Behavior of Ride Sharing in Distinct Boston Areas and External Influences

**Team Members:** Jill Peloquin, Rick Soljour, Tova Donahue

**Project Description:** Ride sharing has become an increasingly important component of city transportation options. It is estimated that Uber and Lyft drivers take ~ $20 million dollars away from the MBTA in Boston and usage of ridesharing increased 25 percent between 2017 and 2018, with 81.3 million trips. Given that usage has been increasing over time and provides a viable occupation, as a team we want to analyze data from ride share companies to understand where in Boston rides are generated and under what circumstances.

**Research Questions to Answer:** We were able to find a data set from an approximately 3 week period of time (<https://www.kaggle.com/brllrb/uber-and-lyft-dataset-boston-ma>) from the Boston area in 2018 subdivided by district with price and climate-related data included. We are most interested in the location of ride request and price by district and the influences of weather (temperature, precipitation), day of the week, and time of day.

***We hope to use this analysis and produce a series of recommendations for a ridesharing driver; where should you be, at what time for what sort of weather when you are working in Boston?***

Key to this analysis is answering questions such as the following:

1. What districts of Boston had the highest density of ridesharing trips in Nov/Dec 2018?
2. What is the average price for trips by district?
3. Are request patterns for these districts influenced by factors like time of day and day of week?
4. Do ridesharing requests change with climate variability? Does temperature or precipitation affect requests or pricing?

Datasets to Be Used:

* Kaggle: <https://www.kaggle.com/brllrb/uber-and-lyft-dataset-boston-ma>)
* Open Weather (potentially)

General Tasks:

* Import Data – Rick
* Clean Data Set – Jill
* Big Picture Summary / General Analysis of data set
* Detailed Graphs using MatPlotLib – Tova
* Statistical / Correlation Analysis
* Summarizing Write Up for series of recommendations.
* Presentation organization

Brainstorming Sesssion

Write this as a brief summary of your interests and intent, including:

* The kind of data you'd like to work with/field you're interested in (e.g., geodata, weather data, etc.)
  + As a group we wanted to continue our work with weather data,. We also have an Uber driver on our team, so there is a personal interest component as well.
* The kinds of questions you'll be asking of that data
  + Does extreme weather make for longer / shorter rides?
  + Daily temps correlation to rides
  + What ride share company generates the most/least rides from different geographical areas in Boston
  + How is ride sharing distributed in Boston – Heat Map density
    - Let’s us drill down to top neighborhoods to focus on
  + Peak times for ride requests
  + Does climate effect number of rides requested
  + Day of the week distribution
  + Correlation to temp and origin /destination
  + Price per mile
* Possible source for such data
  + Kaggle
  + Public Weather API