The Battle of Neighborhoods

Description of the Data

In this section each data source is briefly described. For the evaluation of the City of Toronto there are the following data sources used.

1.1 Dataset for neighborhood boundaries

For a good estimation of the data it is helpful to visualize the data and the results with a choropleth map. For this map type a geometry file is necessary.

The City of Toronto operates a data portal which provides such a file:
https://open.toronto.ca/dataset/neighbourhoods

The second important use is to compute the land area from the boundaries data. The data is given in the World Geodetic System 1984 (WGS84).

For a direct access the data was downloaded and was placed on github:
https://github.com/DirkFritz/Coursera_Capstone/blob/master/Neighbourhoods.geojson

IMPORTANT: For the use of the file with jubyter and folium the browser Firefox is necessary. If you uses chrome the map wouldn't plot. The file size is seemingly to big.

1.2 Dataset for Crime Statistics

The Police of Toronto has public safety data portal. The service provides a detailed listing for all crime types for each neighborhoods which happens in the last years.

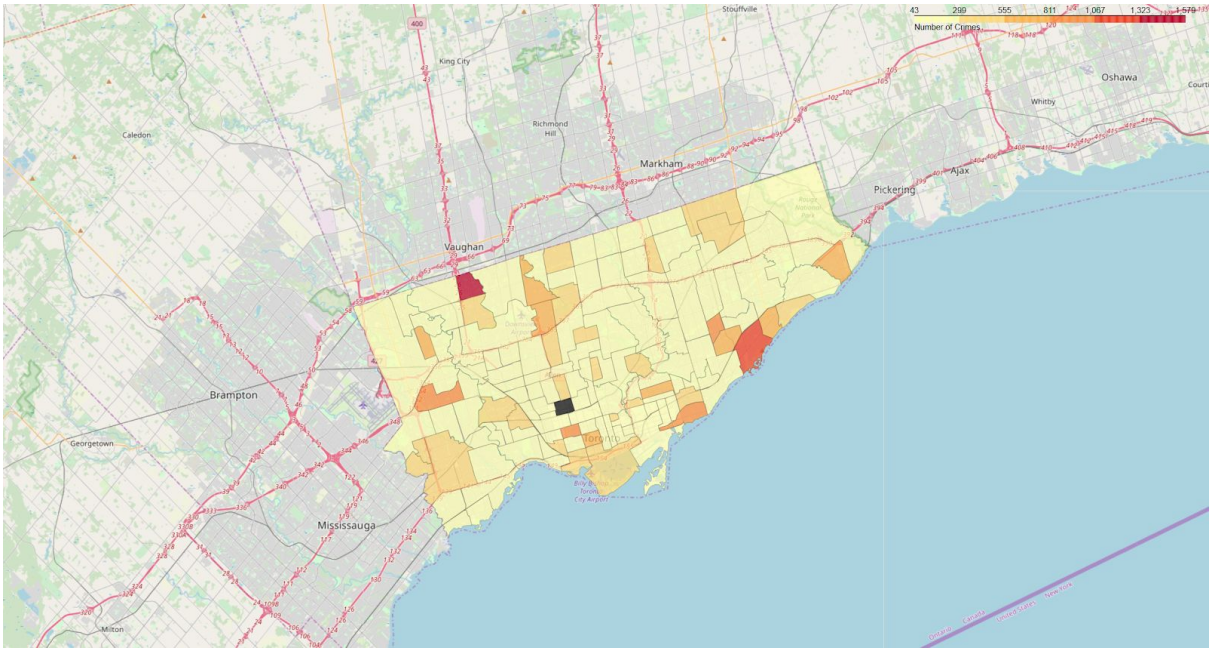
The link to the data portal is:
https://data.torontopolice.on.ca/datasets/neighbourhood-crime-rates-boundary-file-/data?geometry=-80.334%2C43.526%2C-78.479%2C43.874

The provided dataset has the following crime types:

1. Assaults

2. Auto Thefts
3. Breaks and Enters
4. Robberies
5. Theft-Overs
6. Homicide

For the evaluation the sum of all crime types is used. The download of the dataset can accessed over the following link:
https://github.com/DirkFritz/Coursera_Capstone/blob/master/Neighbourhood_Crime_Rates_Boundary_File.csv

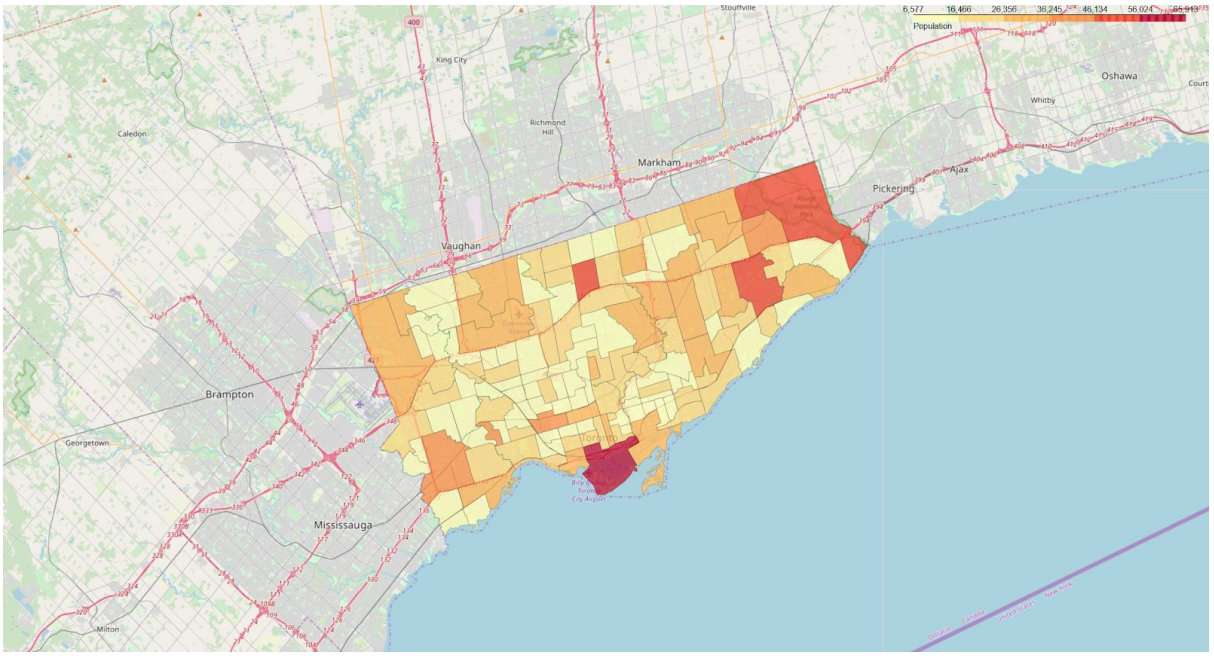


Map with the overall number of crimes

1.3 Dataset for the Neighborhood profile

For the criteria population there is also data provided by The City of Toronto's Open Data Portal.
https://open.toronto.ca/dataset/neighbourhood-profiles/

In the dataset are a lot further information of each neighborhood.



Map of the population for each neighborhood

The complete file can be directly accessed here:
https://raw.github.com/DirkFritz/Coursera_Capstone//master/neighbourhood-profiles-2016-csv.csv

Dataset for the Number of Competitors

The number of competitors is important criteria if a further business at a specific location shall be economically. For getting this data the Foursquare API is used:
https://developer.foursquare.com/places

Foursquare is a location-based recommendation service for restaurants and other places.

For the analysis the API function explore is used. In the following evaluation the restaurant type “pizza place” is chosen as an example. With the Foursquare API places can be clearly identified by an id.
The category “Food” the correspond to 4d4b7105d754a06374d81259 and a pizza place is 4bf58dd8d48988d1ca941735.

Venue Category	Venue Id
Chinese Restaurant	4bf58dd8d48988d145941735
Sandwich Place	4bf58dd8d48988d1c5941735
Fast Food Restaurant	4bf58dd8d48988d16e941735
Pizza Place	4bf58dd8d48988d1ca941735
Fried Chicken Joint	4d4ae6fc7a7b7dea34424761
Wings Joint	4bf58dd8d48988d14c941735

Result of a Foursquare request for category food