

## Chapters 1-4

Stat 128: Elementary Statistics

A Document By:  
**Nathan Warner**



July 03, 2023  
Computer Science  
Joliet Junior College  
United States

## Contents

<b>1</b>	<b>Learning Outcomes</b> . . . . .	<b>3</b>
<b>2</b>	<b>Chapter 1:</b> . . . . .	<b>4</b>
2.1	1.1: Introduction to the Practice of Statistics . . . . .	4
2.2	1.2: Observational Studies versus Designed Experiments . . . . .	9

# 1 Learning Outcomes

## Chapter 1:

1. Define data collection techniques including observational studies and design of experiments.
2. Identify appropriate sampling methods.

## Chapter 2:

1. Differentiate qualitative and quantitative data graphically. This includes graphs such as bar plots, histograms, and dot plots.

## Chapter 3:

1. Calculate measures of central tendency for data.
2. Explain the concept of resistance.
3. Decide which measure of central tendency to report for various data sets.
4. Determine measures of dispersion for data.
5. Determine standard scores, percentiles, and quartiles.
6. Identify outliers using quartiles.
7. Interpret boxplots.

## Chapter 4:

1. Evaluate the linear correlation coefficient for bivariate quantitative data.
2. Evaluate whether the coefficient is significant at a given level.
3. Explain the difference between correlation and causation.
4. Determine the least-squares regression equation for a given set of bivariate data.
5. Predict values of the dependent variable using the least-squares regression equation.
6. Interpret the slope and intercept of the least-squares regression equation.
7. Test the requirements of the least-squares regression model using residual analysis.
8. Determine and interpret the coefficient of determination.
9. Graphically analyze bivariate quantitative data for outliers and influential observations.
10. Describe the association between two qualitative variables using conditional distributions.
11. Explain Simpson's Paradox.

## 2 Chapter 1:

### 2.1 1.1: Introduction to the Practice of Statistics

#### Objectives for this section.

1. Define Statistics and Statistical Thinking
2. Explain the Process of Statistics
3. Distinguish between Qualitative and Quantitative Variables
4. Distinguish Between Discrete and Continuous Variables.

#### Define Statistics and Statistical Thinking:

Statistics is the science of collecting, organizing, summarizing, and analyzing information to draw conclusions or answer questions. In addition, statistics is about providing a measure of confidence in any conclusions.

**Note:** We must report a measure of our confidence in our results because we do not have 100% certainty our answers are correct.

The information referred to in the definition above is *data*. **Data** are a "fact or proposition used to draw a conclusion or make a decision." Data describes characteristics of an individual.

One crucial thing to understand about **data**, is that it **varies**. One thing that makes an interesting study is the fact that the data within the study varies. A study about number of hearts a human has is not only uninteresting but not worth doing. This is because the data does not vary.

#### **Two Major Goals:**

1. Describe Variability.
2. Understand sources of Variability.

In Statistics, the same approach to solving a problem can still lead to different results. This does not happen in a math class.

#### Explain the Process of Statistics.

First let's define some vocabulary:

- **Population:** The entire group to be studied is called the population.
- **Sample:** In statistics, it is often impractical or impossible to get access to the entire **population**, which is why we only look at a **sample**. A sample is a **subset** of the population being studied.
- **Individual:** An individual is a person or object that is a member of the population being studied.
- **Statistic:** A statistic is a numerical summary of a sample.
- **Descriptive Statistics:** Descriptive statistics consist of organizing and summarizing data. Descriptive statistics describe data through numerical summaries, tables, and graphs.
- **Inferential Statistics:** inferential Statistics uses methods that take a result from a sample, extend it to the population, and measure the reliability of the result.
- **Parameter:** A parameter is a numerical summary of a population.

### The process of statistics:

1. Identify the problem to be solved. It is important to clearly lay out the questions that the researcher wants answered, along with clearly specifying which population the study applies.
2. Collect the data.
3. Describe the data.
4. Perform inference.

**Example: The AP - National Constitution Center conducted a national poll to learn how adult Americans feel existing gun-control laws infringe on the second amendment to the U.S Constitution**

**The Following statistical process allowed the researchers to conduct their study.**

1. **Identify the research objective.:** The researchers wished to determine the percentage of adult Americans who believe gun-control laws infringe on the public's right to bear arms.
2. **Collect the information needed to answer the question posed in (1).:** It is unreasonable to expect to survey the more than 200 million adult Americans to determine how they feel about gun-control laws. So the researchers surveyed a sample of 1007 adult Americans. Of those surveyed, 514 stated they believe existing gun-control laws infringe on the public's right to bear arms.
3. **Describe the data.:** Of the 1007 individuals in the survey, 51% believe existing gun-control laws infringe on the public's right to bear arms. This is a descriptive statistic, because its value is determined from a sample.
4. **Perform inference.:** The researchers at the AP - National Constitution Center wanted to extend the results of the survey to **all** adult Americans. When generalizing results from a sample to a population, the results are **uncertain**. To account for this uncertainty, researchers reported a 3% *margin of error*. This means that the researchers feel fairly certain (in fact, 95% certain) that the percentage of *all* adult Americans who believe existing gun-control laws infringe on the public's right to bear arms is somewhere between 48% and 54%

### Distinguish between Qualitative and Quantitative Variables

First let's define some vocab:

- **Variables:** The characteristics of the individuals in a study. Variables vary, which means they can take on different values.
- **Constants:** Variables that do not vary. Inferential statistics is not necessary with constants.

One goal of research is to learn the causes of variability.

Variables can be classified into two groups: qualitative and quantitative.

- **Qualitative, or categorical variables** allow for the classification of individuals based on some attribute or characteristic.
- **Quantitative variables** provide numerical measures of individuals. The values of a quantitative variable can be added or subtracted and provide meaningful results.

**Example:** Determine whether the following variables are qualitative or quantitative.

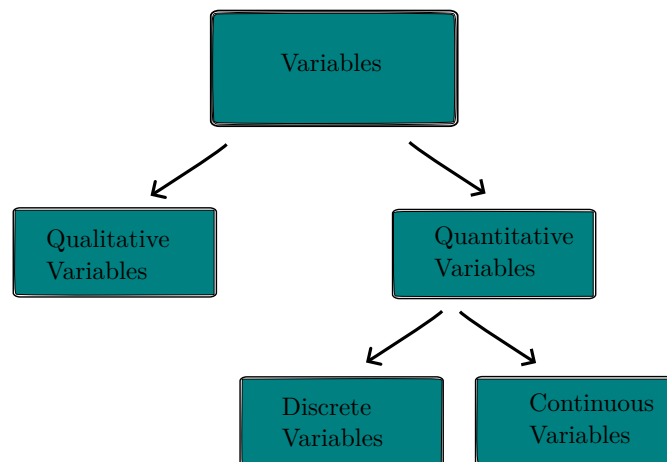
- a.) **Gender.:** Qualitative
- b.) **Temperature.:** Quantitative
- c.) **Number of days during the past week that a college student studied.:** Quantitative
- d.) **Zip Code.** Qualitative

**Caution:** A numeric value does not automatically suggest a variable is quantitative.

***Distinguish between Discrete and Continuous Variables.***

- A **discrete variable** is a quantitative variable that has either a finite number of possible values or a countable number of possible values. A discrete variable cannot take on every possible value between any two possible values.
- A **continuous variable** is a quantitative variable that has an infinite number of possible values that are not countable. A continuous variable may take on every possible value between any two values. Continuous variables typically result from measurement. Continuous variables are often rounded. If a certain make of car gets 24 miles per gallon (mpg) of gasoline, its miles per gallon must be greater than or equal to 23.5 and less than 24.5, or  $23.5 \leq mpg < 24.5$

This Figure illustrates the relationship among qualitative, quantitative, discrete, and continuous variables.



**Example:** Distinguish whether the quantitative variables are discrete or continuous.

- a.) **The number of heads obtained after flipping a coin five times.:** Discrete
- b.) **The number of cars that arrive at a McDonald's drive through between 12:00 PM and 1:00 PM:** Discrete
- c.) **The Distance a 2011 Toyota Prius can travel in city driving conditions with a full tank of gas.:** Continuous

**Vocab:**

- The list of observed values for a variable is **data**.
- **Qualitative data** are observations corresponding to a **qualitative variable**.
- **Quantitative data** are observations corresponding to a quantitative variable.
- **Discrete data** are observations corresponding to a discrete variable.
- **Continuous data** are observations corresponding to a continuous variable.

**Example: Distinguish between Variables and Data**

The following table presents a group of selected countries and information regarding these countries.

Identify the individuals, variables, and data.

Country	Government Type	Life Expectancy (Years)	Population (in millions)
Australia	Federal parliamentary democracy	81.63	21.3
Canada	Constitutional monarchy	81.23	33.5
France	Republic	80.98	64.4
Morocco	Constitutional monarchy	75.47	31.3
Poland	Republic	75.63	38.5
Sri Lanka	Republic	75.14	21.3
United States	Federal Republic	78.11	307.2

**Qualitative:** Government Type

**Quantitative:** Life Expectancy and Population

**Continuous:** Life Expectancy

**Discrete:** Population

**Data:** Everything under Government Type, Life Expectancy, and Population.

## All Vocab / Concepts From Section 1.1

---

- **Population:** The entire group to be studied is called the population.
- **Sample:** In statistics, it is often impractical or impossible to get access to the entire **population**, which is why we only look at a **sample**. A sample is a **subset** of the population being studied.
- **Individual:** An individual is a person or object that is a member of the population being studied.
- **Statistic:** A statistic is a numerical summary of a sample.
- **Descriptive Statistics:** Descriptive statistics consist of organizing and summarizing data. Descriptive statistics describe data through numerical summaries, tables, and graphs.
- **Inferential Statistics:** inferential Statistics uses methods that take a result from a sample, extend it to the population, and measure the reliability of the result.
- **Parameter:** A parameter is a numerical summary of a population.
- **Variables:** The characteristics of the individuals in a study. Variables vary, which means they can take on different values.
- **Constants:** Variables that do not vary. Inferential statistics is not necessary with constants.
- **Data:** The list of observed values for a variable.
- **Qualitative data** are observations corresponding to a **qualitative variable**.
- **Quantitative data** are observations corresponding to a quantitative variable.
- **Discrete data** are observations corresponding to a discrete variable.
- **Continuous data** are observations corresponding to a continuous variable.

### Concepts:

- Statistics and Statistical Thinking.
- Describe Variability
- Understand Sources of variability
- Statistical studies are concerned with both describing the variability in the data and understanding the sources of variability in data. Understanding the sources allows researchers to control it and reach better conclusions.
- The process of statistics
- Inferential/Descriptive Statistics
- Variables
  - Qualitative (Categorical) / Quantitative
  - Discrete / Continuous
- Data
  - Qualitative (Categorical) / Quantitative
  - Discrete / Continuous



## 2.2 1.2: Observational Studies versus Designed Experiments