Section 1.1 Types of Data

Get Started – What is the difference between information and knowledge?

- How is qualitative data different from quantitative data?
- How is a population different from a sample?
- How do you tell a statistic from a parameter?
- What is the difference between discrete data and continuous data?

Get Started – What is the difference between information and knowledge?

What is it that separates good students from poor students?

Are good students just born smarter?

They say that hard work can overcome talent, but many students claim to spend hours studying and still do poorly. Even worse, after spending all that time studying, such a student might get discouraged and give up. They think that they aren't a "math person". They must not be if they studied 100's of hours for the test and still failed…

Well, it's not JUST about hard work. It's not even necessarily about working efficiently. What researchers have discovered is that what you are thinking about is extremely important for learning effectively.

Metacognition is thinking about your thinking. In our case, we will be thinking about our learning. We will be learning about our learning process. As we improve our process, we will improve as students (we learn better), which will hopefully improve our grades!

Most Math Students have the bad habit of only studying the day before a test. There are two serious disadvantages attached to this method of study.

The first reason is that this does not allow for regular practice of mathematics. Just like becoming a good athlete or musician or cook or artist, to become good at mathematics you must practice regularly.

One recent study found that it takes 10,000 hours of deliberate practice to achieve mastery in a field. Deliberate practice is a highly structured activity engaged in with the specific goal of improving performance.

The second reason cramming before a test isn't a good idea has to do with short term vs long term memory. It has been found that within 24 hours - on average – you will forget up to 80% of what you just learned. If the study material is reviewed regularly, you remember more, and it takes longer to forget what you have learned. You turn the short-term memory (Snapchat) into a

long-term memory (flash drive). You goal should be long term knowledge, not short-term success.

Building up your mathematical ability is like building a house. You need a strong foundation, because future math subjects are built upon the material that has come before.

When you cram for a test and get a decent grade, you may deceive yourself into mistakenly believing you have mastered an important skill.

You knew the material for a short time (until the test was over), but you won't remember it three weeks from now when it comes up in class again. So, you really didn't KNOW it.

According to the Forgetting Curve, you forget 80% of what you've learned in the first 24 hours. This is like trying to build a house with a cardboard foundation. After 24 hours, the cardboard gets wet and loses 80% of its strength. You try to build on top of this weak foundation and eventually the whole thing comes tumbling down.

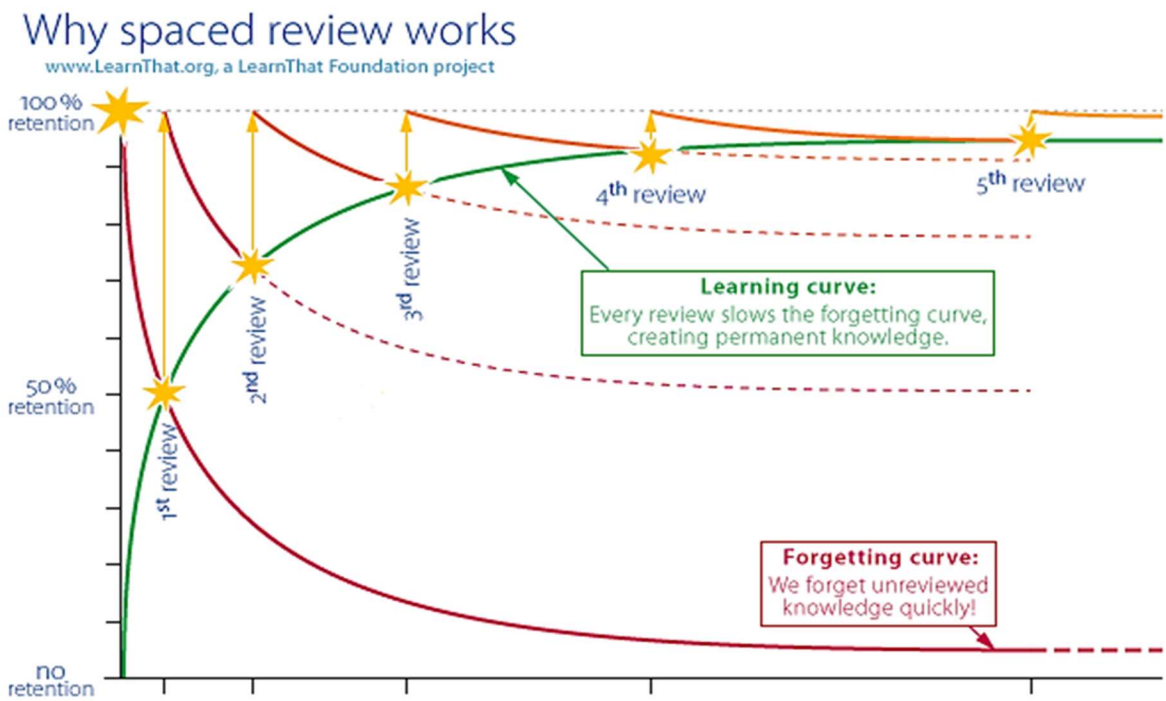
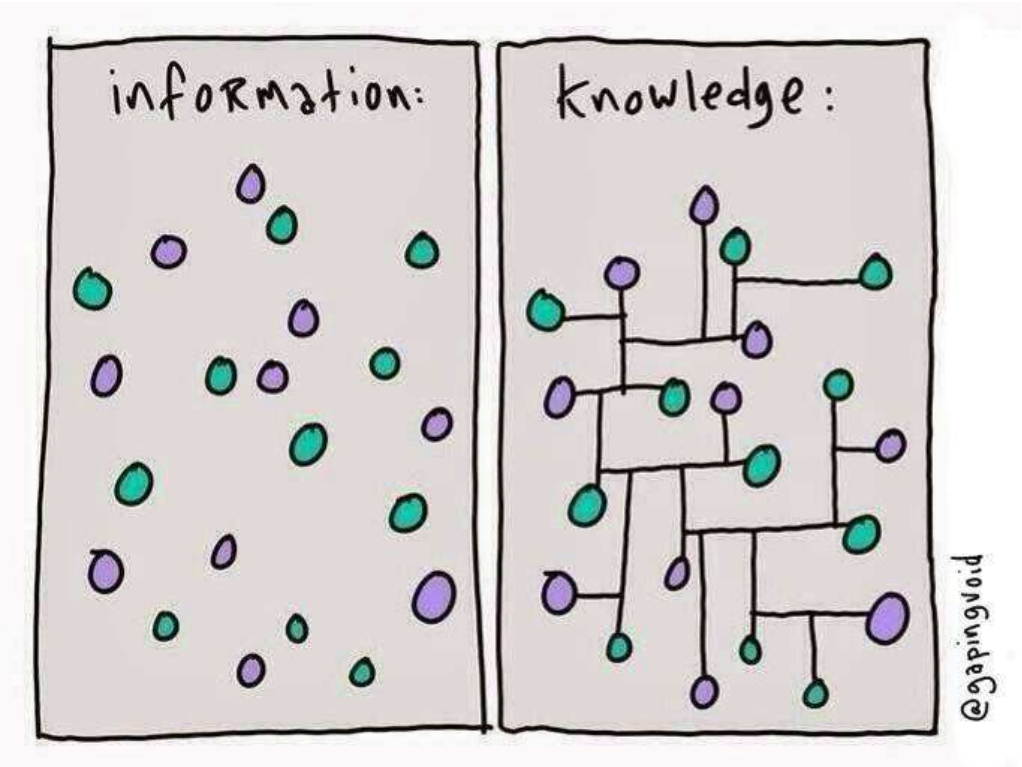


Figure 1- The Forgetful Curve

Believing you have mastered a skill that you have mostly forgotten is an example of poor metacognition (you do not have an accurate gauge of your true ability). There is a HUGE difference between knowing that you know it and believing that you know it (but forgetting most of it a week later).

The hardest part of learning is making the transition from knowing a bunch of individual facts to being able to put the pieces together in a sensible, useable way.



The challenge you will have is discovering how to organize the information in the chapter into knowledge. Near the end of this chapter, we'll examine how concept maps can help you to organize the information in the chapter and prove to yourself that you understand the concepts in the chapter.

How is qualitative data different from quantitative data?

Key Terms

Qualitative data Quantitative data Categorical data

Summary

Suppose I poll 60 elementary school students and ask them: "What is your favorite ice cream flavor?" Ice cream flavor is NOT a numerical quantity. We call data this type of data **qualitative data**. (Instead of thinking about "quantities" we are thinking about "qualities". Qualities include things like gender, major, car brands, names, places, etc.)

In practice, one could have many, many qualities that result from a single poll question. For example, "What is your name?", is a question that results in a lot of qualitative data. Unfortunately, analyzing this type of data is often difficult because almost every piece of data is different.

What we typically do is break qualitative data down into a manageable number of categories that we can analyze. For this reason, the terms "qualitative data" and **categorical data** are frequently used interchangeably in the context of statistics.

Suppose I poll 100 college students and ask: "How many times did you go to office hours to seek help when you had difficulty in class?" The type of data that results from this type of question is **quantitative data**. ("quant" is the root of the word "quantity" which measures "how much of something we have." Quantitative data measures the amount of something and results in a number.

Sometimes, data made up of numbers is actually qualitative data. For example, "The 210 freeway" is qualitative data even though it has a number in it, because 210 is not measuring the amount of anything it is just serving as a name. Zip codes and social security numbers are also not quantitative data.

A good way for you to tell whether a number is acting as quantitative data is to ask yourself – "Is this number measuring or counting the amount of anything?" If it is, then the number is quantitative data.

Sometimes we can take qualitative data and turn it into quantitative data. For example, customer satisfaction ratings of "Poor", "Fair", "Good", and "Excellent" could be represented with numbers as 1, 2, 3, and 4 respectively. By converting the qualitative data to quantitative data, we can apply various statistical techniques.

Notes

Guided Example 1	Practice
Classify each of the following as quantitative or qualitative data. a. The president of the US is tall. **Solution** Tall is not a numerical quantity son this data is qualitative. b. The fire fighter is 5 ft 10 inches tall. **Solution** We can convert this data to inches, 70 inches. Since this data measures an amount of height and is a number, this is quantitative data. c. The area code is 626. **Solution** Although this data is a number, it does not measure an amount of something. So, this is qualitative data.	Classify each of the following as quantitative or qualitative data. a. Sam Dean has blue eyes. b. Dave Reynolds wears a size 13 shoe. c. A student has completed 3 math classes.

How is a population different from a sample?

Key Terms

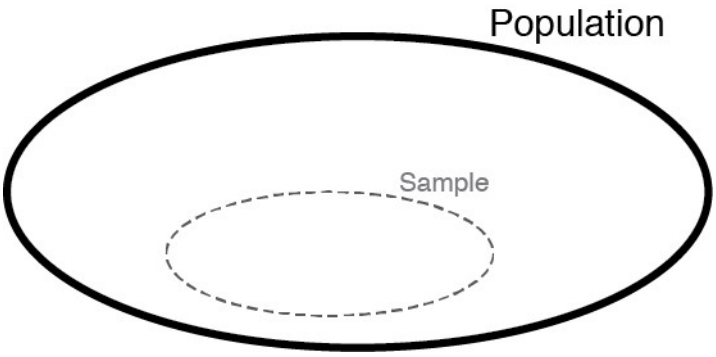
Population Sample

Summary

In statistics, we frequently want to understand data that comes from a particular group. For example, suppose I want to have a better understanding of the student body at a college.

The **population** is the entire group we are looking to study. In this case, our population is all students at the college.

Since it is generally not possible to collect data on an entire population (we don't have the time to survey all students) we collect data from a smaller group or subset taken from the population. This smaller group is called the **sample**.



An example of a sample for a student population would be all the students in a particular math class. In general, we would like our sample to do a good job of representing our population. Using students from a particular math class as our sample probably won't do a good job of representing all of the students at the college.

Notes

Guided Example 2	Practice
For each pair determine which is the sample and which is the population. a. The professional women's basketball team Phoenix Mercury and all basketball teams. **Solution** All basketball teams is the population. A portion of this population is the professional women's basketball team the Phoenix Mercury. Since this is a subset of the population, the Phoenix Mercury is a sample. b. People with college degrees and people who have attended college. **Solution** The larger population is people who have attended college. A subset of this population is the sample people with college degrees.	For each pair determine which is the sample and which is the population. a. Students enrolled in college and students enrolled at the University of Arizona. b. Registered voters in the United States and registered voters likely to vote in the United States.

How do you tell a statistic from a parameter?

Key Terms

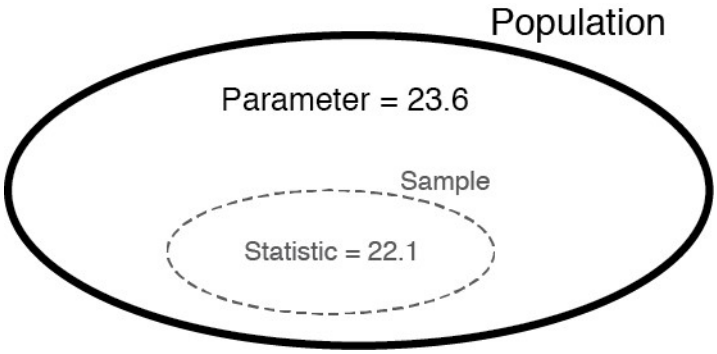
Statistic Parameter

Summary

A parameter is distinguished from a statistic on the basis of whether it corresponds to a population or sample. A number that represents a characteristic of the population is called a **parameter**. A number that represents a characteristic of the sample is called a **statistic**.

For example, we could define the population to be all college students. If we calculated the average age of all college students and found it to be 23.6 years old, this average would be a parameter, because it was calculated using the ages of ALL college students.

If I randomly selected 100 students at the University of Nevada and calculated that their average age was 22.1 this would be a statistic, because we only used the ages of a sample of the whole population.



Notes

Guided Example 3	Practice
Determine which **bold** phrase is a statistic and which is a parameter. a. In the Substance Abuse and Mental Health Services Administration survey, **13.2%** of respondents said they have driven under the influence of alcohol. **Solution** A survey is administered to a group of people which is presumably the population. From this group, only a portion will respond so the respondents are a sample. This means the number 13.2% is a statistic. b. In a Gallup poll of 1,023 renters, it was found that **58.2%** of them said that they had only wireless phones. **Solution** This poll was conducted with 1023 renters which is a sample of the group all renters. This means that 58.2% is a statistic.	Determine which **bold** phrase is a statistic and which is a parameter. a. In the 2014-2015 school year the retention rate in Math 150 was **89.6%** while the retention rate in all math courses at PCC was **79.4%.** b. The **average height of Americans** is more than then **average height of women in America.**

What is the difference between discrete data and continuous data?

Key Terms

Discrete data Continuous data

Summary

Quantitative Data can be further broken two into two types: discrete or continuous.

Discrete data is data that you can count on your fingers (assuming you had an infinite supply of fingers). For instance, we can count the number of eggs a hen lays similar to how we count on our fingers. Or we could count the number of people in a room on our fingers. **Continuous** data is data generated from measurements. We don't count how tall someone is, we measure how tall they are. Heights are continuous data.

Discrete data is quantitative data that is counted, not measured. We count the number of people or eggs. We don't measure the number of eggs.

Continuous data is quantitative data that is measured, not counted. We measure heights, weights, and volume. We don't count them.

For example, suppose I want to measure precisely how tall someone is. Our convention is to give our height in feet and inches, but this is really an **approximation**. It gives our height to the nearest inch. Two people who claim to be 5 feet 4 inches tall are probably not *exactly* the same height.

Notes

Guided Example 4	Practice
Classify each of the following pieces of quantitative data as either continuous or discrete. a. Number of math classes you have taken. **Solution** Since you count the number of math classes you have taken, this is discrete data. b. The speed of the Gold Line Metro train. **Solution** The speed of the Gold Line Metro train is measured so it is continuous data.	Classify each of the following pieces of quantitative data as either continuous or discrete. a. The length of your foot. b. Number of children in a household.