# Crime in the city:
# Spatial planning for crime reduction

Ruben Sharpe

March 16, 2021

## Contents



## 1 Introduction

It is intuitively clear that public venues, such as bars and restaurants, but also public service locations such as transport hubs or banks impact the functioning of local communities. However, it is not straightforward to predict whether this impact will be positive or negative.Bars and restaurants can enliven a neighbourhood and may contribute to greater safety because of increased presence of people in the streets as potential witnesses deter crime. On the other hand, the presence of bars and restaurants is also correlated with alcohol abuse related crime. Similarly, there is a duality in the social impact for the presence of shops and banks. These may have a positive impact in the daytime, but may also attract criminal activity at night because of reduced traffic compared to residential areas.

The complexity and the interrelatedness of these issues carries over to the complexity of city planning. The difficulty for city planners, moreover, is that

they rarely get to design a neighbourhood from scratch. Typically, they have to deal with neighbourhoods with a mostly fixed set of public facilities, and which may be suboptimal from a design point of view. The question for a city planner, therefore, is often : "What would be, *for a certain neighbourhood* and given all its other attributes, the impact of building (or removing) a shopping mall? Or a park? Or adding, or removing a bus stop?"

There is, of course, a lot of scientific literature on this topic but that is out of the scope of this project. This report is aimed at city planners and is intended to show the feasibility of machine learning as a tool for spatial planning.
Machine learning allowes an automated approach to compare neighbourhoods both in terms of their public venues and in terms of social markers, such as crime rates. This provides city planners, apart from their professional expertise, with examples of good and bad practices that may help them in their decision making. For reasons of convenience, the analysis in this report will look at the city of Toronto (Ontario, Canada). Crime rates will be used as the proxy for the health status of communities, although other markers such as employment rate or physical health could also be used.

# 2 Data

Location data, information about public venues and crime statistics will be used to classify neighbourhoods (k-means classifier) and map them with respect to the prevalance of crime (choropleth). From this, generalised conclusions will be drawn with respect to the makeup of 'high-crime' and 'low-crim' neighbourhoods.

## 2.1 Location data

Location data for the city of Toronto can be acquired by scraping postal codes, boroughs and neighbourhoods from Wikipedia.[1] This data needs to be enriched with the GPS coordinates of the respective neighbourhoods (Table 1). These coordinates can be obtained from the internet in the form of a '.csv' file from the Coursera website for the Data Science Capstone.[2]

| |
|---|
| Postal Code |
| Borough |
| Neighbourhood |
| Latitude |
| Longitude |

Table 1: Toronto location data (103 postal codes for 140 neighbourhoods in 11 boroughs).

[1] https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
[2] https://cocl.us/Geospatial_data

## 2.2 Public venues

Foursquare is used to obtain information about the public venues, specifically their category, in each neighbourhood (Table 2). For this purpuse, the neighbourhood is arbitrarily defined as the area within a 1.500 m radius around a postal code's geographical centre. The top 10 most prevalent venues will be used to characterise the neighbourhood.

| |
|---|
| Neighbourhood |
| Neighbourhood Latitude |
| Neighbourhood Longitude |
| Venue |
| Venue Latitude |
| Venue Longitude |
| Venue Category |

Table 2: Venue data (6884 venues, 349 unique venue categories for 140 neighbourhoods in 99 postal code areas).

## 2.3 Crime data

The city of Toronto provides a lot of open data related to the city.[3] This data ranges from polls conducted by the government, to inventory lists of street furniture, to 'bicycle count and locations' and is ever increasing. From this repository, a list of crime rates can be obtained.[4] This list contains, for each neighbourhood the number of incidences of certain types of crime (Table 3). Population information from the 2016 census, included in the file, can be used to normalise crime rates relative to population size. Conveniently, the file also contains geographical information so that this information can be easily mapped.

| |
|---|
| Assault |
| Auto Theft |
| Breaking and Entering |
| Homicide |
| Robbery |
| Theft |

Table 3: Crime data categories. Data is available for each year from 2014–2019.

[3] https://www.toronto.ca/city-government/data-research-maps/open-data/
[4] https://open.toronto.ca/dataset/neighbourhood-crime-rates/