Project Report guidelines

The project reports can include several different formats: a simulation exercise, a data analysis, a demonstration of software features, function/macro writing, or a discussion of a method not studied in class. These topics require different approaches to report-writing, and we'll try to cover all the bases in what follows. Not all of what follows will be applicable to any one project–you will have to pick and choose.

Project Selection
Of course, some students' choices of project topic are arbitrary, but I encourage students to choose a topic directly relevant to their personal/work/professional interests. This can mean revisiting old data sets/analyses, or applying methods from this class to current/future projects. Alternatively, students often select a method not discussed in class because it is closely related to a method covered in class. Do not worry if your project ideas don't seem like a good "fit" to the course; we try to be flexible in the scope of our projects.

Many graduate students include simulation studies as part of their research, so the STAT 540 project presents a natural opportunity to make forward progress on material that is or will be directly relevant to the student.

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, coding problem or model that precipitated your interest in your project. If you choose to report on a new method or model, you will find that describing them 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 to write.

You have the option of introducing specific code or data sets at this point; the discussion of a specific example can help explain why you became interested in a method or application in the first place. Use your discretion.

Software demonstration/coding In many cases, you may be demonstrating features of SAS's `DATA` step, or special output/graphics features in both **R** and **SAS**. In these cases, you will want to review introductory approaches or related methods that we learned in class. Include examples of code and typical output. Discuss these approaches critically, and introduce the procedures or approaches you would like to illustrate. Be sure to discuss their advantages compared to other available procedures. In some cases, this could be difficult if we didn't learn comparable procedures in class; I suggest that you focus on motivating the method on its own merits instead.

If your software demonstration includes methods and models, please review the additional material below.

Modelling–Methodology
Any alternative methods should be introduced at 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 detail on methodology derivations; straightforward presentation of results will be fine. You should also decide how you will compare methods.

Modelling–Simulation Study
For a simulation study, you could describe the factors you plan to vary. In any given study, there are typically too many factors to vary at once: number of simulations, test size, sample size, distribution, distributional parameters (e.g., scale), methodology, etc. Limit your study to only two or three factors and select factor levels accordingly.

For a stochastic simulation, your factors will be quite different from those in a typical simulation study, but you should consider a similar number of variations (e.g., for *Gambler's Ruin*, you could vary probability of winning, opening stake, etc). Start modestly and in such a way that you could easily add factor level combinations if you find you have the time.

Simulation, Analysis, and Results
At this point, you can present your data, if you have not already done so. 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 source of 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.

If introducing new software procedures, be sure to discuss features and options in some detail. An iterative approach can be useful here, showing a basic graph or table, and then discussing how you enhanced the output. More than one data set may be necessary to capture the full range of the software procedures you are studying.

When presenting results from a simulation study or analysis, 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.

New code, output, and graphics should generally be included in an Appendix, unless they fit nicely within the flow of your narrative.

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