Crime in Toronto

Carlo Galli

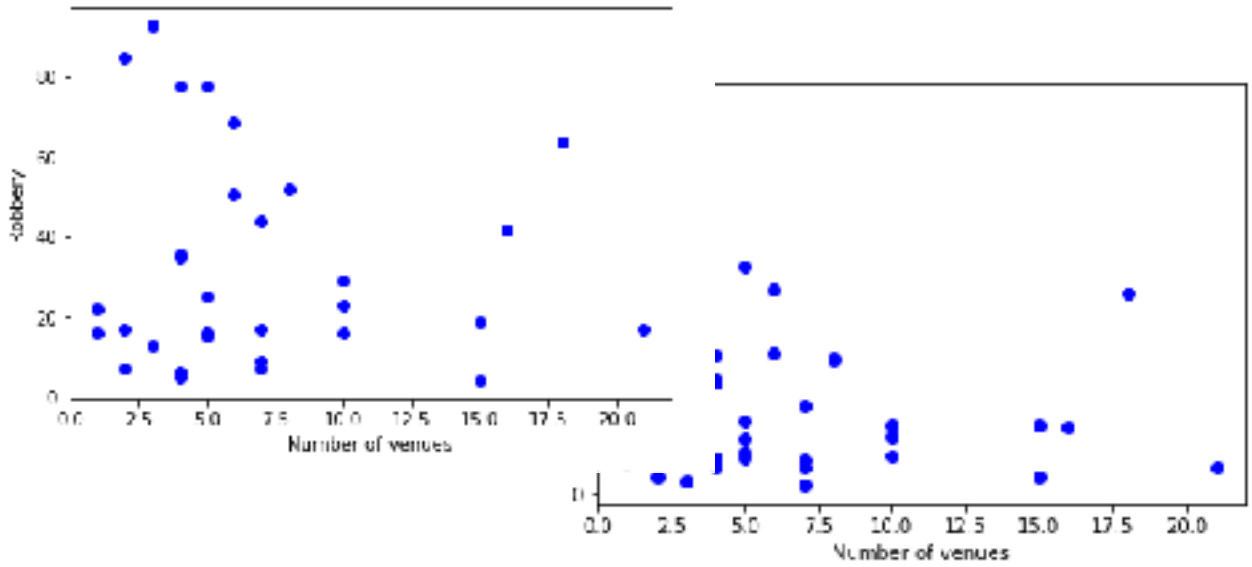
November 2019

Background

Crimes often concentrate in specific areas or neighborhoods of a city, which, for specific socio-economic factors, may be more conductive to illicit activities. The availability of geo-spatial data allows for computational analysis of such factors and a deeper understanding of patterns that most often associate with increasing crime rates. Such data is of interest to policy-makers who may have to plan and promote the economic development of neighborhoods or design specific interventions aiming at decreasing crime rates indirectly by acting on the social fabric of poor or depressed areas. The purpose of the present project is to assess whether criminal activities correlate with the presence of a specific social pattern, as it can be inferred from the commercial activities on the territory of the city of Toronto. By developing a model or highlighting specific associations, this analysis may help local administration to better plan the city development.

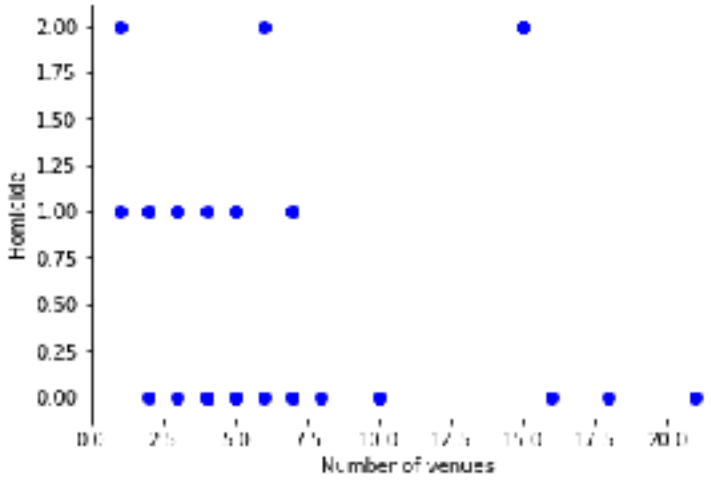
Data

To carry out this analysis, I relied on geo-spatial data on the city and neighborhoods of Toronto provided by Foursquare and as used in our previous assignments in this course. These data comprise the number and type of commercial activity divided by neighborhood in Toronto. Furthermore I relied on the crime data provided by the Toronto Police and publicly available online as a csv file at http://data.torontopolice.on.ca/datasets/neighbourhood-crime-rates-boundary-file-/data. These data comprise the number of event classified as Assault, Break and Enter, Auto Theft, Robbery and Homicide during the years 2014-2018. To carry out the project I investigated what features in the neighborhood (number of stores, restaurant etc) best correlate with the criminal activity, if any and assess whether it is possible to predict a relation between the two, using multiple linear and non-linear regression.



Methodology

To analyse the data, I started from the database of Toronto neighbourhoods that was elaborated for our previous assignment in Coursera and I merged it with the crime database provided by the Toronto Police. I merged the two databases obtaining a pandas data frame containing neighbourhoods and crime data associated to specific postal codes, which I then geographically located using a database of geospatial data, which we were previously provided by coursera. Using these longitudinal and latitudinal data I obtained venues data from Foursquare, connecting through the API, using my credentials. Grouping by venues I obtained the number of venues per neighbourhood and I was able to start analysing the relation between venues and crime. The crime data I decided to include in my analysis were Assault, Auto theft, Break & Enter, Robbery, Theft and Homicide. I decided to include only the number of crime events for the year 2018 and exclude the previous year stats.



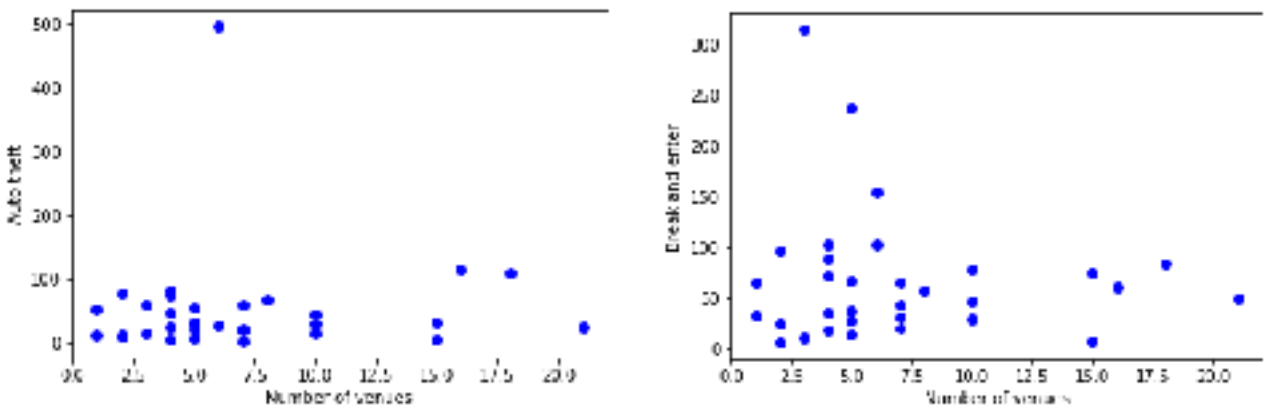
Results

A first view of a scatter plot of venues and number of homicides reveals how the number of homicides in Toronto in 2018 was generally low, with a maximum number of 2

homicides in 3 neighbourhoods. As these were very few I decided not to consider them in the following analysis. Although 0 homicides were recorded in several neighbourhoods regardless of the number of venues present there (mean=7.45 venues), all the neighborhoods where 1 homicide occurred had less than 7.5 venues (mean=3.7). Actually, the difference in the mean amount of venues had p=0.05. This suggests that there may be safe neighbourhoods with fewer venues or more numerous venues, but most homicides happen in neighbourhoods with fewer venues.

Similarly, assaults and robberies appeared to have some correlation with the number of venues, as most assaults occurred in neighbourhoods lacking a high number of venues.

By comparison, a scarce correlation could be found between venues and Break and enter crimes and auto thefts did not appear to correlate with the number of venues.



Discussion

As crime represents a serious challenge for local administrators, it is reasonable to wonder what the best ways to face this challenge are. The first and foremost strategy is maintaining the rule of the law via policing. However, an alternative approach may be represented by supporting those social activities that are conductive to the development of the neighbourhood and may inhibit illicit activities. One of such approaches may actually be promoting the establishment of new venues in a neighbourhood may be beneficial to the social fabric of that neighbourhood and help reduce crime or rather if new businesses and venues may actually be victim of crime and be better established elsewhere. This preliminary analysis only investigated the association between factors and can not establish any cause-effect relation, but its results are suggestive that the

number of venues in a given neighbourhood do correlate with some crime typologies. More specifically, although auto theft seem independent of the composition of nature of a neighbourhood, crimes such as Assaults, Homicides and Robberies mostly concentrate in neighbourhoods where the number of venues are low. This may be easily interpreted as a confirmation that the availability of commercial businesses and socially active areas represent a way to discourage criminal activity, possibly by the mere presence of people in the area. Though this cannot be sufficient to restore legality in an affected area, in the long run it may reinforce the social fabric of an area, making it less susceptible to fall prey to crime.