Section 1.1 Introducing Statistics

Statistical Study – a study in which data are collected from a sample to answer an investigative question about a population.

Investigative Question – a question that can be answered using data and is designed to explore variability in a population.

Population – the entire group of individuals or objects that we want information about.

Sample – a smaller group selected from the population from which we collect data.

Variable – a characteristic that can take on different values for individuals in a study.

Data – the values collected for each variable from the individuals in a sample.

Variability – the extent to which data values differ from one another.

Section 1.2 Variables

Observational Unit – the individual person, object, or thing on which a variable is measured.

Parameter – a fixed, usually unknown value that describes a population.

Statistic – a value calculated from a sample, used to estimate a parameter.

Categorical Variable (Qualitative) – a variable that places an individual into one of several groups or categories; summarized using proportions.

Quantitative Variable – a variable that takes numerical values for which arithmetic operations make sense; summarized using numerical measures.

Discrete Data – quantitative data that results from counting; can only take specific, separate values (usually whole numbers).

Continuous Data – quantitative data that results from measuring; can take any value within a range.

Section 1.3 Tabular Representation and Summary Statistics for One Categorical Variable

Frequency – the number of times a particular category occurs in a data set; also called a count.

Relative Frequency – the proportion or percentage of the total data set that falls into a particular category; calculated by dividing frequency by the total sample size.

Proportion – the ratio of a part to the whole.

Frequency Table – a table that displays the count of observations in each category of a categorical variable.
Relative Frequency Table – a table that displays the percentage of observations in each category of a categorical variable.

Section 1.4 Graphical Representations for One Categorical Variable

Frequency Bar Graph – a graph that displays the count of observations in each category of a categorical variable.

Relative Frequency Bar Graph – a graph that displays the percentage of observations in each category of a categorical variable.

Circle Graph (Pie Chart) – a graph that displays relative frequencies as proportional slices of a circle, where the entire circle represents 100% of the data.

Double Bar Graph – a graph that compares the distribution of a single categorical variable across two or more groups, using side-by-side bars.

Segmented Bar Graph – a graph that compares the distribution of a categorical variable across two or more groups, where each bar is divided into segments representing relative frequency.

Mosaic Plot – a variation of a segmented bar graph where the width of each bar is proportional to the sample size of that group, allowing both distribution and group size to be displayed in a single graph.

Distribution – the pattern of values a variable takes, including which categories are more or less common.