Final Project Guideline

General Information

As a part of the assigned work for our course, students are to engage in a final project. Project will be carried out in groups of 3 (or 4 if the project is sophisticated enough, subject to instructor's approval). The final project must be original work and you should use the real-world data sets. Ideally, you should try to do something a bit interesting, like studying a data set that has not been thoroughly evaluated, or using a different approach. You should make sure that your analysis is not trivial. This is an application-based project but you should not worry about "the best" result. The goal of the final project to use algorithms we learn during classes to analyze the data and discover something interesting. You are free in choosing both the area of research (data, topic) and the algorithms to use.

Credits: 20% of total grade.

Deadlines

- May 10 *Proposals due*
- May 22th *Final reports due*
- May 27th *In-class presentations*

Proposal

You must submit a proposal for the instructor's approval.

Your project one-page proposal should tell the instructor about the problem you aimed to solve, ideas about how you might begin: general approach and methods (decision trees, regressions etc.) you intend to perform, data you are going to use and approximate effort required. It must consist of a preliminary title, a high-level summary of your project and outline your main goals, a brief description of what you plan to do. The purpose of the proposal is to make sure that you are on the right track and obtain the useful feedback from the instructor.

Final Report

The required element of the final report:

- Title
- Introduction: (the main scientific concept of the problem, summarizes the paper, and the goal of the project)
- Model: (the description of the data sets, the evaluation metrics, the data mining tools used, and any other details related to the main work)
- Results (the experiment result, the performance metrics should be given and be compared with theoretic results, etc.)
- Related Work: A brief description of related work, with citations to relevant papers.

- Conclusion: (the quality of your results, material on future work, etc.)

Presentation

The last week of classes will be dedicated to project presentations, each group will give short presentations on what they have accomplished (10-15 minutes each). You will only receive a grade at the end, based on your final projects and presentation. Extensions will be granted only in cases of extreme necessity. Late days will not apply to the final reports.

Software

You should use R programming language to do your final project, for the presentation you should use R Shiny.

Grading Policy

The final project grade will be based on the following components:

Criteria	Description	%	Points
Curiosity	• Ability to creatively apply old tools in new ways • The data or the analytical tool used is authentic • The goal of the research	30%	1.2
Skepticism	• Literature review • Appropriate visualization and usage of statistical analysis methods • Clarity of hypothesis • Questions and answers	40%	1.6
Organization	• Clarity and quality of R code • Clarity of summary, findings, recommendations • Presentation quality	30%	1.2

Note: Trying to avoid similar projects/themes, compare your topics with topics of people who have already decided on a project topic.