

Finding the best location to open a veterinary clinic in Toronto

Nebojsa Gogic

February 5, 2021

Table of Contents

1. Introduction	3
1.1. Background	3
1.2. Business Problem	3
2. Data	3
2.1. The licensed dogs and cats report	3
2.2. Pets Data cleaning	4
2.3. Cats and Dogs Data Exploration	5
2.4. Toronto Neighborhoods Data	6
2.5. Neighborhoods Data cleaning	6
2.6. Using FourSquare API to get veterinary clinics in Toronto	7
3. Methodology	8
4. Results	13
5. Discussion	15
6. Conclusion	15
6.1. Future Works	15
7. References	16

1. Introduction

1.1. Background

Starting a business is a difficult endeavor. One of the challenges is finding a good location. Depending on the size of the office, location, and other factors, expect to spend a large amount of money on renovations, equipment, fixtures, and more. This report will be targeted to veterinarians interested in opening a veterinary clinic in Toronto.

1.2. Business Problem

The number of registered dogs in Toronto is around 80,000 which is approximately one third of total dogs in Toronto and it is continuously growing. And not only dogs, the number of cats and other pets is increasing.

There are many veterinary clinics in Toronto but there are areas in Toronto that have a relatively small number of veterinary services compared to number of pets.

Choosing an area that is not saturated with clinics and with lots of pets will help ensure regular business. We will try to find the locations with no veterinary clinics or low number of veterinary services compared to number of pets. We will calculate the average year over year pets number increase rate as an additional parameter for detecting a good location.

Using data science, we will detect the best neighborhoods to open a veterinary clinic based on the above criteria. We will then explore and highlight advantages of few neighborhoods so that the stakeholders can choose the best location for the veterinary clinic.

2. Data

To solve the business problem, we used the following data:

1. The licensed dogs and cats report by forward station area (FSA) from the Toronto Open Data Portal [4]
2. List of Toronto neighborhoods with FSA from Wikipedia [3]
3. Number and location of existing veterinary clinics in each neighborhood, obtained using the FourSquare API [2]

2.1. The licensed dogs and cats report

This report contains data by forward station area from 2013 to September 2020. The following table displays sample data from 2019.

FSA	CAT	DOG	Total
M1A	0	1	1
M1B	236	511	747
M1C	264	700	964
M1E	415	803	1,218
M1G	169	325	494

We only need the FSA and Total columns. We merged all the data into a data frame:

```
df_cats_dogs.head()
```

	FSA	Total_2013	Total_2014	Total_2015	Total_2016	Total_2017	Total_2018	Total_2019	Total_2020
0	M0M	1.0	1.0	NaN	NaN	NaN	NaN	NaN	NaN
1	M1B	967.0	961.0	905.0	919.0	912.0	872.0	747.0	441.0
2	M1C	1354.0	1332.0	1227.0	1190.0	1072.0	1053.0	964.0	591.0
3	M1E	1621.0	1574.0	1494.0	1447.0	1430.0	1386.0	1218.0	666.0
4	M1G	709.0	654.0	655.0	609.0	605.0	572.0	494.0	291.0

The data frame contains 96 rows and 11 columns

2.2. Pets Data cleaning

We removed the outliers. These were the rows with very low cats and dogs numbers. Removing these rows, the null values were automatically removed.

```
df_cats_dogs.describe()
```

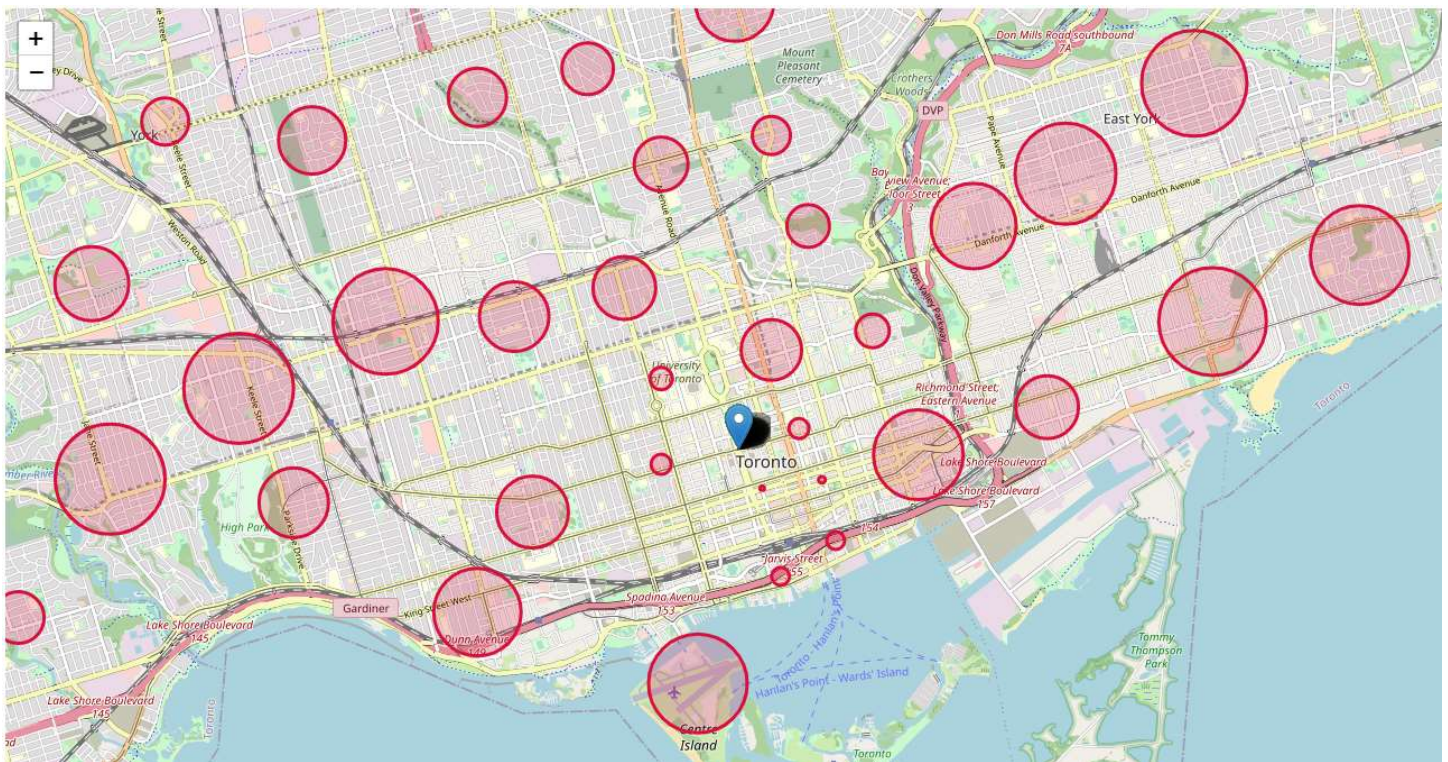
	Total_2013	Total_2014	Total_2015	Total_2016	Total_2017	Total_2018	Total_2019	Total_2020
count	96.000000	96.000000	96.000000	96.000000	96.000000	96.000000	96.000000	96.000000
mean	839.593750	834.375000	831.156250	831.437500	843.208333	806.833333	726.510417	449.625000
std	460.999431	455.704943	451.604935	452.87121	460.164888	442.141598	401.971903	253.004608
min	30.000000	34.000000	34.000000	33.000000	44.000000	49.000000	46.000000	43.000000
25%	519.000000	517.250000	515.750000	499.750000	511.500000	486.000000	429.500000	246.500000
50%	805.000000	787.500000	786.500000	787.500000	803.500000	744.000000	650.000000	394.000000
75%	1056.250000	1106.000000	1135.750000	1190.500000	1200.750000	1133.250000	995.250000	628.250000
max	1880.000000	1879.000000	1904.000000	1869.000000	1864.000000	1805.000000	1694.000000	1022.000000

2.3. Cats and Dogs Data Exploration

Using Google Geolocation API, we added the latitude and longitude of each FSA to the dataset:

	FSA	Total_2013	Total_2014	Total_2015	Total_2016	Total_2017	Total_2018	Total_2019	Total_2020	Latitude	Longitude
0	M1B	967.0	961.0	905.0	919.0	912.0	872.0	747.0	441.0	43.806686	-79.194353
1	M1C	1354.0	1332.0	1227.0	1190.0	1072.0	1053.0	964.0	591.0	43.784535	-79.160497
2	M1E	1621.0	1574.0	1494.0	1447.0	1430.0	1386.0	1218.0	666.0	43.763573	-79.188711
3	M1G	709.0	654.0	655.0	609.0	605.0	572.0	494.0	291.0	43.770992	-79.216917
4	M1H	545.0	545.0	512.0	500.0	464.0	437.0	405.0	212.0	43.773136	-79.239476

We used Folium maps to visualize the number of pets:



2.4. Toronto Neighborhoods Data

We used BeautifulSoup library to extract the data from Wikipedia

	Postal Code	Borough	Neighborhood
1	M1A\n	Not assigned\n	Not assigned\n
2	M2A\n	Not assigned\n	Not assigned\n
3	M3A\n	North York\n	Parkwoods\n
4	M4A\n	North York\n	Victoria Village\n
5	M5A\n	Downtown Toronto\n	Regent Park, Harbourfront\n

2.5. Neighborhoods Data cleaning

We removed “Not assigned” Neighborhoods from the dataset and used the Google Geolocation API to add latitude and longitude to the dataset. Then we calculated the distance from each neighborhood to Downtown and filtered out neighborhoods that are not within 7,000 meters range from downtown:

	FSA	Borough	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Distance to Downtown
2	M5A	Downtown Toronto	Regent Park	43.654260	-79.360636	2232.736527
4	M7A	Downtown Toronto	Queen's Park	43.662301	-79.389494	838.466754
9	M5B	Downtown Toronto	Garden District	43.657162	-79.378937	799.642642
15	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	1102.638211
16	M6C	York	Humewood-Cedarvale	43.693781	-79.428191	5394.533092

2.6. Using FourSquare API to get veterinary clinics in Toronto

We used FourSquare API to find all venues in radius of 7 km of downtown Toronto. Then, we found all the venues within 1.5 km range from the center of each neighborhood. This is the venue count that we used as one of the good location candidate criteria.

```
import numpy as np
all_veterinary_venues = get_venues("4d954af4a243a5684765b473", 43.6548046, -79.3883031, CLIENT_ID, CLIENT_SECRET, VERSION, radius=7000, limit=500)
arr = np.array(all_veterinary_venues)
df_all_veterinary_venues = pd.DataFrame(arr, columns = ['Venue Id', 'Venue Name', 'Venue Category', 'Venue Latitude', 'Venue Longitude'])
```

```
df_all_veterinary_venues.head()
```

	Venue Id	Venue Name	Venue Category	Venue Latitude	Venue Longitude
0	4ae06ddef964a5203b7f21e3	Toronto Humane Society	Animal Shelter	43.65769802359708	-79.3565828808708
1	4b81644af964a52049a330e3	VEC Veterinary Emergency Clinic	Veterinarian	43.67397762493825	-79.38973071994448
2	4bd1dc14046076b0f0a77271	Spadina Animal Hospital	Veterinarian	43.64722300237428	-79.39565838078587
3	4db09bf70437a93f7f770755	Wellesley Animal Hospital	Veterinarian	43.66494103924063	-79.38458123431751
4	516ff685e4b03e954dae2613	Sherbourne Animal Hospital	Veterinarian	43.653999	-79.369174

We calculated the average year over year rate of change and venue density. Then, we merged neighborhoods and pets data into a single dataset:

```
df_neighborhoods_pets.head()
```

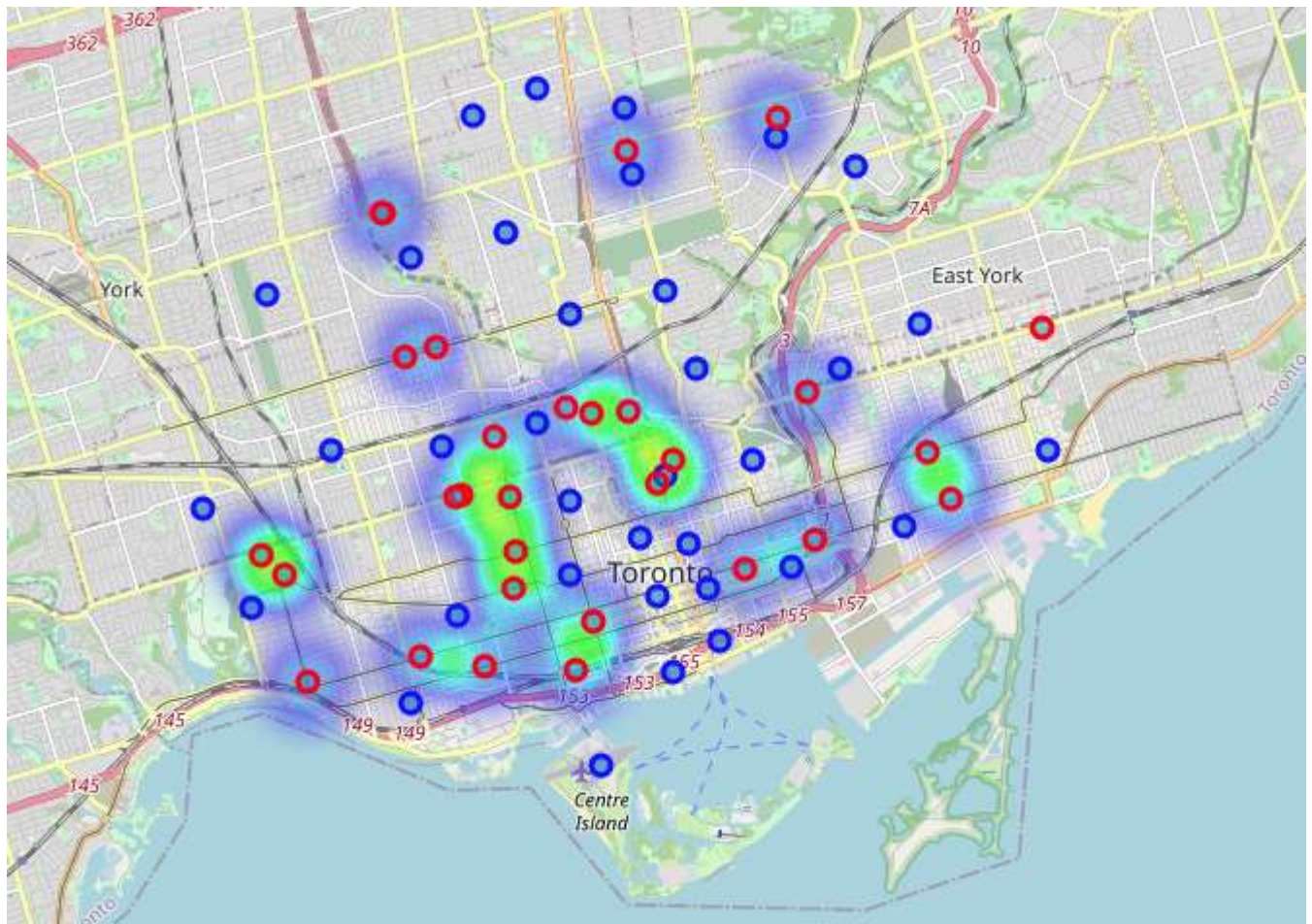
	FSA	Borough	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Distance to Downtown	Total_2013	Total_2014	Total_2015	Total_2016	Total_2017	Total_2018	Total_2019	Total_2020	Venue Count	Venue Density	avg_roc
0	M5A	Downtown Toronto	Regent Park	43.654260	-79.360636	2232.736527	966.0	1069.0	1192.0	1219.0	1365.0	1355.0	1307.0	837.0	2.0	0.238949	5.355950
1	M5B	Downtown Toronto	Garden District	43.657162	-79.378937	799.642642	196.0	212.0	225.0	248.0	265.0	254.0	259.0	186.0	3.0	1.612903	4.864994
2	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	1102.638211	78.0	76.0	70.0	82.0	91.0	88.0	73.0	52.0	1.0	1.923077	-0.447088
3	M6C	York	Humewood-Cedarvale	43.693781	-79.428191	5394.533092	983.0	934.0	939.0	964.0	964.0	939.0	812.0	531.0	4.0	0.753296	-2.984232
4	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306	1645.178245	138.0	150.0	187.0	223.0	229.0	244.0	246.0	176.0	1.0	0.568182	10.445688

3. Methodology

We used the following features to detect areas in Toronto that are good candidates for opening a veterinary service:

- Number of existing veterinary services in each neighborhood.
We found all the venues within 1.5 km range from the center of the Neighborhoods
- Venue density - the ratio between the number of existing veterinary services and total number of cats and dogs in the area.
- The average annual growth rate of number of pets.
We used the data that we collected from the cats and dogs reports for 2013 to 2020 to calculate this value for each neighborhood.

We displayed neighborhoods and venues on a map. We used Folium heat map feature to display density of the veterinary venues. Areas with low density are candidates for good locations.

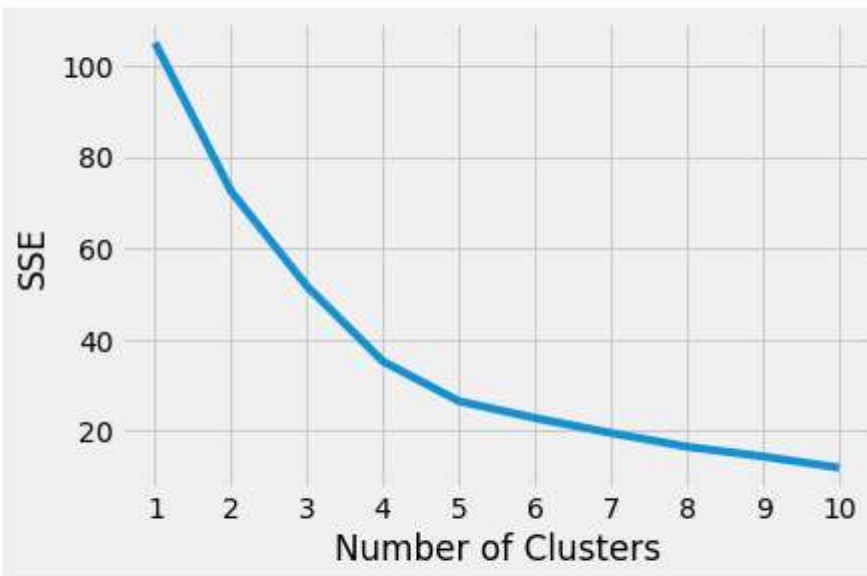


Since we had unlabeled data that needed to be categorized, we decided to use the K-Means clustering to find the groups in the data.

We used the total number of pets in 2020 and number of venues in each neighborhood as the features for clustering.

Using the “Elbow” method we determined the best value for number of clusters:

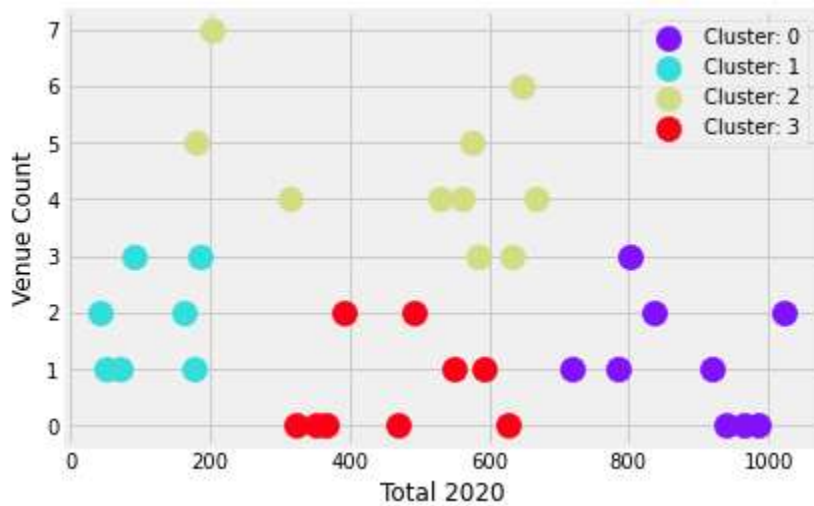
```
import matplotlib.pyplot as plt
plt.style.use("fivethirtyeight")
plt.plot(range(1, 11), sse)
plt.xticks(range(1, 11))
plt.xlabel("Number of Clusters")
plt.ylabel("SSE")
plt.show()
```



```
!pip install kneed
from kneed import KneeLocator
kl = KneeLocator(
    range(1, 11), sse, curve="convex", direction="decreasing"
)
print('Optimal number of clusters: ', kl.elbow)
```

Optimal number of clusters: 4

Let's examine the clusters:



Cluster 0 contains the largest number of pets and low number of venues (between 0 and 3)

Cluster 1 contains small number of pets and small number of venues

Cluster 2 is between clusters 0 and 1 in terms of pets numbers but contains the highest number of venues of all clusters

Cluster 3 is similar to cluster 2 in terms of pet numbers but contains low number of venues (between 0 and 2)

Let's see each cluster in more details:

Cluster 0:

	Cluster Labels	FSA	Borough	Neighborhood	Total_2020	Venue Count	Venue Density	avg_roc
26	0	M4S	Central Toronto	Davisville	720.0	1.0	0.138889	-1.506251
15	0	M4K	East Toronto	The Danforth West	784.0	1.0	0.127551	-2.541566
16	0	M6K	West Toronto	Brockton	801.0	3.0	0.374532	1.638185
0	0	M5A	Downtown Toronto	Regent Park	837.0	2.0	0.238949	5.355950
31	0	M5V	Downtown Toronto	CN Tower	918.0	1.0	0.108932	5.613540
12	0	M4J	East York	East Toronto	939.0	0.0	0.000000	-2.918603
11	0	M6H	West Toronto	Dufferin	966.0	0.0	0.000000	-1.241449
17	0	M4L	East Toronto	India Bazaar	985.0	0.0	0.000000	-1.675971
22	0	M6P	West Toronto	High Park	1022.0	2.0	0.195695	-0.319605

As we can see, the density of veterinary venues is very low compared to number of pets. Some neighborhoods have good positive year over year rate of change in pets registration and some have no veterinary venues in the range of 1,500m from the center of the neighborhood. The following neighborhoods are good candidates for the veterinary service location: CN Tower, India Bazaar, Dufferin, East Toronto

Cluster 1:

	Cluster Labels	FSA	Borough	Neighborhood	Total_2020	Venue Count	Venue Density	avg_roc
10	1	M5H	Downtown Toronto	Richmond	43.0	2.0	4.651163	8.161113
2	1	M5C	Downtown Toronto	St. James Town	52.0	1.0	1.923077	-0.447088
9	1	M4H	East York	Thornccliffe Park	70.0	1.0	1.428571	-1.833908
7	1	M5G	Downtown Toronto	Central Bay Street	92.0	3.0	3.260870	4.581226
13	1	M5J	Downtown Toronto	Harbourfront East	162.0	2.0	1.234568	3.395224
4	1	M5E	Downtown Toronto	Berczy Park	176.0	1.0	0.568182	10.445688
1	1	M5B	Downtown Toronto	Garden District	186.0	3.0	1.612903	4.864994

The density of veterinary venues is much higher compared to cluster 0, due to the low pets number. The **Berczy Park** has the average year over year rate of change in pets registration over 10%. We will take a look at this neighborhood later in this report.

Cluster 2:

	Cluster Labels	FSA	Borough	Neighborhood	Total_2020	Venue Count	Venue Density	avg_roc
29	2	M5T	Downtown Toronto	Kensington Market	179.0	5.0	2.793296	-0.757754
27	2	M5S	Downtown Toronto	University of Toronto	202.0	7.0	3.465347	1.018999
33	2	M4X	Downtown Toronto	St. James Town	314.0	4.0	1.273885	-1.737794
3	2	M6C	York	Humewood-Cedarvale	531.0	4.0	0.753296	-2.984232
34	2	M4Y	Downtown Toronto	Church and Wellesley	561.0	4.0	0.713012	3.946533
24	2	M5R	Central Toronto	The Annex	575.0	5.0	0.869565	-0.227871
18	2	M4M	East Toronto	Studio District	584.0	3.0	0.513699	0.975394
25	2	M6R	West Toronto	Parkdale	632.0	3.0	0.474684	-2.662169
8	2	M6G	Downtown Toronto	Christie	646.0	6.0	0.928793	-2.343140
14	2	M6J	West Toronto	Little Portugal	668.0	4.0	0.598802	0.210017

Cluster 2 contains more veterinary venues than the previous two clusters and the number of registered pets is similar to the number of pets in cluster 1. There are three neighborhoods in this cluster with over 600 registered pets in 2020 but the year over year is close to 0 or negative. This cluster of neighborhoods does not look like a good choice for a new veterinary service location compared to the previous two clusters.

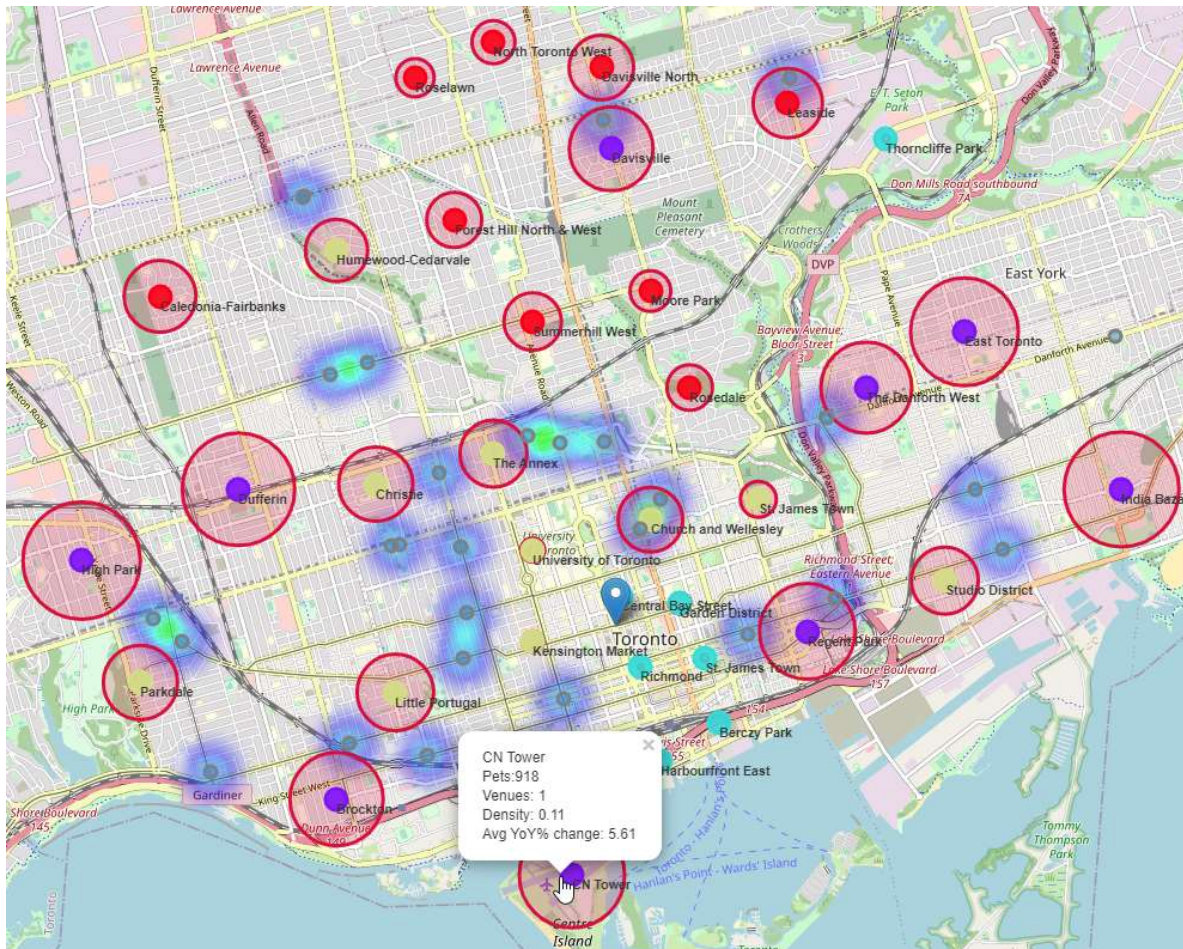
Cluster 3:

	Cluster Labels	FSA	Borough	Neighborhood	Total_2020	Venue Count	Venue Density	avg_roc
19	3	M5N	Central Toronto	Roselawn	324.0	0.0	0.000000	-1.178214
28	3	M4T	Central Toronto	Moore Park	352.0	0.0	0.000000	-2.740607
23	3	M4R	Central Toronto	North Toronto West	367.0	0.0	0.000000	-0.237512
32	3	M4W	Downtown Toronto	Rosedale	392.0	2.0	0.510204	-2.670048
21	3	M5P	Central Toronto	Forest Hill North & West	470.0	0.0	0.000000	0.020979
30	3	M4V	Central Toronto	Summerhill West	492.0	2.0	0.406504	-0.676786
20	3	M4P	Central Toronto	Davisville North	550.0	1.0	0.181818	2.401422
6	3	M4G	East York	Leaside	594.0	1.0	0.168350	-2.209854
5	3	M6E	York	Caledonia-Fairbanks	627.0	0.0	0.000000	-4.049466

In average, this cluster has the lowest venue count and the venue density is slightly higher than in cluster 0. There are a few neighborhoods with no veterinary venues. One of them, **Caledonia-Fairbanks** has 627 registered pets, but the average year over year is -4%. Other candidates from this cluster are the **North Toronto West** and **Roselawn** neighborhoods. We will display these locations on the map and compare with other neighborhoods. We will also take a look at few other neighborhoods in this cluster.

4. Results

Let's visualize the clusters on a map. This will help us to better understand the results:



Let's take a look at the location candidates from cluster 0:

CN Tower neighborhood: Located in downtown core, close to the major highways Gardiner and Don Valley Parkway, with 918 registered pets in 2020, only one existing veterinary venue and average YoY rate of change of 5.61% is definitely a good location for a veterinary service.

India Bazaar is another great neighborhood to open a veterinary venue. The average YoY rate of change in pet registration is -1.68% but there are no other venues nearby and the number of registered pets is 985. This neighborhood is also close to the major highways, Gardiner and DVP, which makes it easily accessible.

Dufferin neighborhood is very similar to India Bazaar. The number of registered cats and dogs in 2020 was 966. The average YoY rate of change in pet registration is -1.24% and there are no other venues nearby. This is clearly a good location to open a veterinary venue.

East Toronto neighborhood has slightly lower YoY rate of change in pet registration than the previous two neighborhoods that we mentioned above (-2.92%), but a large number of registered cats and dogs in 2020 (939) and no other venues nearby make it also a good location for a veterinary service.

Cluster 1:

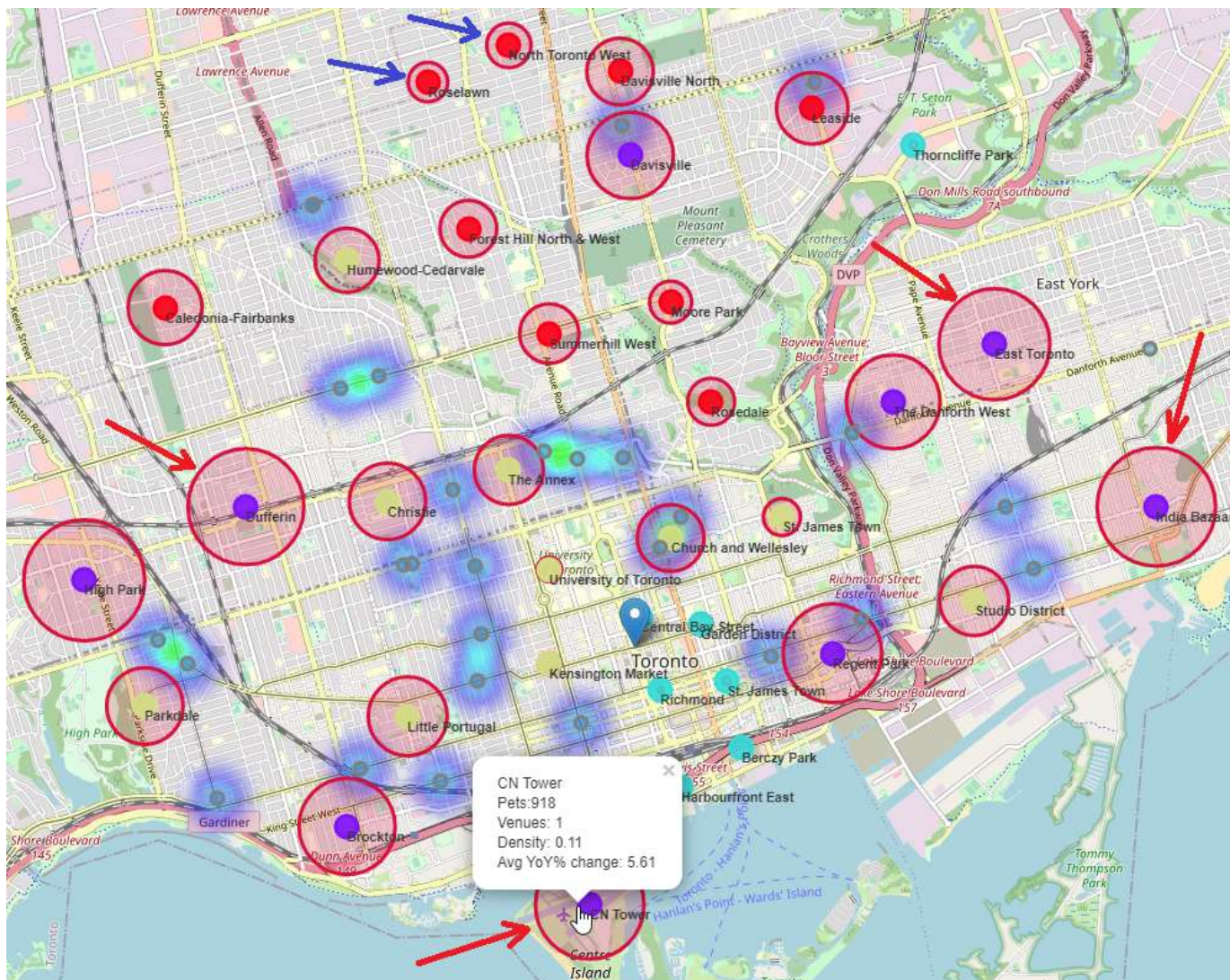
We identified **Berczy Park** as a possible candidate due to its average YoY rate of change of 10.45%. This neighborhood is located at downtown core, close to the major highways. It is an area with many parks and a fountain with statues of 27 dogs and a cat.

Cluster 2 does not look like a good choice for a veterinary service location, so let's take a look at neighborhoods from Cluster 3 that we identified as candidates for a veterinary venue:

Caledonia-Fairbanks has no veterinary venues nearby. The number of registered pets in 2020 was 627 which is a decent number but the average YoY rate of change in pets registration is -4.05% and the distance from downtown Toronto is longer than from some other good locations that we found (i.e. Dufferin neighborhood)

North Toronto West and **Roselawn** both have no veterinary venues nearby. The number of registered pets in 2020 is 367 and 324 respectively and the average YoY rate of change in pet registration is negative but close to 0, specially for the **North Toronto West** neighborhood. Both areas are good locations for opening a veterinary venue in Central Toronto borough, but it is relatively far from downtown Toronto compared to other good locations that we found.

Let's see the best location candidates on the map:



5. Discussion

The results indicate that despite many existing veterinary venues, there are neighborhoods in Toronto with only a few and with no veterinary services at all with a large number of registered pets.

Examples: East Toronto, Dufferin, India Bazaar in cluster 0 and Roselawn, Moore Park, North Toronto West and Caledonia-Fairbanks in cluster 3.

The average year over year increase rate of registered cats and dogs is highest in the downtown core. This might be surprising, but this is the pet friendly area with many parks near the Lake Ontario. This is also the area that is developing fast and the population is growing.

Based on data analysis that we performed in the previous section and location of the neighborhoods the following neighborhoods stand out as good candidates for a veterinary venue:

CN Tower neighborhood, India Bazaar, Dufferin, East Toronto, Berczy Park, Caledonia-Fairbanks, North Toronto West and Roselawn neighborhoods.

Berczy Park is a fast-developing neighborhood with high rate of increasing pet registration. It is a pet friendly area with parks and a fountain with statues of 27 dogs and a cat. The number of pets is relatively low and despite a fast-growing pets registration trend, we would not recommend this area as a perfect location, but the stakeholders should keep an eye on it.

North Toronto West and **Roselawn** are the second-best neighborhoods to open a veterinary venue. They both do not have any veterinary venues nearby, but they are relatively far from the downtown core. The blue arrows indicate the position of these locations on the above map.

Based on the venue density, pets number growth and geo location, we recommend the following neighborhoods as the best candidates for a veterinary venue: **CN Tower neighborhood, India Bazaar, Dufferin, and East Toronto**. The red arrows indicate the position of these locations on the above map.

6. Conclusion

Using the licensed dogs and cats report by forward station area (FSA) from the Toronto Open Data Portal, list of Toronto neighborhoods with FSA from Wikipedia, number and location of existing veterinary clinics in each neighborhood, obtained using the FourSquare API, K-Means clustering algorithm and data exploration and analysis we identified the locations that the stakeholders should be using as a starting point for the final exploration and decision.

6.1. Future Works

Improve the accuracy of the model by including the commercial rent and demographics data that can be obtained from Statistics Canada [2]. Commercial rents indexes for retail, office and industrial buildings is important information as well as the population density, structure, and growth.

7. References

1. Developer. (2020, September 09). Retrieved February 7, 2021, from <https://developer.foursquare.com/>
2. Government of Canada, S. C. (2021, February 07). Statistics Canada: Canada's national statistical agency. Retrieved February 7, 2021, from <https://www.statcan.gc.ca/eng/start>
3. List of postal codes of Canada: M. (2021, February 05). Retrieved February 7, 2021, from https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
4. Open Data Dataset. (2020, October 2). Retrieved February 7, 2021, from <https://open.toronto.ca/dataset/licensed-dogs-and-cats-reports/>