

## Population and sample

### Population and sample

- We are in sections 1.1, 1.2 of the textbook
  - *Statistics* is the science of analyzing and understanding data to make decisions in an ethical way.
  - The overall group of things we are interested in is called the *population*.
  - If we study part of a population to understand the whole population, the smaller group we study is called a *sample*.
  - A *parameter* is a numerical value that describes an entire population.
  - A *statistic* is a numerical value that describes a sample.
    - We often want to use these sample statistics, which we can measure, to estimate population parameters that we can't measure.
  - Ethical considerations - from the American Statistical Association:
    - Professional integrity
    - Integrity of data and methods
    - Responsibilities to Science/Public/Funder/Client
    - Responsibilities to Research Subjects
    - Responsibilities to Research Team Colleagues
    - Responsibilities to Other Statisticians or Statistics Practitioners
    - Responsibilities Regarding Allegations of Misconduct
    - Responsibilities of Employers
- See Ethical Guidelines for Statistical Practice, URL:  
<https://www.amstat.org/ASA/Your-Career/Ethical-Guidelines-for-Statistical-Practice.aspx>, ASA, 2018.
- *Experimental design* is the process of designing a statistical study so that the results will be meaningful.
    - Observational study - measure or observe members of the sample without trying to affect them
    - Experimental study - a treatment is randomly assigned to members of the sample
  - Sampling techniques
    - Convenience sampling
    - Simple random sampling
    - Stratified sampling
    - Cluster sampling
    - Systematic sampling
    - Multistage sampling

1. In a certain class, exam grades were compared with attendance to see whether attending class had an effect on the exam grades.

- (a) Is this an experiment or an observational study? Explain your answer in two to three sentences.

---

---

---

---

---

---

---

---

---

---

- (b) What is the implied population in this study?

- (c) Would it be feasible to study the entire population in this study? Why or why not?

2. Explain in your own words the difference between a population parameter and a sample statistic. Try to write a definition you could use to explain the concept to someone else.

---

---

---

---

---

---

---

---

3. Suppose that, in order to estimate the average height of Marshall students, we measure the height of 100 students and average the heights that we measured.

(a) What is the implied population?

(b) What is the sample?

(c) What is the population parameter in this situation?

(d) What is the sample statistic in this situation?

(e) Would it be feasible to study the entire population in this study? Why or why not?

4. Classify each sample as simple random, systematic, cluster, or stratified.
- (a) In a large school district, all teachers from two randomly selected buildings are interviewed to determine their opinion about student homework load.
  - (b) Every seventh customer entering a shopping mall is asked to select their favorite store.
  - (c) Nursing supervisors are selected at random to determine their annual salaries.
  - (d) Professors at Marshall are divided into groups according to their rank (instructor, assistant professor, etc) and then several are selected from each group to form a sample.
5. Classify each study as observational or experimental.
- (a) A study of the effects of education of employees on sales volumes involved obtaining information on highest degree acquired and sales volume of each of the selected 55 employees.
  - (b) Subjects were randomly assigned to two groups. One group was given an herbal supplement and the other group a placebo. After 6 months, the numbers of respiratory tract infections each group had were compared.