Evaluating Vulnerable Areas of New York City: Data

Jack Breingan

25/08/2020

1 Data

The data in this NYC Open Data (`https://data.cityofnewyork.us`). Specifically, we used the NYPD Complaint Data found here `https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Historic/qgea-i56i` for the bulk of our data, and the geospatial data on New York's neighbourhoods found here `https://data.cityofnewyork.us/City-Government/Neighborhood-Names-GIS/99bc-9p23` for our spatial analysis.

We have used the most recent data available at the time of writing, August 2020, and confined ourselves to a one year period - that is, we have used data from 2019. Crime data changes significantly year by year, as police and other agencies act to suppress it and offenders concentrate around new methods, so earlier data is not likely to be as relevant, although certainly some kinds of crime, particularly domestic, are likely to remain in similar places with similar trends.

Our location data contained the name of each neighbourhood in NYC, the corresponding borough it was part of, and the latitude/longitude coordinates of the neighbourhood.

The crime data provided locations, dates, times and types of crimes. This data contained the following useful information:

1. CMPLNT_NUM: Randomly generated persistent ID for each complaint
2. CMPLNT_FR_DT: Exact date of occurrence for the reported event (or starting date of occurrence, if CMPLNT_TO_DT exists)
3. CMPLNT_TO_DT: Ending date of occurrence for the reported event
4. CMPLNT_FR_TM: Exact time of occurrence for the reported event (or starting time of occurrence, if CMPLNT_TO_TM exists)
5. OFNS_DESC: Description of Offense
6. PD_DESC: More granular description of Offense

7. CRM_ATPT_CPTD_CD: Indicates whether the crime was successfully completed
8. LAW_CAT_CD: Level of offense: felony, misdemeanor or violation
9. LOC_OF_OCCUR_DESC: Where a crime occurred in relation to the premises - in front of, behind, etc
10. BORO_NM: The name of the borough in which the incident occurred
11. JURISDICTION_CODE: Jurisdiction responsible for incident. Either internal, like Police(0), Transit(1), and Housing(2); or external(3), like Correction, Port Authority, etc.
12. ADDR_PCT_CD: The precinct in which the incident occurred
13. PREM_TYP_DESC: Specific description of premises; grocery store, residence, street, etc.
14. VIC_AGE_GROUP: Victims Age Group
15. VIC_RACE: Victims Race Description
16. VIC_SEX: Victims Sex Description
17. Latitude of incident
18. Longitude of incident

We combined this with our location data, calculating which neighbourhood centre was closest to each crime, to associate each incident with a neighbourhood, then evaluated crimes on a neighbourhood level. We also examine some trends across the city as a whole. Using Foursquare, we examined the kind of locations prominent in these neighbourhoods, to get an idea of the most at-risk types of premises, and whether the crimes reflected the general trend of locations present or disproportionatley target something specific. This data allowed us to make profiles of the types of crimes that affect each neighbourhood, and also who they targeted. We were also able to make observations about the times crimes were likely to occur.

Using this data, we outline the most vulnerable to crime neighbourhoods in New York, and the types of crime which are prominent there. This data provides advice for anyone looking to handle these offences, whether government, law enforcement, or private businesses considering new locations.