

## Chapter 1 Introduction to Statistics Notes

### STA 2023 SECTION 1.1 AN OVERVIEW OF STATISTICS

#### Learning Outcomes:

- 1) Define statistics
- 2) Identify populations and describes data sets.
- 3) Distinguish between populations and samples
- 4) Distinguish between parameters and statistics
- 5) Distinguish between descriptive and inferential statistics

“Statistics is the grammar of science.”

---Karl Pearson

Statistics – the study of procedures for collecting, describing, and drawing conclusions from information.

“**Statistics** is the science of collecting, organizing, analyzing, and interpreting data to make decisions.”

#### Sources of Information:

##### A. Population –

- Census –
- Parameter –

##### B. Sample –

- Sample Data –
- Statistic –

\*Ideally, we want the sample we choose to represent the population appropriately. Since this is rarely possible, we use methods of sampling (choosing a sample) that give a better chance of the sample reflecting the major characteristics of the population under consideration.

## Chapter 1 Introduction to Statistics Notes

### Distinguishing Between a Parameter and a Statistic:

1. In a survey of automobile owners, 6% said they had to change their engine control module at least once.  
Ans: \_\_\_\_\_
2. In a recent year, the average math score on the ACT for all graduates was 20.2.  
Ans: \_\_\_\_\_
3. In a random check of several hundred retail stores, the Food and Drug Administration found that 34% of the stores were not storing fish at the proper temperature. Ans: \_\_\_\_\_
4. Last year, a small company spent a total of \$5,150,694 on employees' salaries.  
Ans: \_\_\_\_\_

### Popular Methods of Sampling:

Simple Random Sample—A simple random sample of size  $n$  is chosen by a method in which each collection of  $n$  population items is equally likely to make up the sample.

Examples:

\*For this course, a sample will refer to a simple random sample (SRS).

statistics has two major branches:

Descriptive statistics:

Inferential statistics:

## Chapter 1 Introduction to Statistics Notes

**Exercise:** A study of 513 Internet-wide survey respondents found that 97% said music is important to them, and 83% said they actively look for new music.

- a) Identify the population and the sample.

Population-

Sample-

- b) Determine which part of the study represents the descriptive branch of statistics.

- c) What conclusions might be drawn from the study using inferential statistics?

## Chapter 1 Introduction to Statistics Notes

### STA 2023 SECTION 1.2 DATA CLASSIFICATION

#### Learning Outcomes:

- 1) Distinguish between quantitative and qualitative data.
- 2) Classify data with respect to the four levels of measurement: nominal, ordinal, interval, and ratio

#### Notes:

When collecting information, we get a \_\_\_\_\_. The items or subjects in a sample or a population are called the \_\_\_\_\_. Examples:

The characteristics of the individuals about which we collect information are called \_\_\_\_\_.  
Examples:

The \_\_\_\_\_ that we obtain are called \_\_\_\_\_.

#### Types of Variables:

##### 1. Qualitative or Categorical Variables –

a. Ordinal Variables –

b. Nominal Variables –

##### 2. Quantitative or Numerical Data –

a. Discrete Variables –

b. Continuous Variables –

## Chapter 1 Introduction to Statistics Notes

**Example 1:** Classify the data for each attribute as quantitative or qualitative, then by level of measurement.

- a) Hair color
- b) Letter grades (A, B, C, D, F)
- c) SAT scores
- d) Zip codes
- e) The final standings for the teams in the Pacific Division of the National Basketball Association.
- f) A collection of phone numbers.
- g) Sizes of drinks (small, medium, large, etc.)
- h) Heights in inches
- i) Place in a race (1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, etc.)
- j) Time of day (using a 12-hour clock)
- k) Salary

## Chapter 1 Introduction to Statistics Notes

### STA 2023 SECTION 1.3 DATA COLLECTION AND EXPERIMENTAL DESIGN

#### Learning Outcomes:

- 1) Distinguish between observational and experimental study
- 2) Analyze experimental designs
- 3) Identify and analyze different sampling techniques
- 4) Analyze survey questions and identify bias

The goal of every statistical study is to collect data and then use the data to make a decision.

A statistical study can usually be categorized as an observational study or an experiment.

1. **Observational study:** a researcher observes and measures characteristics of interest of part of a population but does not change existing conditions.

**Example:** Cancer rate due to smoking, brain problems after a head injury, automobile crashes due to Speed, etc.

2. **Randomized Experiment** – the investigator assigns treatments to the experimental unit at random. If there are large differences in outcomes among the treatment groups, we can conclude that the differences are due to the treatments.

**\*Confounding Variables –**

#### Data Collection

There are several ways to collect data.

- A **simulation** is the use of a mathematical or physical model to reproduce the conditions of a situation or process. Collecting data often involves the use of computers.
- A **survey** is an investigation of one or more characteristics of a population. Most often, surveys are carried out on *people* by asking them questions.

## Chapter 1 Introduction to Statistics Notes

### Experimental Design

The goal of sampling is to get a subset of the population that appropriately represents the population's characteristics.

If a procedure produces this result, it is said to be \_\_\_\_.

If the sample overestimates or underestimates a population value, it is said to be \_\_\_\_.

### **\*\*Sampling types:**

Most of the time, it's tough and costly to collect data from the entire population. why? as the population size is much bigger than we can handle properly. Ideally, we want the sample we choose to represent the population appropriately. Since this is rarely possible, we use sampling methods (choosing a sample) that give a better chance of the sample reflecting the significant characteristics of the population under consideration.

#### Methods of Sampling:

1. **Simple Random Sample** – A simple random sample of size  $n$  is a sample chosen by a method in which each collection of  $n$  population items is **equally likely** to make up the sample, just as in a lottery.

Examples:

2. **Samples of Convenience** – A sample that is not drawn by a well-defined random method.

Examples:

\* A sample of convenience may be okay if the sample does reflect the population appropriately.

## Chapter 1 Introduction to Statistics Notes

3. **Stratified Sampling** – A sample in which the population is divided into groups, called \_\_ \_\_\_\_, where the members of each \_\_ \_\_\_\_ have similar characteristics. Then a random sample is drawn from each stratum.

\*\* Example:

4. **Cluster Sampling** – A sample in which items are drawn from the population in \_\_ \_\_\_\_ or \_\_ \_\_\_\_\_. This type of sample is useful when a straight forward simple random sample is difficult to take. Examples:

The difference between cluster sampling and stratified sampling:



## Chapter 1 Introduction to Statistics Notes

5. **Systematic Sampling** – A sample in which population items are ordered in some way and then it is decided how frequently to take items.

Examples:

6. **Voluntary Response Sample** – A sample in which the respondents themselves decide to be involved in a study or process. Examples: