**Data Science Program Final Project**

**Executive Summary**

For my final project I’m going to use data to see if the same meteorite can be in a different part of the world

**Business Objectives**

To showcase the skills I have have acquired through the Data Science program. I will be using R, Python wrangle, analyze, and visualize the “meteorite data” dataset made available by UCI Machine Learning on Kaggle.

At the end of the project, I should be able to explain my work in layman’s term, and present their findings to the students, faculty, staff, and potential employers, along with other interested parties via Zoom.

**Background**

As a way to activate and put practical use to what the students have learned, doing a final project is a good way to demonstrate that.

I have chosen the “” dataset from NASA’S open data portal, because I am interested in aerospace and outer space.

**Scope**

I will be using the software taught in the program to complete the project. I will be intentional on using tools of their interest or tools that may aid finding a job.

**Functional requirements**

Data Wrangling: I will be cleaning up missing data, getting rid of string data, dropping missing and non needing data.

Data Analysis: For the analysis I will be doing T-tests

Data Visualization: I will be using graphs to show how many times a meteorite can be broken apart and be in two different parts of the world.

Presentation: I will be using slides to show my work and my findings. I will also be making sure that my presentation is easy to read and see.

**Personnel requirements**

Once a week, I will review the past week work load and plan out the next week. And report my progress to my instructor

Once a week, I will meet with my instructor. I will be prepared to ask questions and seek guidance for the next steps.

If I run into problems with coding I will seek help from mentor

**Delivery schedule**

Week 1: Import dataset into preferred software to begin data wrangling. Any unnecessary columns should be removed. Also Educate myself on meteorites and how they break up in earth's atmosphere.

Week 2: Study the dataset and ask questions. What are some possible correlations? Is the data normally distributed? What are some predictive models I can make from it? Visualize the data to see if there is any interesting findings.

Week 3: Modeling/Optimization (Combined Stepwise - Forward and Backward Selection) and Machine Learning (Random Forest.)

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

Week 5: Compile findings into a Power Point slideshow. Go over it with my instructor and friend/family member to ensure that the presentation is clear and logical. Work on the style and layout of the presentation so it is delightful on the eyes.

Week 6: Make final touches to the Power Point presentation. IF I do come up with a brand-new analysis. If there is enough time to verify my findings I will do it. I will practice my presentation a couple of times and with my instructor.

**Other requirements**

All programs used should be free of charge.

**Assumptions**

The software programs and platforms I will use should be available, up-to-date, and not broken.

**Limitations**

If something should come up for me during this six-week period, the project may be delayed. If the instructor or mentor have scheduled or unscheduled time-off, the project may be delayed as well. I may experience a roadblock in my work, which may push back the completion date.

**Risks**

The risks that may arise are such as natural disasters, power outages, family emergencies or broken software/hardware. I am eager to complete the program so there should be no motivation issues. The instructor and mentor are phenomenal so there is no concern of no help from them. The risk of this project being incomplete is minimal. I will be successful in completing this project!