Rubric for Final Projects (36-315)

Zach Branson

April 21, 2020

1 Presentation on May 6

Presentations should be around 15 minutes, with up to 10 minutes for follow-up questions. All team members must be present for the presentation and participate. Teams may use slides, walk through the .html file, or another format of the team's choice. The TAs and I will be looking for the following things during your presentation:

- Do they provide a clear description of the dataset? (10pts)
- Do they state thoughtful research questions they would like to answer with their dataset? (10pts)
- Do they showcase graphs that are of high quality (correctly labeled, interpretable, and informative)? (20pts)
- Are they able to clearly explain how to interpret their graphs, as well as pinpoint what the main takeaways of each graph are? (20pts)
- Are the graphs (and their main takeaways) relevant to the team's research questions? (20pts)
- Do they state the main conclusions and takeaways that can be made from the project overall? (10pts)
- Do they clearly communicate a cohesive narrative for the project? And if there are any follow-up questions, do they answer them clearly and correctly? (10pts)

2 .html file, due May 4 12pm

This file (produced from RMarkdown) will act as a substitute for a poster (which is typical for 36-315). Though due on May 4, this file will be made public for the statistics community to see on May 6. After that date, we will leave the files publicly available unless any member of the team asks for the files to be taken down after May 6 - if that's the case, we will make the file unavailable after May 6.

Typically, posters start with a short paragraph and/or bulleted list describing the data, followed by a short paragraph and/or bulleted list describing the main research questions of the project. Your .html file should do the same thing. After that, the middle part of your file should consist of your most important graphs (5 to 8), bookended by statements motivating how the graphs fit into the overarching narrative of the project and their main takeaways. For example, a poster could take the following format:

> We wanted to learn about <Research Question 1>, which suggests we should examine <such and such variables.>
>
> [insert your amazing graph]
>
> The above graph suggests that <your main takeaways, interpretations, and conclusions>.
>
> We also wanted to better understand <Research Question 2>. To do this, we plotted <such and such variables> using <graph type of your choice.>
>
> [insert your next amazing graph]

And so on. (Please do not just copy-and-paste the above format; it's supposed to be a broad idea of how you should structure your document, rather than literal instruction.) At the end of your file, you should include a short paragraph discussing the main conclusions of your project and possibly future work (like in a poster).

The TAs and I will be looking for the following things during your presentation:

- Do they provide a clear description of the dataset? (10pts)

- Do they state thoughtful research questions they would like to answer with their dataset? (10pts)
- Do they showcase graphs that are of high quality (correctly labeled, interpretable, and informative)? (20pts)
- Do they provide concise explanations (1-3 sentences) for how we should interpret each graph, what their main takeaways are, and how they fit into the overarching narrative of the project? (20pts)
- Are the graphs (and their main takeaways) relevant to the team's research questions? (20pts)
- Do they state the main conclusions and takeaways that can be made from the project overall? (10pts)
- Is the document well-organized, and does it clearly communicate a cohesive narrative for the project? (10pts)

3 Final Project Report (due May 13 by 5pm)

Each team will submit a final report that should be 5-10 pages long, excluding graphs. (Font should be 11pt or 12pt in size, double-spaced.) Furthermore, each team member should submit an individual 1-2 paragraphs discussing what they did as an individual for the project.

I will be grading final reports based on "basic items" and "detailed items." The "basic items" are in bold below, and will generally be graded in an all-or-nothing fashion. For example, the first "basic item" is, "Does the project provide a description of the dataset?" worth 5pts. If the answer to that question is "Yes," then the team is awarded 5pts; if the answer is "No," the team is deducted 5pts. The basic items are meant to set a minimum grade that every team should receive based on very minimum effort; the basic items add up to 60pts. The remaining 40pts are divided into "detailed items" that I will rate on a 0-to-5 scale. Each "detailed item" is under a "basic item" and is worth 5pts each. Thus, you'll find 8 "detailed items" below, adding up to the remaining 40pts.

Does the project provide a description of the dataset? (5pts)
Rated 0 to 5: What is the quality of this description? Does it clearly communicate what the rows and columns (i.e., subjects and variables) are in the dataset in a way that is understandable to a CMU statistics undergraduate?

Does the project have at least three clearly stated research questions? (9pts; minus 3pts for each question that is missing or not clearly stated)
Rated 0 to 5: What is the quality of the research questions? Are they well-motivated by real-world/scientific interests? Or are they shallow? For example, a shallow question would be in the form, "What does the distribution of this variable look like?" A more interesting question would motivate why we would like to inspect particular distributions with real-world context.

Does the project follow the "Graph Guidelines" described in Homework9 (6 to 10 graphs, at least three of the different kinds of graphs, no more than two univariate graphs, no more than 3 of the same type of graph)? (16pts; each criterion is worth 4pts)
Rated 0 to 5: What is the quality of these graphs? Are these graphs easily readable, interpretable, and properly labeled?
Rated 0 to 5: Are the graphs well-motivated, given the research questions of the project? In other words, do the graphs address/answer the research questions of the project, or do they only provide tangential (or even irrelevant) details?

Does the project provide further descriptions/interpretations for each graph? (10pts; minus 2.5pts for each description that is missing, up to 10pts)
Rated 0 to 5: What is the quality of these descriptions? After reading these descriptions, is it crystal clear what is being displayed in the graph and what the main takeaways are?

Does the project include any kind of statistical analyses and/or summary statistics/tables to complement their graphs? (5pts)
Rated 0 to 5: Are the statistical analyses appropriate given the type of data and research questions of interest? Are the statistical analyses interpreted correctly?

Does the project provide some clear conclusions that can be made from

their graphs and analyses? (10pts)
Rated 0 to 5: What is the accuracy of the claims being made? Are the claims well-supported by the graphs and analyses presented? Are the conclusions well-aligned with the research questions of the project?

Does the project discuss questions that have not been answered by the project, but could be answered with future work? (5pts)
Rated 0 to 5: Did the team provide adequate reasons as to why these questions were left as future work (e.g., they need more data, need more nuanced statistical techniques they haven't learned, etc.)? Are these future-work questions well-motivated given what the team has completed for this project?