STA314H: Final Project Report Outline

This document provides a basic outline of your final report of the course project as well as 1-2 sentences suggesting what you are expected to include (feel free to include and explore more relevant things).

1. The **title** of your report shoud contain

   - the data set
   - team members with names and student IDs
   - Kaggle team name with your final ranking and prediction score

2. **Problem Statement**

   - Describe the specific problem or research question addressed by the project.
   - Explain the relevance and importance of the problems.

3. **Statistical analyses**: essentially, describe the statistical analysis conducted to address your identified research problems. The following are just a few examples.

   - Data preprocessing or feature engineering: steps for data cleaning, transformation, feature selection, and feature engineering techniques applied.
   - Your choice of machine learning algorithms, give a description of the model and your justification / reasoning of this model. How did you train the model? Did you consider other types of models? Is your model reasonable to use in your context?
   - Specify the metrics chosen to evaluate model performance (e.g., accuracy, precision, recall, F1 score) and explain their appropriateness.

4. **Results and conclusion**: your findings and conclusion of the specified project problems.

5. **Discussion** on any limitations of your work or potential future directions you feel inspired based on your project results

6. **References** such as papers, datasets, and tools that are crucially used in the project

7. **Code** to reproduce your results

The grades are roughly distributed as: item 2 (3 pts), item 3 (10 pts), item 4 (5 pts) and items 5-7 (2 pts).