2.1: Organizing Qualitative Data

Data is easier to interpret if it is organized into a visual display. Useful displays for qualitative data include tables, bar graphs, and pie charts.

Tables:

A *frequency distribution* lists each category and the number of data points (occurrences) corresponding to that category. When arranged in the form of table, this is called a *frequency table*.

Tables often include the *relative frequency*.

The *relative frequency* of a category is the proportion (or percentage) of the total observations that fall within that category.

$$\text{Relative frequency} = \frac{\text{Frequency}}{\text{Sum of all frequencies}}$$

A *relative frequency distribution* lists the categories along with their relative frequencies.

Example 1: Let's make a frequency distribution summarizing the birth countries of the current Houston Astros.
https://www.espn.com/mlb/team/roster/_/name/hou/houston-astros

Bar charts (bar graphs):

A *bar graph* is a visual display in which the category names are along the vertical or horizontal axis, and the frequencies (or relative frequencies) are on the other axis.

Note: Category name = value of the nominal variable

In a *side-by-side bar graph*, the relative frequencies for values of the nominal variable are simultaneously displayed for two or more subgroups (or for two or more years, or for two or more values of some other variable).

Example 2: Use the information below to create a side-by-side bar graph for headcount at LSC campuses.

Campus	Students Enrolled Fall 2017	Students Enrolled Fall 2018
LSC-CyFair	22,125	22,453
LSC-Kingwood	12,849	12,740
LSC-Montgomery	15,914	15,266
LSC-North Harris	15,845	15,660
LSC-Tomball	9,306	9,703
LSC-University Park	13,273	13,328

Pie charts:

A pie chart is a circle divided into sectors, in which each sector represents a category. For each category, the relative frequency is equal to the ratio of the sector area to the total circle area. (In other words, if one-third of the observations are in a given category, then one-third of the circle area will be in the sector corresponding to that category.)

Example 3: Sketch a pie chart to represent the following data for voters in a particular county.

	Number of Voters
Conservative	2,311
Liberal	1,402
Independent	822

Example 4: Create a dotplot representing the manufacturers of vehicles owned by the class. (Dotplots can be used to represent quantitative data also.)