

2.1: Introduction

Once a sample is collected, we can organize and present the data in tables and graphs. These tables and graphs help summarize, interpret and recognize characteristics within the data more easily than raw data. There are many types of graphical summaries. We will concentrate mostly on the ones that we can use technology to create.

A population is a collection of all the measurements from the individuals of interest. Remember, in most cases you cannot collect data on the entire population, so you have to take a sample. Now you have a large number of data values. What can you do with them? Just looking at a large set of numbers does not answer our questions. If we organize the data into a table or graph, we can see patterns in the data. Ultimately, though, you want to be able to use that table or graph to interpret the data, to describe the distribution of the data set, explore different characteristics of the data and make inferences about the original population.

Some characteristics to look for in tables and graphs:

1. Center: middle of the data set, also known as the average.
2. Variation: how spread out is the data.
3. Distribution: shape of the data.
4. Outliers: data values that are far from the majority of the data.
5. Time: changing characteristics of the data over time.

There is technology that will create most of the graphs you need, though it is important for you to understand the basics of how they are created.

Qualitative data are words describing a characteristic of the individual. Qualitative data is graphed using several different types of graphs, bar graphs, Pareto charts, and pie charts. Quantitative data are numbers that we count or measure. Quantitative data graphed using stem-and-leaf plots, dotplots, histograms, ogives, and time series.

The bar graph for quantitative data called a **histogram** looks similar to a bar graph for qualitative data, except there are some major differences. First, in a bar graph the categories can be put in any order on the horizontal axis. There is no set order for these data values. You cannot say how the data is distributed based on the shape, since the shape can change just by putting the categories in different orders. With quantitative data, the data are in specific orders since you are dealing with numbers. With quantitative data, you can talk about a distribution; the shape changes depending on how many categories you set up. This shape of the quantitative graph is called a **frequency distribution**.

This leads to the second difference from bar graphs. In a bar graph, the categories are determined by the name of the label. In quantitative data, the categories are numerical categories, and the frequencies are determined by how many categories (or what are called classes) you choose. There can be many different classes depending on the point of view of the author and how many classes there are. The third difference is that the categories touch with quantitative data, and there will be no gaps in the graph. The reason that bar graphs have gaps is to show that the categories do not continue on, as they do in quantitative data.

This page titled [2.1: Introduction](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Rachel Webb](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.