Best Neighborhood for Italian restaurant in Toronto

Data Science Project

04/23/2020 – Balint Pataki

Introduction

The main objective of the project is to find the best location for a new, Italian restaurant in Toronto. This information can be useful for entrepreneurs based in Canada and I am designing this study for them.

Ideas and methods

The location resolution will be by neighborhood and the parameters I am going to observe in order to recommend the best option(s) are:

- number of restaurants nearby

- number of Italian restaurants nearby

I applied an unsupervised Machine Learning technique called K-means in order to create groups of Neighborhoods based on the feature parameters (number of restaurants).

Data

We need the following information:

- List of neighborhoods in City of Toronto, Ontario, Canada
- Geographical coordinates of the neighborhoods (latitude, longitude)
- Venue data related to:
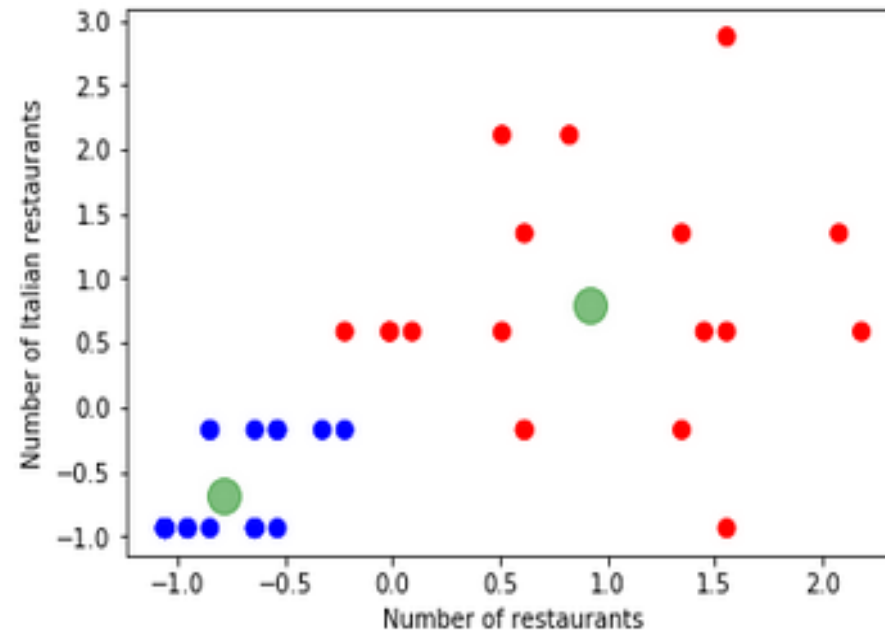 - Restaurants and
  - Italian restaurants
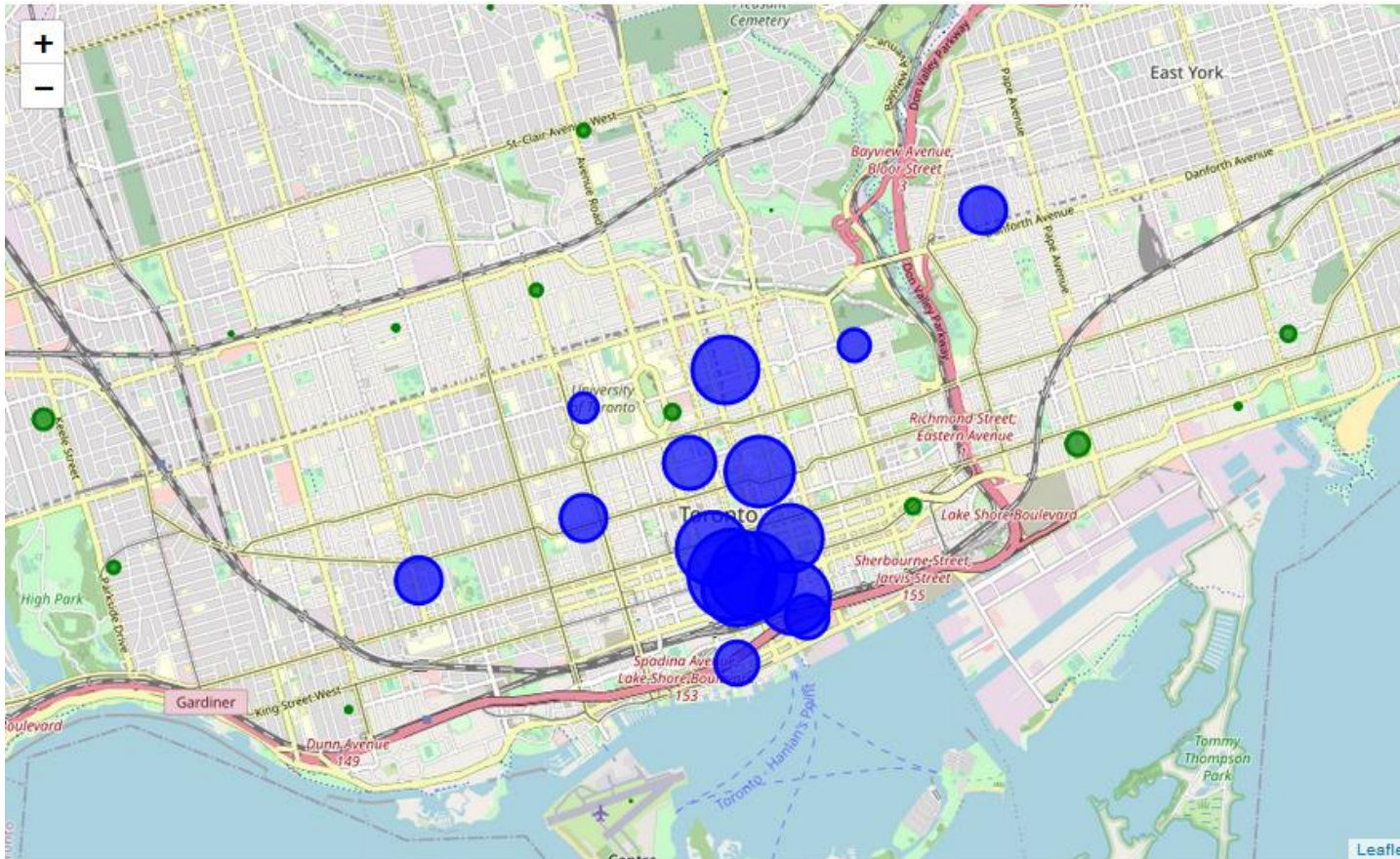
Our Master Dataframe on a map

Neighborhoods of Toronto

Clusters in the data

▶ We basically found two separated groups with K-Means unsupervised machine learning algorithm, applied for number of restaurants and Italian restaurants per neighborhoods:
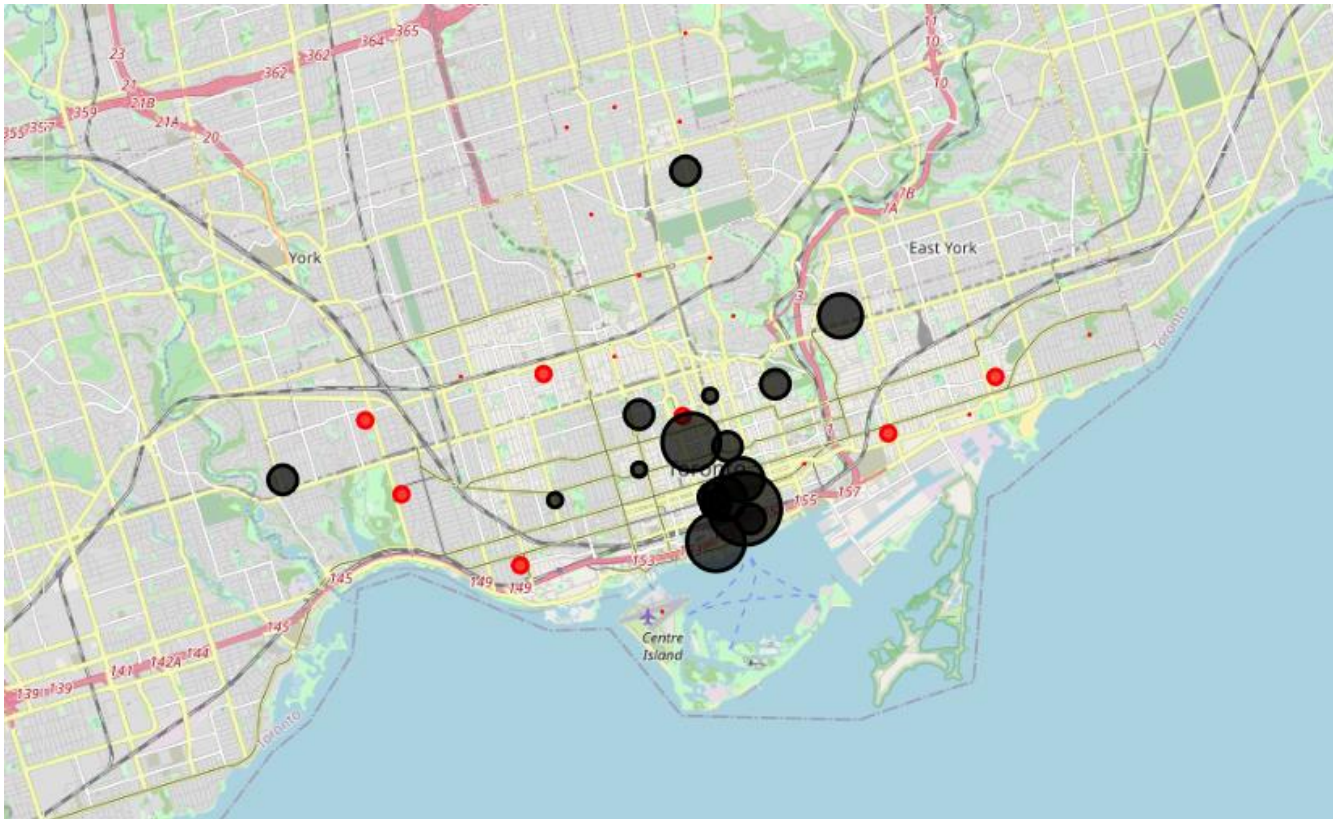
Mapping Restaurants in Toronto by Neighborhoods

Cluster 0 – neighborhoods with fewer restaurants marked with green

Mapping Italian Restaurants in Toronto by Neighborhoods

Cluster 0 – neighborhoods with fewer Italian restaurants marked with red

Conclusion

We need to look at **Cluster 0 with less restaurants** for potential placement but **we shouldn't forget about more crowded neighborhoods** - from there we can filter for the places where we found zero Italian restaurants - these two options can identify promising neighborhoods for our business.

Please consider that this study gives us a good starting point, but further examination is needed, adding population, net income, tourism data to our model.

Thank you for your attention!