

What is Statistics?

Chapter 1





Learning Objectives

- LO1** List ways statistics is used.
- LO2** Know the differences between descriptive and inferential statistics.
- LO3** Understand the differences between a sample and a population.
- LO4** Explain the difference between qualitative and quantitative variables.
- LO5** Compare the differences between discrete and continuous variables.
- LO6** Recognize the levels of measurement in data.

Why Study Statistics?

Why Study Statistics?

1. Numerical information is everywhere
2. Statistical techniques are used to make decisions that affect our daily lives
3. The knowledge of statistical methods will help you understand how decisions are made and give you a better understanding of how they affect you.
4. No matter what line of work you select, you will find yourself faced with decisions where an understanding of data analysis is helpful.

Some examples of the need for data collection.

1. Research analysts for Merrill Lynch evaluate many facets of a particular stock before making a “buy” or “sell” recommendation.
2. The marketing department at Colgate-Palmolive Co., a manufacturer of soap products, has the responsibility of making recommendations regarding the potential profitability of a newly developed group of face soaps having fruit smells.
3. The United States government is concerned with the present condition of our economy and with predicting future economic trends.
4. Managers must make decisions about the quality of their product or service.

What is Meant by Statistics?

- In the more common usage, statistics refers to numerical information

Examples: the average starting salary of college graduates, the number of deaths due to alcoholism last year, the change in the Dow Jones Industrial Average from yesterday to today, and the number of home runs hit by the Chicago Cubs during the 2007 season.

- We often present statistical information in a graphical form for capturing reader attention and to portray a large amount of information.

STATISTICS The science of collecting, organizing, presenting, analyzing, and interpreting data to assist in making more effective decisions.

Types of Statistics – Descriptive Statistics and Inferential Statistics

LO2 Know the differences between descriptive and inferential statistics.

Descriptive Statistics - methods of organizing, summarizing, and presenting data in an informative way.

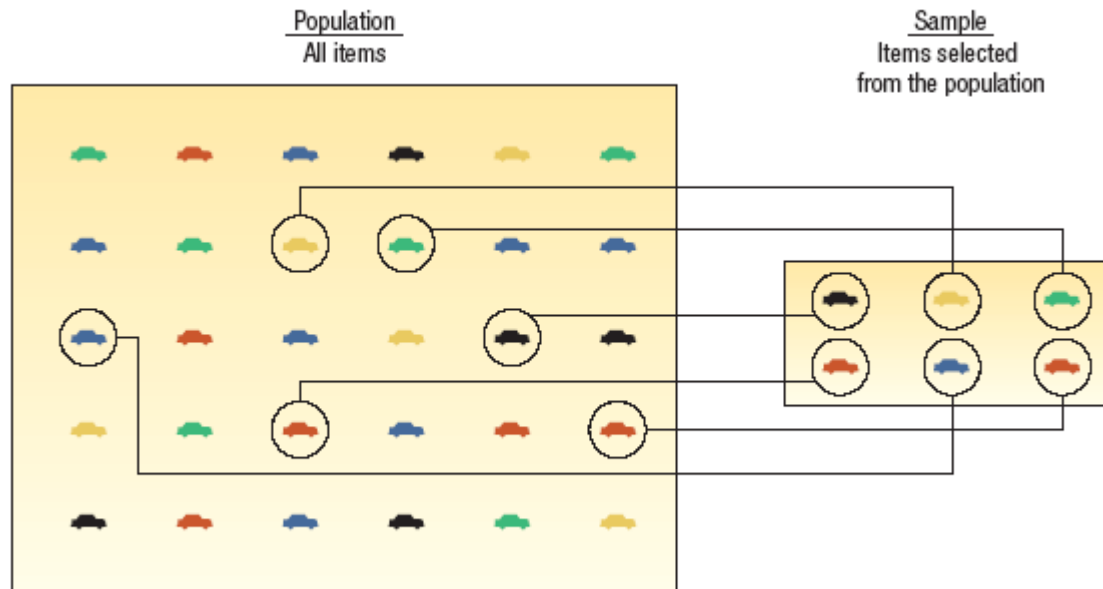
Inferential Statistics: A decision, estimate, prediction, or generalization about a **population**, based on a **sample**.

Note: In statistics the word *population* and *sample* have a broader meaning. A population or sample may consist of ***individuals*** or ***objects***

Population versus Sample

A **population** is a **collection** of **all** possible individuals, objects, or measurements of interest.

A **sample** is a **portion**, or **part**, of the population of interest



Why take a sample instead of studying every member of the population?

1. Prohibitive cost of census
2. Destruction of item being studied may be required
3. Not possible to test or inspect all members of a population being studied

Using a sample to learn something about a population is done extensively in business, agriculture, politics, and government.

EXAMPLE: Television networks constantly monitor the popularity of their programs by hiring Nielsen and other organizations to sample the preferences of TV viewers.

LO4 Explain the difference between qualitative and quantitative variables.

Types of Variables

Qualitative or **Attribute variable** - the characteristic being studied is *nonnumeric*.

EXAMPLES: Gender, religious affiliation, type of automobile owned, state of birth, eye color are examples.

Quantitative variable - information is reported *numerically*.

EXAMPLES: balance in your checking account, minutes remaining in class, or number of children in a family.

LO5 Compare the differences between discrete and continuous variables.

Quantitative Variables - Classifications

Quantitative variables can be classified as either *discrete* or *continuous*.

Discrete variables: can only *assume certain values* and there are *usually “gaps”* between values.

EXAMPLE: the number of bedrooms in a house, or the number of hammers sold at the local Home Depot (1,2,3,...,etc).

Continuous variable can *assume any value* within a specified range.

EXAMPLE: The pressure in a tire, the weight of a pork chop, or the height of students in a class.

Summary of Types of Variables

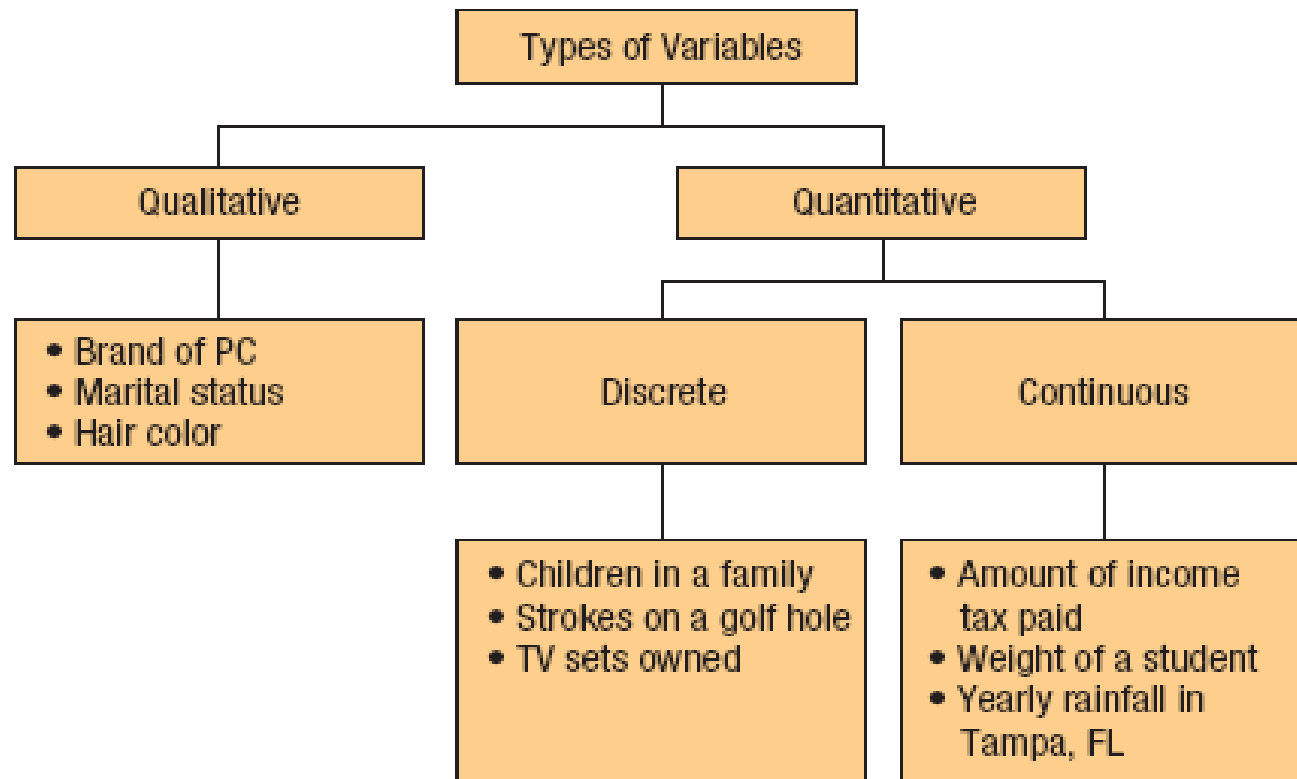


CHART 1–2 Summary of the Types of Variables

Four Levels of Measurement

Nominal level - data that is classified into categories and cannot be arranged in any particular order.

EXAMPLES: eye color, gender, religious affiliation.

Interval level - similar to the ordinal level, with the additional property that meaningful amounts of differences between data values can be determined. There is no natural zero point.

EXAMPLES: Temperature on the Fahrenheit scale.

Ordinal level – data arranged in some order, but the differences between data values cannot be determined or are meaningless.

EXAMPLE: During a taste test of 4 soft drinks, Mellow Yellow was ranked number 1, Sprite number 2, Seven-up number 3, and Orange Crush number 4.

Ratio level - the interval level with an inherent zero starting point. Differences and ratios are meaningful for this level of measurement.

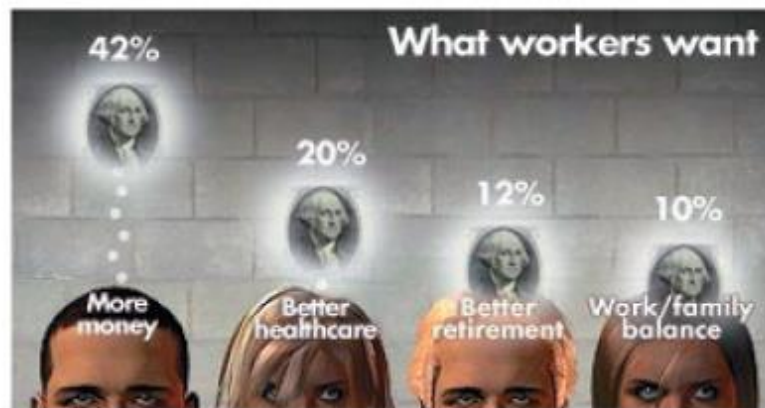
EXAMPLES: Monthly income of surgeons, or distance traveled by manufacturer's representatives per month.

Nominal and Ordinal Level Data

NOMINAL LEVEL DATA

Properties:

1. Observations of a qualitative variable can only be **classified** and **counted**.
2. There is **no particular order** to the labels.



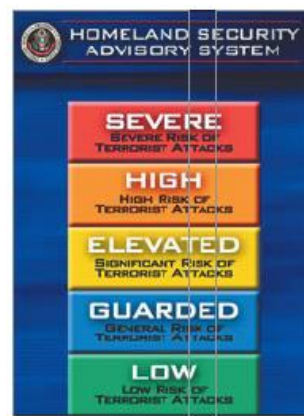
By Anne R. Carey and Chad Palmer, USA TODAY
Source: hudson-index.com



ORDINAL LEVEL DATA

Properties:

1. Data classifications are represented by sets of labels or names (high, medium, low) that have **relative values**.
2. Because of the relative values, the **data classified can be ranked or ordered**.



Interval and Ratio Level Data

INTERVAL LEVEL DATA

Properties:

1. Data classifications are ordered according to the amount of the characteristic they possess.
2. Equal differences in the characteristic are represented by equal differences in the measurements.

Size	Bust (in)	Waist (in)	Hips (in)
8	32	24	35
10	34	26	37
12	36	28	39
14	38	30	41
16	40	32	43
18	42	34	45
20	44	36	47
22	46	38	49
24	48	40	51
26	50	42	53
28	52	44	55

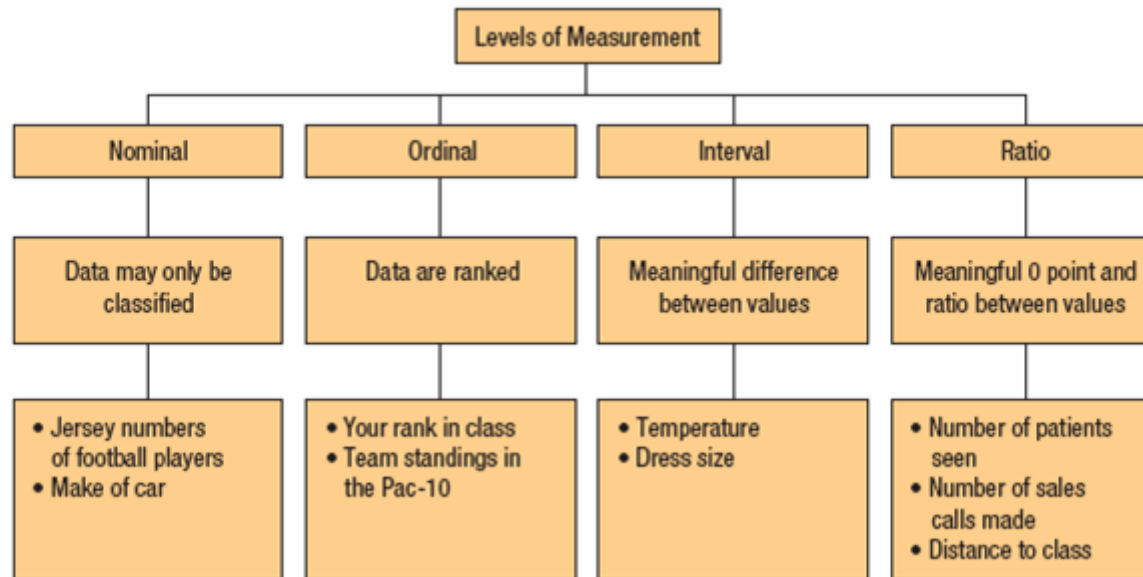
RATIO LEVEL DATA

Properties:

1. Data classifications are **ordered** according to the amount of the characteristics they possess.
2. Equal differences in the characteristic are represented by equal differences in the numbers assigned to the classifications.
3. The zero point is the absence of the characteristic and the ratio between two numbers is meaningful.

Practically all quantitative data is recorded on the ratio level of measurement. Ratio level is the “highest” level of measurement.

Summary of the Characteristics for Levels of Measurement



Why Know the Level of Measurement of a Data?

- The level of measurement of the data dictates the calculations that can be done to summarize and present the data.
- To determine the statistical tests that should be performed on the data