6101 - Introduction to Data Science Midterm Project Proposal

Team Members

First Name	Last Name
Aaron	Yang
Jianjun	Gao
Luhuan	Wang

Possible Research Topics

The two main factors affect the final price of Uber and Lyft in Boston.

Smart Questions

- Is there any relationship between price and hour?
- Is there any other factors to affect the final price of Uber and Lyft?
- Do the weather factors, like temperature, wind, and sunset affect the final price?

Processes

1. Create a GitHub Repository;
2. Confirm an appropriate topic;
   - Understand the project purpose
   - Seek for a potential dataset/database/data sheet
   - Confirm the topic based on the dataset
3. Confirm the whole project processes;
4. Make a clear plan and timeline;

5. Prepare the first draft and slides;

6. Make a summary and final modification.

Possible Data

The data is from Kaggle.com. The data is about Uber and Lyft Dataset Boston, MA. This is a very beginner-friendly dataset. It does contain a lot of NA values. It is a good dataset if you want to use a Linear Regression Model to see the pattern between different predectors such as `hour` and `price`.

A really amazing part of this dataset is that I have included the corresponding `weather data` for that `hour` with a short summary of the `weather`. Other important factors are `temperature`, `wind`, and `sunset`.

The original data can be available on the Kaggle.com.

Link to GitHub Repo

DATS 6101- Midterm Project- Group 1