Springboard Data Science Capstone Project Proposal

Cecilia Lee

**Topic: An analysis of CitiBike sharing system**

Citi Bike is US's largest bike share program, with 10,000 bikes and 600 stations across Manhattan, Brooklyn, Queens and Jersey City. It was designed for quick trips with convenience in mind. However, bikes tend to pile up in some bike stations and run out in other bike stations, because of commuting patterns. In this project, we would like to find out the pattern of CitiBike users, and get insights about how to improve user experiences. Some particular goals of this project are as follows:

* Find out the behaviors of Citi Bikers riders? -When and Where do Citi Bikers ride? How far do they go?
* Find out the top 50 popular stations. Find out any insufficient station
* Predict the number of free bikes at each station in the future
* Understand how much does the weather influence the CitiBike usage

The client of this project will be the CitiBike system operating company and also the CitiBike users. For the system operating company, this project can help them make efficient operation decisions. For the bike users, they can know which station may have more available bikes in advance and help them decide their bike route and save their commute time.

The datasets for this project can be acquired from the CitiBike website, and also the weather data from Kaggle as fallows:

**CitiBike System Data:** <https://www.citibikenyc.com/system-data>

**NYC Weather Data:** <https://www.kaggle.com/mathijs/weather-data-in-new-york-city-2016>

The approach I planned to solve the problem will be finding out the behaviors of CitiBike riders of different age groups by examining the CitiBike Histories data set, eg., finding the most popular routes and stations for different age groups. Then, the weather data will be analyzed to see how the weather influences the CitiBike usage. Finally, by providing the weather and the age group, the possible bike routes and bike usage will be predicted.

I plan to use Python for this project, and put the project reports and codes and all relevant files on Github.