STAT 113: KEY BIG PICTURE CONCEPTS

Key ideas

- **Data beats anecdotes** (and yes, "data is" is grammatical; not "data are", despite what you may have been told)
- **Variability** is natural and is also predictable and quantifiable.
- **Random sampling** allows results of surveys and experiments to be extended to the population from which the sample was taken.
- **Random assignment** in comparative experiments allows cause and effect conclusions to be drawn.
- **Association** is not causation (see also: **confounding variables**)
- **A graph of the data is a first step** in analyzing data (and sometimes is enough to answer the question of interest).

Basic ideas of statistical inference

- Seeing **a difference or association in a dataset doesn't mean** that difference or association reflects **a structural fact** about the world (it might just be sample noise)
- The concept of a **sampling distribution** and how it applies to making statistical inferences based on samples of data (including the idea of **standard error**).
- The concept of **confidence interval**, including the **interpretation** of confidence level and margin of error.
- The concept of **statistical significance** including significance levels and P**-values**.
- The null and alternative hypotheses are not statements about the data; they are statements about the structure of the world.
- **The P-value is not the probability that the null hypothesis is true**, it is the probability of data like ours *if* the null hypothesis were true (Compare: chance that an old white guy is in the U.S. Senate, vs chance that a member of the U.S. Senate is an old white guy)
- **Statistical significance does not necessarily imply practical importance**, especially for studies with large sample sizes.
- **Finding no statistically significant difference or relationship does not necessarily mean there is no difference or no relationship** in the population, especially for studies with small sample sizes.