**Data Set Project**

**Project Title:** Point A to B

**Team Members:** Steven Wang  
Jenny Armour  
Jana Simashkevich  
Depa Ramcharran

**Project Summary:**

Analyze the impact of weather on Toronto Bike Share usage over a four years period between 2014-2017

**Research Questions:**

1. How do temperatures and precipitation levels impact bike share usage for the members
2. How do temperatures and precipitation levels impact bike share usage for non-members
3. How do temperatures and precipitation levels impact bike share trip distance for non-members
4. How do temperatures and precipitation levels impact bike share trip distance for non-members
5. How are locations impacted, does core downtown vs further away impacted the same way?
6. Research bike lane rollout and tourism numbers for each of the years to see if this has additional impact on bike usage

**Datasets to be used:**

Bike Share Data: <https://open.toronto.ca/dataset/bike-share-toronto-ridership-data/>  
Open Weather Map: <https://openweathermap.org/api>  
Google Geolocation API: <https://developers.google.com/maps/documentation/geolocation/intro>

**Rough Break Down of Tasks:**

1. Import Bike Share datasets (csv files)

Available Datapoints: date, time, pickup location, drop off location, trip duration, user type

1. Get historical weather temperatures and precipitation details for Toronto over 2014-2017 period from the Open Weather
2. For all unique pick up / drop off locations from the Bike Share data set use geolocation API to determine the latitude and longitude and calculate trip distance, store this data in CSV file for reference to avoid calling API’s
3. Analyze the data and check the data quality
4. Clean the data as needed
5. Think of usage patterns:
   1. Trips shorter than X minutes, do we need to keep those? Did the person change their mind or are these errors?
   2. Pick up / drop off location is same vs different locations? Does it tell as anything about commute vs leisure trips?
   3. Can we bucket the usage based on time?
6. Create visualizations to analyze the following patterns
   1. Temperatures vs Number of Trips (for members vs non-members)
   2. Precipitation levels vs Number of Trips (for members vs non-members)
   3. Temperatures vs Trip Durations (for members vs non-members)
   4. Precipitation levels vs Trip Durations (for members vs non-members)
   5. Create location heatmaps per season to see the impact of weather in different locations