MSDS 596 Regression & Time Series Fall 2018

Final Project Guideline

Due dates: various - see below

The project is to be carried out by a team of two investigators. Imagine the upper management is contemplating to disband your team, and this is your last chance to show how useful you are in order to save the job. You can choose to do (but are not limited to) one of the following: find a strategy that will potentially generate positive return; demonstrate your ability to evaluate risk more accurately, including risk calculation and stress testing; demonstrate your ability to help other teams/lines of business in generating useful information and future strategy from data.

The final project counts towards 20% of your grade. It consists of a project proposal (10%), an interim report (optional), a project presentation (40%) and a final project report (50%). Their due dates and respective contents are as follows.

Project proposal (10%**, due Tues 11/6 @ 6:40pm).** The *project proposal* describes what you plan to do for the project. What is the problem you aim to address, why it is important to the business/economy/environment/society, what kind of data you are going to use, and a list of possible methodologies you plan to use.

Interim report (optional; due Tues 11/27 @ 6:40pm). While the *interim report* is optional, it is a great opportunity for your team to obtain feedback from the instructor and/or make sure you're on the right track. The interim report should include data description, preliminary analysis, the methodologies you are using, and the results you expect to get.

Project presentation (40%**, Tue 12/11 in class).** The *project presentation* is limited to 7 minutes (or shorter, depending on the number of groups in class), plus Q&A with the audience. Your presentation should cover the major components in the final report (see below). Both team members should speak a roughly comparable amount during the presentation. A clearly outlined, concisely styled, well-paced presentation with thoughtful use of visual displays is highly valued.

You are also expected to ask questions and provide constructive feedback during the Q&A part following other teams' presentations. A portion of your presentation evaluation will be based on the quality of your participation and feedback as audience.

Final report (50%**, due Thur 12/13 @ 11:59pm).** The *final report* is a written report explaining your work. It should contain the following components, including supporting graphics or tables:

1. An introduction to the problem addressed by the report, its background, and its significance to the company you're working for/economy/environment/society, etc.;
2. A description of the data used in the analysis;
3. The quantitative method(s) you've employed to address the problem, including the rationale for your choices, as well as model assumptions relevant to real-life interpretability;

4. The results of your analysis. This should include introspective comments on the performance of methods employed, such as model diagnostics and sensitivity analysis.

5. Conclusions, implications, and suggestions for the future based on your analysis and findings.

In addition, you should acknowledge all sources of external information referenced within your work, using a proper bibliography style system.

While it's important that you organize the report to clearly and comprehensively communicate the above things, you do not need to insert extra written filler to meet a particular minimum page limit. Submitting more than 5 double-spaced pages of writing, not including figures and tables, will probably decrease rather than increase your grade. Conciseness and clarity are highly valued.

Submission instruction. For all written submissions, **one and only one** of the team members will submit the file on Canvas, while clearly indicating both members' names in the writeup.

For the final report, you will submit (i) the written report in `.pdf` format, in which all supporting figures and tables are included, but no code nor raw data is shown. Separately, submit (ii) dataset(s) used for your project, and (iii) `R` code that can reproduce all the results in your report by directly reading the dataset.