STA 2023

STATISTICS

1

Chapter 1

STATS STARTS HERE

2

WHAT IS STATISTICS?

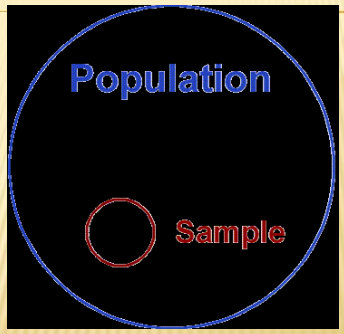
- The study of the methods used to collect and analyze data.
- Data can be collected through experimentation or by inquiry.
- We want to use the data to draw conclusions about the general population (based upon the information we obtained from our sample population).
- We want to interpret and present our data in an easily understood manner!

3

WHAT IS DATA?

- Observations that have been collected!
  - e.g., measurements or survey responses
- Raw data is information that has been collected but not manipulated in any way.
- A population consists of all members of a group of interest.
- A sample consists of a subgroup taken from the population.

4



Retrieved from http://www.uwsp.edu/psych/stat/1/popsamp.gif, January 10, 2010.

5

A subset of the population.

Retrieved from http://www.nedarc.org/nedarc/collectingData/participantSelection/images/subsetPopulation.gif, January 10, 2010.

6

THE SCIENCE OF STATISTICS

- Descriptive statistics are used to organize and summarize information in order to make that information easy to understand
- Inferential statistics are used to infer properties about the population based on measurements from a sample

7

MORE ABOUT DATA

- A variable is a characteristic that differs from person to person (e.g., gender, height, gpa, etc.)
- Quantitative data are numbers representing counts or measurements
- Categorical data or (qualitative data) can be separated into different classes

8

QUANTITATIVE AND CATEGORICAL

- Incomes of college graduates
  - Is quantitative
- Genders of college graduates
  - Is categorical
- Ethnicity of college graduates
  - Is categorical
- Heights of college graduates
  - Is quantitative

9

MORE ABOUT DATA

- Discrete data means that the number of possible values each data point can be is countable
  - The number of eggs that hens lay is discrete data
- Continuous data doesn't just take on countable values
  - The amount of milk that cows produce is continuous data

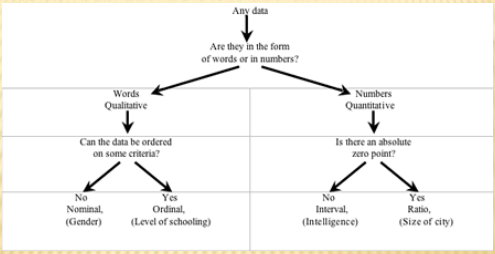
10

MORE ABOUT DATA

- Data can be classified by four levels of measurement:
  1. Nominal—data consists of names, labels, or categories only (can't be arranged in an ordering scheme)
     - Names of TV shows watched at 7 p.m. tonight
  2. Ordinal—data can be arranged in some order but the differences between data values can't be determined or are meaningless
     - Magazine ranking of the "safest city to live in"
  3. Interval—like ordinal, except the difference between any two data values is meaningful. Here there is no natural zero starting point.
     - A collection of temperatures or a collection of years
  4. Ratio—like interval, but includes the natural zero starting point (zero means none is present)
     - Weights (in carats) of diamond rings or prices of college textbooks

11

LEVELS OF MEASUREMENT



- Your text differentiates between categorical and quantitative!

Retrieved from http://srmdc.net/images/chapter6/63.gif, January 10, 2010.

12

EXAMPLES

- My SAT score
  - Is Interval
- A set of countries
  - Is nominal
- The first, third, and fifth person in a race
  - Is ordinal
- The number pinned on a sports person
  - Is nominal
- The number of pizzas a person can eat before getting sick
  - Is ratio

Modified from http://changingminds.org/explanations/research/measurement/types_data.htm, retrieved January 10, 2010.