Name _____

Chapter 1 Getting Started

Section 1.1 What is Statistics?

Objective: In this lesson you learned how to identify variables in a statistical study, distinguish between quantitative and qualitative variables, identify populations and samples, determine the level of measurement, and compare descriptive and inferential statistics.

Important Vocabulary			
Statistics	Individuals	Quantitative Variable	Qualitative Variable
Population Data	Sample Data	Nominal Level	Ordinal Level
Interval Level	Ratio Level	Descriptive Statistics	Inferential Statistics

I. Introduction

Statistical methods enable us to:

What is **statistics**?

Focus Points:
- Identify variables in a statistical study
- Distinguish between quantitative and qualitative variables
- Identify populations and samples
- Distinguish between parameters and statistics

What must be remembered about properly applied statistical procedures?

Statistical results should be interpreted by whom?

Define the following terms:

1. Individuals –

2. Variable –

3. Quantitative Variable –

4. Qualitative Variable –

Note: Qualitative variables can also be called _____ variables.

5. Population Data –

6. Sample Data –

7. Population Parameter –

8. Sample Statistic –

II. Levels of Measurement: Nominal, Ordinal, Interval, Ratio

Another way to classify data is according to one of the following four *levels of measurement*. These levels indicate:

Focus Point:
- Determine the level of measurement

The **nominal level of measurement**:

Data at the **ordinal level** may be arranged in some order, but:

The **interval level** applies to data:

The **ratio level** applies to data:

To determine the level of measurement of data:

Level of Measurement	Suitable Calculation
Nominal	
Ordinal	
Interval	
Ratio	

III. **Critical Thinking**

Descriptive Statistics:

Focus Point:

- Compare descriptive and inferential statistics

Inferential Statistics:

Section 1.1 Examples – What is Statistics?

(1) Television station QUE wants to know the proportion of TV owners in Virginia who watch the station's new program at least once a week. The station asked a group of 1000 TV owners in Virginia if they watch the program at least once a week.

a. Identify the individuals of the study and the variable.

b. Do the data comprise a sample? If so, what is the underlying population?

c. Is the variable qualitative or quantitative?

d. Identify a quantitative variable that might be of interest.

e. Is the proportion of viewers in the sample who watch the new program at least once a week a statistic or a parameter?

(2) The following describe different data associated with a state senator. For each data entry, indicate the corresponding *level of measurement* and explain your reasoning.

a. The senator's name is Sam Wilson.

b. The senator is 58 years old.

c. The years in which the senator was elected to the Senate are $1980, 1986, 1992,$ and 1998.

d. The senator's total taxable income last year was $\$878{,}314$.

e. The senator surveyed his constituents regarding his proposed water protection bill. The choices for response were strong support, support, neutral, against, or strongly against.

f. The senator's marital status is "married."

g. A leading news magazine claims the senator is ranked seventh for his voting record on bills regarding public education.

Section 1.2 Random Samples

Objective: In this lesson you learned the importance of random samples, how to construct and simulate random samples, and describe types of sampling.

Important Vocabulary		
Simple Random Sample	Simulation	Stratified Sampling
Systematic Sampling	Cluster Sampling	Convenience Sampling

I. Simple Random Samples

Focus Points:
- Explain the importance of random samples
- Construct a simple random sample using random numbers
- Simulate a random process

A **simple random sample** of n measurements:

Important features of a simple random sample are:

-
-
-

An easy way to select random numbers is to use a(n) _____.

The term random should not _____!

How to draw a random sample:

1.
2.
3.

Another important use of random-numbers is in _____. The word simulation refers to:

A **simulation**:

Sampling with replacement means:

II. Other Sampling Techniques

Focus Point:

- Describe stratified sampling, cluster sampling, systematic sampling, and convenience sampling

Stratified sampling:

Systematic sampling:

The advantage of a systematic sample is that it is easy to get. However, one danger in using systematic sampling is:

Cluster sampling:

Cluster sampling is primarily used by:

Convenience sampling:

In some cases, this may be all that is available, and in many cases, _____. However, convenience sampling does run the risk of _____. It is good advice to be very _____when the data come from the method of convenience sampling.

III. Critical Thinking

Define the following terms:

1. Sampling Frame –

2. Undercoverage –

3. Sampling Error –

4. Nonsampling Error –

Section 1.2 Examples – Random Samples

(1) Is open space around metropolitan areas important? Players of the Colorado Lottery might think so because some of the proceeds of the game go to fund open space and outdoor recreational space. To play the game, you pay one dollar and choose six different numbers from the group of numbers 1 through 42. If your group of six numbers matches the winning group of six numbers selected by simple random sampling, then you are a winner of a grand prize of at least $\$1.5$ million.

a. Is the number 25 as likely to be selected in the winning group of six numbers as the number 5?

b. Could all the winning numbers be even?

c. Your friend always plays the numbers $1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6$. Could she ever win?

(2) Use a random-number table to simulate the outcome of tossing a balanced (that is, fair) penny 10 times.

a. How many outcomes are possible when you toss a coin once?

b. There are several ways to assign numbers to the two outcomes. Because we assume a fair coin, assign an even digit to the outcome “heads” and an odd digit to the outcome “tails.” Then starting at block 3 of row 2 of Table 1 in the Appendix, list the first 10 single digits.

c. What are the outcomes associated with the 10 digits?

d. If you start in a different block and row of Table 1 in the Appendix, will you get the same sequence of outcomes?

Section 1.3 Introduction to Experimental Design

Objective: In this lesson you learned how to plan a statistical study.

Important Vocabulary			
Census	Sample	Observational Study	Experiment
Placebo Effect	Control Group	Randomization	Replication

I. Planning a Statistical Study

What are two essential components for obtaining reliable information?

Focus Point:
- Discuss what it means to take a census

Basic guidelines for planning a statistical study.

1.

2.

3.

4.

5.

6.

7.

One issue to consider is whether to use _____

_____.

In a **census**:

If we use data from only part of the population of interest, you have a(n) _____.

In a **sample**:

II. Experiments and Observations

Focus Point:
- Describe simulations, observational studies, and experiments
- Identify control groups, placebo effects, completely randomized experiments, and randomized block experiments

When gathering data for a statistical study, you want to:

In an **observational study**: In an **experiment**:

Statistical experiments are commonly used to determine the effect of a treatment. However, the design of the experiment needs to _____ for other possible causes of the effect. For instance, in medical experiments, the _____ is the improvement or change that is the result of patients just believing in the treatment, whether the treatment itself is effective or not.

The **placebo effect**:

A common way to assign patients to treatment and control groups is by using a random process. This is the essence of a _____.

A **completely randomized experiment:**

A **block:**

Many experiments are also _____, meaning neither the individuals in the study nor the observers know which subjects are receiving the treatment.

There is a **control group**:

Randomization:

Replication:

III. Surveys

Focus Point:
- Discuss potential pitfalls that might make your data unreliable

A common means to gather data is to ask people questions. This process is the essence of _____. Sometimes the possible responses are simply 'yes' or 'no'. Other times the respondents choose a number on a scale, called a(n) _____.

Some potential pitfalls of a survey

Nonresponse:

Truthfulness of response:

Faulty recall:

Hidden bias:

Vague wording:

Interviewer influence:

Voluntary response:

Sometimes our goal is to understand the cause-and-effect relationships between two or more variables. Such studies can be complicated by _____ or _____.

A **lurking variable:**

Two variables are **confounded** when:

IV. Choosing Data Collection Techniques

Which technique is best? The answer depends on:

- Surveys:

- Observational studies:

- Experiments:

A data collection plan identifies:

-
-
-
-
-
-

Section 1.3 Examples – Introduction to Experimental Design

(1) Which technique for gathering data (sampling, experiment, simulation, or census) do you think might be the most appropriate for the following studies?

a. Study the effect of stopping the cooling process in a nuclear reactor.

b. Study the amount of time college students taking a full course load would spend watching TV.

c. Study of the effect on bone mass of a calcium supplement given to young girls.

d. Study of the credit hour load of *each* student enrolled at your college at the end of the drop/add period this semester.

(2) In 1778 Captain James Cook introduced goats to the Hawaiian Islands. It was later observed that the Silver Sword plant appeared to be less and less common. Botanists suspected the goats to be the cause and conducted a statistical study. They set up stations around the island with similar climate and soil conditions. Each station consisted of two plots of land, one with a fence around it to keep the goats out. Identify the following:

a. the treatment –

b. the experimental group –

c. the control group –

(3) Comment on the usefulness of the data collected as described.

a. A uniformed police officer interviews a group of 20 college freshmen. She asks each one his or her name and then if he or she has used an illegal drug in the last month.

b. Jessica saw some data that show that cities with more low–income housing have more homeless people. Does building low-income housing cause homelessness?