Guidelines for Statistical Reporting

- Now that you know all the major components of statistical inference and the linear model, let's talk about writing up your results.
- First, all the guidelines we discussed for exploratory data analysis still hold:

1. A statistical analysis is a written argument.
2. If you don't have something nice to say (about your output), don't display it at all.
3. Document your decisions.
4. Identify features that should be reflected in statistical models.
5. Remember the difference between sample and population.
6. Show us the code (a guideline for this class).

Guideline Seven

Explain how your research question is connected to the specific hypotheses you test.

- How are your concepts operationalized?
- What are your dependent and independent variables?
- Are there issues that the reader should bear in mind for interpreting results?

Guideline Eight

Draw connections between your EDA and your modeling choices.

- Modeling is often an iterative process.
- Statistics is often presented as a linear process:
    - Examine data structure → clean → EDA → run regression → done
    - In practice, we often move in cycles.
        - Build a model → test assumptions → alter the model …

Guideline Eight (cont.)

- Your writing should guide the reader through this process.
    - Remember that modeling choices have trade-offs—the key is to be transparent.
- Be responsive to features you find in EDA.
    - Remember that specification is more than just log transforms.
    - Think creatively about ways to model features using all the tools you know: interaction terms, indicators variables, polynomial terms, etc.

Guideline Nine

Don't disregard a model because you don’t like the results.

- This is dishonest.
  - It's really a type of data mining.
  - But in the classical context, you can think of it as torturing the data.
    - Changing the specification until you get what you want
- Remember: Negative results are just as important as positive results.
  - The reader needs to know about both negative and positive results to form a conclusion.
  - When you have multiple specifications, report them all.
  - Let the reader see how robust the effect you’re describing is.
- By the way, you should know this by now: A negative result does not mean you can accept the null. You just fail to reject it.

Guideline Ten

Explain the practical significance of results.

- Statistical significance is important, but you can’t end there.
- Guide the reader to understand the magnitude of an effect.
- Remember that R-squared is overused—look for better measures of effect size: slope coefficients, standardized slope coefficients, etc.