

CHAPTER 2

DATA

What are Data?

- Data are values, including their context.
- Without their context, the values do not have a meaning.
- Examples:
 - names, gender, age, height

The Five W's

- Who
- What
- When
- Where
- Why
- How

WHO

- Individual of whom we record some characteristic.
- Surveys - Respondents
- Experiments
 - Experimental Units (animals, plants, etc..)
 - Human Subjects
- Archives - Records

WHAT

- Variables
 - characteristics about each individual
- Categorical variables (qualitative)
 - used as labels
 - ex: gender, home state, fav basketball team
- Quantitative variables
 - used as numbers with their measurement units
 - ex: age, height, hours you study statistics

QUANTITATIVE OR CATEGORICAL??

- Age
 - Number of Years - Quantitative
 - Adult or Senior - Categorical
- Education
 - HS, Bachelors, Advanced Degree - Categorical
 - Number of Years total - Quantitative

Where, When, Why, and How

- All of these questions facilitate understanding of the data
- 95% of ISU students consume less than 3 drinks per week.
 - Where - probably not at pre-game of ISU football game
 - When - probably not Saturday night
 - Why - potentially to create a positive image of ISU
 - How - were the individuals randomly asked??

Random Variables

- When describing *quantitative* data, we usually assign a variable to represent the topic in discussion. This is called a **random variable**.
- The capital letter represents that variable in general
- The lower case letter represents a specific/observed value of the variable

Random Variables

- Ex. Let X = age in years. I am 23, so for me, $x = 23$
- Ex. Let Y = points per game. Dwayne Wade is averaging $y = 28.6$ points per game.
- There is no reason for choosing X and Y , they just have become the most common choices for random variables