Project and Presentation Requirements for STAT 540, Spring 2018

- Important dates:
  1. February 27 (in class): Submit hardcopy of your project proposal. This is a 1 page detailed summary of your project plan.
  2. (Tentatively) April 16 - April 27 (last two weeks of classes): Project presentation, during class time or by appointment. More information on this will be provided later. Submit slides *by midnight the day before your presentation.*
  3. Monday, April 30: Submit final project reports on Canvas by midnight
  4. Wednesday, May 2 by midnight: Submit evaluation of other students' projects/presentations on Canvas by midnight
- Project expectations:
  1. Please talk to me about potential project topics. You are welcome to work on something related to your dissertation *as long as it is distinct from what you were already planning to include in your thesis.*
  2. Projects must contain some original work. They should include (i) the implementation and study of computer code and a computational algorithm **and** (ii) *at least one* of the following:
     - A thorough simulation study. This is especially appropriate if the focus of the project is one or more papers. Examples of original work here include conducting a simulation that was not done in the manuscript and/or a comparison with an algorithm that was not studied in the manuscript. Replicating a simulation study that is already published is a reasonable starting point but you need to do more.
     - An application of existing methods or models to a new data set. The methods must be directed towards answering a carefully formulated scientific question. In general it is not a good idea to work on a data set unless you have well defined questions and some knowledge of the data source.
     - An application of a new method to a data set that has already been studied before (via other methods).
     - Some theoretical development: either rewriting/explaining a proof clearly or going a little beyond what is done in the paper.
  3. Attribution of sources: It is very important that you attribute sources correctly. Do not copy material from the web – you can use web searches to find appropriate references and then cite them. Do not use any writing from papers without citing them. This is not just a

guideline for this project but is also to ensure that you do not violate principles of academic honesty. When in doubt, cite more rather than less.

4. Submit `R code` (or Matlab/C etc.) that you used to implement the project.
5. Your writeup:
   (a) Your **writeup** must be typed. The main portion of the manuscript should be no longer than 7 pages, including figures and tables, but excluding references (that is, references do not count against your word limit). The audience for this writeup are your classmates and instructor. You are allowed to have technical supplementary material but this material must also be carefully edited and no longer than 5 pages.
   (b) Introduction/motivation (description of problem): This should be very concise and clear to someone who does not know any of the specific terminology used in your field. Some of the main points to convey: (i) what is the main problem being solved? (ii) outline of the methodology (algorithms, models); (iii) what, in particular, you are investigating in this project; (iv) how the rest of your manuscript is organized. This may also include necessary background material such as related methods or previous analyses done on the data you are studying.
   (c) Description of model and algorithm: make sure you clearly explain the model and/or algorithm that you are studying. Do not simply copy what is in a paper – you have to translate that information into an easy-to-understand version.
   (d) Description of data and/or simulation study: Where were the data obtained and what do you know about how they were collected? If the data are publicly available, provide enough information (e.g. website) to make it easily accessible. If you are describing a simulation study, explain how you decided to carry it out. You may say that you are re-using simulation design from another paper, but you have to explain why those parameter setting are reasonable or interesting.
   (e) Results: This should consist of: (a) Conclusions about methodology: What, if anything, have you learned about the applicability of the algorithms you tried to the problem you were considering? (b) What are useful guidelines for someone else who might choose to implement the algorithm(s) presented? (c) Scientific conclusions: If this is relevant, describe the implications of your analysis in the context of your particular scientific domain. Again, this should be written so a non-expert can understand it.
   (f) Future work: As briefly as possible, explain what still remains unresolved. For instance, is there other statistical methodology

or algorithm that you would like to apply to this problem? Are there other questions you would like to pursue with the same data set or statistical problem?

(g) Supplementary material: This is not necessary but it is likely to be a useful place for you to write down technical details that may not fit into the main manuscript. I will look over this – if it is hard to read or not edited carefully, you will lose points. You may also include extra figures, tables here.

6. Project scoring:
   (a) Proposal: 5
   (b) Organization/clarity of writing: 10
   (c) Correctness of work: 5
   (d) Depth (how deep you went into the subject): 5
   (e) Presentation: 10
   (f) Attendance/evaluation of other students' projects: 5
7. Presentation Guidelines:
   (a) All talks will be for a precisely scheduled time (to be determined by the number of projects submitted, usually between 12-20 minutes long). You cannot convey too many details – your focus should be on explaining some basic ideas. It is fine to finish in slightly less time but it is not okay to go over.
   (b) You are required to send me your slides by midnight *the day before your talk*. Please ensure that your slides are in pdf format. If you are going to use any other format, you need to get special permission.
   (c) You will be allowed a prescribed number of slides (usually 1 fpr every two minutes). Given the short amount of time allotted to each student, I will delete any slides that go beyond the prescribed limit.
   (d) Your goal should be to communicate in as effective a way as possible what you did in the project. The presentations are primarily meant to be educational for the entire class so you will also be graded on the clarity of your explanation.
   (e) Rehearse your presentations: giving a very short talk is not easy. Here are some suggestions: `http://personal.psu.edu/muh10/540/shorttalk.html`
   (f) Attendance is compulsory throughout the project presentation period. I will take attendance at the beginning of each class.
   (g) At the end of each class, if we need extra time for discussion, it is possible that the class will go over the allotted time. Please allow for an extra 15 minutes at the end of each of these days.
   (h) Depending on the number of projects, a few of you may present your work to me in my office at a specially scheduled time.