

Final Project Guidelines for STATS/DATASCI 551 Bayesian Data Analysis

Yixin Wang
University of Michigan

The final project is an integral part of this course. There is no better way to learn about Bayesian data analysis than by applying it to a research question of your own. The final project is an individual project. You have a lot of freedom in choosing a topic for your final project. The only criterion is that it deeply involves applying Bayesian data analysis to a real-world problem. You choose a dataset, an interesting question about it, and address it with Bayesian modeling.

The project involves project report and project presentation. For the report, **please use 12-point font, single-spaced, and 1-inch margins. Page limits are without figures**; include as many pages of figures as needed; please place all figures and tables at the end of the report with proper labels and captions. We offer a LaTeX template; you are encouraged to use it but not required.

Project Report

The project report is due at the end of the semester. The report should have a similar format to that of an academic paper. Notebooks or markdowns with annotated code which generates the results from the final project should also be submitted as supplementary materials. You can further include a set of appendices (of any length) to which you can banish any details. (We may not read the appendix when grading your work.)

The project report should (at least) include

- Introduction: Clear description of the problem. Describe the problem you are trying to solve, why it is important or useful, and summarize any important pieces of prior work that you are building upon.
- Dataset: Clear description/visualization of the data. Describe the dataset or datasets you are working with. Show examples from the datasets. If you collected or constructed your own dataset, explain the process you used to collect the images and labels, and why you made the choices you did in the data collection process.
- Method: Clear and thorough description of statistical analysis. Describe the method you are using; this may also contain parts of the implementation of your model, or other components along with sanity checks to ensure that those components are correctly implemented.
- Results: Clear and thorough interpretation of results. Describe the key results and figures that you obtained. This may interleave explanations of the methods you run and figures you generate as a result of those methods.
- Limitations / Conclusion / Future work: What did you learn in doing this project? What are the shortcomings or failure cases of your work? If you had more time or resources, how would you continue or expand upon the work you have already done?

You should turn in a .pdf file containing your final report, together with the notebooks/markdowns containing all the code and the generated results (tables, figures etc) that are included in the report. You must run all cells in your notebook to receive credit; we will not rerun your notebook. Please combine the report and the notebook/code and submit a single pdf file.

The project report must be submitted through gradescope by the deadline. The project report is 2-3 pages long, excluding figures or references. For students enrolled in STATS/DATASCI 651, the project report is 4-6 pages long, excluding figures or references. (A longer report is not necessarily better; we prefer concisely written reports.)

The deadline of the final project report is strict, and late days cannot be applied. All late final projects receive a score of zero. In case of requesting exceptions due to severe medical reasons, a doctor's note and a signature from your graduate (or undergraduate) advisor is needed.

Project Presentation

Based on your final report, you should prepare a 5 to 7-minute recording of your project presentation. You can use Zoom to record your presentation and save as a .mp4 file; see instructions [here](#). The project presentation must be submitted through gradescope by the deadline.

Project Evaluation

We evaluate the project on ambition, significance, originality, technical depth, results, relevance, and writing quality. Two good books about writing are [Strunk Jr and White \(2007\)](#) and [Williams \(1990\)](#).

The project is graded 80% based on the project report and 20% based on the project presentation.

The following grading criteria for the **final project report** is based on 100% maximum.

- Solid understanding and description of problem (+5%).
- Quality of data analysis.
 - Appropriateness of **Bayesian** statistical models and efficiency of the computational algorithms adopted (+20%).
 - Evaluation, expansion/adjustment, and discussions of potential limitations of **Bayesian** statistical models (+20%).
- Presentation of results, including visualization (+15%).
- Proper interpretation of results (+10%).
- Well-organized and clean code (+5%).
- Clarity of written report (+15%).

- Irrelevance with course materials (-40%). (Again, the focus should be on Bayesian inference. Be aware that applying a technique that has the word “Bayes” in it, e.g., “naive Bayes classifier”, does not mean that you are practicing the Bayesian methodology).
- Plagiarism of code or written report (-100%).

We provide an **example final project report** on Canvas. However, we note that this example is **not** a perfect project report according to the grading criteria. For example, it does not perform any model checking.

The grading criteria for the **final presentation** is the following:

- Concise Delivery: the presentation length is 5 to 7 minutes and you should be able to dedicate enough time to each section.
- Clarity: the goals of the project should be clearly presented.
- Organization: your slides and the presentation flow should be easy to follow and tell a coherent story.
- Effectiveness: your findings and their importance should be reported in an effective way, i.e. you should convince the listener why this result is useful.
- Relevance: the techniques and methodology used in the project are covered in the course or closely related.
- Critical thinking: your presentation should take into account potential limitations to your work and possible areas to improve.

References

Strunk Jr, W. and White, E. B. (2007). *The Elements of Style Illustrated*. Penguin.

Williams, J. (1990). Toward clarity and grace. *Chicago: The University of Chicago*.