AP Statistics

2019-01-14 1.1 Analyzing Categorical Data and Displaying Quantitative Data with Graphs

Notes taken by: **Noah Overcash**

The values of categorial variables are category labels, such as "male" and "female."

The distribution of a categorical variable lists the categories and gives either the *count* or *percent* of individuals in each category.

A frequency tables shows the counts/frequencies of data

A relative frequency table shows a table of the percents or relative frequencies of each.

When percents are rounded, **roundoff error** can occur. This does not mean a mistake was made.

- These type of errors typically occur in pie charts.

Pie charts should be used only to emphasize each category's relation to the whole.

Good bar graphs have:

- Bars of the same width
- Area varying in proportion to height
- Express the right impression to our eyes
- A starting number of zero

Bad bar graphs replace bars with pictures; they boost **eye appeal** but not data readibility.

Two-way tables show two different variables/categories (often on different axes)

- These tables will show totals for each row/column, allowing us to calculate aggregate **marginal distribution**

Marginal distributions tell nothing about the relationships between two variables

To describe relationships between categorial variables, certain percents from the body of the table must be used (**conditional distribution**)

- Conditional distributions are a value out of the whole

Segmented bar graphs show the distribution of categories within each bar

Organizing a statistical problem (four-step process):

State the question to answer

Plan how to go about the question and what statistical techniques need to be used

Do create graphs and carry out calculations

Conclude with a practical conclusion in the real-world problem.

Remember: **S**tatistics **P**roblems **D**emand **C**onsistency