Quantitative Data: Numerical data, in which the numbers indicate quantity

2.2: Organizing Qualitative Data

Overview: Types of Data

Qualitative data (does not measure the amount of anything)

Nominal (categorical): Values are names or categories. Values do not measure the amount of any characteristic. Values do not have a meaningful order. (It does not make sense to rank the values in an order.)

Ordinal: Values can be arranged in a meaningful order, from smallest to largest. A small value indicates the object has less of the characteristic being measured; a large value indicates the object has more of the characteristic being measured.

Could be qual or quant

Quantitative

Interval: Values are ordinal, having a meaningful order. In addition, the difference between values has meaning. In other words, an equal difference between values represents an equal difference in the characteristic being measured. Thus, it is meaningful to add and subtract values. However, there is not a meaningful zero (a value of zero does not indicate the absence of the characteristic being measured).

Ratio: Values are ordinal (have a meaningful order) and interval (have a meaningful difference between values). In addition, the value of zero has meaning (a value of zero indicates absence of the characteristic being measured). Thus, ratios between values have meaning (if one value is three times as large as another value, then it has three times as much of the characteristic being measured). For this reason, it is meaningful to multiply and divide values. (And, because ratio data is also interval data, it is meaningful to add and subtract values.)

Examples of nominal variables:

First name, student ID, ethnicity

Examples of ordinal variables:

Truck size ($\frac{1}{2}$ ton, $\frac{3}{4}$ ton, 1 ton)

Examples of interval variables:

Temperature

Examples of ratio variables:

Weight, height, distance

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:

#2.28 Colleges of Students

Eng	Eng	Bus	Bus	Eng
Lib	Lib	Eng	Eng	Eng
Bus	Bus	Eng	Bus	Eng
Lib	Bus	Bus	Bus	Eng
Eng	Eng	Lib	Eng	Bus

College	Frequency		Relative Frequency
Engineering	𝍸 𝍸 ||	12	12/25 = 0.48
Liberal Arts	||||	4	4/25 = 0.16
Business	𝍸 ||||	9	9/25 = 0.36
		n = 25	1.00

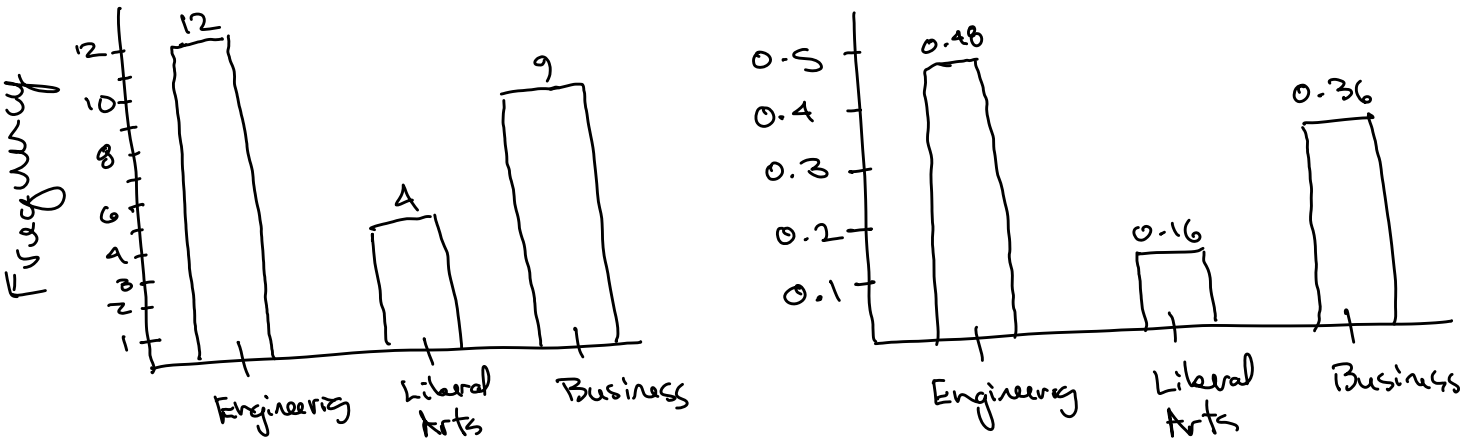
Bar charts (bar graphs):

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: Using College Data from Ex 1.



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: