Lorenzo Rappuoli

July 16, 2021

Toronto

Battle of neighborhoods

Clustering Toronto's neighbohroods

- This project seeks to group neighborhoods into homogenous groups that indicate the likelihood of being a victim of a firefight.

- The results are based on data collected by the police and data on population, population density and average income of the inhabitants.

- Once the neighborhoods have been ranked, it will be easy for the reader to identify the best neighborhoods to live in

Data acquisition and cleaning

The data sources for the project are the following ones:

- **Demographics of Toronto neighbourhoods**:
  https://en.wikipedia.org/wiki/Demographics_of_Toronto_neighbourhoods

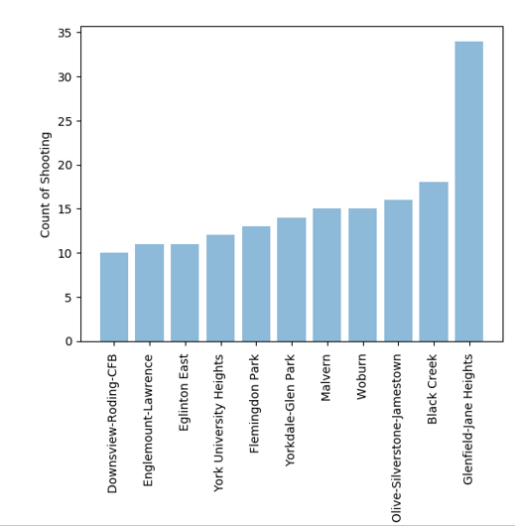
  In this Wikipedia page there is a list of demographic data on each Toronto neighbourhood from the Canadian census. It is useful because it contains the information about population, density and average income for every neighborhood.

- **Shootings & Firearm Discharges:**

  https://open.toronto.ca/dataset/shootings-firearm-discharges/

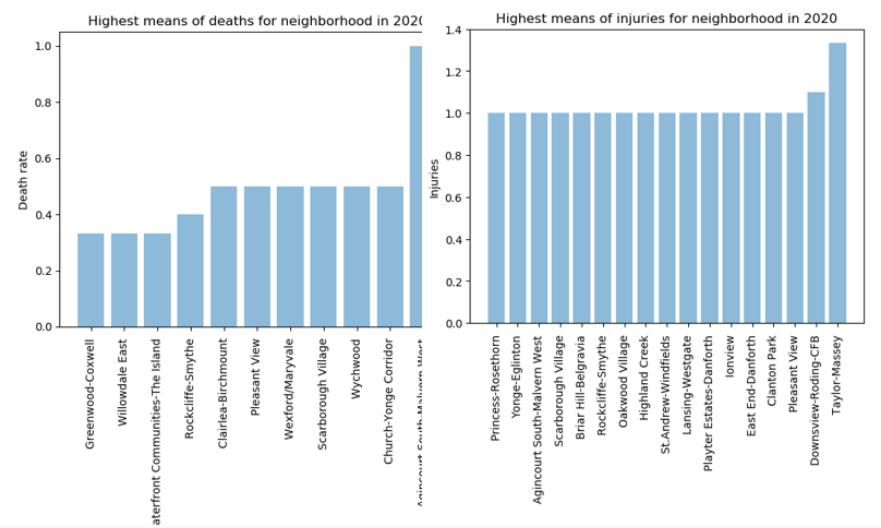
  This dataset contains all shooting-related occurrences reported to the Toronto Police Service.

Data analysis (1)



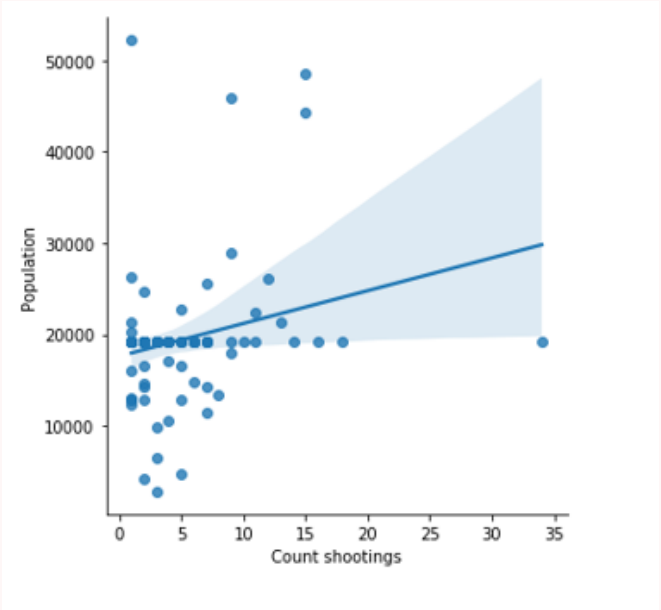
Neighborhoods that had more than ten shootings in 2020

Neighborhoods that had the most deaths and injuries



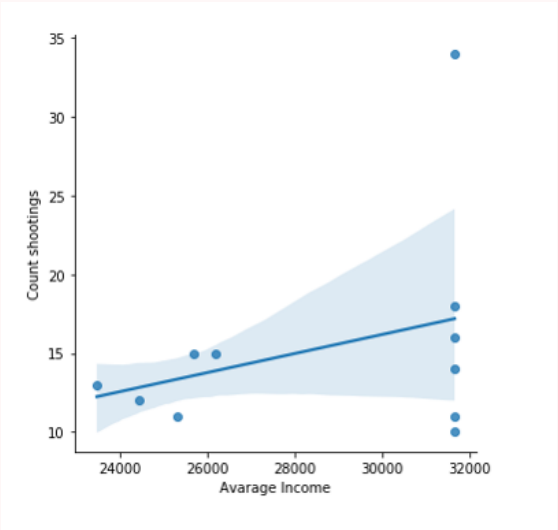
This indicates that more shootings do not always result in more deaths and injuries

Data analysis (2)



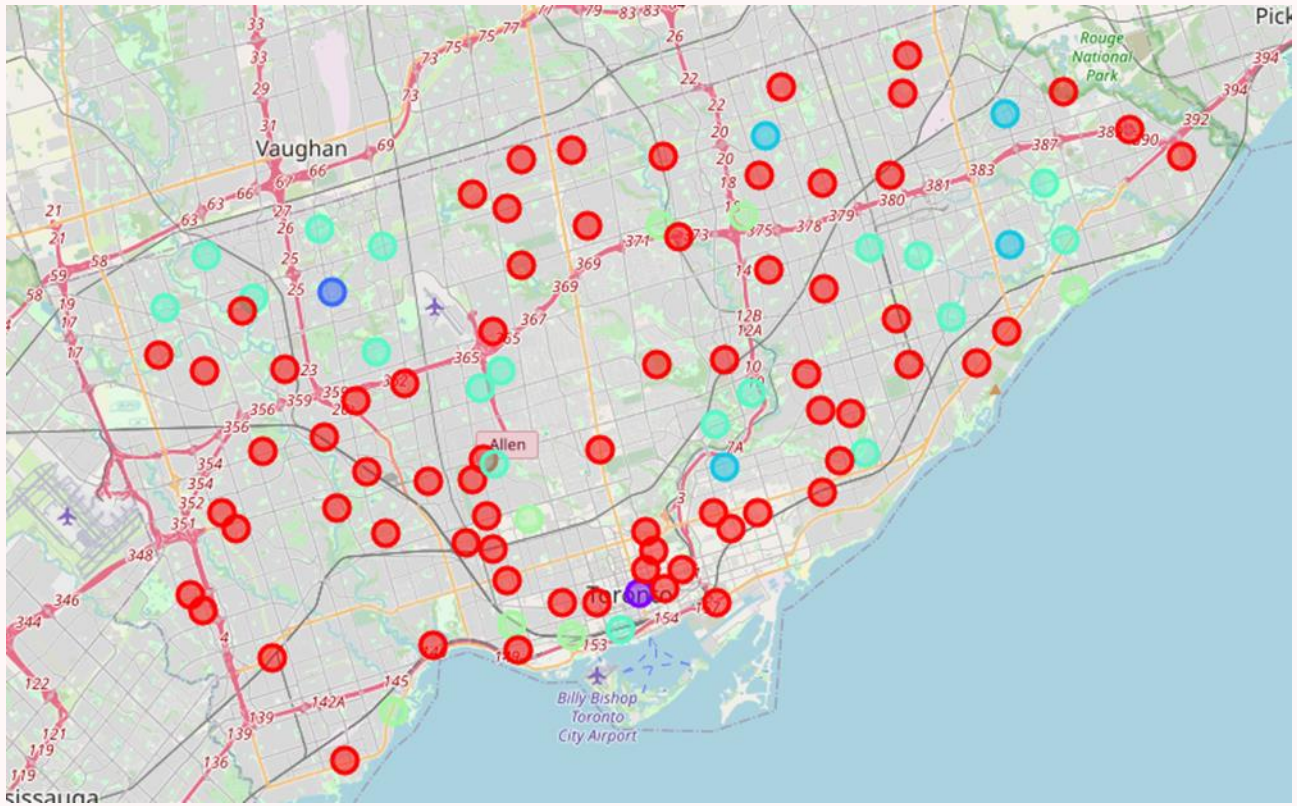
Positive correlation between Population and number of shootings

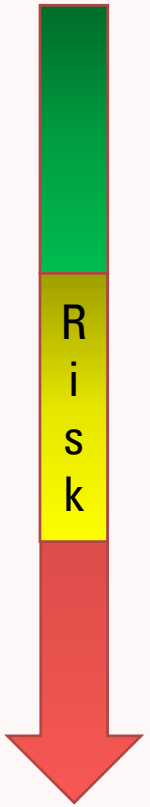
Positive correlation between average income and number of shootings (only for neigh. With more than 10 shootings in 2020)



A rich neighborhood isn’t automatically more safe than a poor one

Result of the clustering





Cluster 5 (light green)

Cluster 0 (red)

Cluster 1 (purple)

Cluster 4 (green)

Cluster 3 (dark green)

Cluster 2 (dark blue)

Conclusion

- Based on the results, the best neighborhoods from a safety perspective were found to be: Bayview Village, Guildwood, Henry Farm, Mimico, Niagara, Roncesvalles, Wychwood.
- Among the worst instead are: L'Amoreaux, Malvern, Old East York, Woburn.
- The worst of all is Glenfield-Jane Heights.
- In general, it is not possible to say that the average income and population density are sufficient indices to predict whether a neighborhood is dangerous. But it is possible to say that the higher the average income, the greater the chance of not getting into a firefight.