# CIS 9650 – Term Project Proposal

### Team 2:

Cynthia Rojas – 14101049

Pongyin Hui – 23054197

Leo Cohen Gussack – 23976155

Joann Lin – 16006991

### Problem Statement:

Our project will be focused on analyzing crime data in the five boroughs of New York City (NYC) from 2018 and 2019. Unfortunately, under the current political and social climate, this topic is increasingly relevant as we see crime rates increase in NYC. As many of us are from New York and commuting throughout the city either for work or school, we find this topic both interesting and relevant to our daily lives.

Our hypothesis is that there is a negative correlation between crime and foot traffic.  We believe that both the prevalence and severity of crimes in NYC is directly correlated to the amount of “traffic” in a specific region.  This may be due to the number of potential witnesses available. In order to select the areas of higher traffic, we decided to analyze the locations of public Wi-Fi and streetlights, both of which would draw more people traffic to that area. We expect to see less crime in an area that experiences more Wi-Fi connectivity and has more streetlights than in areas without.  In order to test our hypothesis, we will be comparing data and creating charts/graphs to illustrate our findings.

### Data Sets Used:

We will be using 3 sets of data to test our theory.

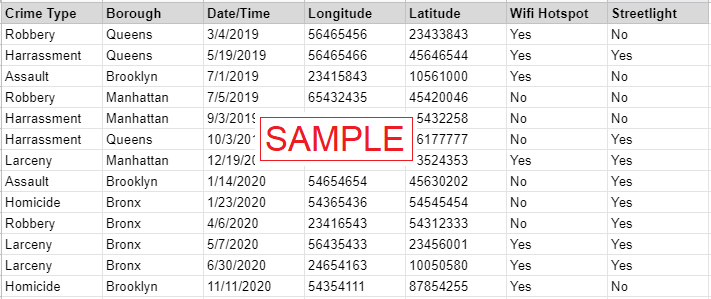
1. NYPD Complaint Data:
   * <https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Historic/qgea-i56i/data>
   * This data set includes crime reports for the 5 boroughs of NYC.  The data notes 4 key points that we will be using in our project - Crime Type, Crime Location, Borough, and Date/Time.
2. NYC Wi-Fi Hotspot Locations:
   * <https://data.cityofnewyork.us/City-Government/NYC-Wi-Fi-Hotspot-Locations/yjub-udmw/data>
   * Includes locations of publicly available Wi-Fi.
3. NYC Mobile Telecommunications:
   * <https://data.cityofnewyork.us/City-Government/Mobile-Telecommunications-Franchise-Pole-Reservati/tbgj-tdd6/data>
   * Includes locations of streetlights.

### How Datasets Will Be Used:

We will analyze both Wi-Fi hotspots and streetlight locations.  Pros and cons of both datasets below:

|  |  |  |
| --- | --- | --- |
| **Dataset** | **Pros** | **Cons** |
| Wi-Fi | All Wi-Fi hotspots available | Unsure of strength of correlation by itself |
| Streetlights | Seems like it may be a stronger correlation, even on its own | Only streetlights used by telecommunication companies are available |

We are looking to combine the 3 datasets in order to analyze the correlation (if any).  Sample combined data below.



### Analysis proposals:

1. Correlation chart between crime occurrence and Wi-Fi / streetlight locations.
2. Calculate rate of crime reduction / increase (e.g., an additional streetlight results in a 0.9% reduction in violent crime).
3. Specifically focus on violent crime occurrences when witnesses matter most.
4. Break down of crime between night and day (maybe Wi-Fi during daytime, and streetlight during nighttime).

### Expected Result:

We hope to have a better understanding of ways cities can reduce crimes.  If our hypothesis is correct, the simple solutions of adding more streetlights and providing more public Wi-Fi hotspots could result in a safer city for us and our fellow New Yorkers.