Exploratory Data Analysis of 2020 NYPD Crimes and NYC Subway Stations
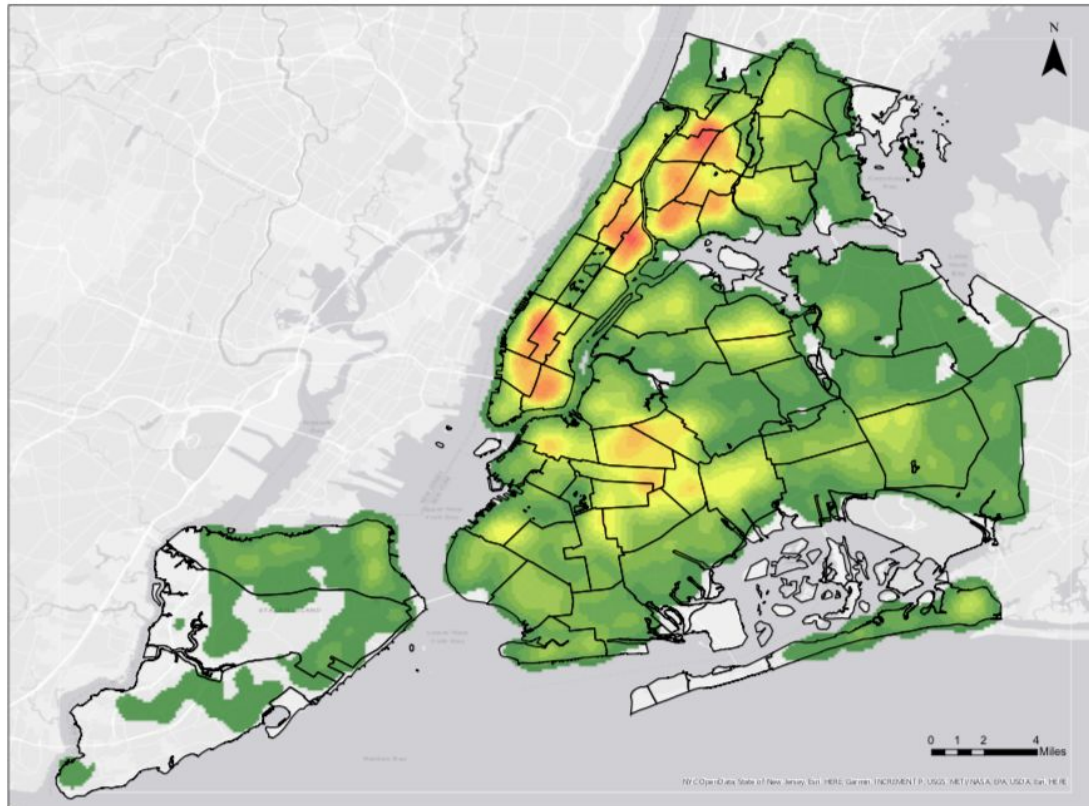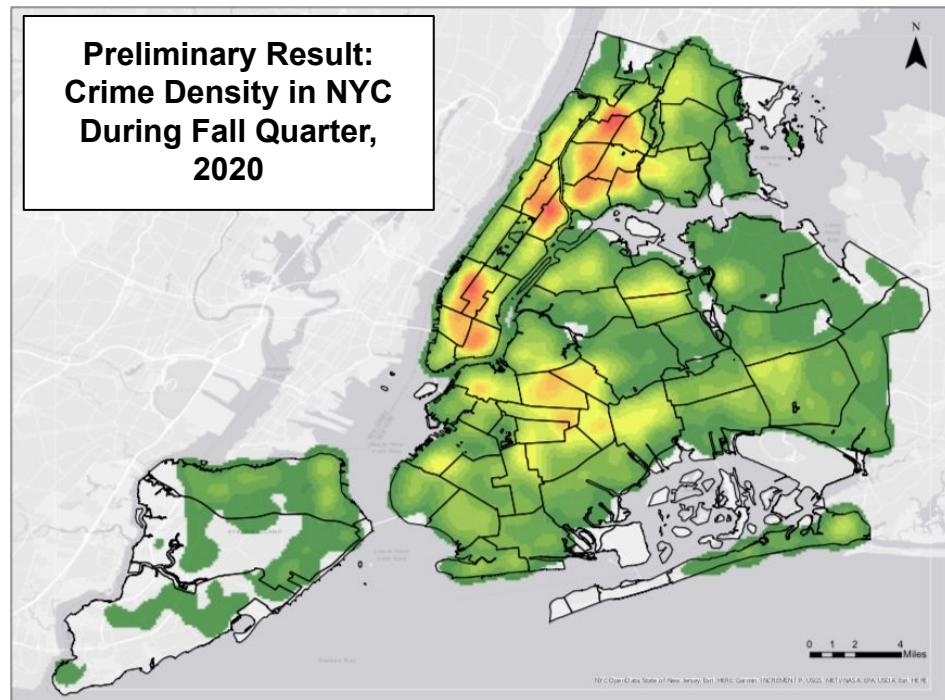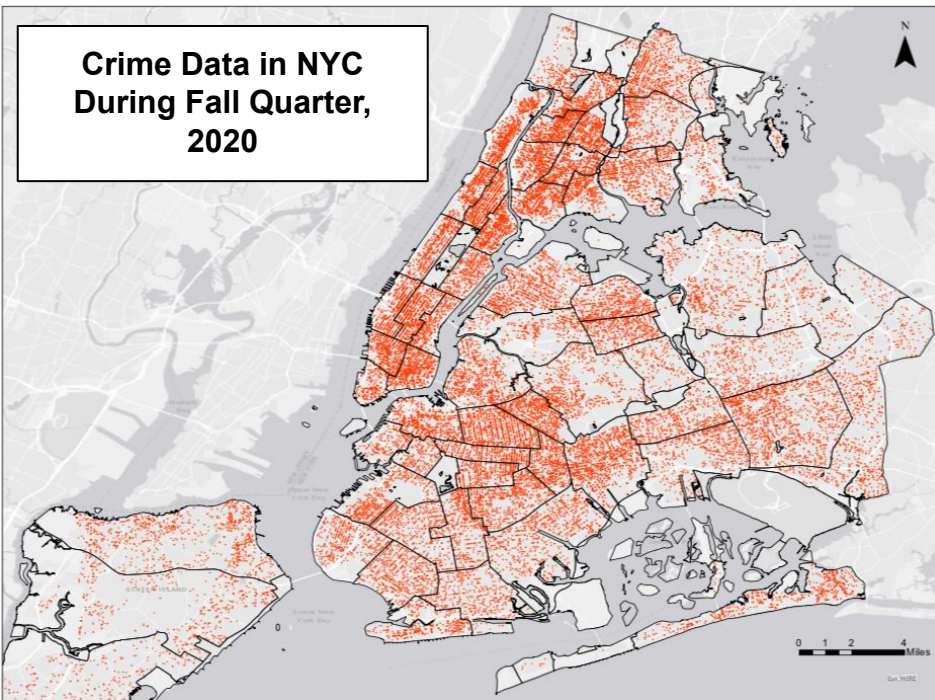

Rachel Hausmann

Introduction

- Context: NYC has an investment in protecting tourists traveling through the city.

- Objective: Provide the City of New York with an analysis of where there is a correlation between high foot traffic and crime density.

- Goal: Provide the City of New York with specific subway stations that may benefit from an increase of police presence or other means of crime reduction- particularly heavily traveled stations

Map of the 25 Most Heavily Used Subway Stations for Fall Quarter, 2020
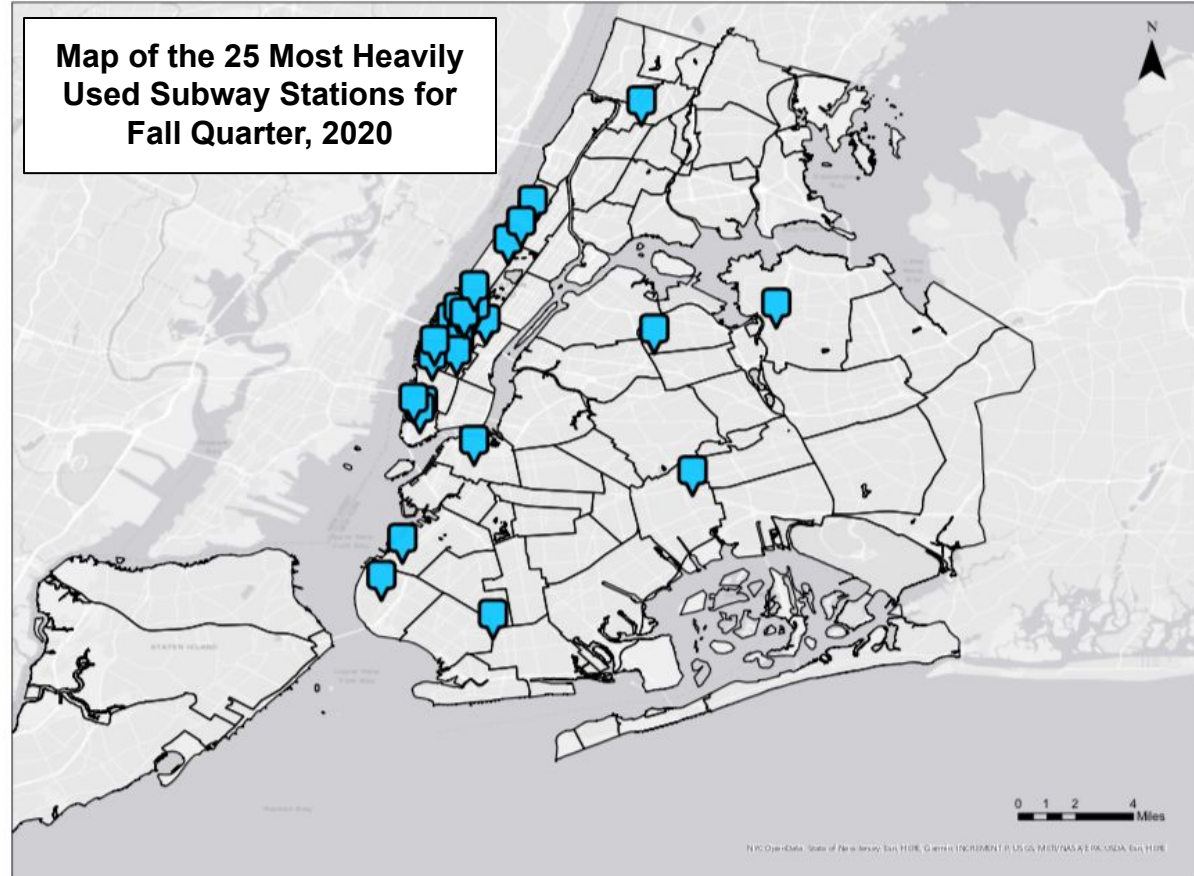
Methodology

- Clean and aggregate NYC MTA and NYPD Crime data
- Merge the two datasets
- Run density analysis of crimes and subway stations via a spatial buffer join

Crime Data in NYC During Fall Quarter, 2020

Preliminary Result: Crime Density in NYC During Fall Quarter, 2020

** Crime Data in NYC During Fall Quarter 2020 consists of reported crimes and mta data between August 8, 2020 and December 26, 2020
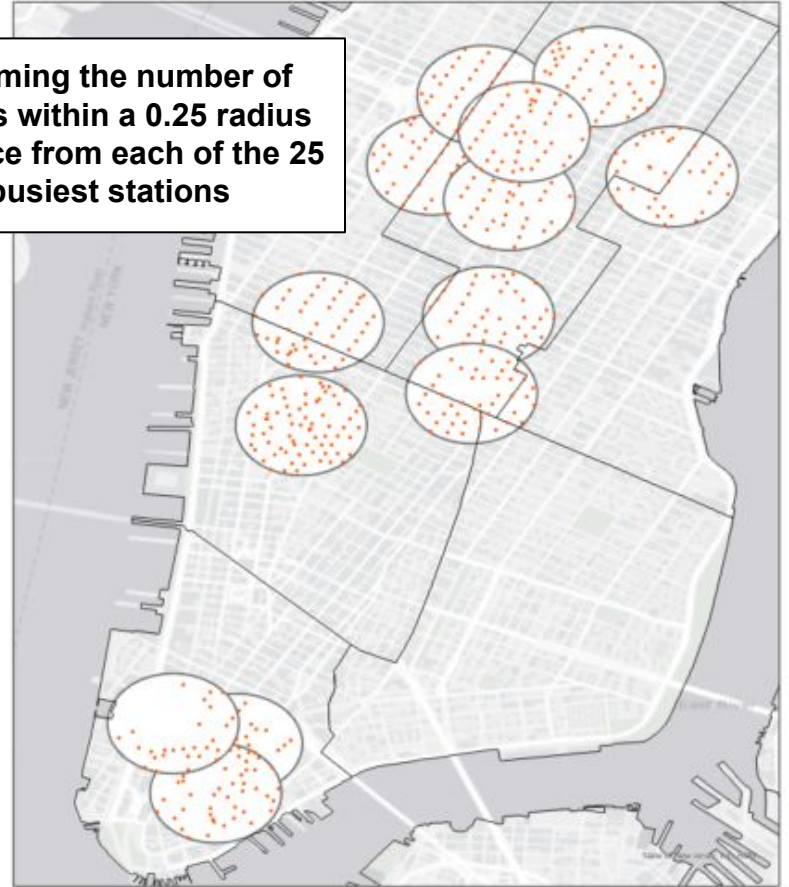
Spatial joining toolbox:
<u>Buffer</u>, Overlay, Clip

0.5 miles

Map of the 25 Most Heavily Used Subway Stations for Fall Quarter, 2020
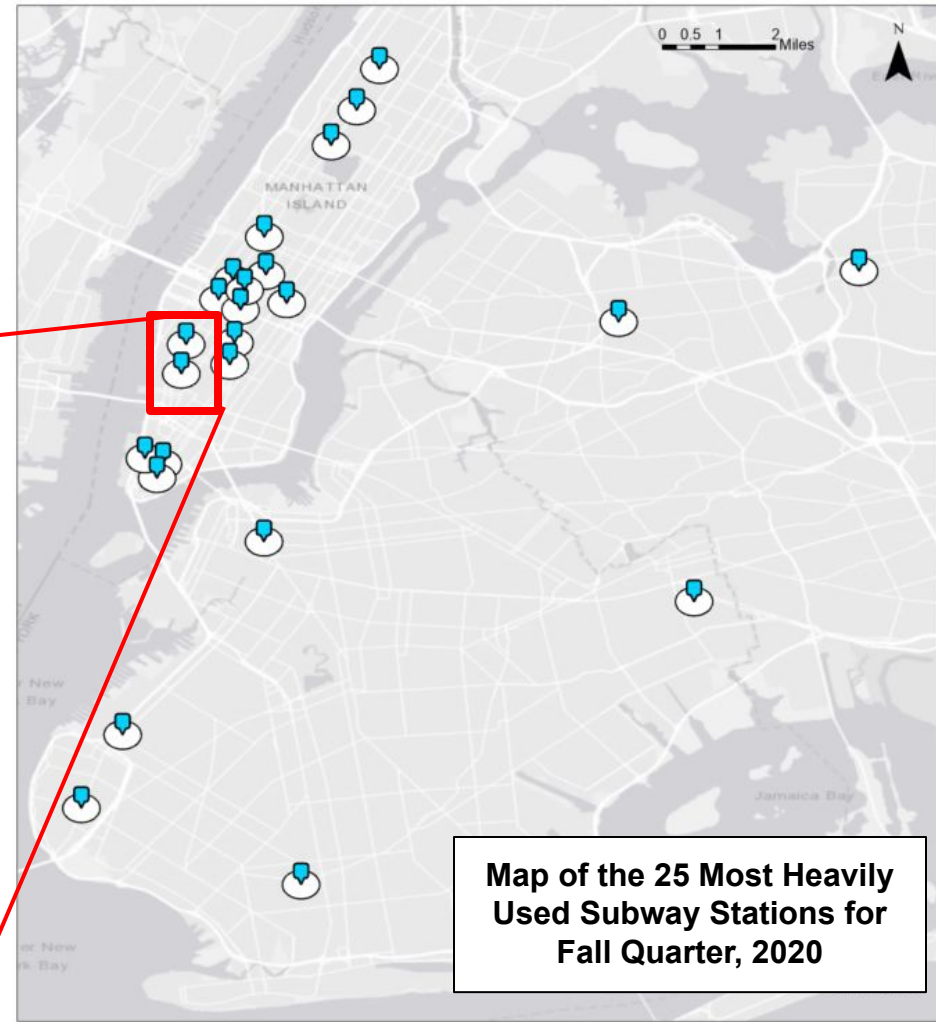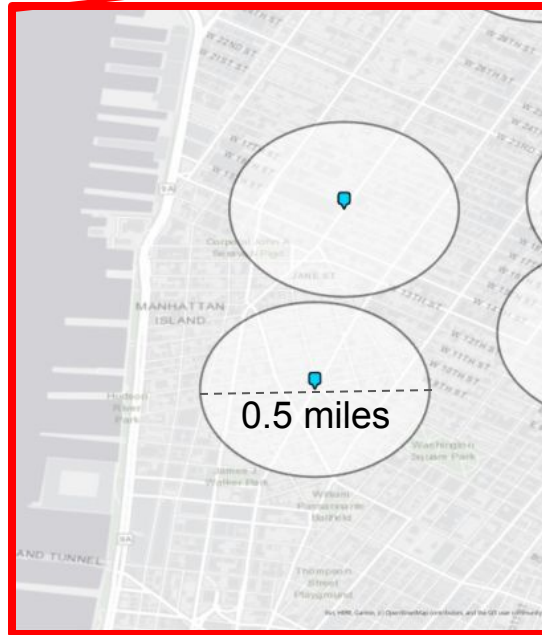
# Spatial joining toolbox: Buffer, **<u>Overlay</u>**, **<u>Clip</u>**


Summing the number of crimes within a 0.25 radius distance from each of the 25 busiest stations

Results

Of the 25 busiest NYC stations, a higher density of crimes surround 3 of the top 5 stations with the highest number of daily entries: Penn Station, Fordham, and Port Authority
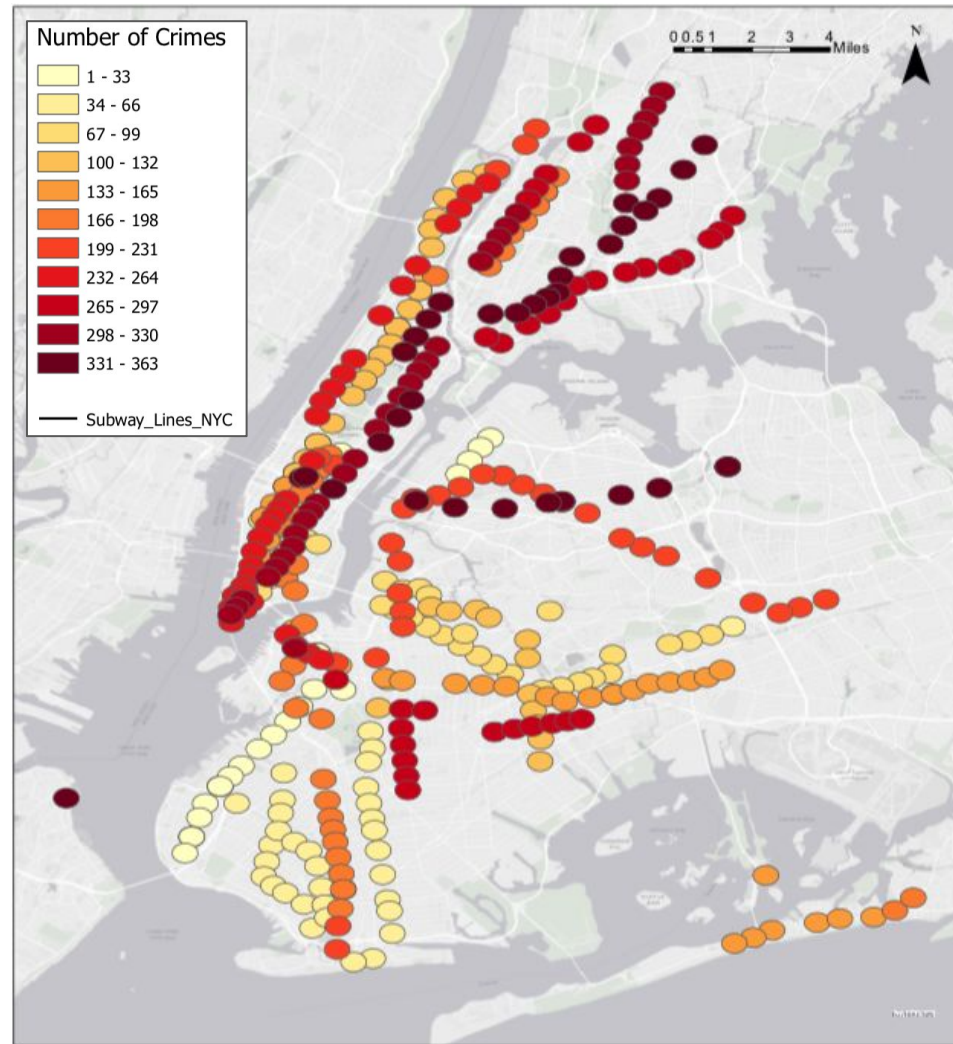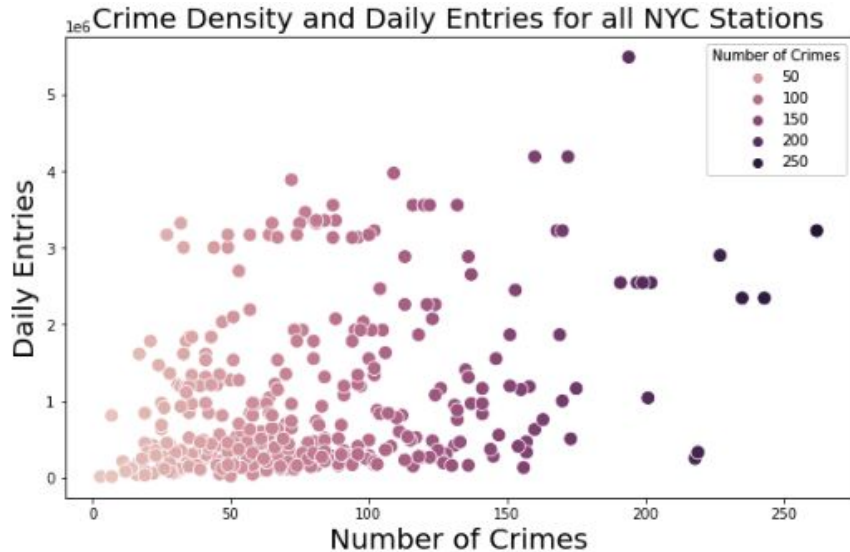
No correlation between foot traffic and crime density.

Correlation between Crime Density and Daily Entries
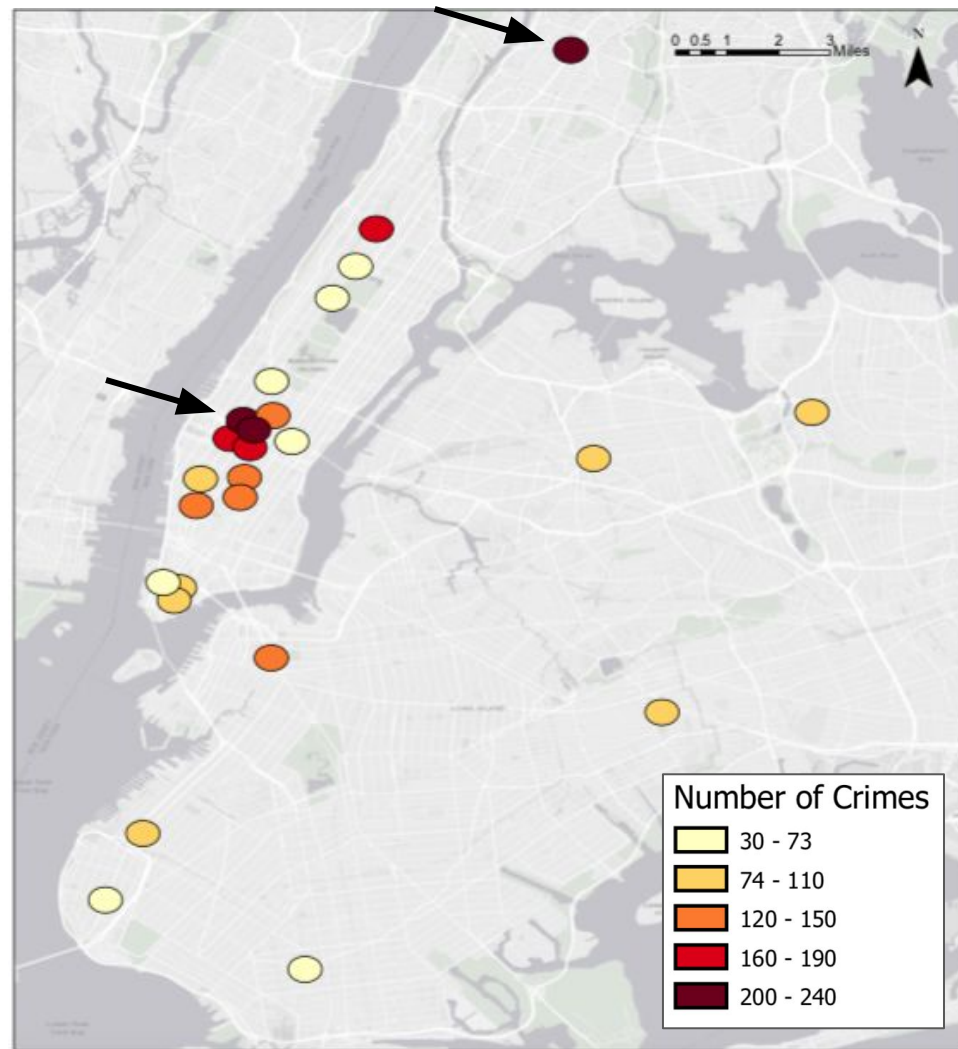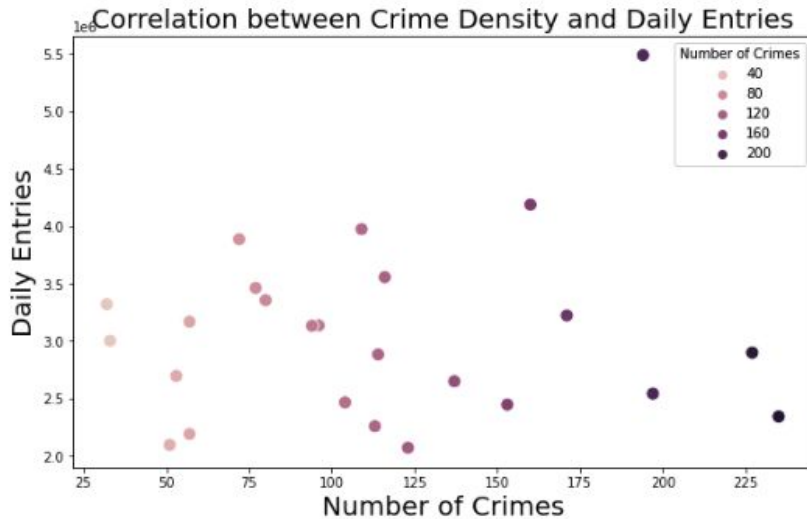
Conclusion & Future Work

Possible connection between subway lines and density of crimes along that particular line.

Weak positive correlation between daily foot traffic and crime density.

Crime Density and Daily Entries for all NYC Stations

Conclusion & Future Work

Possible connection between density of crimes and subway lines.

Are certain subway lines "safer" than others?