Battle of Neighbourhoods

Opening coffee-shop in Toronto

Alex Honcharuk. May 11, 2020

Introduction

New «Co-ffee» restaurant in Toronto

«Co-ffee» - is an old coffee-shops brand, that have open branches all over the world. They specialise on small coffee-points near places of sightseeing and rest.

For some time they were building a project of opening another branch in Toronto.

It was decided to use modern data science methods to determine appropriate place for new restaurant.

Interest

Criteria are milestone parts of project

1. *Number of restaurants nearby*. Company is interested to place their branch in free of other restaurants neighbourhood.

2. *Availability of city sights*. Company is sure that this would attract more customers.

3. *City transport nearby*. This would allow to reach restaurants easier.

4. *Access to apartments*. Tourist accommodation places will grant the flow of some new customers.

Data extraction

Reasonable data - reasonable result

1. https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M - database of Toronto neighbourhoods by postal code.

2. https://cocl.us/Geospatial_data - database with all geospatial data by postal codes.

3. https://api.foursquare.com/ - foursquare API with data regarding all venues in Toronto.

All data was cleaned and aggregated into one proper database with features selected due to the criteria.
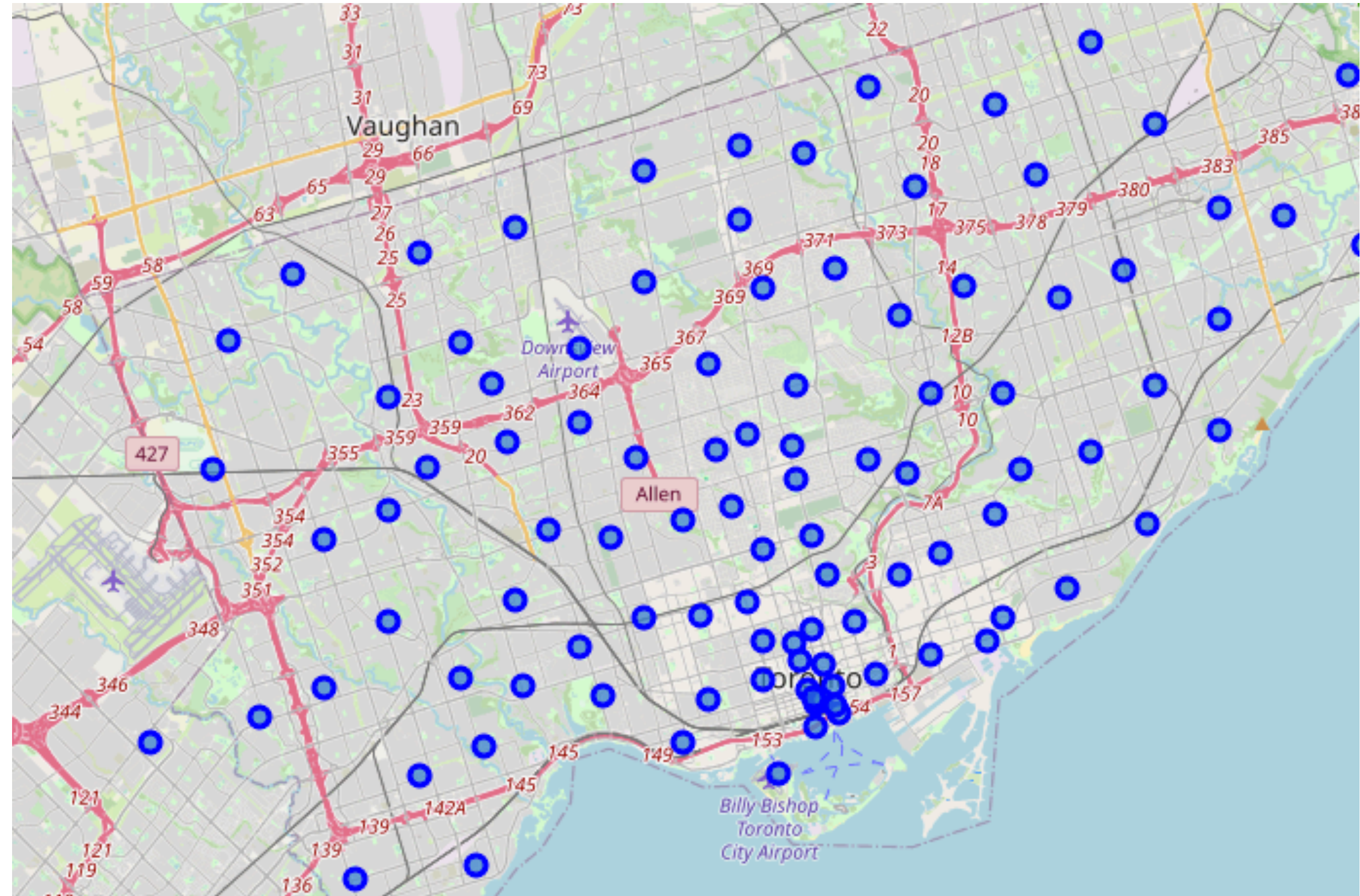
Toronto Neighbourhoods

Looking at the case

As we can see from the map, most of Toronto neighbourhoods are situated in the center.

They are spread equally over other regions of the city.

Visually, there are no possibilities to group those neighbourhoods on any clusters using only geospatial data.

Clustering Neighbourhoods

First step of analysis

Using criteria of «Co-ffee» company, we can choose appropriate cluster for further investigation.

As we can see, clusters 1-4 had A LOT of restaurants - more then 1 per neighbourhood. Looks like those are not suitable.

Let's check cluster 0. Number of restaurants is 0.9 per neighbourhood - good chance of finding neighbourhood without restaurant. 26 sights is huge! This is very satisfying parameter. 1

0 transport and 14 apartments only adding confidence to our choice.

```
         Shops  Restaurants  Sights  Apartments  Transport
Cluster
0         62.0         49.0    26.0        14.0       10.0
1         24.0         16.0     4.0        22.0        6.0
2        100.0         77.0     8.0         4.0        3.0
3         19.0         60.0     2.0         7.0        0.0
4        104.0         48.0     3.0         1.0        3.0
```

	Latitude						
	count	mean	std	min	25%	50%	75%
Cluster							
0	55.0	43.709159	0.051154	43.602414	43.670283	43.706876	43.7493
1	5.0	43.690448	0.051680	43.640816	43.662696	43.665860	43.7127
2	25.0	43.713771	0.057025	43.605647	43.669542	43.709060	43.7543
3	6.0	43.651651	0.008350	43.644771	43.647364	43.649249	43.6512
4	12.0	43.697038	0.050996	43.648429	43.657946	43.674281	43.7222

Picking Neighbourhoods

Discover neighbourhoods using criteria

So, now we have cluster and criteria. We can combine those to find perfect places for a restaurant. Let's check if there are any neighbourhoods without restaurants.

Great, we have not only one, but three of them! That's always good to have a choice. First even have access to 3 sights, that's more than OK. Each has access to transport. Those are just perfect.

|	Postal Code	Borough	Neighborhood	Latitude	Longitude	Shops	Restaurants	Sights	Apartments	Transport	Cluster
14	M4C	East York	Woodbine Heights	43.695344	-79.318389	2.0	0.0	3.0	0.0	1.0	0
74	M5R	Central Toronto	The Annex / North Midtown / Yorkville	43.672710	-79.405678	0.0	0.0	1.0	1.0	1.0	0
101	M8Y	Etobicoke	Old Mill South / King's Mill Park / Sunnylea /...	43.636258	-79.498509	0.0	0.0	1.0	1.0	1.0	0

Toronto Neighbourhoods Recommendation

Illustrating results

So, I chose few neighbourhoods (some of them are nearby to each other) - Woodbine Hights, The Annex/North Midtown/Yorkville and Old Mill South/King's Mill Park/Sunnylea.

All this neighbourhoods are situated in the center, which means more tourists and visitors. Also, those were neighbourhoods without any restaurant, and with access to transport and a lot of sights (mostly parks).

Looks like those are really promising places to open coffee-shop. Most interesting is the The Annex/North Midtown/Yorkville one, that can have business places nearby, and also would be accessible from Annex Apartments and Queens park.
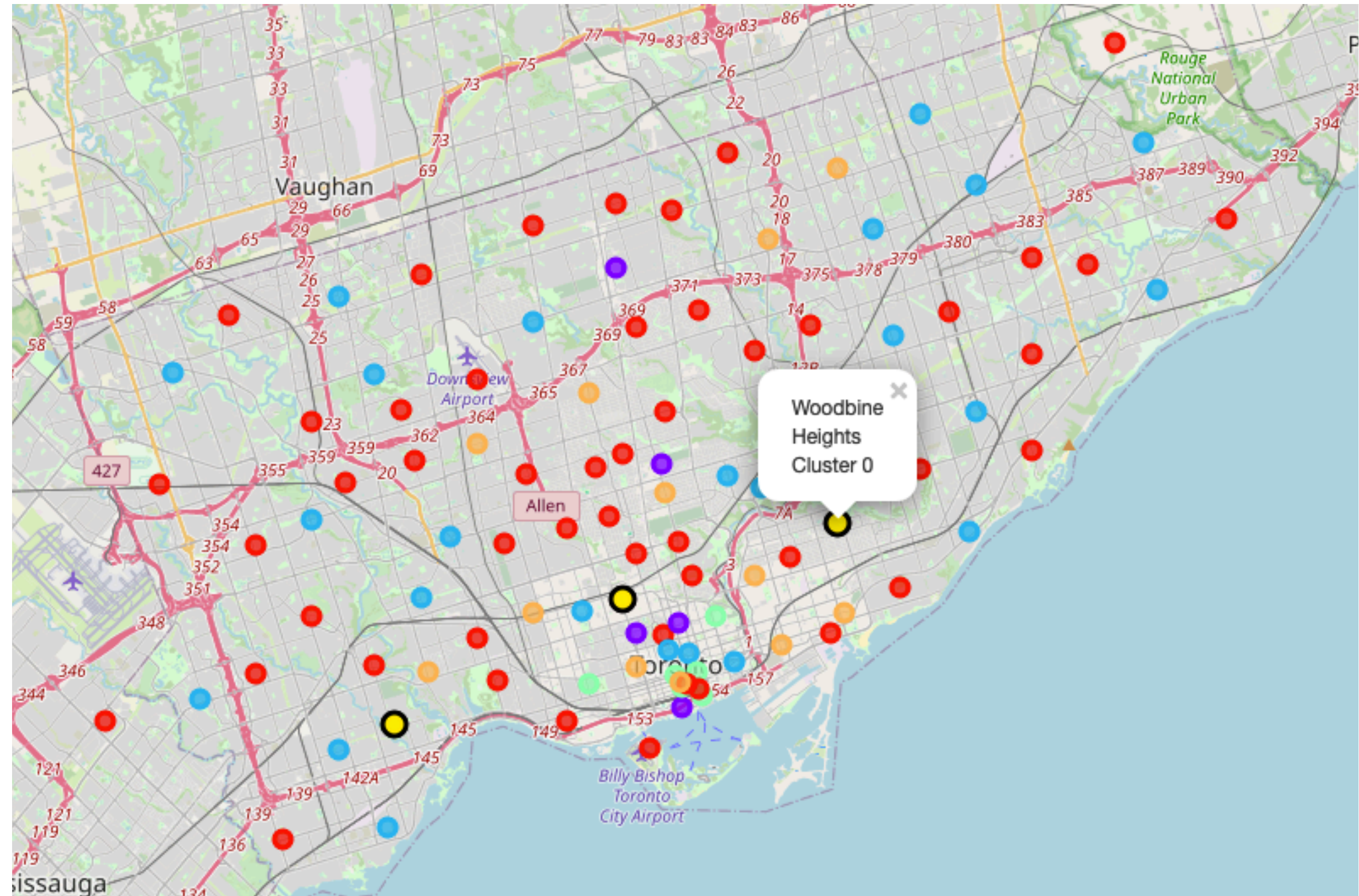
Conclusions

What we achieved

Purpose of this project was to identify good places to open a coffee-shop in Toronto, following some criteria and using Python data analysis methods.

During the research, was built a database, that included information regarding all venues in Toronto, grouped by Neighborhoods. K-cluster analysis helped to choose promissing group of places, which were analyzed farther.

After additional research, 3 candidates were chosen. All those places satisfied criteria.

Using the research material, was built a map with the results.