

Choosing a location for a pet care center in Toronto

IBM Data Science Professional
Certificate

Capstone Project

Business problem

Company

A Pet Company is planning to establish a new pet service in **Toronto**, offering:

- Grooming, washing
- Pet sitting
- Dog walking
- Pet supplies sale
- And more

Context

An ideal location for the venue:

- In a neighbourhood with the largest number of pets
- Dog runs nearby
- Less competitor venues nearby

Problem statement

Using available public data, conduct analysis and suggest 5 best locations for *A Pet Company* center in Toronto.

Data

Neighborhoods

A [Wikipedia page](#), scraped with BeautifulSoup to obtain the table of postal codes and to transform the data into a pandas dataframe.

Registered Pets

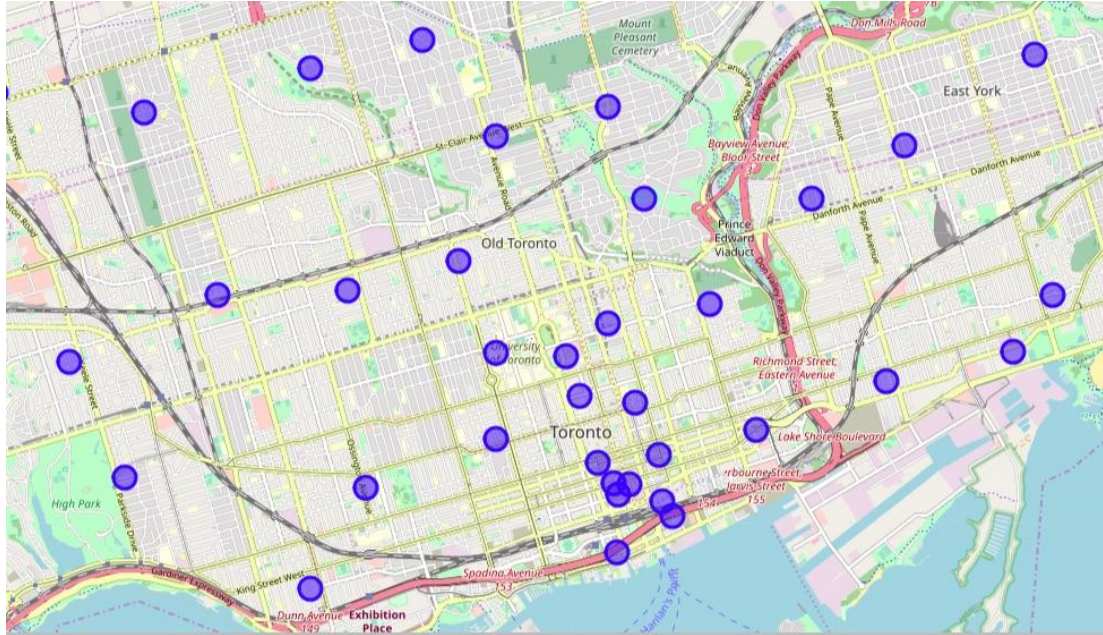
Public dataset from Toronto city government:
[Licensed Dogs and Cats Reports](#)

Existing venues

Number of pet shops, pet services and vets and their location in every neighborhood will be obtained using **Foursquare API** and then used for k-means clustering.

Methodology and Insights

The neighborhoods are grouped by zip code



Total number of postal code areas: 103

Neighborhoods with the highest no. of pets

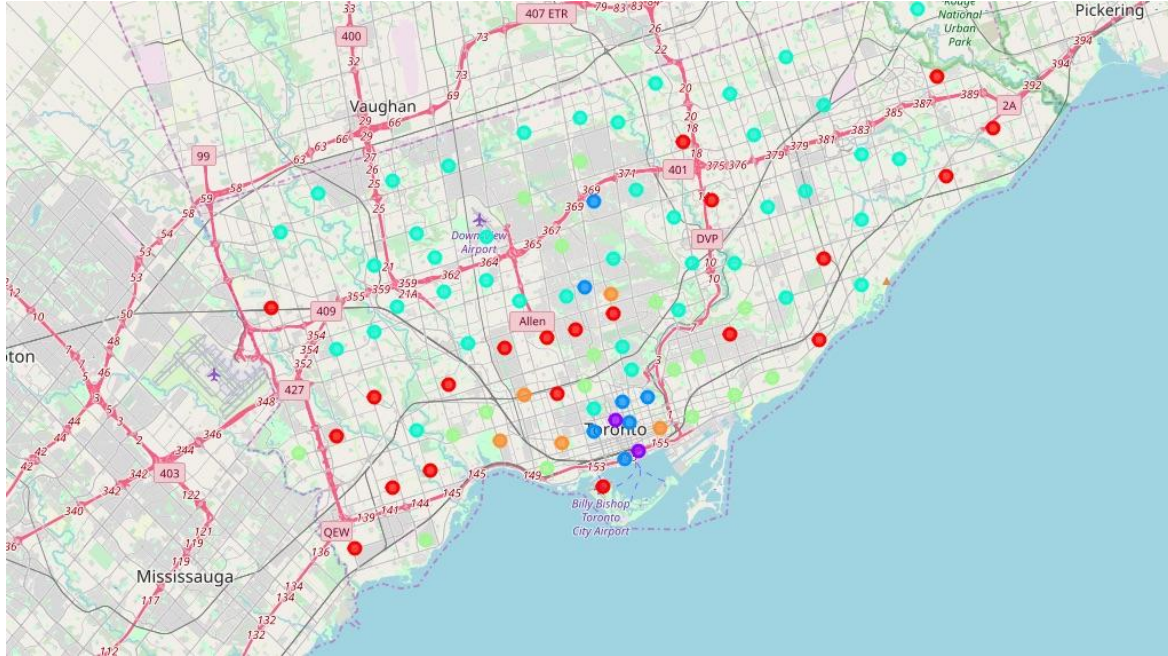
	PostalCode	Borough	Neighborhood	Latitude	Longitude	Pets
0	M4C	East York	Woodbine Heights	43.695344	-79.318389	1864
1	M8V	Etobicoke	Mimico South, New Toronto, Humber Bay Shores	43.605647	-79.501321	1848
2	M4L	East Toronto	The Beaches West, India Bazaar	43.668999	-79.315572	1828
3	M6P	West Toronto	High Park, The Junction South	43.661608	-79.464763	1809
4	M6H	West Toronto	Dovercourt Village, Dufferin	43.669005	-79.442259	1761
5	M6S	West Toronto	Swansea, Runnymede	43.651571	-79.484450	1693
6	M4J	East York	East Toronto	43.685347	-79.338106	1681
7	M4E	East Toronto	The Beaches	43.676357	-79.293031	1611
8	M2N	North York	Willowdale South	43.770120	-79.408493	1547
9	M6N	York	The Junction North, Runnymede	43.673185	-79.487262	1487

No. of venues and parks (Foursquare data)

	PostalCode	Borough	Neighborhood	Latitude	Longitude	Pets	NumberOfVenues	NearbyParks	VenuesPerPet
0	M4C	East York	Woodbine Heights	43.695344	-79.318389	1864	1.0	0.0	0.000536
1	M8V	Etobicoke	Mimico South, New Toronto, Humber Bay Shores	43.605647	-79.501321	1848	2.0	0.0	0.001082
2	M4L	East Toronto	The Beaches West, India Bazaar	43.668999	-79.315572	1828	4.0	0.0	0.002188
3	M6P	West Toronto	High Park, The Junction South	43.661608	-79.464763	1809	3.0	1.0	0.001658
4	M6H	West Toronto	Dovercourt Village, Dufferin	43.669005	-79.442259	1761	2.0	4.0	0.001136

- Most neighborhoods in downtown Toronto have very less or even no pets. We won't consider them for our business location.
- The most pet populated neighborhoods are in Etobicoke and West Toronto.
- All the pet populated neighborhoods have at least one pet shop or pet service.
- Not all pet populated neighborhoods have dog parks. We won't consider them for our business location.

K-means clustering



The model produced 6 clusters of neighborhoods.

K-means clustering

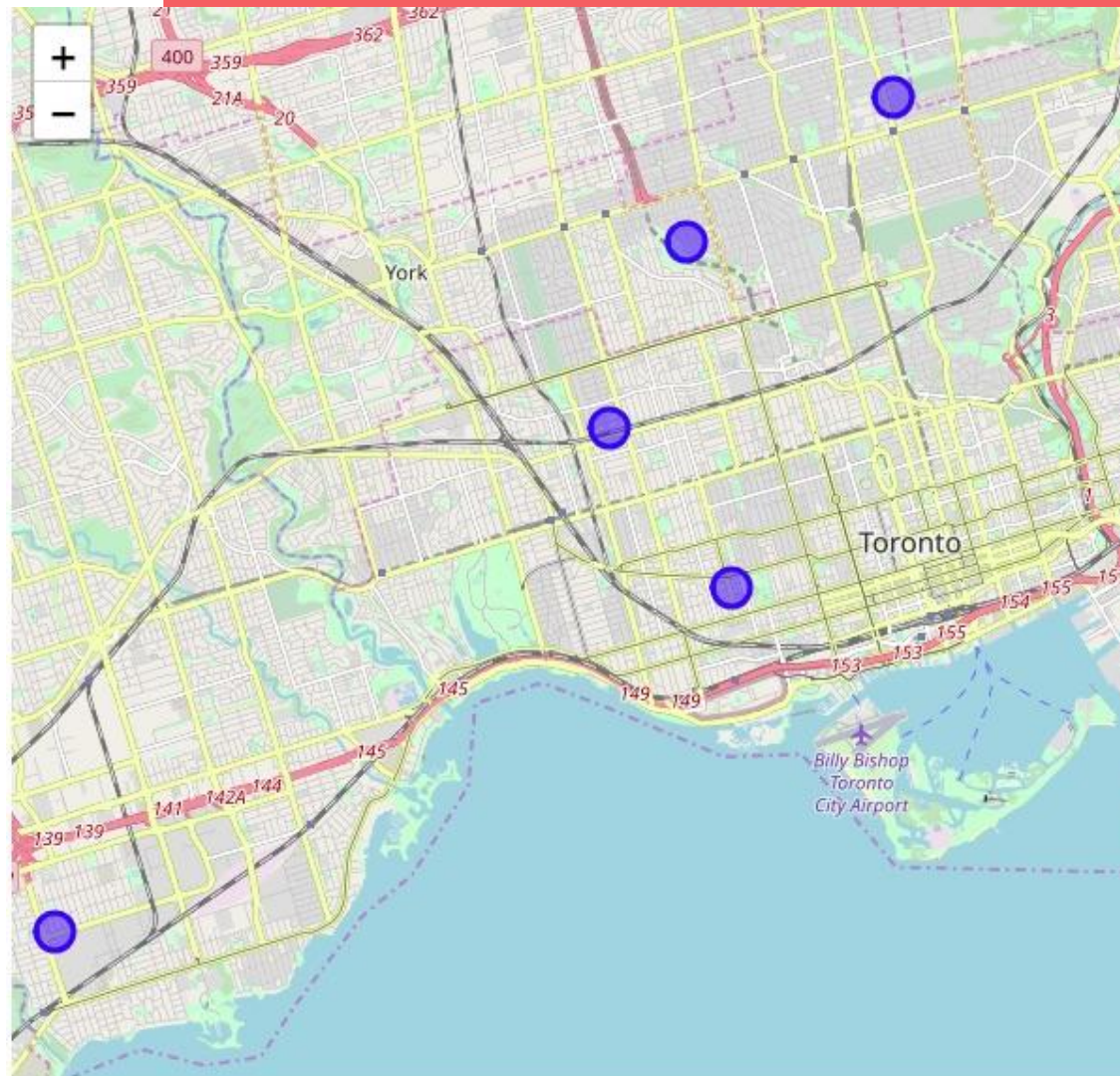
	Cluster Labels	Latitude	Longitude	Pets	NumberOfVenues	NearbyParks	VenuesPerPet	
0	0	43.697682	-79.395005	1174.904762	0.714286	0.238095	0.000664	←
1	1	43.651362	-79.380345	180.000000	3.500000	2.000000	0.020184	
2	2	43.679022	-79.388186	459.000000	3.714286	1.428571	0.008312	
3	3	43.739974	-79.395566	550.404762	0.309524	0.142857	0.000607	
4	4	43.682428	-79.402669	1357.588235	3.294118	0.294118	0.002612	←
5	5	43.666580	-79.413833	1274.200000	2.800000	3.400000	0.002284	←

We take a closer look at clusters 0, 4 and 5 and select the best neighborhoods within them.

- Have park
- Less competitors
- Low venues per pet ratio

Best Neighborhoods

- ★ Alderwood (Etobicoke)
- ★ Long Branch (Etobicoke)
- ★ Humewood-Cedarvale (York)
- ★ Dovercourt Village (West Toronto)
- ★ Dufferin (West Toronto)
- ★ Little Portugal (West Toronto)
- ★ Trinity (West Toronto)
- ★ Davisville North (Central Toronto)



Conclusion and future directions

- Built quick and useful model to cluster neighborhoods.
- Accuracy and freshness of the pet data has room for improvement.
- For better model accuracy and more insights, other features can be included in the future studies:
 - Population by neighborhood/pet per capita
 - Walk score of neighborhood
 - Foot traffic
 - Real estate prices