Behavioral Data Science: Project Template

April 10, 2025

Aims

This project is an opportunity to demonstrate your understanding of behavioral data science through an end-to-end analysis. Your goal is to apply the ideas and tools from this course to a dataset and a problem of your choice, and to carry out an analysis that reflects methodological understanding.

Your submission should consist of two key components: a **written report** and the **accompanying code**. The written report must follow the structure outlined below and should address the majority of the listed guiding questions.

The focus is not on obtaining the "best" performance, but rather to demonstrate an understanding of workflows and methods.

Introduction

What is the problem? Why does this problem require statistical analysis? What are the key variables $\mathcal{D}$? What are your initial hypotheses? What kinds of insights do you hope to gain from this analysis (e.g., modeling latent patterns, testing hypotheses, predicting behavior)?

Methods

Which model(s) did you choose to represent the behavioral process? These could be:

- Classical statistical models (e.g., regressions),
- Machine learning models (e.g., neural networks, random forests),
- Latent variable or computational models (e.g., factor analysis, diffusion model).

It is important to clearly state the assumptions of the model and the types of the variables involved (e.g., discrete, continuous). Also discuss:

- What software and libraries did you use?
- How did you evaluate **model fit** and **predictive performance**?

Results

Use this section to present your results. The more aspects of the analysis you explicate, the stronger this section becomes. Use tables and visualizations to support your findings wherever possible.

Discussion

Reflect on what the analysis revealed about the behavioral problem.

- What insights did you gain?
- Did the results confirm or challenge your expectations?
- What were the main difficulties or trade-offs in your modeling decisions?
- What could be done in future work to extend or improve the analysis (e.g., richer data, alternative models, more domain knowledge)?

General Remarks

- Your report should include a title page with your name, institution, and any other relevant information.
- Clear visualizations and tables will improve the clarity of your report. High-quality figures are expected, especially for graduate students.
- You may (but are not required to) cite academic literature, online materials, or tools you relied on.
- You are encouraged to compare at least two different models or modeling strategies (required for graduate students, as every analysis needs a baseline and a benchmark).