

# STA/OPR 9750 Final Project Requirements

Timothy Teräväinen

Spring 2021

## Structural Requirements

- An R Markdown-based document with PDF output. Include and submit all code written and used, either as embedded R Markdown segments with `echo = FALSE` if used for plotting or analysis, or separately if used for retrieving and pre-processing data).
- At least 2 output pages of exploration (including figures), using tidyverse.
- At least 4 output pages of analysis and conclusion (including figures).
- At most 10 pages.

## Figure Requirements

- At least 2 exploratory figures (at most 70% figures by volume of output).
- At least 2 model/analysis figures (at most 70% figures by volume of output).

## Modeling Requirements

- At least 2 fitted advanced models/simulations. See me if this is an issue.

## Final Project Evaluation<sup>1</sup>

### Presentation (20% of total):

- **Exploration of Data, and Written Analysis, 10%:** Summary of dataset (subgroups,  $n$ , number of variables, any limitations of the data, source of data, etc.) before modeling. Also, explanation and justification of analyses used.
- **Hypothesis and Conclusion, 10%:** A statement of what you are looking for in the data, how you addressed missing or corrupt data, what you discovered, whether you were correct, etc. This grade is based on the quality of the hypothesis and presentation of findings, not just correctness.

### Analysis (60% of total):

- **Appropriateness, 20%:** Are the analyses motivated by your hypothesis and the form of the data? You should be able to answer why you did what you did, for example, a linear regression instead of random forests or a time series.
- **Technical Quality, 40%:** Did the code do what you claim it did? Are there errors or bugs in the code (programming or statistical)? Did you do model validation?

### Figures (20% of total):

- **Appropriateness, 10%:** Figures are relevant to the exploration and analysis of the data. They should illustrate highlights and conclusions from your exploration and analysis.

---

<sup>1</sup>NOTE: Any descriptive text (not in bold type) is given only as an example of some details associated with the criterion (in bold type). It is not intended as an exhaustive list of evaluation metrics.

- **Quality, 10%:** Titles and subtitles. Axes are labeled and correctly-scaled. Multiple plots in the same figure have color-coding and a legend explaining them. Conclusions should be at least as clearly visible in a graph, as they are in the writing.