**Project Proposal**

**Big Idea / Question:** What are good indicators/predictors of crime?

In this project, we will be focusing particularly on the geographic location of New York City. We will attempt to assess how potential factors (e.g. ridesharing pickup/dropoff data, time of day, day of week, housing data) may be related to the frequency of crimes in this area.

For example, we hypothesize that ridership and use of ridesharing applications is likely more frequent in areas of high crime. This can be attributed to safety concerns that individuals may have when using public transportation or walking from location to location. With the rise of apps like Uber or Lyft, we believe that individuals may turn to these alternative means of transportation.

**Visualizations/Methodology**

1. Crime Data & Ride Sharing: Visualization that shows where the density of crimes are located in the NYC area segmented by borough, census tract, or neighborhood, with an overlay of data points that display the location of rider pick-up and drop-off locations or density of these pick-up drop-off locations.
2. Time of Day: Animated map visualization that shows how crime data varies based on across time of day.
3. Day of Week: Animated map visualization that shows how crime data varies based on day of the week.

If possible, we will also try to run statistical t-tests to assess whether the relationship between these variables are statistically significant.

**Data Sources**

This is one potential directory we may use to assess ridesharing data in NYC. The data set contains information on over 4.5 million Uber pickups in New York City from April to September 2014, and 14.3 million more Uber pickups from January to June 2015. Our analysis would be confined to this time period.

* [**https://www.kaggle.com/fivethirtyeight/uber-pickups-in-new-york-city**](https://www.kaggle.com/fivethirtyeight/uber-pickups-in-new-york-city)

This dataset includes all valid felony, misdemeanor, and violation crimes reported to the New York City Police Department (NYPD) from 2006 to 2017. We would likely filter this data set so that the time frame analyzed corresponds more closely to the ridesharing data when assessing the relationship between these two data points. This data set also includes relevant information on time of day and day of week.

* [**https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Historic/ qgea-i56i/data**](https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Historic/qgea-i56i/data)

Incorporating the crime data from NYPD, we would examine the hotspot areas of higher crimes and examine if there was a relationship to housing value. The data would be confined to present-day values according to specific boroughs in NYC.

* [**https://a836-pts-access.nyc.gov/care/search/CommonSearch.aspx?mode= ADDRESS**](https://a836-pts-access.nyc.gov/care/search/CommonSearch.aspx?mode=ADDRESS)