

## Project Report guidelines

The project reports can include several different formats: a simulation exercise, a data analysis, a discussion of an on-going debate, a discussion of a method not studied in class. The first and last will be the more typical choices, and those are the formats we will discuss.

### Project Selection

Many Master's and Ph.D. students include a simulation study as part of their thesis, so the STAT 740 project presents a natural opportunity to make forward progress on material that is directly relevant to the student. Likewise, students often select a method not discussed in class because it is closely related to a method covered in class. Of course, a student's selection of a methods project could be related to their thesis work, just as in the case of a simulation study.

### Introduction

The first part of your report can include your motivation for the project (as suggested by the preceding discussion). It is perfectly fine for the discussion here to be personal or anecdotal. Do not feel you have to leave context out of the report. The first part of your report should also introduce the data, method, or model that precipitated your interest in your project. You will find that describing a method or model in full detail is tedious—students often leave out important information and notation. Be patient and get all the details down—this will make the rest of the report that much easier.

You have the option of introducing specific data sets at this point. In a typical academic paper, the motivating example could be introduced later, but sometimes discussion of a specific example can help explain why you became interested in a method or application in the first place. Use your discretion.

### Modelling

Having introduced the method or study, you can discuss estimation or inference for your study, and describe the study's limitations. I would not recommend discussing the analysis or simulation at this point; this can be saved for later.

### Modelling–Methodology

With the motivating method and the estimation/inference problems introduced, you can now discuss the related method(s) you plan to explore. The methods should be introduced in about the same level of detail as used for the introductory material. Deviations from the introductory material could include different methods, different models, different test statistics, etc. You do not need to go into a great deal of elaboration on methodology derivations, but some attempt at deriving important or critical results should be made. E.g., if you study a method for adjusting the step direction of the Newton-Raphson algorithm, then the new step direction should not simply be presented as a fait accompli; instead some insight (perhaps geometrical) into the derivation of the result should be made.

You should also decide how you will compare the methods. For two density estimation methods, for instance, you can compare AIMSE or some other measure.

### **Modelling–Simulation Study**

For a simulation study, you could now describe all the factors you plan to vary. In any given study, there are typically too many factors to vary at once: test size, sample size, distribution, distributional parameters (e.g., scale), methodology, etc. Limit your study to only three or four factors and select factor levels accordingly. Start modestly and in such a way that you could easily add factor level combinations if you find you have the time.

You also have to decide how you will make comparisons. For a confidence interval study, you have to decide whether you will compare coverage probabilities, confidence interval length, or some other measure.

### **Simulation and Analysis**

At this point, you can present your data. If working with an actual data set, finding one with interesting subject matter is difficult; finding one with interesting subject matter, and an interesting analysis is even harder. For this reason, I am not placing a premium on the data—you can make up the example and the data itself in order to expedite your analysis. The purpose of the data set is to highlight your new methods—an exhaustive analysis is not necessary.

When presenting results, rely on graphical methods as much as possible. It is common statistical practice to summarize simulation results in a table format, and these results are often unreadable; use graphs whenever possible.

If you construct new code, I hope to place it on my homepage. I have pretty loose standards for documentation, but it would be useful if you included a modest header describing your function/source code. New code should be included in an Appendix.

Be sure to write a brief conclusion summarizing your analysis. Many times, students simply present an analysis and then the reader is left to draw their own conclusions. The analysis is only a means to an end, though it is easy to lose sight of that in stat classes that focus so heavily on analysis.