Writing a Statistical Report for STAT 217

I. Introduction

- Give a brief background of the research problem and how the data were collected.
- Clearly outline the question(s) of interest you will address with the statistical analysis. The more specific you define the question of interest here, the easier the rest of the analysis is. It should start with,"Is there evidence.." and it should be as specific as possible. Your *Summary of Statistical Findings* should directly answer the question you pose here.
- Example for a two sample t-test: Is there evidence to suggest that the true mean number of cavities in people living in Bozeman is less than the true mean number of cavities in people living in Billings?

II. Statistical Procedures Used

- Describe the response and explanatory variables.
- Define null and alternative hypotheses for the appropriate statistical test.
- Use plots to visualize the raw data and describe what you see. **Include pertinent plots in the body of the report when you refer to them**. Give an initial impression about the answer to the question of interest.
- Describe how you checked all assumptions, what your conclusions are regarding the assumptions, and justify your conclusions. Refer to (and include) the appropriate plots.
- Describe the appropriate statistical procedures used to answer the research question(s) (e.g. two-sample t-test, analysis of variance, etc). Be specific about which procedures are answering which questions of interest (if there are multiple questions).
- This section should lay out the steps, decisions, and logic leading to the final model used to answer the question of interest.

III. Summary of Statistical Findings

- This section should consist of three sentences:
  1. **Decision Sentence**: Base a decision off of the strength of evidence provided by the p-value (reject or fail to reject).
  2. **Evidence Sentence**: Write a conclusion communicating your results in context. The conclusion should start with "There is _____evidence that..." Make sure you include the value of the test statistic and the distribution it follows under the null hypothesis. The conclusion should be worded very similar to the question of interest.
  3. **Estimate Sentence**: Give the estimate for the statistic of interest and be sure to include a confidence interval.
- Example for a two-sample t-test: At a significance level of 0.05, we reject the null hypothesis. There is moderate evidence that the true mean number of cavities in people living in Bozeman is less than the true mean number of cavities in people living in Billings (p-value=0.0263 from two-sample t-stat=2.39 on 21 df). The true mean number of cavities in Bozeman residents is estimated to be 1.4 cavities less than the true mean number of cavities in Billings residents, with a 95% confidence interval from 0.4 to 2.4.

IV. Scope of Inference

- Write a brief Scope of Inference. Follow the examples below. Specifically, answer these two questions and comment on their implications:
  1. Were the units randomly selected from some larger population? *(i.e. What, if any, larger population can you infer the results to?)*
  2. Were the units randomly assigned to groups? *(i.e. Are cause-and-effect statements justified?)*
- Example: The 11 Bozemanites and 12 Billingites were randomly selected from their communities, so we can infer our results to all residents of Bozeman and Billings. The 11 Bozemanites and 12 Billingites were not randomly assigned to fluoride treatments, so we cannot conclude that having fluoride in the water causes a decrease in the mean number of cavities.
- Make sure you write the Scope of Inference specific to the language of the problem (not just generic statements). For example, DO NOT say, "There was no randomization so cause and effect statements cannot be made. Instead, for example, say, "The lizards were not randomly assigned to diets, therefore we can not conclude that the fly diet caused the increase in average weight over the lizards on the tofu diet."
  1. Say what was (or was not) randomly assigned to what, therefore...
  2. Say what was (or was not) randomly selected from what, therefore...
  3. No more than 2 sentences!

V. Appendix

- Include R-code and relevant output in an appendix at the end of the report.

VI. OTHER

- **Style**: Report must be written in complete sentences in paragraph format. Use the above five headings in your report so that it is clearly organized.
- The report should be double spaced.
- Some projects will request an additional section. Follow instructions on the project sheet.
- Part of the grade for each project is based on writing (organization, grammar, appropriate use of statistical terms, etc.)
- I expect you to use the comments from previous homeworks and projects to improve subsequent homeworks and projects!