Critically Evaluating Scientific Writing

Kristine Dinh

Summary: Here is the paper rank from strongest to weakest: exam 2, exam 11, exam 5, and exam 1. Below is the detail critically evaluating of each paper with six criterion including accuracy, clarity, logic, organization, scope, and figures and tables.

Exam 1: This paper started out with a clear introduction and purpose of finding the factors that would affect high school students' test score for college. It is unclear in understanding all the methods used in this analysis. In terms of accuracy, there are some part that jumped too quick to conclusion. For instant, on the "Analysis of Parameter Estimates" table, the writer said to eliminate all the variables that shows insignificant relationship. However, there is no sign or explanation saying those variables are not significant. In general, the variables can be insignificant due to multicollinear problems, variables factor levels, or outliers. On the last paragraph, the author concludes the analysis saying that better test score would increase chances of getting into college. However, there are no clear evidence that the author would explain beside showing a table of parameter estimates. As a suggestion, an interpretation of the table would be nice to have in the analysis. As mentioning tables, there are no captions and references to explain the information in the tables, which makes the readers feel difficult to understand where the writer gets information from.

The paper is written in languages that is easy to understand. There is no terminology that made the paper difficult to follow. However, the abbreviation MCAS should be explain somewhere in the analysis since I have no idea what MCAS stands for. The author organized this paper by introducing, giving results, and then conclude the paper with vary strong argument. However, there are some addition steps like method and variable exploratory analysis should also be include in the paper to a deeper understanding of the data. As for scoping, there are more that the writer could have been done with the data. If I do the analysis, I would add variable exploratory section to explain the significant of the variables, explain the correlation between each variables, reassure the assumption of error terms to be normal approximate and random, validate the data using R-square, MSE, etc., and diagnostic to see the quality of the model.

Exam 2: This writer did a good job in introducing the topic and methods of statistical analysis. I can see that the writer has all parts needed for the data analysis including data introduction and summary, variable exploratory analysis, methods, statistical analysis, clear results, and diagnostic to test for model quality. However, there the writer did not consider the normality of target variable, which is an important factor to consider in a multiple linear regression. There are data in the Appendix to support the writer's arguments. However, I would suggest composing a summary table of the results and insert it in the text. This way, the readers don't have to scroll all the way to the Appendix to check the results while reading the analysis.

The paper is very well written and easy to follow. Sometimes, the sentences are too long and tacky which made the paper confusing. But other than this, the paper is well written. The paper is properly organized with each part separated to their own subheading. However, the plots take too much space in the report which makes it seems like plots are abundant. Although plots take a lot

of spaces, some plots are difficult to see due to the small font. I would suggest cutting down the number of plots presenting in the report and explain the plots in detail.

There is more that the writer could have been done with the data. For example, trying different variable selection methods to select new/more variable to increase R-square. In addition, the writer can also add one more metric to validate the model as R-square alone cannot be reliable.

Exam 5: In terms of accuracy, this report seems to not have any mathematical errors. However, the authors did delete four observation from the dataset and did not state why beside saying they are zeros. The paper, overall, is understandable and easy to read. The author seems to be very nice in the paper by saying "please" and "I believe" which sounds like the author is begging the readers to do something. Some parts are harder to understand due to the run-on sentences. In addition, there are too many raw column names that are hard to understand like SATM, SATPartRate, CostPupil, and SCounselRatio.

The data and calculation in this report support the conclusions drawn by the author. There are many figures and tables to support the author's conclusion. However, the figures and tables take too much space from the report making the paper look unprofessional. Figures and tables do have clear and detail captions which is very easy for readers to follow along. For this part, I would recommend the author to pick the plots/tables that would represent the data analysis best to present in the report.

Of course, there is more that could have been done with the data. The author did a great job in checking for multicollinearity issue and the normality of the error terms. However, more analysis could have been done on this part like checking for outliers and influential points. In addition, the variable exploratory analysis part could have been longer to have more details of each variables along with their hypothesis.

Paper 11: All methods, analysis, and statement in this paper seems to be correct. The paper is quite difficult to follow. The author seems to shift between models often, which made it hard to read the results accurately. The data and calculation support the conclusion drawn by the authors. However, I feel like there are some repetitive information in the paragraph. The organization of this paper is not as pleasing. All paragraphs are too short making the report seems like a list of information. In terms of figures and tables, there are no captions and references. This leads readers to be confused when reading the paragraph and trying to find the results in plots and tables at the same time. Just a minor suggestion, I think it would look nicer to write the formula in a nice format instead of just writing p/(1-p).

There is more that could have been done with the data. For example, in addition to presenting multiple models, the author could have dive into the model diagnostics to check for the quality of the model. Details explanation of diagnostic would be more helpful to evaluate the models. In addition, I would also suggest the author to use multiple validation metric to ensure the robustness of the models. This way, the audience would have a clear understanding of the important of the variables to the target variable.