

# DEFINITIONS

- **Statistics: Conducting studies to collect, organize, analyze, and understand data**
- **Population: The entire group that we are interested in**
- **Sample: Subset of the population**
- **Ways of collecting data:**
  - **Census: Collecting data from the whole population (impractical and expensive)**
  - **Survey (poll): Collecting data from a sample**

# **IMPORTANT POINT!**

- **The sample must be selected properly (i.e. all members of the population must have the same chance of being selected)**
- **If the sample is biased (i.e. some members of the population are more or less likely than some others to be selected), then the survey is not scientific. This means:**
- **THE RESULTS OF THE SURVEY ARE MEANINGLESS**

# DEFINITIONS continued

- **Kinds of Statistics**
  - **Descriptive: Information about the sample**
  - **Inferential: Information about the population**
- **Reminder: We can use inferential statistics only if the sample is not biased!**

# DEFINITIONS continued

- **Variable: A characteristic about which we collect data**
- **Variable examples:**
  - We want to know the favorite color of the members of a population
  - We want to know the height of the members of a population
  - We want to know the number of children the members of a population have
- **Data examples:**
  - Person X's favorite color is red (qualitative)
  - Person X's height is 63.2 inches (quantitative continuous)
  - Person X has 3 children (quantitative discrete)

# DEFINITIONS continued

- **Kinds of Studies**
  - **Observational – Researcher observes what is happening**
    - **Examples: Ages, incomes, preferences of the population**
  - **Experimental – Researcher manipulates a variable and tries to determine how that affects the other variable**
    - **Examples: Drug testing, exercise and diet studies**
  - **Quasi-experimental – Groups are not randomly assigned**

# **MORE ON EXPERIMENTAL STUDIES**

- **Variables**
  - **Independent** – The variable that is being manipulated
  - **Dependent** – The one that is being observed
- **Problems**
  - **Confounding variables**
  - **Artificial conditions**
  - **Non-random selection of subjects**
  - **Hawthorne effect**

# **SAMPLING**

- **RANDOM SAMPLING**
  - **RANDOM NUMBERS**
- **SYSTEMATIC SAMPLING**
  - **FIRST SUBJECT IS SELECTED RANDOMLY**
  - **THEN EVERY K-TH NUMBER**
- **STRATIFIED SAMPLING**
  - **DIVIDE POPULATION INTO GROUPS**
  - **SUBJECTS WITHIN GROUPS ARE RANDOMLY SELECTED**

# **SAMPLING continued**

- **CLUSTER**
  - **USE AN INTACT GROUP THAT IS REPRESENTATIVE OF A PORTION OF THE POPULATION**
- **CONVENIENCE**
  - **NOT SCIENTIFIC!**
- **CENSUS**
  - **NOT PRACTICAL**

# MISUSES OF STATISTICS

- **SUSPECT SAMPLES**
- **AMBIGUOUS AVERAGES**
- **IMPLIED CONNECTIONS**
- **MISLEADING GRAPHS**
- **FAULTY SURVEY QUESTIONS**
  - **PUSH POLL**