# SDS 291: Case Study Rubric

Case studies will be evaluated holistically on the basis of the formal components of the report (introduction, data description, methods/results, and discussion), the statistical methods selected, and the writing style. Each of these components will be graded out of either 5 or 10 points, for a total of 45 points:

- Score of 5 (or 10): work that is exemplary

- Score of 4 (or 8): work that is very good

- Score of 3 (or 6): work that is satisfactory

- Score of 2 (or 4): work that is below expectations

- Score of 1 (or 2): a reasonable attempt.

The bulleted items below represent guidelines for each of the six components. An exemplary report should generally be able to answer YES to most of these items. Please note that case studies are only intended to be 500 to 1,000 words in length, so each item below may reasonably be addressed in only a sentence or two!

| Introduction (5 points) | Grade: |
|---|---|
| YES/NO • Is the topic of the case study clearly explained and some justification/motivation for the analysis provided? | *Comments:* |
| YES/NO • Are the research question and goals of the analysis clearly explained? | |
| YES/NO • Are the outcome and explanatory variables of interest clearly identified? | |

| Data Description & EDA (5 points) | Grade: |
|---|---|
| YES/NO • Is the source of the data clearly explained? | *Comments:* |
| YES/NO • Are any important data wrangling steps (e.g., removing observations with missing values, constructing new variables) clearly described? | |
| YES/NO • Are the number of observations in the dataset (and the number of observations removed from analysis) clearly identified? | |
| YES/NO • Are the distributions of the outcome and relevant explanatory variables, as well as the relationships between them, clearly and accurately described? Using appropriate summary statistics and at least one data visualization? Are these unpacked and interpreted for the reader? | |

| Statistical Methods (10 points) | Grade: |
|---|---|

YES/NO    • Are the statistical methods and modeling approach appropriate and logical? Are they motivated by the research question and exploratory data analysis?

YES/NO    • Are these methods described succinctly and in sufficient detail that someone knowledgeable could replicate the analysis and evaluate the analytic decisions made?

YES/NO    • Is the mathematical representation of the final population regression model provided without errors?

YES/NO    • Is the notation used for the outcome and explanatory variable(s) in this model clearly and succinctly defined?

*Comments:*

---

| Results (10 points) | Grade: |
|---|---|

YES/NO    • Are the estimates of the model parameters and their accompanying standard errors provided?

YES/NO    • Are the relevant estimated slopes, population means, hypothesis tests, and/or predictions (as appropriate for the stated research question and goals) correctly interpreted *in context*?

YES/NO    • Are confidence intervals or prediction intervals provided for each of these interpreted quantities?

YES/NO    • Are the final results of the analysis logical and explained in a clear and straightforward manner? Are these results clearly tied back to the research question and goals?

YES/NO    • Are the regression assumptions and model goodness-of-fit investigated and appropriately assessed? These items may be included in the R appendix rather than the case study write-up, but should be clearly marked.

*Comments:*

---

| Discussion (5 points) | Grade: |
|---|---|

YES/NO    • Are the case study objectives, results, and conclusion clearly restated?

YES/NO    • Are these results discussed in the broader context of the case study (i.e., are the results and conclusion connected back to the larger motivation for the analysis)?

YES/NO    • Are potential limitations of the data and/or analysis acknowledged and discussed?

*Comments:*

| Writing Style (10 points) | Grade: |
|---|---|
| YES/NO    • Overall throughout the report, are ideas expressed precisely and written in a manner that is almost entirely free from spelling and/or grammatical error? | *Comments:* |
| YES/NO    • Overall throughout the report, is the writing and exposition geared towards someone with a similar level of statistical training and domain knowledge? | |
| YES/NO    • Do the data visualizations and tables help to complement the writing and support the themes/arguments of the case study? | |
| YES/NO    • Is statistical evidence (data visualizations, tables, summary statistics, estimates, hypothesis tests) skillfully woven throughout the report? | |
| YES/NO    • Is the output from any R code appropriately styled and polished? Are tables nicely formatted as tables? Are data visualizations shown with descriptive axis labels and titles? | |
| YES/NO    • Is the R code otherwise hidden from the reader, and is the report free from messages, warning messages, and error messages? | |
| YES/NO    • Does the R appendix at the end of the report contain code for all aspects of the analysis with clear documentation? | |

**Total Score:**