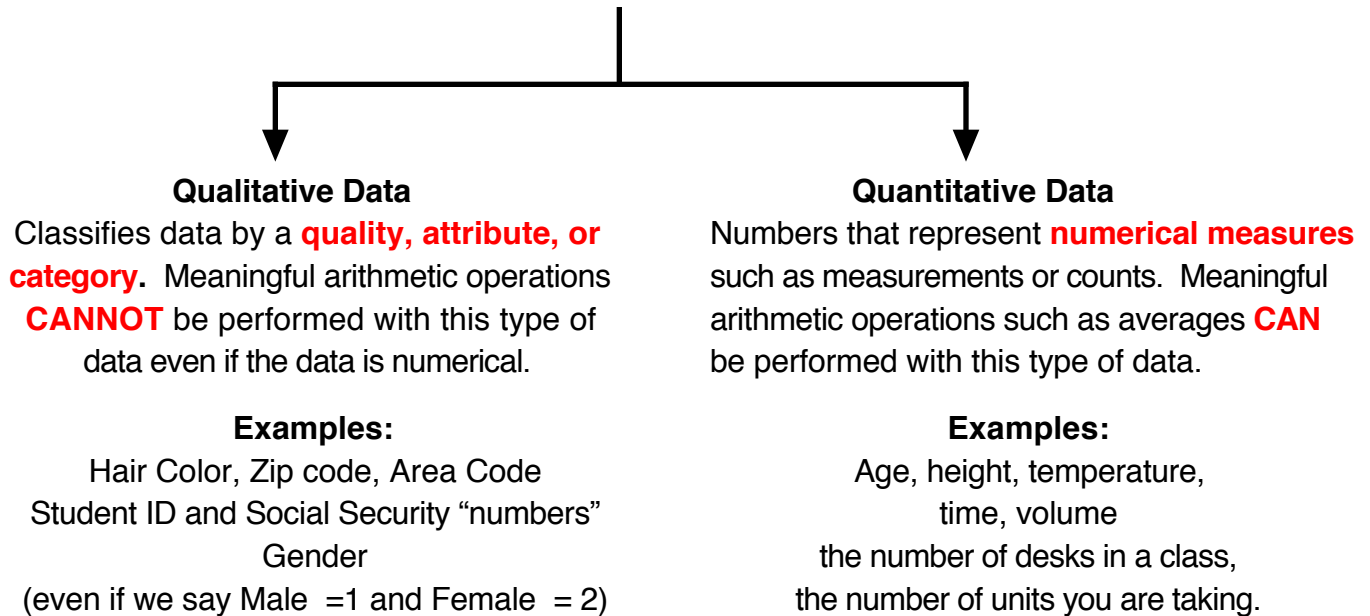


Chapter 1 Review

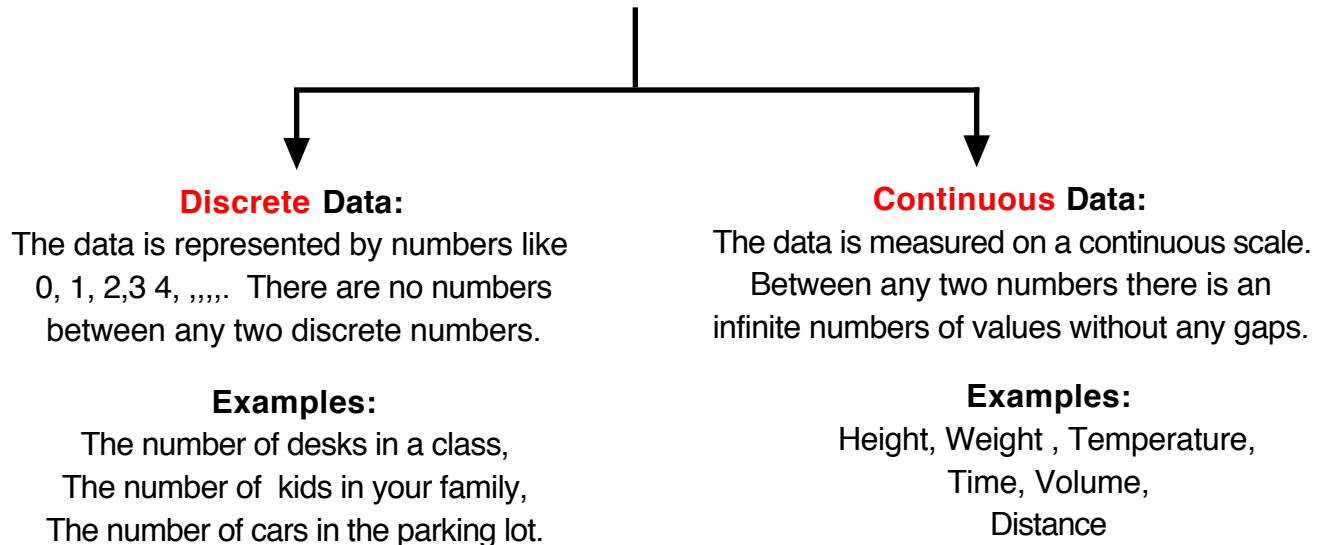
Data

Observations, Measurements or Responses that are collected

There are Two Types of Data



There are two types of
Quantitative Data



The Two Types of Statistics

Descriptive Statistics is the act of Collecting, Organizing, Displaying and Summarizing data.

Inferential statistics involves the use of methods that allow us to take known numerical properties about a **sample** of the population data and then use those numerical properties to **infer what numerical properties MIGHT be true** about the **entire population**.

Population versus Sample

A **Population** is the **entire collection** of things that will be studied. A **Census** is a collection of information about a characteristic that was collected from **the entire population**.

Sample is a collection that consists of only **part of the entire population**. A sample is considered to be a **Random Sample** if every individual member of the population has an **equal chance of being selected**.

Parameter versus Statistic:

A **Parameter** is a **measurement** describing a **numerical property** about a **population**.

A **Statistic** is a **measurement** describing a **numerical property** about a **sample**.

Common Errors in the Collection of Data

Self Selected Sample, Self Interest Study, Samples without basic controls, Small Sample Size, Not a Random Sample, Loaded Question, Missing Data.

Accurate versus Precise

Accurate Numbers count or measure **exactly what the true count or measure is**.

Precise Numbers: **The number of decimal places in the number** is a measure of how precise a number is.

Observational Studies

Observational Study: The subject is observed. **No attempt to modify the subject** in any way is allowed. You can measure and record information about the characteristics being observed.

A **Cross Sectional Study** occurs when subjects are observed at the current **point in time**.

A **Retrospective Study** occurs when data on the subjects is collected **from the past** by reviewing old records, interviews or other studies previously performed and recorded.

A **Longitudinal Study** is a study where the subject or subjects are observed at regular intervals over a period of time and measurements of a selected trait are recorded at each time period **starting from the present time**.

Experiments

Experiment: The subject is given a **treatment** and then the subject is observed to see what effect the treatment has on a specific trait.

The **Experimental Group** (or treatment group) is the group of subjects that are given the treatment.

The **Control Group** (or Placebo Group) is a group of subjects that are given a “fake “ treatment called a **placebo**. The control group thinks it is being treated but in fact they are not. effective.

Blinding

The **Control Group** (or Placebo Group) is “blind” to the fact that they do not get the real treatment.

Double Blind Experiments

Both the persons conducting the treatment and the subjects being treated are “blind” to who is in the **Experimental Group** (or treatment group) and who is in the **Control Group**.

Correlation and Causality Correlation does not imply Causality

Correlation: An event is **correlated** with an outcome if when **the event is present then the outcome is also present**.

Causality is the relationship between an action and an outcome in such a way that the **action is the cause the outcome**.

Confounding occurs when **more than one action may have an affect on an outcome**.

Non Sampling Error

A **Non Sampling Error** occurs in the collection phase of the experiment. The data was recorded incorrectly. The sample data was collected from a sample that was not random. The data was collected but the measuring device was not working correctly. Some data was lost during the collecting phase or recording phase. These types of errors create errors in the data itself. Any sample with this type of error cannot be used and must be ignored.

Sampling Error (the difference in sample and population parameters)

The numerical properties about a sample cannot be expected to be the same as the numerical properties of the entire population. The difference between the numerical properties about the **sample** and the numerical properties of the **entire population** is called **Sampling Error**.

A **Sampling Error** does not mean that an actual error was made in the techniques used in the experiment. It is a simple acknowledgment that the numerical properties about a **sample cannot be expected** to be an exact match for the numerical properties of the **entire population**. This type of error cannot be corrected or eliminated but the size of the error can be reduced with large sample sizes.