

Introduction

Latvia is a relatively small country (in the beginning of 2021 the population was 1,8 million) located in Northern Europe. The capital is Riga with estimated population of 627 763 people. Through the years there has been emigration to other countries – United Kingdom, Germany, United States of America, Canada etc. Causes for moving to other countries can be different – job opportunities, financial gain, education or merely a want to change the scenery. So, for people who have never been to places like Toronto, New York, London or Berlin, or who want to choose the best possible place for their new home, it can be hard to choose the right neighborhood/borough.

The goal is to compare Toronto and Riga and see how are they different. Of course, the size of the cities is very different, but may be there are some similarities among the neighborhoods considering activities/amenities.

Also, this kind of information may provide potential travelers with insight into what to expect from Toronto if you are familiar with Riga and vice versa.

Data

Firstly, through Web scraping data about neighborhoods/boroughs and coordinates are to be obtained from:

Riga	Toronto
Neighborhoods/boroughs: https://lv.wikipedia.org/wiki/R%C4%ABgas_apkaimju_uzskait%C4%ABjums	Neighborhoods/boroughs: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
Latitudes/longitudes: Manually created json file. Why? There isn't one source which provides with accurate data about coordinates of the boroughs in Riga. Some of	Latitudes/longitudes: https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-

the coordinates provided were entirely wrong. For example, latitude.to provided with coordinates which resulted in the Baltic Sea, which clearly is not the center of Riga. Due to the size of Riga and Latvia, this is not surprising.

[DS0701EN-SkillsNetwork/labs_v1/Geospatial_Coordinates.csv](#)

Riga

Example of performed Web scraping using BeautifulSoup objects:

```
# The basic information will be gathered from wikipedia and converted into a dataframe
Riga_source = requests.get('https://lv.wikipedia.org/wiki/R%C4%ABgas_apkaimju_uzskait%C4%ABjums').text
soup=BeautifulSoup(Riga_source,'lxml')
table=soup.find('table')
riga_df = pd.read_html(str(table))
r_df = pd.DataFrame(riga_df[0])
```

Raw dataframe was created:

	Attēls	Apkaimē	ApkaimesnovietojumsRīgā	Platība (km²)	Iedz.sk. (2014)[2]	Iedzīvotājublīvums(iedz./km²)	Galvenās ielas	Priekšpilsēta
0	NaN	Atgāzene	NaN	745	1613	21651.0	Graudu iela,Vienības gatve	Zemgales priekšpilsēta
1	NaN	Avoti (arī Avotu iela)	NaN	1815	19710	108595.0	Avotu iela,Aleksandra Čaka iela, Augusta Deglav...	Latgales priekšpilsēta
2	NaN	Āgenskalns	NaN	4613	27923	605311.0	Akmeņu iela,Bāriņu iela,Daugavgrīvas iela,Kaln...	Kurzemes priekšpilsēta/Zemgales priekšpilsēta
3	NaN	Beberbeķi	NaN	1204	411	34136.0	Beberbeķu iela,Beberbeķu 5. līnija,Beberbeķu 9...	Zemgales priekšpilsēta
4	NaN	Bergļi	NaN	5706	2942	5160.0	Berģu iela,Brīvības gatve,Rožu iela,Upsciena ...	Vidzemes priekšpilsēta

There were columns with pictures and other data which was not necessary for the project, so they were dropped and the column names were translated to English, so readers can get a sense of what the columns represent.

```
#Now the dataframe should be transformed and headers should be translated into english
r_df.columns = ['Image', 'Borough', 'Borough placement', 'Area (km2)', 'Population', 'Population density(people/km2)', 'Main streets', 'Neighborhood']
r_df.drop(['Image', 'Borough placement', 'Main streets'], axis=1, inplace=True)
```

Next, previously mentioned json object was created manually. Coordinates were gathered in an excel file and next, the file was converted to json file. The result of the conversion was inserted into the project. Also, latitude and longitude columns were converted to float type, as this is the correct type for these columns.

```
import json
json_string = '[{"Borough": "Atgāzene", "Latitude": "56.91727466", "Longitude": "24.07453804"}, {"Borough": "Avoti (arī Avotu iela)", "Latitude": "56.952469", "Longitude": "24.140854"}]'
a_json = json.loads(json_string)
riga_geo = pd.DataFrame.from_dict(a_json, orient="columns")
#Coordinates have to be transformed to float
riga_geo["Latitude"] = pd.to_numeric(riga_geo["Latitude"], downcast="float")
riga_geo["Longitude"] = pd.to_numeric(riga_geo["Longitude"], downcast="float")
riga_geo.head(15)
```

	Borough	Latitude	Longitude
0	Atgāzene	56.917274	24.074537
1	Avoti (arī Avotu iela)	56.952469	24.140854
2	Āgenskalns	56.933331	24.066666
3	Beberbeķi	56.950001	23.950001
4	Bērģi	56.983330	24.299999
5	Bieriņi	56.919724	24.058611
6	Bišumuiža	56.909698	24.136101
7	Bolderāja	57.033333	24.049999
8	Brasa	56.973110	24.141119
9	Brekši	56.966362	24.303980
10	Bukulti	57.000000	24.267000
11	Centrs	56.946285	24.105078
12	Čiekurkalns	56.983330	24.166666
13	Daugavgrīva	57.045120	24.039288
14	Dārziems	56.944546	24.177010

To create the final coordinate dataframe for Riga, previously mentioned two dataframes were merged, using column Borough. The resulting dataframe contains information about Boroughs, Neighborhoods and their respective Population, Latitudes and Longitudes.

	Borough	Area (km2)	Population	Population density(people/km2)	Neighborhood	Latitude	Longitude
0	Atgāzene	745	1613	21651.0	Zemgales priekšpilsēta	56.917274	24.074537
1	Āgenskalns	4613	27923	605311.0	Kurzemes priekšpilsēta/Zemgales priekšpilsēta	56.933331	24.066666
2	Beberbeķi	1204	411	34136.0	Zemgales priekšpilsēta	56.950001	23.950001
3	Bērģi	5706	2942	5160.0	Vidzemes priekšpilsēta	56.983330	24.299999
4	Bieriņi	4274	9443	22094.0	Zemgales priekšpilsēta	56.919724	24.058611
5	Bišumuiža	2243	2360	105216.0	Zemgales priekšpilsēta	56.909698	24.136101
6	Bolderāja	8329	12994	156009.0	Kurzemes priekšpilsēta	57.033333	24.049999
7	Brasa	1741	13241	76054.0	Vidzemes priekšpilsēta	56.973110	24.141119
8	Brekši	2039	1751	85875.0	Vidzemes priekšpilsēta	56.966362	24.303980
9	Bukulti	5183	675	13023.0	Vidzemes priekšpilsēta	57.000000	24.267000
10	Centrs	3732	35274	945177.0	Centra rajons/Latgales priekšpilsēta/Vidzemes ...	56.946285	24.105078

Toronto

Data about Toronto was obtained in a similar manner – through Webscraping with BeautifulSoup objects. The difference was in how the coordinate data was gathered – through csv file which was provided by Coursera course Applied Data Science Capstone.

```
#Now the coordinates will be gathered from resource which was kindly provided by Coursera course Applied Data Science Capstone
toronto_geo = pd.read_csv("https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs_v1/Geospatial_Coordinates.csv")
```

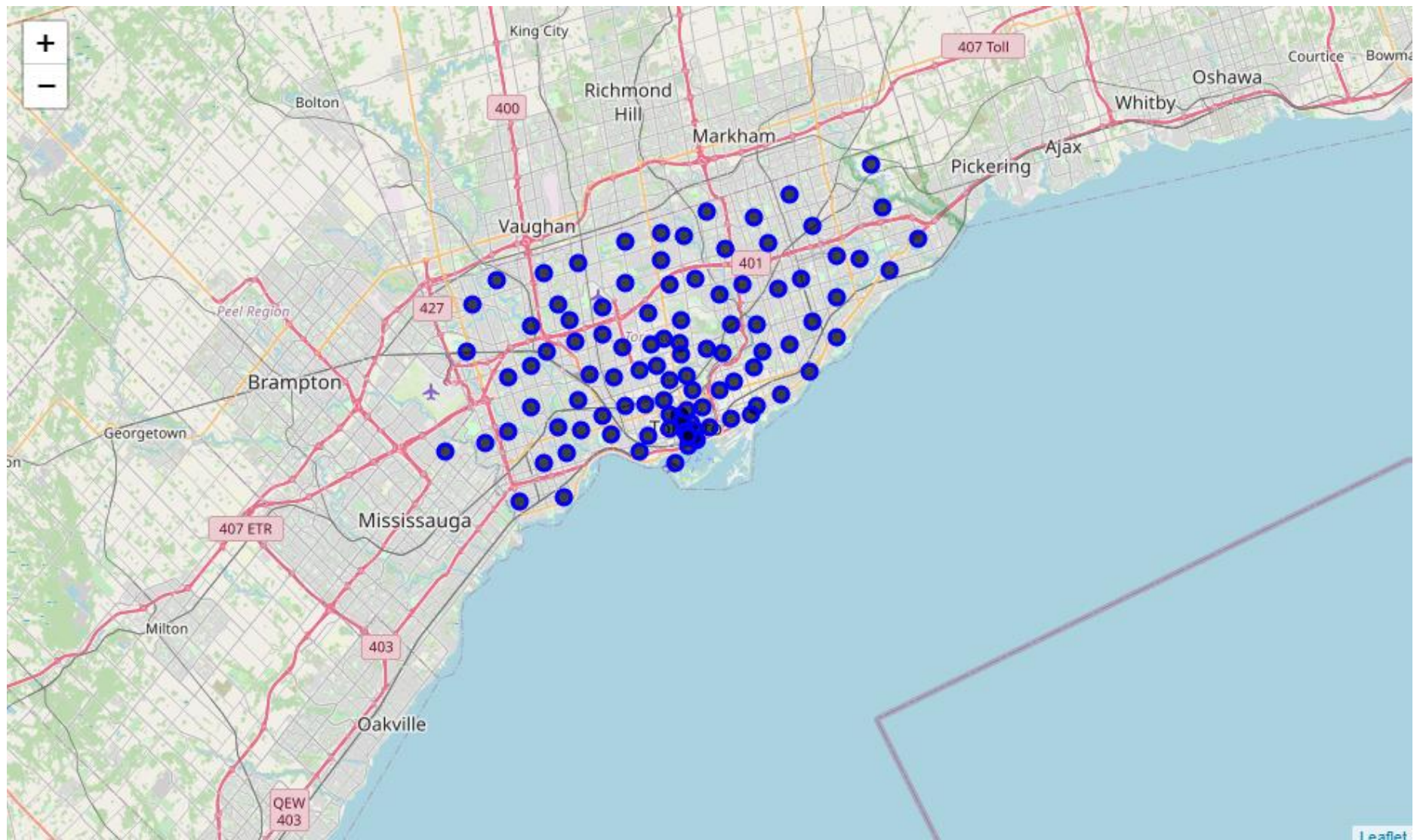
Next the dataframes were merged using Postal codes and the resulting dataframe had all the necessary data about Boroughs, Neighborhoods and their Latitudes and Longitudes. This procedure was similar to the one performed on the data about Riga, but was far more simpler as the coordinates were already formed into a csv file and I didn't have to manually validate them and create a json object.

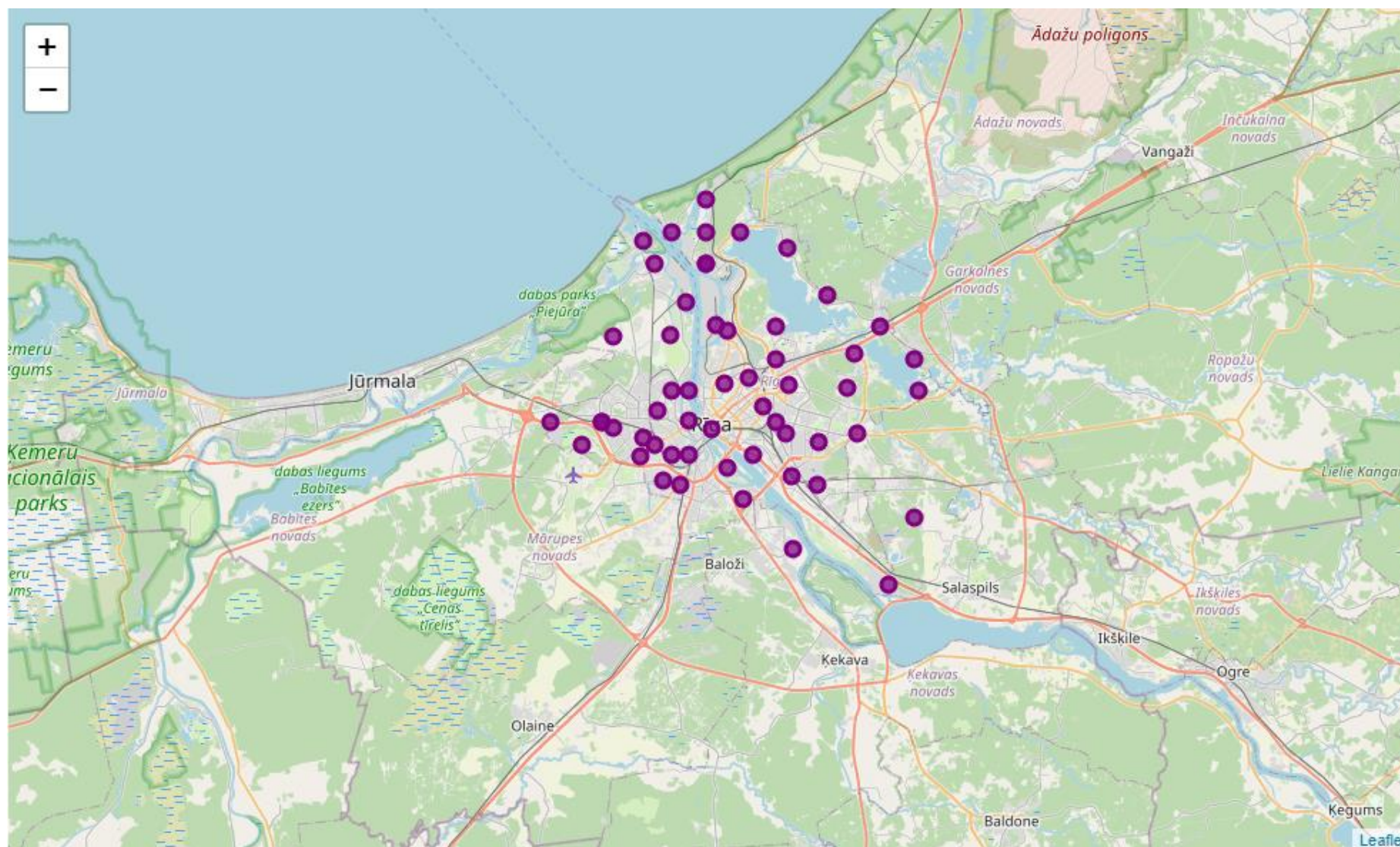
	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Queen's Park	Ontario Provincial Government	43.662301	-79.389494
5	M9A	Etobicoke	Islington Avenue	43.667856	-79.532242
6	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353
7	M3B	North York	Don Mills North	43.745906	-79.352188
8	M4B	East York	Parkview Hill, Woodbine Gardens	43.706397	-79.309937
9	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937
10	M6B	North York	Glencairn	43.709577	-79.445073

Exploratory data analysis

City maps

Gathered data was used to visualize the boroughs in the chosen cities – Riga and Toronto. The visualization was created using folium maps. Folium takes coordinates and can be told what color markers and what labels should be populated on a map constrained by user defined longitude and latitude. This kind of visualization can demonstrate similarities and dissimilarities of the structure and geographical locations of cities. From the two maps illustrated in the following pages, one can observe that both cities have access to a body of water and Riga in comparison to Toronto seems to be ‘greener’ and surrounded by lakes and rivers. Also, the markers show how the center of Toronto is densely packed with boroughs, while Riga seems to be more sparsely populated in terms of boroughs.





Data about Venues on Foursquare

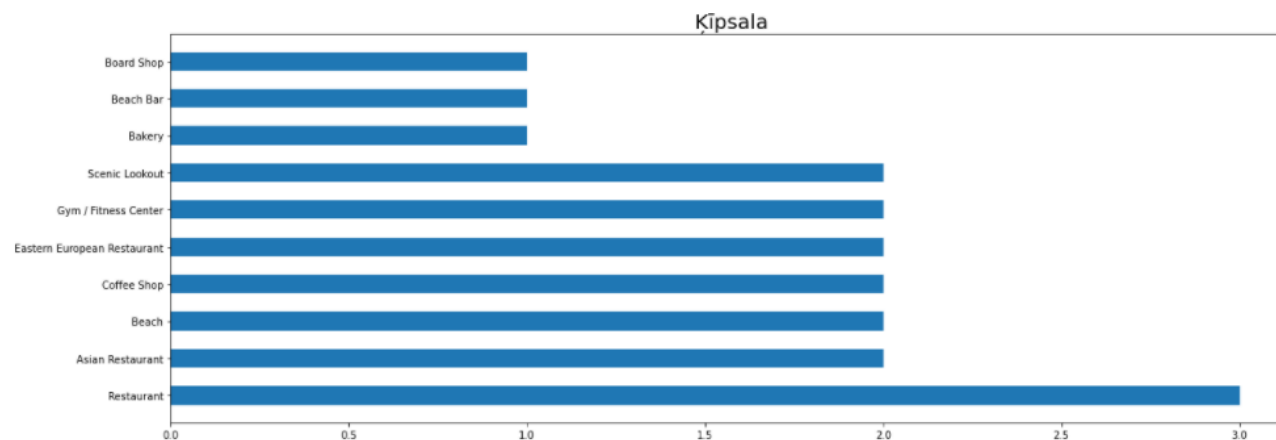
The next step would be to get venues near each borough. Foursquare is a great service for that. A developer account was created in order to have opportunity to utilize the service. To get the nearby venues, limit, radius and version were defined. From the developer's profile client ID and client secret were obtained. These defined values were used in a function, which calls Foursquare and retrieves data about venues. At this point, venue coordinates and categories were of interest.

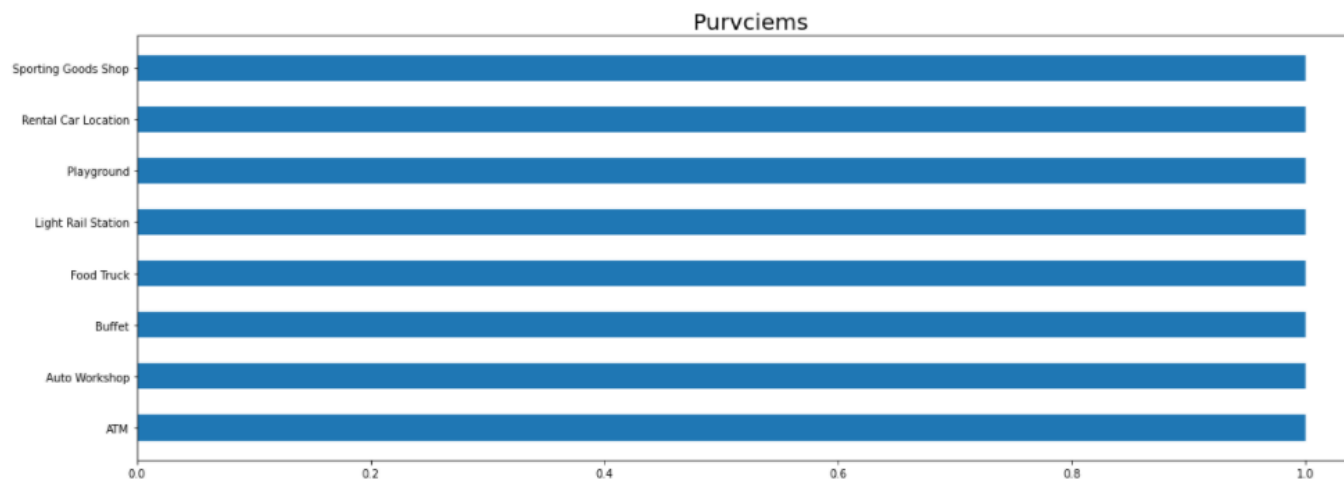
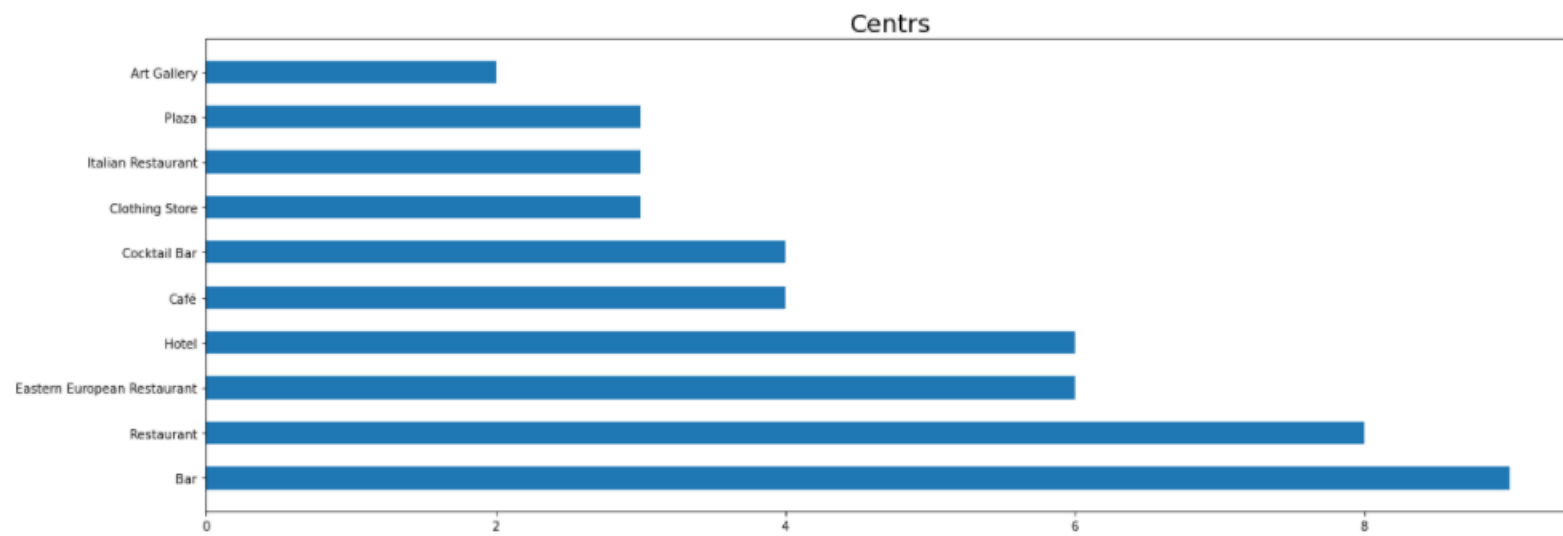
Results for the cities were very different even if just looking at the volume of the gathered data:

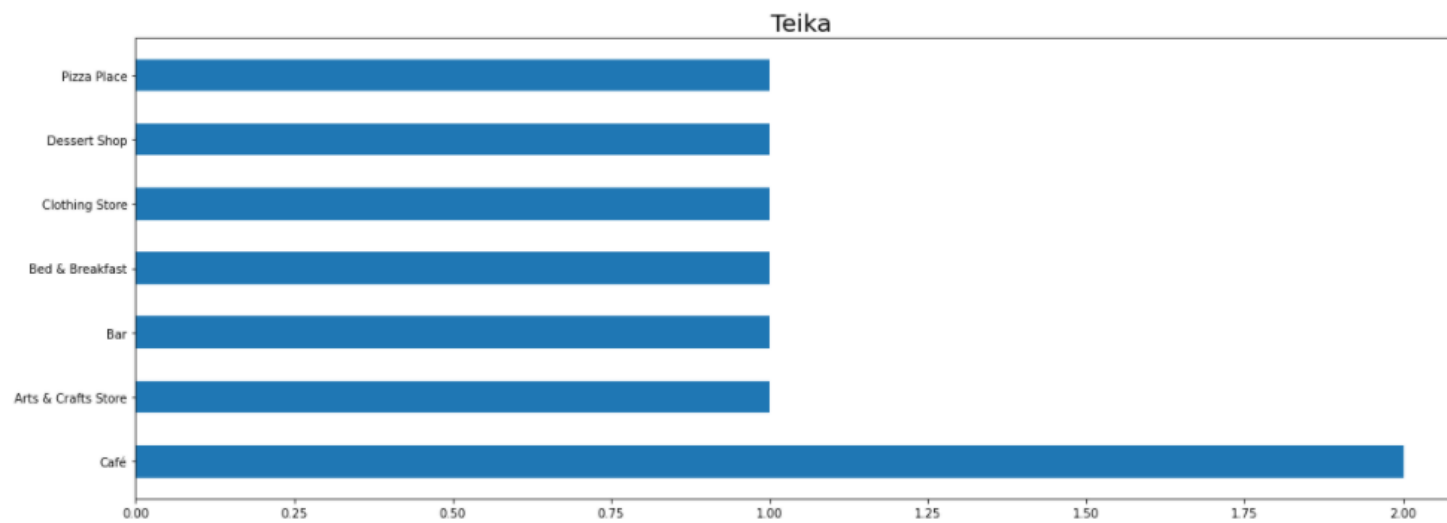
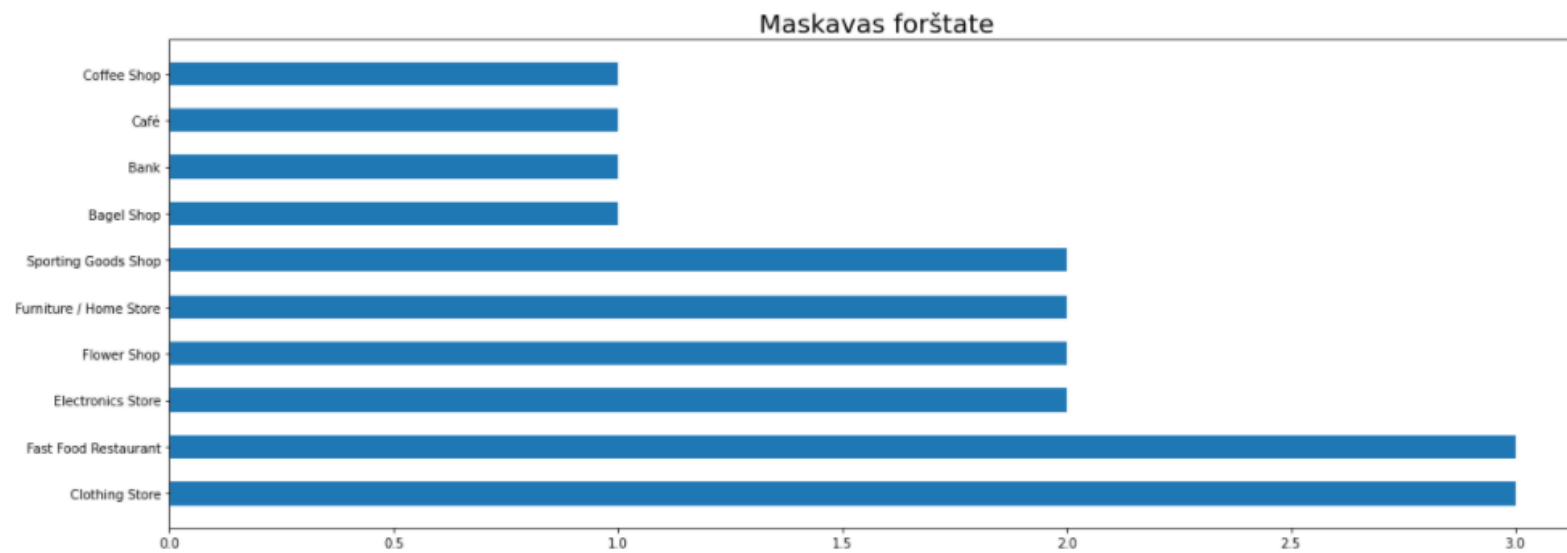
<code>Riga_venues.shape</code>	<code>Toronto_venues.shape</code>
<code>(627, 7)</code>	<code>(2128, 7)</code>

To see what are the differences between the types of venues in the boroughs, matplotlib pyplot was used. Here are some examples for both cities (more can be seen in the notebook, if you are interested).

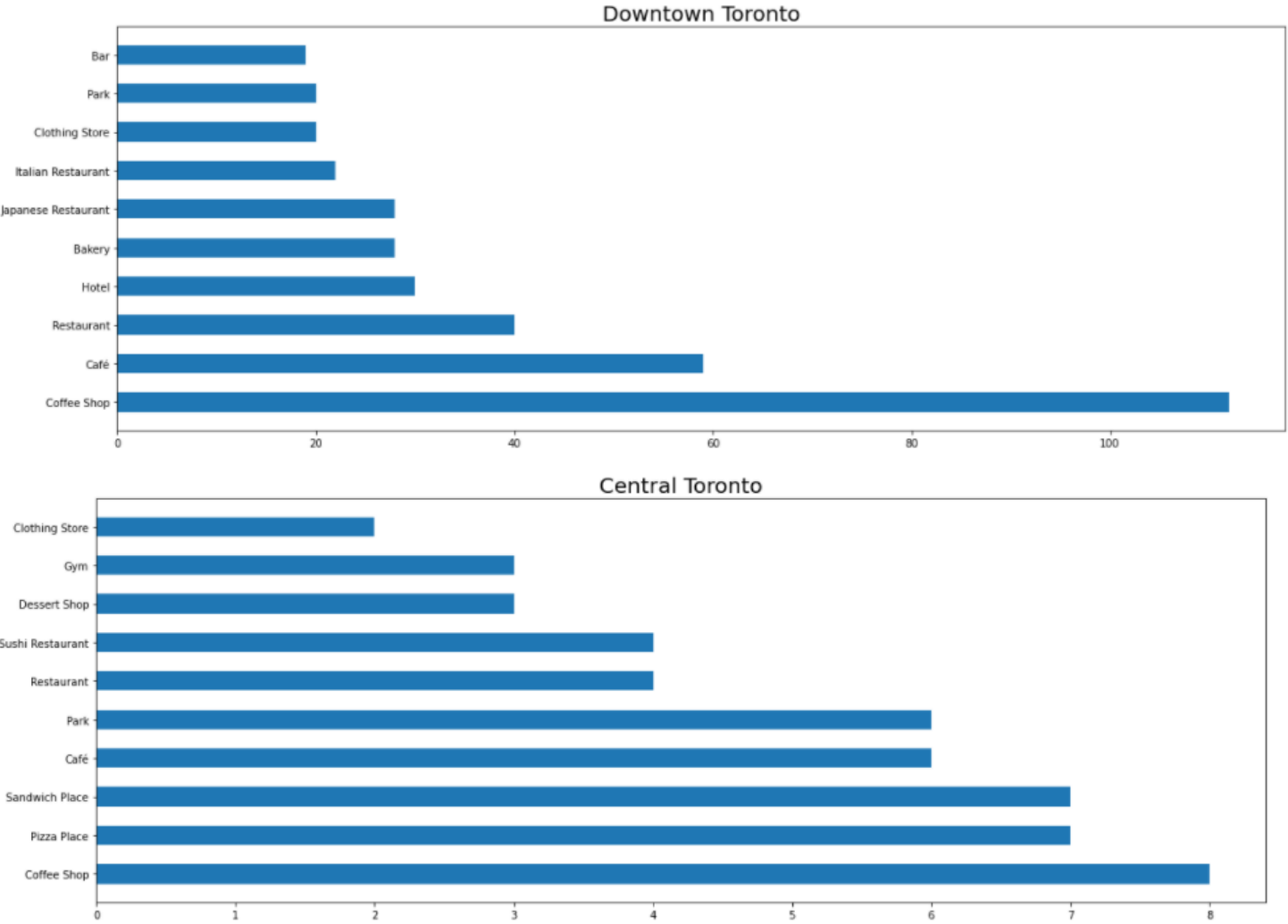
Riga

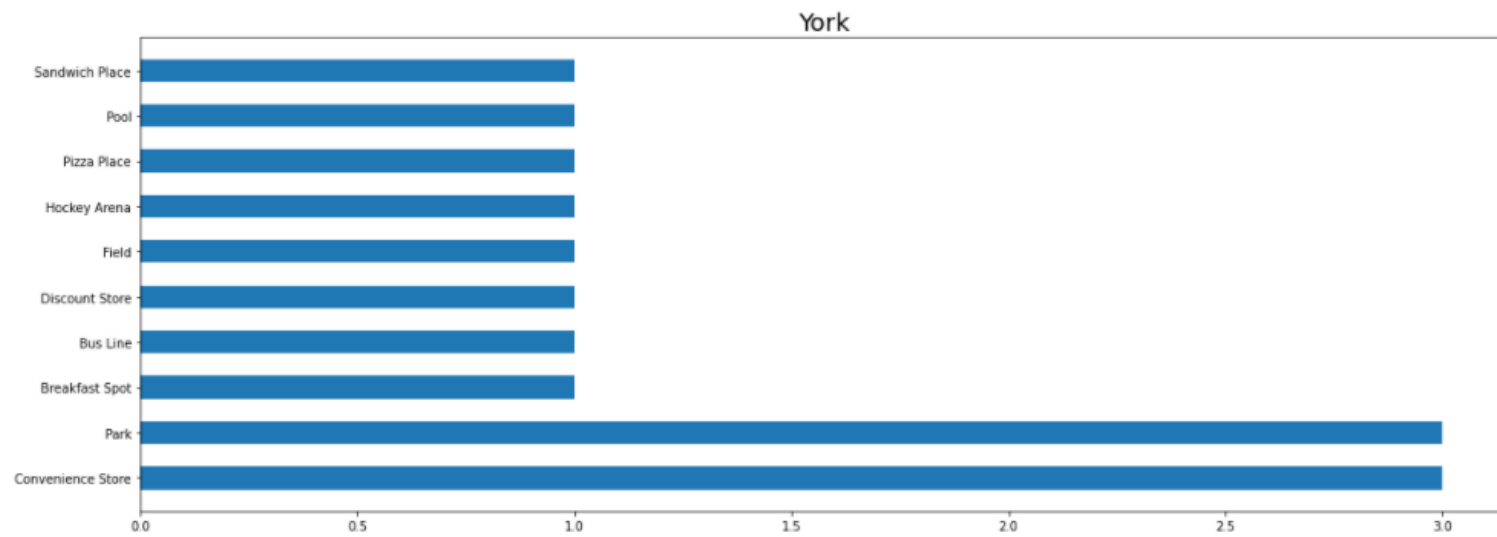
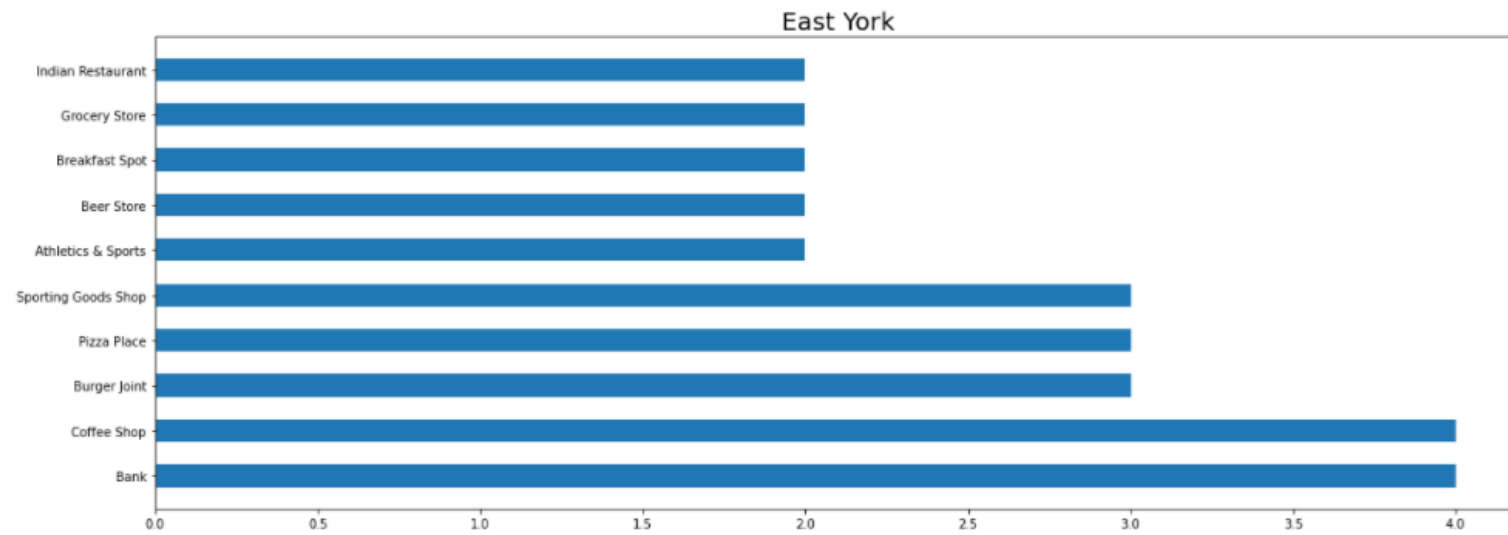






Toronto





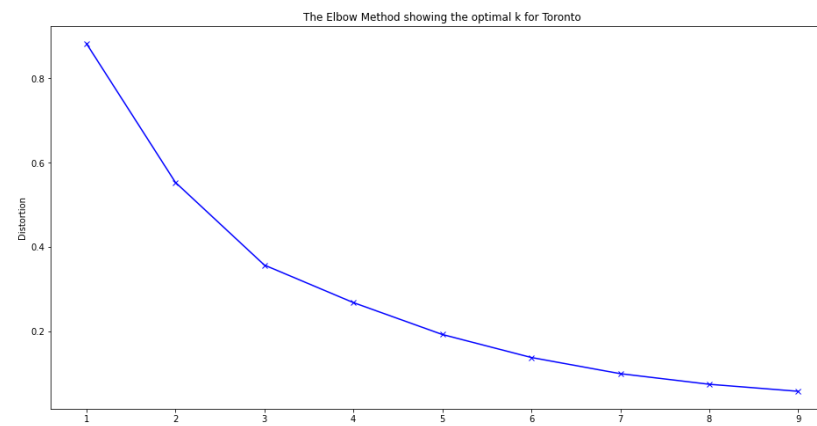
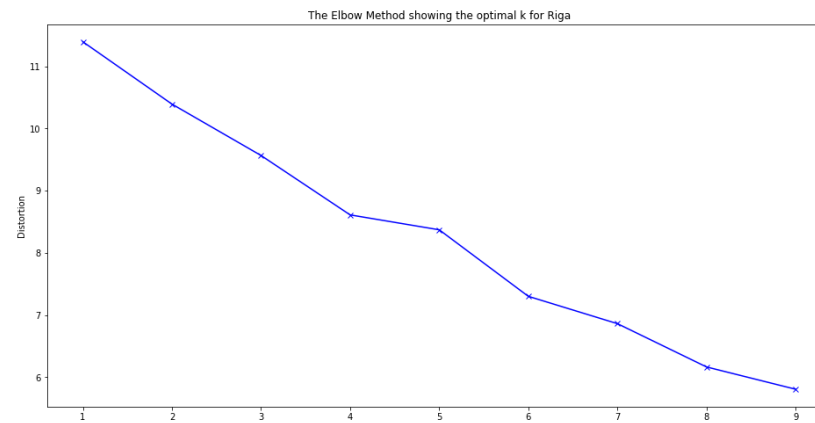
Next, to explore the venue data in a more detailed way and perform clustering analysis, the data was rearranged using one-hot encoding for each category, and finally dataframes with ten most common venues for each borough.

	Borough	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Atgāzene	Bus Station	Grocery Store	Bus Stop	Pool	Arts & Crafts Store	Train Station	Athletics & Sports	Forest	Yoga Studio	Farm
1	Beberbeķi	Moving Target	Garden Center	Bus Stop	Bike Trail	Yoga Studio	Discount Store	Fast Food Restaurant	Farmers Market	Farm	Event Space
2	Berģi	Bus Station	Beach	Auto Garage	Lake	Yoga Studio	Fish Market	Fast Food Restaurant	Farmers Market	Farm	Event Space
3	Bieriņi	Trail	Bus Stop	Museum	Sculpture Garden	Neighborhood	Bakery	Track	Outdoor Sculpture	Bike Rental / Bike Share	Department Store
4	Bišumuiža	Karaoke Bar	Auto Garage	Yoga Studio	Diner	Fish Market	Fast Food Restaurant	Farmers Market	Farm	Event Space	Electronics Store
5	Bolderāja	Bar	Pizza Place	Pharmacy	Turkish Restaurant	Food & Drink Shop	Fast Food Restaurant	Steakhouse	Gym	Café	Eastern European Restaurant
6	Brasa	Park	Gym / Fitness Center	Clothing Store	Bar	Pet Store	Convenience Store	Cosmetics Shop	Bus Stop	Restaurant	Bus Station
7	Brekši	Auto Garage	Art Gallery	Bar	Yoga Studio	Discount Store	Fish Market	Fast Food Restaurant	Farmers Market	Farm	Event Space

	Borough	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Central Toronto	Coffee Shop	Pizza Place	Sandwich Place	Park	Café	Sushi Restaurant	Restaurant	Gym	Dessert Shop	Italian Restaurant
1	Downtown Toronto	Coffee Shop	Café	Restaurant	Hotel	Bakery	Japanese Restaurant	Italian Restaurant	Park	Clothing Store	Bar
2	Downtown Toronto Stn A	Coffee Shop	Seafood Restaurant	Italian Restaurant	Japanese Restaurant	Restaurant	Pub	Café	Bakery	Creperie	Cheese Shop
3	East Toronto	Coffee Shop	Greek Restaurant	Italian Restaurant	Ice Cream Shop	Brewery	Café	American Restaurant	Bakery	Pub	Bookstore
4	East Toronto Business	Light Rail Station	Comic Shop	Farmers Market	Restaurant	Brewery	Auto Workshop	Garden	Garden Center	Burrito Place	Fast Food Restaurant
5	East York	Coffee Shop	Bank	Burger Joint	Sporting Goods Shop	Pizza Place	Intersection	Park	Beer Store	Athletics & Sports	Skating Rink
6	East York/East Toronto	Intersection	Park	Convenience Store	Eastern European Restaurant	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Drugstore	Dumpling Restaurant
7	Etobicoke	Pizza Place	Sandwich Place	Coffee Shop	Grocery Store	Bakery	Fast Food Restaurant	Café	Gym	Pharmacy	Liquor Store
8	Etobicoke Northwest	Truck Stop	Garden Center	Rental Car Location	Bar	Drugstore	Donut Shop	Discount Store	Distribution Center	Dog Run	Doner Restaurant

Machine learning analysis

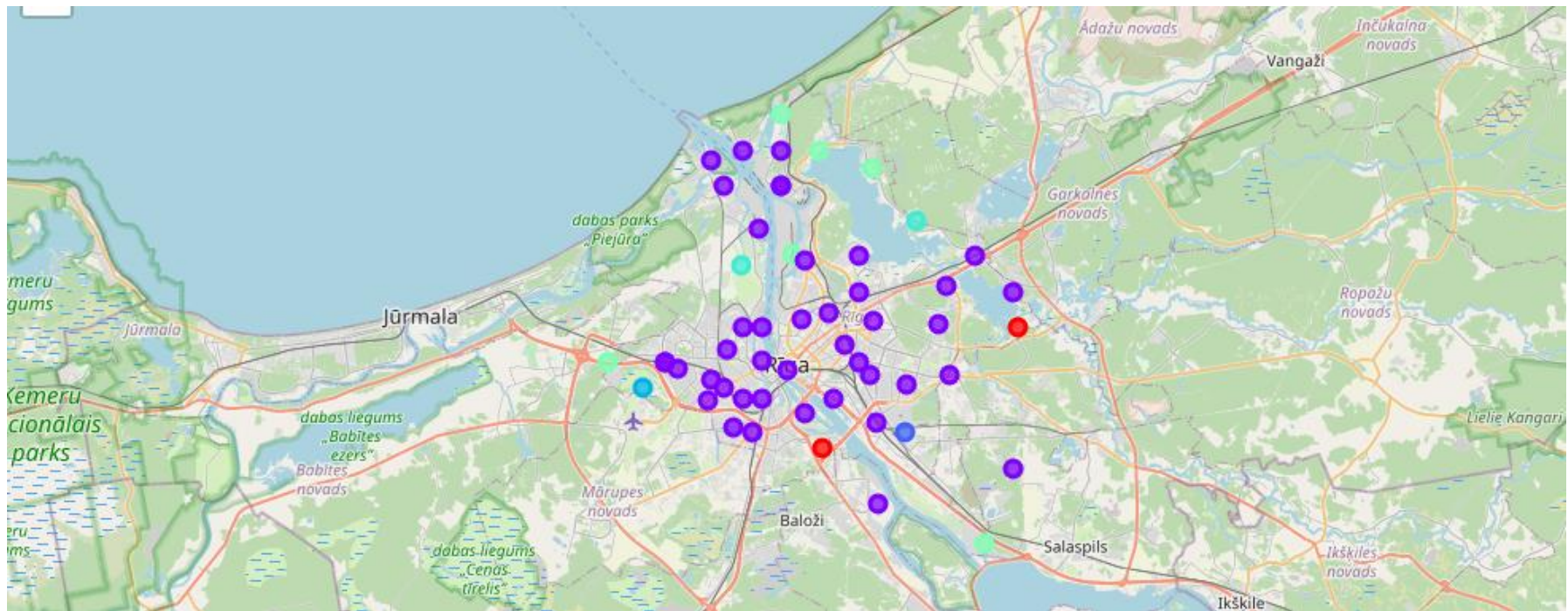
Finally, clustering analysis using the k-Means algorithm was performed in order to categorize similar boroughs into clusters. This algorithm needs to know the number of clusters it should form. It is hard to determine appropriate number of clusters. In order to determine number for Riga and Toronto, elbow method was used.

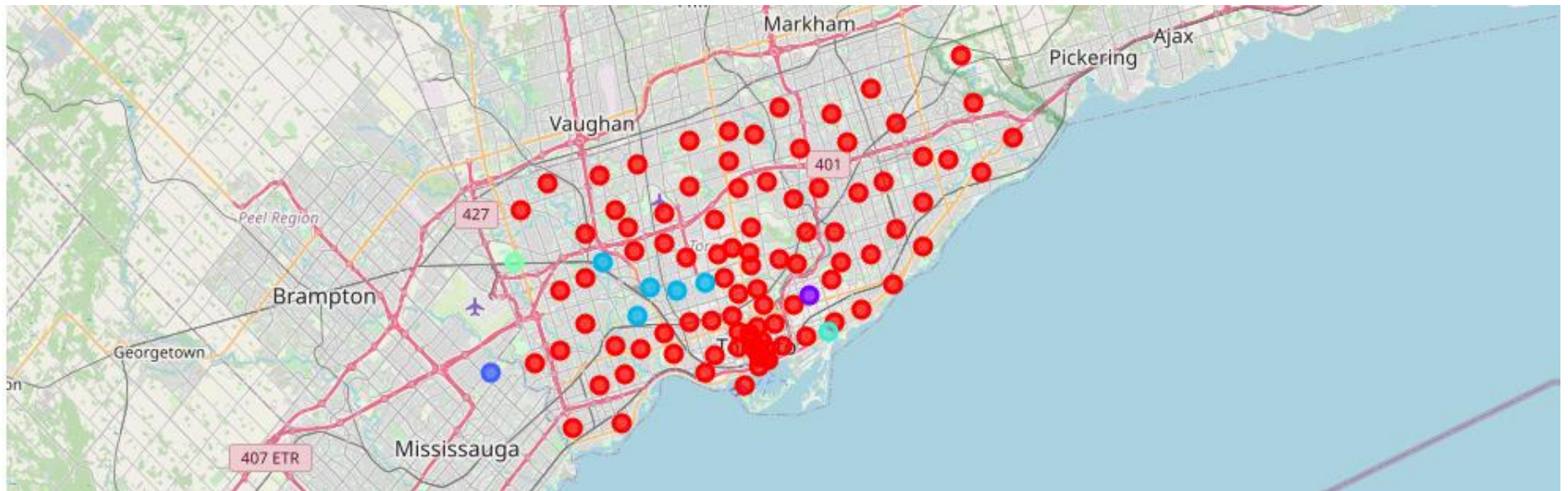


Cluster count was set to 6 for both cities and k-Means algorithm performed.

Results and discussion

The results of work with data and k-Means algorithm were plotted in a map using folium. As one can see in this and the following page, k-Means algorithm indeed clustered 6 different sets of boroughs together. In the results of Riga, it can be seen that those boroughs, which are in the center, are in one cluster. In the outer parts of the city are located different clusters, which indicate that there is a difference between the center of the city and the outskirts. For Toronto, it seems that there is one dominating cluster which is located not only in the center, but further in the city.





It could be argued that a smaller count of K could be selected, as it seems that there is the main cluster and a few minor ones. Let's look at what kind of venues are more popular in Riga and what – in Toronto.

If we look at the main cluster for Riga, which the algorithm clustered together and sort it by Population in descending order, then one can see that there are boroughs in which the most common venue is Bus Stop (Kengarags, Jugla, Vecmīlgrāvis). For two of those boroughs second comes Pizza places. It could mean that these boroughs have good public transportation coverage and there are only few places where one could go and eat out. Also, in boroughs which can be categorized as residential areas (Purvciems, Kengarags, Imanta, Pļavnieki, Teika and others) at most one can find a pizza place, buffet, diner, café, fast food and similar places, but for restaurants one should look at Centrs (this is the city center) as the most common types of venues are bars, different kinds of restaurants, hotels and cafés.

	Cluster Labels	Borough	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
29	1	Purvciems	ATM	Sporting Goods Shop	Playground	Rental Car Location	Food Truck	Auto Workshop	Light Rail Station	Buffet	Diner	Farm
50	1	Kengarags	Bus Stop	Train Station	Electronics Store	Shopping Mall	Lounge	Grocery Store	Pedestrian Plaza	Gym / Fitness Center	Clothing Store	Convenience Store
16	1	Imanta	Casino	Tennis Court	Hockey Arena	Gym	Photography Studio	Café	Cosmetics Shop	Motel	Light Rail Station	Playground
31	1	Pļavnieki	Grocery Store	Electronics Store	Café	Farmers Market	Food & Drink Shop	Park	Bank	Gym	Fast Food Restaurant	Casino
9	1	Centrs	Bar	Restaurant	Eastern European Restaurant	Hotel	Café	Cocktail Bar	Plaza	Italian Restaurant	Clothing Store	Modern European Restaurant
46	1	Ziepniekkalns	Casino	Tennis Court	Hockey Arena	Gym	Photography Studio	Café	Cosmetics Shop	Motel	Light Rail Station	Playground
23	1	Maskavas foršate	Fast Food Restaurant	Clothing Store	Electronics Store	Flower Shop	Sporting Goods Shop	Furniture / Home Store	Café	Cosmetics Shop	Sushi Restaurant	Shopping Mall
38	1	Teika	Café	Pizza Place	Clothing Store	Arts & Crafts Store	Bed & Breakfast	Dessert Shop	Bar	Yoga Studio	Dive Spot	Fish Market
48	1	Āgenskalns	Bakery	Hotel	Bus Stop	Spa	Fast Food Restaurant	Bookstore	Farmers Market	Eastern European Restaurant	Café	Medical Center
19	1	Jugla	Bus Stop	Pizza Place	Kebab Restaurant	Fast Food Restaurant	Fish Market	Beach	Café	Bakery	Recreation Center	BBQ Joint
17	1	Ilģuciemš	Spa	Market	Gym / Fitness Center	Café	Bakery	Fast Food Restaurant	Farmers Market	Farm	Event Space	Electronics Store
42	1	Vecmīlgrāvis	Bus Stop	Pizza Place	Pharmacy	Park	Dance Studio	Bakery	Farmers Market	Café	Scenic Lookout	Performing Arts Venue
47	1	Zolitūde	Pool	Supermarket	Food & Drink Shop	Soccer Field	Bus Line	Comedy Club	Comfort Food Restaurant	Farm	College Cafeteria	Event Space

But, if we look at the Toronto data, the dissimilarity between the two cities is evident. Most common type of venue among the boroughs is coffee shop, followed by different kinds of restaurants, pizza places, parks, banks, bakeries and breweries. It is evident that venue types across the table of Toronto is more homogenous than in Riga. In Riga, the city center is populated with entertainment and plethora of dining options, while in Toronto, these options are common in both the center and in outskirts of the city. It is possible that this is common in more populated areas, as the businesses can depend on a lot of residents which may choose to go for a coffee in a local café or coffee shop rather than travel and endure the challenges of traffic to visit a more central coffee shop. Also, the number of those options near to home or work, provides with a choice of a venue rather than just going to a place just because it's the only one in the close proximity. Another factor that can contribute to these results is the prosperity of the whole country. In the 2020 World Happiness Report which considers GDP, social support, health life expectancy, freedom to make life choices, generosity and perceptions of corruption, Canada ranked 11 while Latvia ranked only 57 out of 153 countries. It can be concluded that citizens of Canada are more prosperous than citizens of Latvia, which could explain why there is dissimilarity between the most common venue types. It is possible that may be people in Riga do not have the financial means to, for example, frequently buy coffee outside of home, and they see that it is more reasonable to make coffee at home rather than spend the extra money. It would be interesting to see how frequently people in Riga and people in Toronto visit restaurants and see if that has impact on the number of restaurants in the city. Another factor might be cultural differences. It is possible that people in Riga see a trip to the restaurant as a event for special occasions not just for the experience and fun, while people in Toronto frequent restaurants on a more regular basis and don't see it as something special.

	Borough	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Central Toronto	Coffee Shop	Pizza Place	Sandwich Place	Park	Café	Sushi Restaurant	Restaurant	Gym	Dessert Shop	Italian Restaurant
1	Downtown Toronto	Coffee Shop	Café	Restaurant	Hotel	Bakery	Japanese Restaurant	Italian Restaurant	Park	Clothing Store	Bar
2	Downtown Toronto Stn A	Coffee Shop	Seafood Restaurant	Italian Restaurant	Japanese Restaurant	Restaurant	Pub	Café	Bakery	Creperie	Cheese Shop
3	East Toronto	Coffee Shop	Greek Restaurant	Italian Restaurant	Ice Cream Shop	Brewery	Café	American Restaurant	Bakery	Pub	Bookstore
4	East Toronto Business	Light Rail Station	Comic Shop	Farmers Market	Restaurant	Brewery	Auto Workshop	Garden	Garden Center	Burrito Place	Fast Food Restaurant
5	East York	Coffee Shop	Bank	Burger Joint	Sporting Goods Shop	Pizza Place	Intersection	Park	Beer Store	Athletics & Sports	Skating Rink
6	East York/East Toronto	Intersection	Park	Convenience Store	Eastern European Restaurant	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Drugstore	Dumpling Restaurant
7	Etobicoke	Pizza Place	Sandwich Place	Coffee Shop	Grocery Store	Bakery	Fast Food Restaurant	Café	Gym	Pharmacy	Liquor Store
8	Etobicoke Northwest	Truck Stop	Garden Center	Rental Car Location	Bar	Drugstore	Donut Shop	Discount Store	Distribution Center	Dog Run	Doner Restaurant
9	Mississauga	Coffee Shop	Hotel	Gym	American Restaurant	Fried Chicken Joint	Gas Station	Sandwich Place	Middle Eastern Restaurant	Intersection	Burrito Place
10	North York	Coffee Shop	Clothing Store	Fast Food Restaurant	Restaurant	Pizza Place	Bank	Park	Japanese Restaurant	Sandwich Place	Shopping Mall
11	Queen's Park	Coffee Shop	Sushi Restaurant	Yoga Studio	College Cafeteria	Beer Bar	Spa	Smoothie Shop	Sandwich Place	Burrito Place	Café
12	Scarborough	Fast Food Restaurant	Bank	Bakery	Coffee Shop	Chinese Restaurant	Intersection	Breakfast Spot	Pizza Place	Pharmacy	Cosmetics Shop
13	West Toronto	Café	Bar	Coffee Shop	Restaurant	Italian Restaurant	Bakery	Breakfast Spot	Pizza Place	Park	Vegetarian / Vegan Restaurant

Conclusion

Analysis of city data was performed using Webscraping. Further, exploratory analysis, visualizations and data summaries were created using matplotlib pyplot, pandas and folium libraries. Data about venues was obtained from Foursquare API. Clustering analysis was performed on the gathered data using k-Means algorithm. Analysis showed how dissimilar the city of Riga and Toronto are.

One can conclude that if you travel from Riga to Toronto, you can expect to see a lot of coffee shops, different kinds of restaurants and a densely populated city compared to Riga. If you choose a hotel that is not in the center, you shouldn't be worried that there won't be any places to go to get coffee or dinner. If you seek to relocate from Riga to Toronto and you are living currently in a sparsely populated borough, then you should be prepared to encounter a lively city and even if you don't live in the center of the city, you will still be able to go out. On the other hand, if you prefer a more peaceful setting, maybe it is worth looking at the near-Toronto area. For travelers from Toronto to Riga it is advised to look for hotels near city center, as it would be more convenient in terms of entertainment and going out. There is a good public transportation infrastructure, but for foreigners it can be difficult to navigate. Regarding relocation, compared to Toronto, a lot of boroughs in Riga are not that densely populated, so for seekers of peace and quiet, it can be a possibly attractive option.

Overall, these conclusions are very broad and high-level rather than a detailed report, but from the data we can make sense of overall tendency in both cities. More data should be gathered in order to conclude concrete statements - data about income, population, transportation structure, economy, rent etc.