Describing univariate and bivariate data

February 14, 2018

Contents

Many of the figures that we will be creating and analyzing during the course will be representations of univariate (meaning one variable) and bivariate (meaning two variables) data. You will frequently be asked to write a description about a visualization, and you should aim to be precise and consistent in your terms. Use the short summaries below as a guide and a reminder when writing about the features contained in a univariate or bivariate plot.

0.1 Describing univariate data

When describing the visual properties of univariate data, remember to discuss the following traits:

- shape:
  - right-skewed, left-skewed, symmetric (skew is to the side of the longer tail)
  - unimodal, bimodal, multimodal, uniform
- center: mean (`mean`), median (`median`), mode (not always useful)
- spread: range (`range`), standard deviation (`sd`), inter-quartile range (`IQR`)
- unusual observations

For additional guidance, follow this link for a summary of what the above terms mean.

0.2 Describing bivariate data

When describing the visual properties of univariate data, you will frequently be looking at a scatterplot. When describing the shapes of scatterplots we highlight:

- **Direction**: What direction is the data trending? Positive direction or negative direction?
- **Form**: This is analogous to *shape* for univariate data. Is the dataset linear? Is is curved? Does it not have a form?
- **Strength**: How clustered are the data points around the underlying form? Stated another way, what are the strength of the correlations? Typical descriptors are *strong*, *moderate*, or *weak*.