1920S MA482

Criteria for Successfully Completing the Case Study

The **Case Study** is an opportunity to reproduce a research analysis. This is an individual assignment.

Note: you may collaborate with another individual. See the **Guidelines for Appropriate Collaboration** to prevent committing academic misconduct. You are responsible for all answers submitted; the instructor reserves the right to require an oral examination over the answers submitted to ensure the work is that of the student.

General Instructions:

During the course of your study, you may use any resource available to you - including the internet, texts, notes, etc. If any resource has a substantial impact on your report, it should be appropriately cited. You need not cite course notes; this is understood to be common knowledge. However, if you choose to employ a strategy you discover in a textbook, it should be cited. If you consult with an individual who provides an idea you follow-up on, that person should be given credit in the text with an appropriate citation.

The project will result in a report. Your report will summarize a study you choose to reproduce. The research you choose to reporduce must adhere to the following:

- It must be a published article in an academic journal or an internal study conducted at Rose-Hulman (previous Biology thesis, for example).
- You could not have been a collaborator on the initial study.
- Could benefit from methods discussed in the course.

You need not reproduce every aspect of the original research. Instead, you must:

- Reproduce the primary question of interest. It is possible that you will not be able to fully reproduce the work; however, you should be able to replicate the process and conjecture as to why the results disagree.
- Replicate a graphic in the paper or provide one if not included in the original article.
- Critique of the study design implemented.
- You must implement a suggested improvement to the original analysis. This could be exploring a secondary question not originally addressed in the article but must make use of material discussed in our course.

If you are having trouble finding a research article with corresponding data to reproduce the analysis, I suggest browsing articles on the following site: **https://datadryad.org/**

You will summarize your findings through the inclusion of an Introduction reviewing the questions of interest, the original study design, and highlighting the results; a Methods section detailing your analysis and improvements; and a Results section discussing the main results and limitations of the study. As this **case study** is meant to mimic a scientific article, the criteria below are meant to produce a paper which is understandable by a graduate of Rose-Hulman, even if they have not taken a course in statistics. Specifically, while the Methods section may not be understood by this general audience, the introduction and conclusions should be clear. They should walk away from your report understanding the "punchline." The specifications also guide you through a high-quality presentation.

As you read through the following criteria, the following two articles can be a helpful starting point:

- **Kuiper: How to Write a Scientific Paper or Poster**
- **Miller: How to Communicate Statistical Findings: an Expository Writing Approach**
- **ASA Recommendations for Poster Presentations**

Criteria:

The criteria stated below ensure that both your analysis as well as the clarity of the presentation are of an acceptable caliber. First, the criteria that apply to the overall report are stated, and then criteria specific to each of the three sections are provided. A project will be considered successful if it adheres to approximately 85% of the criteria below **and** *all* *essential criteria* (highlighted in orange below). The idea is to allow for small mistakes (which were clearly unintentional) which do not distract from the overall quality of the paper.

Overall Criteria:

The submission should be well formatted, which includes the following:

- The report is typed in 11 or 12 point font.
- The report contains 3 distinct sections: Introduction, Methods, Results.
- The report is no more than 3 pages, excluding the appendix; there is no length restriction on the appendix.
- R code / Minitab output should only be placed in the appendix, and not referenced in the text.

R code is not considered readable by a general audience; therefore, it is not appropriate to include it in the body of the report. However, it should be included in the appendix to construct a complete, self-contained, record of the analysis. Note: this code should **not** be referenced in the text; it is included for completeness, not because it helps tell the story of the data.

In addition to formatting, the submission should be publication-ready, which includes being well written. To this end, the report should only require minor editing before it would be suitable for publication, which includes the following:

- The report should have no more than 5 grammatical errors.

It can be very helpful to have someone else proofread your report for grammatical mistakes. Note that artistic choice of wording will not be evaluated.

Introduction Criteria:

The Introduction should briefly, yet clearly, state the question of interest as well as summarize the rationale for the study. For this **Case Study**, you are not the researcher who initiated the study, the introduction serves as a brief summary of what the analysis is trying to accomplish and why. Specifically, your introduction should satisfy the following:

- In your own words (don't just copy and paste from the original article), summarize why the study is of interest.
- State the primary and secondary (if applicable) questions that are to be addressed in the report.
- This section should be at least 3 sentences in length.
- The original paper's citation should be provided.

Methods Criteria:

The Methods section is the "technical" portion of the report. In this section, you should detail the analyses conducted and any summaries constructed. You are striving to give enough detail that a trained statistician could replicate your results; this is not the same as printing R code. The following criteria should be adhered to:

- The Methods section includes an accurate description of the model/analysis constructed.
- Any discrepencies between the original analysis and your replication should be discussed.
- The additional methods employed (for the improvements) should be appropriate for the question.
- Description of the study design for the original paper.
- Critique of the study design for the original paper.

If you are very clear, this can be done without actually stating the model explicitly. However, you *may* include the model along with the description for clarity. However, the model alone is **not** sufficient, and if it is included, it should be formatted using an equation editor (or LaTeX). In addition to the description of the model, you should also include a statement regarding the assumptions:

- The Methods section includes a statement of how any assumptions were assessed.
- If any assumptions were found to be violated, the report should include a statement about how they were addressed.

Again, you want enough detail that I can replicate your work **without the R code**.

Results Criteria:

This component should not be technical. That is, it should be readable by anyone who has graduated from Rose-Hulman. Therefore, our goal is to write up the results of the questions asked in the Introduction in context, avoiding jargon.

- The Results section should **avoid** jargon such as "we rejected the null hypothesis." Every conclusion should be stated in context.
- At least 1 graphic summarizing the data should be included (see criteria for graphics below). The graphic should address a question of interest. This could be the graphic replicated from the paper or one generated if the paper did not include a graphic.
- Any included graphic should be referenced (discussed in context). This discussion often coincides with the discussion of a p-value or confidence interval.
- Address the questions of interest, citing any p-values or confidence intervals relevant to each question.
- Any limitations regarding the study or analysis should be noted.

Criteria for Graphics / Tables:

Graphics help to tell the story presented by the data. Many researchers often "skim" articles by reading the abstract and examining the graphics in the text. Therefore, our graphics need to be of sufficient quality to tell the story of our analysis. Only graphics relevant to the question of interest should be included in the text. Any graphics used solely for assessing assumptions should be relegated to the appendix. The same is true of tables.

- All Figures/Tables should be numbered.
- All Figures/Tables should be appropriately labeled (axis and legends) and captioned (description of graphic).
- Captions are placed just above (or below) the accompanying Figure/Table.

The above specifications ensure the graphic can stand alone. If all you were provided was the graphic (without the accompanying text in the Results section), you should be able to understand the graphic. The implications are discussed in the results section, but the graphic itself should be self-contained. Additional requirements for graphics and tabular summaries:

- Refrain from using too many significant digits on the axis markings.
- Axis labels should include units, where applicable.
- Fractional values should be presented with a leading zero (e.g., 0.3432 and never .3432).
- Graphics for assessing assumptions only should be placed in the appendix (not necessarily referenced in the text).
- Output directly from a computer package should be reformatted to be clear without knowledge of the program itself.

Criteria for Appendix:

The appendix should contain information not relevant for telling the primary story but useful for constructing a complete record of the analysis. The appendix should include the following elements:

- Graphics used to assess assumptions, if applicable.
- Full R code or Minitab output, if applicable.

Example:

While it may at first not seem like it, there is a great deal of flexibility within these specifications. Below are three annotated responses to various case studies. The first would be considered a successful completion (meeting all specifications); the second would be considered a successful completion requiring only minor revisions in order to fully meet all specifications (no essential specifications were violated); the third would **not** be considered a successful completion. All three are annotated to demonstrate how they meet (or fail to meet) each of the above specified criteria. Note: these do not actually replicate an existing analysis, but are meant to show the level of writing expected in the course.

- **Example of Successful Statistical Consulting Report**
- **Example of Successful Statistical Consulting Report which requires some minor modifications**
- **Example of Unsuccessful Statistical Consulting Report**

Last modified: Friday, March 6, 2020, 9:01 AM