distribution affect the mean and median?

9. Answer *Check Your Understanding* questions on page 55.

Describe Spread and Outliers

10. What is the *range*?
11. Is the range a resistant measure of spread? Explain.
12. How do you find *first quartile* Q_1 and *third quartile* Q_3 by hand **and** find on the calculator.

Describe Spread and Outliers (continued)

13. What is the *Interquartile Range* (IQR)?

14. Is the IQR and the quartiles a resistant measure of spread? Explain.

15. How is the IQR used to identify *outliers*?

16. What is the *five-number summary* of a distribution?

17. Explain how to use the five-number summary to make a *boxplot*.

18. Answer *Check Your Understanding* questions on page 61.

19. Do *Technology Corner* (page 61) problem and sketch your graphs.

Describe Spread and Outliers (continued)

20. What does the *standard deviation* (S or S_x) measure?
- b) How do we calculate it by hand **and** find on the calculator.
21. What is the relationship between *variance* (S^2_x or S^2) and *standard deviation* (S_x or S)? Why do we prefer to use standard deviation and NOT variance?
22. What are the 4 *properties* of the standard deviation explained on page 64?
23. Answer *Check Your Understanding* questions on page 64.
24. Do *Technology Corner* (page 65) problem and give the summary statistics
25. How should one go about choosing measures of center and spread?