**Goldman Sachs Stocks**

**Executive Summary**

Students are required to complete a final project of their choice at the end of the program. They are paired up with fellow students and given six weeks to work on the project.

This document is dedicated to Desiree, Kori, Justin, Krista, Mary, and Christophe’s project. It will explain the purpose and scope for the project.

**Business Objectives**

To showcase the skills that Desiree, Kori, Justin, Krista, Mary, and Christophe have acquired through the Data Science program. They will be using R to wrangle, analyze, and visualize the “Goldman Sachs Stocks” dataset made available by Arpit Verma on Kaggle.

At the end of the project, Desiree, Kori, Justin, Krista, Mary, and Christophe should be able to explain their work, and present their findings to other students, faculty, staff, and potential employers, along with other interested parties via Zoom.

**Background**

In order to demonstrate what they have learned, they will be working together on a final project in order to put their newly learned skills into practice.

Desiree, Kori, Justin, Krista, Mary, and Christophe have chosen the “Goldman Sachs Stocks” dataset because they are both interested in financial stocks and savings.

**Scope**

Desiree, Kori, Justin, Krista, Mary, and Christophe will be using the software learned during the program to complete the project. They will be intentionally using tools of their interest.

**Functional requirements**

Data Wrangling: The dataset should be cleaned up for analyzing. Unusable columns should be removed. The datatypes for each column should be converted to a usable format for the needed analysis.

Data Analysis: Desiree, Kori, Justin, Krista, Mary, and Christophe will familiarize themselves with the dataset. They should have a good understanding of what each column means and the values of the data. Working together, they will brainstorm on questions and what they plan to gather from the dataset. Then, they will decide what models, predictions, etc to create.

Data Visualization: Once Desiree, Kori, Justin, Krista, Mary, and Christophe have a comprehensive understanding of information gathered from the dataset, they will work on visualizing the findings. They plan to use Tableau, and compile the visuals and texts in a Power Point slideshow.

Presentation: Working with the instructor, Desiree, Kori, Justin, Krista, Mary, and Christophe will schedule a time to present their findings via Zoom. They should be able to communicate in a clear and easy-to-understand manner. The presentation should be around 20 minutes. They should be dressed professionally for this occasion.

**Personnel requirements**

Desiree, Kori, Justin, Krista, Mary, and Christophe are the six developers. They will need to work closely for this project to succeed. They will touch base once a day via Slack to check in on work progresses and four times a week via Zoom to problem-solve. Once a week, they will plan out the next week. They will take turns being the scrum master, and report their progress to their instructor Joseph Raetano.

Once a week, they will meet with their instructor. They should be prepared to ask questions and seek guidance for the next steps.

They may also consult with their coding mentor, Margaret Martinez.

**Delivery schedule**

Week 1: Import dataset into R to begin data wrangling. Any unnecessary columns should be removed. Educate ourselves on Goldman Sachs Stocks. Set up Github.

Week 2: Study the dataset and ask questions. What stocks correlate with each other? Is the data normally distributed? What predictive models can we make? What interesting findings can we make?

Week 3: Modeling/Optimization (Boxplots, Histograms, Normal Distribution, Linear Regression, Correlation) and Machine Learning (Linear Discriminant Analysis)

Week 4: Review and validate findings from the previous week, and draw conclusions.

Week 5: Compile findings into a Power Point slideshow. Meet with the instructor and fellow students to ensure that the presentation is logical and understandable. Work on presentation style and layout.

Week 6: Make final touches to the Power Point presentation. Desiree, Kori, Justin, Krista, Mary, and Christophe should not attempt to come up with a brand-new analysis. There will not be enough time to verify their findings. They should practice presenting at least a couple times with the six of them, and at least once with their instructor.

**Other requirements**

All programs used should be free of charge. We do not plan on using paid services, but can if agreed upon by the group.

**Assumptions**

The software programs and platforms Desiree, Kori, Justin, Krista, Mary, and Christophe use should be available, up-to-date, and not broken.

**Limitations**

The only limitation that we may have is by not having everyone available or able to work on the project leaving us with a floater. There should not be a possibility of us being late on the presentation, but if there is a chance, we would get a hold of the teacher.

**Risks**

The instructor and mentor are there if we need any help from them. The risk of this project being incomplete is minimal. We will be successful in completing this project!