[YOUR PROJECT TITLE]

DSCI 4780/6780 Final Project Report

*[Group members]*

*[Date]*

# Problem Description [Heading 1 - Font: Georgia 16 pt, bold]

*Briefly describe your project topic business or scientific problem. What is your motivation for selecting this project? How your dataset can be used? What are the important aspects of your analytics design. You are required to identify your target variable (or overall goal of your analytics solution) and talk about its characteristics. It should include the problem description, dataset, and solution subsections.*

**[Text body – Font: Times New Roman 12 pt, single spaced, no indentation for paragraphs, 0.1 inches spacing below the paragraphs] …**

## Problem [Heading 2 - Font: Georgia 14 pt, bold]

Lorem ipsum datum ….

## Dataset

Lorem ipsum datum ….

## Proposed Analytics Solution

Lorem ipsum datum ….

# Data Exploration and Preprocessing [Heading 1]

*The second part of your report will be related to the data understanding. Explore and explain your attributes. Provide bar charts or histograms when necessary (i.e., if you see a pattern after exploration). Next, provide a data quality report (See slides and textbook for creating a data quality report). Describe how you handle missing values and outliers. [Bonus: Describe your feature selection schema and later additional transformations.]*

## Data Quality Report

Lorem ipsum datum ….

## Missing Values and Outliers

Lorem ipsum datum ….

## Normalization

Lorem ipsum datum ….

## Feature Selection and Transformations [Optional]

Lorem ipsum datum ….

# Model Selection and Evaluation [Heading 1]

The third part will be about selecting models and evaluating them. First talk about your performance metrics and how they are related to your problem. Next, briefly describe your models, their input and outputs. Then, provide information on how you create your training and testing dataset and evaluate these models; show, using charts and confusion matrices, the performance comparison of your models. Based on your evaluation, pick a model for deployment. [BONUS: Perform hyper-parameter optimization and show improvement in model performance (5%).]

## Evaluation Metrics

Lorem ipsum datum ….

## Models

Lorem ipsum datum ….

## Sampling and Evaluation Settings

Lorem ipsum datum ….

## *Hyper-parameter Optimization* [Heading 3 - Optional]

Lorem ipsum datum ….

## Evaluation

Lorem ipsum datum ….

# Results and Conclusion [Heading 1]

*In this part, you need to communicate your results by analyzing the results from your analytics models and dataset. Provide a recommendation for the after interpreting your results.*