Sample statistics

- mean, median, quartiles, quantiles, 5-number summary
- standard deviation, IQR, range: Know the formulas, be able to compute
- eye-ball estimates of mean and median, given a histogram or dotplot
- eye-ball estimate of standard deviation, given a bell-shaped distribution
- correlation coefficient, and eye-ball estimates of its value
- sensitivity vs. resistance to outliers: What does this mean? Which statistics are which?
- turning a value into a Z-score: Know the formula, and the meaning of its required components

Multivariate data

- using the wording of a question to decide upon explanatory and response variables
- plots/summary tables that help reveal apparent associations
  - contingency (two-way) tables, usually when both variables are categorical
  - side-by-side bar graphs, particular to two categorical variables
  - side-by-side histograms/dotplots/boxplots, particular to one variable being categorical and the other quantitative
  - scatter plot, particular to two quantitative variables
- What to look for in plots/summary tables that suggests an association
- Linear regression
  - When is it (is it not) appropriate?
  - Interpration of slope
  - Observed values, predicted values, residuals
  - Formulas for slope b_1, intercept b_0

Statistical studies

- the two broad types, observational and experiment; What distinguishes an experiment?
- controlled randomized experiments
  - terminology and components
    - control group
    - placebo, blind study, and double-blind study
    - blocking, matched-pairs
  - able to conclude a causitive impact—beyond mere association—from the explanatory variable. Be able to explain how randomization plays a key role in this, using terms like confounding or lurking variables.

Probability

- terminology: random process, outcomes, sample space, event, complement, independence, mutual exclusivity (or "disjointness")
- conditional probability
- probability rules
  - axioms
  - addition/multiplication rules (in their various forms)
  - Bayes' rule