In this project, you will apply statistical methods and concepts from class to a real-world decision making situation. You will identify a data-driven decision scenario, source appropriate data, analyze the data, and derive actionable insights from your analyses in any industry/sector of your choice. Your final deliverables include a project report (10-12 pages) and an end-of-semester team presentation. Your project report must include the following sections:

(1) **Executive Summary (less than 1 page**): Brief summary of problem, data, analysis, and key findings.

(2) **Table of Contents.**

(3) **Problem Significance:** Who is the target client for this project, and what business problem are you trying to address for this client? Why is this an interesting or important problem? Present some brief industry statistics or background research to make a strong case for your target problem.

(4) **Data Source/Preparation:** Where did you source your data? What variables did it contain and how were they measured? Which variables (DV, IV) did you select for your analysis and why? How did you clean the data (if applicable)?

(5) **Hypotheses:** What are your core hypotheses of interest, and why should the reader be interested in these hypotheses?

(6) **Descriptive Analysis:** What patterns/trends do you see in your data? How did you infer these trends?

(7) **Models:** How did you statistically model your data? You can examine at most 3-4 models, but you must carefully choose your models. Which model is best and why?

(8) **Quality Checks**: How do you know that your analysis is trustworthy, and

(9) **Recommendations:** What recommendations do you have for your client, based on your analysis? Your recommendations must be actionable, i.e., things that managers can act upon.

Note that your report is intended for an executive audience. These are people who have no background or interest in statistical methods or R. Hence, please tone down technical details and write in a layperson's language. Absolutely, no cut-and-paste from R. You can summarize relevant numbers from R output in user-friendly tables by stripping out unnecessary details. You can include your models in a compact manner in the report. If you want to include detailed analysis and R outputs, you can do so in an Appendix section, following your 10-page report.

Sections 3 and 9 are the most important sections in this project. Note that Section 8 requires "actionable" recommendations. This means that the client must be able to act on your insights. For example, if you say that house prices increases with CPI, or a certain disease is more prevalent in women than in men, there's nothing that a manager or a doctor can do about it, since CPI and gender are outside their control. Hence, these insights are not "actionable." You must think about this issue in Sections 3, 4, and 5 (while choosing your research problem, data, and hypothesis); Section 9 is too late to address this very important issue.

Please note that classification problems are NOT acceptable for this project. Even though there are many public data sets on classification problems, I want you to demonstrate what you learned from THIS class, not from the data mining class.

You project will be judged based on the criteria listed above, as well as by the quality of your writing. This project is a test of your writing skills, as much as it is a test of your analytic and problem-solving skills. You have only 10 pages or so to address all of the above sections. Hence, you must be ultra-compact and organized in your use of space, while making sure that you present all important information and not miss out on critical details. Use tables and graphics intelligently to organize information in a compact manner. But just throwing in some graphics and not talking about it in the text is also a bad idea. Please use the best writer in your team to write/supervise/coordinate/organize the project. It is a terrible idea to assign different students to write different sections, in which case, some sections will be well-written and others will be poorly-written, and overall, you will have an incoherent report that does not flow from section to section. Because this is a team project and not an individual project, the entire team will be penalized for poor writing by some of your team members.