

# Final Project Rubric

The final project will be relatively short and sweet compared to the midterm project. In essence you will write up a basic summary of data and it's collection scheme from the midterm project.

## 1 Overview

There are a few things I want this paper to have that will basically be a high quality summary of your data. Major things I would like addressed is

1. What were you researching?
2. What was your research question in particular?
  - The above two points are different. The first is the broad topic of interest, the second is the specific question(s) you tried to answer
3. How did you collect your data?
4. What analytical methods did you use?
  - Regression
  - Hypothesis Test
  - Confidence Interval
5. What were the results?
6. Conclusion?

I would like the paper to be between 3 and 5 pages including graphics which shouldn't be hard. Note your goal is to make sure you address the things I want addressed and not merely length. A group who turns in a 2 page paper that is succinct and accurate will most likely do better than a group that just tells AI to make a half dozen graphs and paste them into a five page paper with near 0 intellectual input (amazingly this has most likely already happened due to not reading the midterm project rubric).

## 2 Deliverables

On the day the final exam occurs I would like....

1. A 3-5 page paper
  - I don't care about spacing or text size or how much of it is graphs. To accurately address the things I want addressed should let you pole-vault over that requirement.
  - I suspect the issue will actually be fitting the information you want into 5 pages
2. RMD file that created any code used for the analysis.
  - Please include the data file if it exists seperately.
  - If you used another program to run the analysis you do not need to submit the RMD file
    - But if so I don't expect to see a standard default ggplot2 graph as I know of no other stat program with that poor aesthetics.

### 3 Grading

#### 1. Research

- Your research topic and question of interest are identifiable
- You can explain why the question you are answering is tied to the research topic
- Who is your population of interest?
- Most of this should be at/near the introduction

#### 2. Data Collection

- Who/what was in our sample?
  - If your sample and population has a mismatch (eg all college students vs 50 Grinnell students) you'll need to mention that
- How did you collect your data?
  - Does your sampling method have a name?
  - How did you ensure randomness when you collected people, assuming you did?

#### 3. Graphic

- I'd like a strong graphical component to your papers with at least one graph (probably a couple, maybe more)
- Absolutely feel free to reuse the graphs you made for the midterm
- Graphs should have captions when in a paper
  - Caption does not the same as alt-text (which is a hidden caption)
- But you do need alt text
- And they must be colorblind friendly somehow such as...
  - Shapes
  - Colorblind friendly pallettes
- Axis, labels, etc... all need to look professional and readable

#### 4. Statistical Analysis

- For STA 209 I'm requiring at least one of the following three:
  - Hypothesis Test
  - Confidence Interval
  - Linear Regression
- Whichever one you choose must have the assumptions checked in some manner
- Most people would either use a hypothesis test in the form of a difference of means (two-sample t-test), two sample proportion test or a multiple linear regression.
- Regardless, I want some quantitative discussion of your data. Examples, not related to hypothesis tests or regression, include...
  - Reporting Pearson or Spearman's correlations (whichever one is most correct to use)
  - A table of counts showing the number of respondents who answered each category
  - Means or 5 number summaries for the different categories
- Big thing is I want a discussion about what your data suggests; how you go about doing that is pretty flexible

#### 5. Discussion of Results

- Briefly discuss what your analysis implies
  - Eg if you went with confidence intervals talk about how we have an idea of the range where we think (PARAMETER) is at
  - Eg what does your small p-value imply about how (CATEGORY A) does compared to (CATEGORY B)?
  - What does the regression imply about the relationship between the explanatory variable(s) and the response?

## 6. Conclusion

- A big part of what I want in your conclusion (besides a brief summary of your results) is...
  - What are the problems with your sampling scheme? Why?
  - What would you do differently when sampling next time if you were to redo this?
- Given any possible problems, do you believe your conclusions are “strong” or should they be taken with a grain of salt?
- Any other variable you’d want to collect or something different you’d do?

## 7. Style

- Does your paper have a natural flow to it? Or are the sections/thinking disjoint?
- Did you write the paper in a “normal” academic manner or was it more “checking the box” for the rubric?
  - eg It is poor form to outright say “Our research question is team A is faster than team B?”.
  - It’s blunt and lacks finesse that we expect when writing
- Does the paper look professional?

## 8. References

- It is unlikely that you will need to formally reference anything in this paper and I will not be surprised if you don’t have a bibliography
  - SIDE NOTE: Short bib’s is a common occurrence in statistics/math to be honest
- If you do reference or cite a work I do want a bibliography
- You might not have a bibliography but I do want you to clearly state in the text...
  - What software you used
  - What version of the software
  - And what was the main function you used for the analysis
  - Eg “...so we performed a t-test using R version 4.3.1 and the t.test() function.”
  - This is actually an important step in allowing others to confirm the analysis/being upfront about what we did.
    - \* SIDE NOTE: It’s getting more and more common for journals to expect you to publish the code you ran the analysis on so others may look it over/use it
- If you do have a bibliography let’s use....\*rolls dice\*.....Chicago format! (or any other APA or MLA format; I honestly don’t care)