Criteria: Written Capstone Project

The Written Capstone Project is an opportunity to design and carry out a study to address a research question of interest to you. As the lead investigators, you are responsible for the design, implementation, and analysis of the study. It is a chance to put the concepts from the course in practice. This is a group assignment and needs to be completed in a group of 3-4 students.

General Instructions:

During the course of your analysis, 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.

Your report will consist of three sections: an Introduction section describing the question(s) of interest, a Methods section detailing the analysis carried out, and a Results section discussing the results and limitations of the study. This report is meant to mimic a brief memo sent in industry summarizing a project. As a consequence, the criteria below are meant to produce a report which is readable 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."

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

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.
- Minitab output should only be placed in the appendix, and not referenced in the text.

Minitab output (that placed in the text window in Minitab, not the graphics) 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 output 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. The introduction serves as a brief summary of what the analysis is trying to accomplish and why. Specifically, your introduction should satisfy the following:

- Summarize why the study is of interest.
- State the primary question(s) that are to be addressed in the report.
- This section should be at least 3 sentences in length.

Note here that jargon is to be avoided. So, the primary question is **not** equivalent to stating a hypothesis. It is, in context, the question you hope to address.

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 Minitab output or giving menu commands for Minitab. The following criteria should be adhered to:

- The Methods section includes a description of the numerical and graphical summaries constructed.

Here, there is a distinction between what was done and what it means. You are simply describing what was done: "the age of our patients is summarized in a boxplot in Figure 1." We are not discussing what we conclude from the plot, simply what the plot is. The plot itself is not included here.

- The Methods section includes an accurate description of the analysis constructed.
- The Methods employed are appropriate for the study.

If you are very clear, this can be done without formally stating the null and alternative hypothesis. For example, you may simply say "we conducted a 1-sample t-test to assess if there was evidence that the age of patients exceeds 28 on average." However, you *may* include the formal hypothesis along with the description for clarity. In addition to the description of the analysis, 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, this is a description, not an information dump of everything you examined. The graphics used to assess the assumptions should be placed in the appendix. You simply include a statement such as "We graphically assessed (with a probability plot) whether the age of patients could be modeled using a Normal distribution, and we found this assumption to be violated. Therefore, we relied on bootstrapping to address our question of interest." The probability plot would be included in the appendix but not part of the main text itself.

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.
- 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.

Note that captions *replace* the need for figure titles (which are automatically generated in Minitab). Figure titles should be removed. 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 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.

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