

Recap Chapters 1 through 3

Need to recognize that there are different types of data

- Different methods for visual assessment

- Qualitative –tables, bar graphs, pie charts
- Quantitative—histograms, boxplots, stem and leaf plots, scatterplots, time series plots,...what does each type of plot show you and when is one preferred over another
 - Shapes of distributions (symmetric, right-skewed, left-skewed)

- Different methods for numeric summaries

- Qualitative –percentages, modes
- Quantitative—mean, median (which one should you use in a particular case?)
 - 5 number summary
 - Standard deviation (or variance); range
 - Detecting potential outliers ($|Z| > 3$; 1.5IQR rule)
 - How should outliers be handled; what impact will they have on your proposed analysis?
 - Measures of relative standing—z-scores; percentiles

Need to understand how data is collected

- Observational studies—what are the potential issues?
 - Volunteer sample
 - Confounding variables
 - Can conclude associations, but not causation
- Designed experiments—what considerations should be made?
 - Blinding (single, double); placebo
 - Randomization to control for confounding variables
 - Replication
 - May be able to conclude cause and effect
- Types of sampling
 - Random
 - Simple random
 - Systematic, cluster, stratified, multistage