

# DATA 410/STAT 538 Winter 2022 Project Report Grading Rubric

|                                  | 8  | 6  | 4  | 0  |
|----------------------------------|--|--|--|--|
| <b>Exploratory data analysis</b> | Establishes excellent understanding of data structure and generation process. Variables are precisely described. Proper numerical and graphical summaries are presented in the report.   | Establishes good understanding of data structure and generation process. Variables are adequately described. Some numerical and graphical summaries are presented in the report.   | Establishes sufficient understanding of data structure and generation process. Variables are described, but not in detail. Few numerical and graphical summaries are presented in the report.                      | Parts of the data are missing or completely ignored. Little to no discussion of variables. Numerical and graphical descriptions are not relevant or ignored.                             |
| <b>Scientific inquiries</b>      | Clearly and thoroughly explores all research questions using appropriate methodology. Makes conclusions on all questions.  | Answers research questions sufficiently using appropriate methodology. Conclusions on only positive results.   | Vaguely answers questions using dubious methodology. Insufficient number of research questions asked (<3). Does not conclude research questions.   | Research questions are not appropriate. Analysis methodology is not appropriate. No conclusions. The report is incomplete, and no scientific knowledge is demonstrated or gained.        |
| <b>Regression analysis</b>       | Provides excellent understanding of statistical inference for regression. Multiple methods are used to answer all questions. Formulae and calculations are clear and concise. Code is reproducible.  | Provides good understanding of statistical inference for regression. Multiple methods are used, but maybe inappropriate or repetitive. Formulae and calculations are understandable. Code is reproducible.                     | Provides sufficient understanding of statistical inference for regression. Regression methods are inappropriate or repetitive. Formulae and calculations are unclear. Code contains errors or is not reproducible. | Limited or poor understanding of statistical inference for regression. Clearly wrong methodology is used. No calculations are explained or provided. Code was not submitted.             |
| <b>Scientific Communication</b>  | No spelling, grammar, or punctuation errors in the text. The paper is organized in a logical manner and concise for the content. Presents tables and figures that are clear and accurate, and each figure/table is referenced and explained in the main body text. | A few (3-5) errors in spelling, grammar or punctuation. The material is presented well, but with some misplaced sections. Presents tables and figures that have good explanations. Most tables/figures referenced in the text. | Many spelling, grammar or punctuation errors (>5). Tables and figures that have lack explanation and do not have sensical captions. Few tables/figures referenced in the text.                                     | Excessive spelling, grammar or punctuation errors (>8). Little concern for presentation of results and no figure/table titles or explanations. No tables/figures referenced in the text. |
| <b>Creativity</b>                | Shows a lot of creative thoughts in addressing the topic, graph design, and data issues. Integration of multiple topics discussed in class.  | Shows some creative thoughts in addressing the topic, graph design, and data issues. Some regression techniques integrated into analysis.  | Shows very little creative thoughts in addressing the topic, graph design, and data issues. Few regression techniques integrated into analysis.  | Shows no creative thoughts in addressing the topic, graph design, and data issues. One regression technique is used, with no consideration of others.                                    |

\*Keep in mind that there can be part marks for each section. For example, if you do an excellent job of describing the data, but maybe you miss a group of variables in your dataset for some reason, you would receive 2.75. When I'm marking, I will be highlighting phrases in this rubric to explain marks.