

SECTION 1.1 – Introduction to the Practice of Statistics

Population – The entire group to be selected.

Individual – A person or object that is a member of the population being studied.

Sample – A subset or part of the population that is being studied.

Parameter – A numerical summary of a population.

Statistic – A numerical summary of a sample.

Descriptive Statistics – Consists of organizing and summarizing data. Descriptive statistics describe data through numerical summaries, tables, and graphs.

Inferential Statistics – Uses methods that take a result from a sample, extend it to the population, and measure the reliability of the result.

Variable – The characteristic of the individual within the population.

Data – The actual values of the variable.

Quantitative variables – Variables that provide numerical measures of individuals. The values of a quantitative variable can be added or subtracted and provide meaningful results.

Qualitative, or categorical, variables – Variables that allow for classification of individuals based on some attribute or characteristic.

Discrete variable – A quantitative variable that has either a finite number of possible values or a countable number of possible values. The term *countable* means that the values resulting from counting, such as 0, 1, 2, 3, and so on. A discrete variable cannot take on every possible value between any two possible values.

Continuous variable – 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.

The Process of Statistics

- 1) **Identify the research objective.** A researcher must determine the question(s) he or she wants answered. The question(s) must clearly identify the population that is to be studied.
 - 2) **Collect the data needed to answer the question(s) posed in (1).** Conducting research on an entire population is often difficult and expensive, so we typically look at a sample. This step is vital to the statistical process, because if the data are not collected correctly, the conclusions drawn are meaningless. Do not overlook the importance of appropriate data collection. We discuss this step in detail in Sections 1.2 through 1.6.
 - 3) **Describe the data.** Descriptive statistics allow the researcher to obtain an overview of the data and can help determine the type of statistical methods the researcher should use. We discuss this step in detail in Chapters 2 through 4.
 - 4) **Perform inference.** Apply the appropriate techniques to extend the results obtained from the sample to the population and report a level of reliability of the results. We discuss techniques for measuring reliability in Chapters 5 through 8 and inferential techniques in Chapters 9 through 15.
-

SECTION 1.1 – Introduction to the Practice of Statistics

☺ **Exercises:**

- 1) Parking at a large university has become a very big problem. University administrators are interested in determining the average parking time (e.g. the time it takes a student to find a parking spot) of its students. An administrator inconspicuously followed 150 students and carefully recorded their parking times. Identify the population of interest and the sample of interest to the university administration

For exercises 2 through 5, determine whether the bolded & underlined value is a parameter or a statistic.

- 2) **Calculus Exam.** The average score for a class of 28 students taking a calculus midterm exam was **72%**.
- 3) **Public Knowledge.** Interviews of 100 adults 18 years of age or older, conducted nationwide, found that **44%** could state the minimum age required for the office of the President.
- 4) **Geoscientists.** According to the Bureau of Labor Statistics, as of 2016, there are **32,000** jobs in the U.S. related to the field of geoscience.
- 5) **Civil Engineering.** A civil engineering firm has developed a new fastening mechanism for attaching steel beams in the construction of highway bridges. To assess the strength of the static load the fasteners can withstand, the firm subjected 60 randomly selected test structures to a static load test and measured the load-to-failure (LTF), i.e. the load at which the structure failed. The mean of these 60 measurements was **45,568 pounds** and the standard deviation was **2,248 pounds**.

For exercises 6 through 11, classify the variable as qualitative or quantitative; if quantitative, further classify it as discrete or continuous.

- 6) Number of students enrolled at Orange Coast College.
- 7) Student ID number.
- 8) Internet connection speed in kilobytes per hour.
- 9) Number of Sequoia trees in a randomly selected acre of Yosemite National Park.
- 10) Birthplace.
- 11) Phone Number.

☺ **Multiple Choice Questions:**

- 12) Which one of the following would be considered a qualitative variable?
A) height B) weight C) blood type D) temperature E) salary
- 13) Which one of the following would be considered a discrete variable?
A) age B) occupation C) time D) gender E) number of siblings
-

SECTION 1.1 – Introduction to the Practice of Statistics

☺ **Exercises:**

For exercises 14 & 15, classify the data in each column as either qualitative or quantitative; if quantitative, further classify it as discrete or continuous. Also, identify the variable under consideration in each case.

- 14) **World's Highest Temperatures.** Information provided by the World Meteorological Organization revealed the following data on the highest recorded temperature rounded to the nearest degree for each continent.

Rank	Continent	Place	Temp (°F)
1	North America	Furnace Creek, Death Valley, CA	134
2	Africa	Kebili, Tunisia	131
3	Asia	Mitribah, Kuwait	129
4	Australia	Oodnadatta, S. Australia	123
5	South America	Rivadavia, Argentina	120
6	Europe	Athens, Greece	118.4
7	Antarctica	Esperanza Research Station	63.5

Source: <https://wmo.asu.edu/content/world-meteorological-organization-global-weather-climate-extremes-archive>

- 15) **Earnings from the Crypt.** According to an article from Forbes, the top 10 highest earning dead celebrities in 2018 are as shown in the following table.

Rank	Name	Earnings (\$millions)
1	Michael Jackson	400
2	Elvis Presley	40
3	Arnold Palmer	35
4	Charles Schulz	34
5	Bob Marley	23
6	Theodor Geisel	16
7	Hugh Hefner	15
8	Marilyn Monroe	14
9	Prince	13
10	John Lennon	12

Source: <https://www.forbes.com/sites/zackomalleygreenburg/2018/10/31/the-highest-paid-dead-celebrities-of-2018/#1ab5482720cc>