

Capstone Project: The Battle of Neighborhoods

Finding the Optimal Location for a Veterinary Clinic in Toronto

Introduction: Description of the business problem

- Business goal: opening a veterinary clinic in Toronto and finding the optimal location
- Importance of detailed market analysis
- Collect and process the data relevant to the business

Introduction: Description of the business problem

- Key questions:
- How many veterinary clinics exist in Toronto?
- Where they are located?
- How are they distributed, i.e. how many clinics there are in each city area?
- What is the population density?
- What is the correlation between the distribution of the existing veterinary clinics and the population density?

Introduction: Description of the business problem

- Assumptions:
- For a small start-up business it is best to have:
- (i) small number of similar businesses in the same area, and
- (ii) large population

Data: What data to use and where to get it from?

- List of FSAs (Forward Sortation Areas) assigned in Toronto:

https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

- Geospatial data: pair the FSAs with latitude and longitude

- Veterinary clinic information extracted from Foursquare:

https://api.foursquare.com/v2/venues/explore?&client_id={} &client_secret={} &v={} &ll={} &radius={} &limit={} &query={}

- Population by FSA available on the Statistics Canada website:

<https://www12.statcan.gc.ca/>

Results:

- FSAs assigned to Toronto (partial screenshot):

M1A <i>Not assigned</i>	M2A <i>Not assigned</i>	M3A North York (Parkwoods)	M4A North York (Victoria Village)	M5A Downtown Toronto (Regent Park / Harbourfront)	M6A North York (Lawrence Manor / Lawrence Heights)	M7A Queen's Park (Ontario Provincial Government)	M8A <i>Not assigned</i>	M9A Etobicoke (Islington Avenue)
M1B Scarborough (Malvern / Rouge)	M2B <i>Not assigned</i>	M3B North York (Don Mills) North	M4B East York (Parkview Hill / Woodbine Gardens)	M5B Downtown Toronto (Garden District, Ryerson)	M6B North York (Glencairn)	M7B <i>Not assigned</i>	M8B <i>Not assigned</i>	M9B Etobicoke (West Deane Park / Princess Gardens / Martin Grove / Islington / Cloverdale)
M1C Scarborough (Rouge Hill / Port Union / Highland Creek)	M2C <i>Not assigned</i>	M3C North York (Don Mills) South (Flemingdon Park)	M4C East York (Woodbine Heights)	M5C Downtown Toronto (St. James Town)	M6C York (Humewood-Cedarvale)	M7C <i>Not assigned</i>	M8C <i>Not assigned</i>	M9C Etobicoke (Eringate / Bloordale Gardens / Old Burnhamthorpe / Markland Wood)
M1E Scarborough (Guildwood / Morningside / West Hill)	M2E <i>Not assigned</i>	M3E <i>Not assigned</i>	M4E East Toronto (The Beaches)	M5E Downtown Toronto (Berczy Park)	M6E York (Caledonia-Fairbanks)	M7E <i>Not assigned</i>	M8E <i>Not assigned</i>	M9E <i>Not assigned</i>
M1G Scarborough (Woburn)	M2G <i>Not assigned</i>	M3G <i>Not assigned</i>	M4G East York (Leaside)	M5G Downtown Toronto (Central Bay Street)	M6G Downtown Toronto (Christie)	M7G <i>Not assigned</i>	M8G <i>Not assigned</i>	M9G <i>Not assigned</i>
M1H Scarborough (Cedarbrae)	M2H North York (Hillcrest Village)	M3H North York (Bathurst Manor / Wilson Heights / Downsview North)	M4H East York (Thorncilffe Park)	M5H Downtown Toronto (Richmond / Adelaide / King)	M6H West Toronto (Dufferin / Dovercourt Village)	M7H <i>Not assigned</i>	M8H <i>Not assigned</i>	M9H <i>Not assigned</i>
M1J	M2J	M3J	M4J	M5J	M6J	M7J	M8J	M9J

Results:

- Toronto FSAs dataframe:

	PostalCode	Borough	Neighbourhood
0	M1B	Scarborough	Malvern, Rouge
1	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek
2	M1E	Scarborough	Guildwood, Morningside, West Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae

- Toronto FSAs dataframe, with lat and long data:

	PostalCode	Borough	Neighbourhood	Latitude	Longitude
0	M1B	Scarborough	Malvern, Rouge	43.8113	-79.1930
1	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek	43.7878	-79.1564
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.7678	-79.1866
3	M1G	Scarborough	Woburn	43.7712	-79.2144
4	M1H	Scarborough	Cedarbrae	43.7686	-79.2389

Results:

- Venue information downloaded through dynamically built API calls to Forsquare:

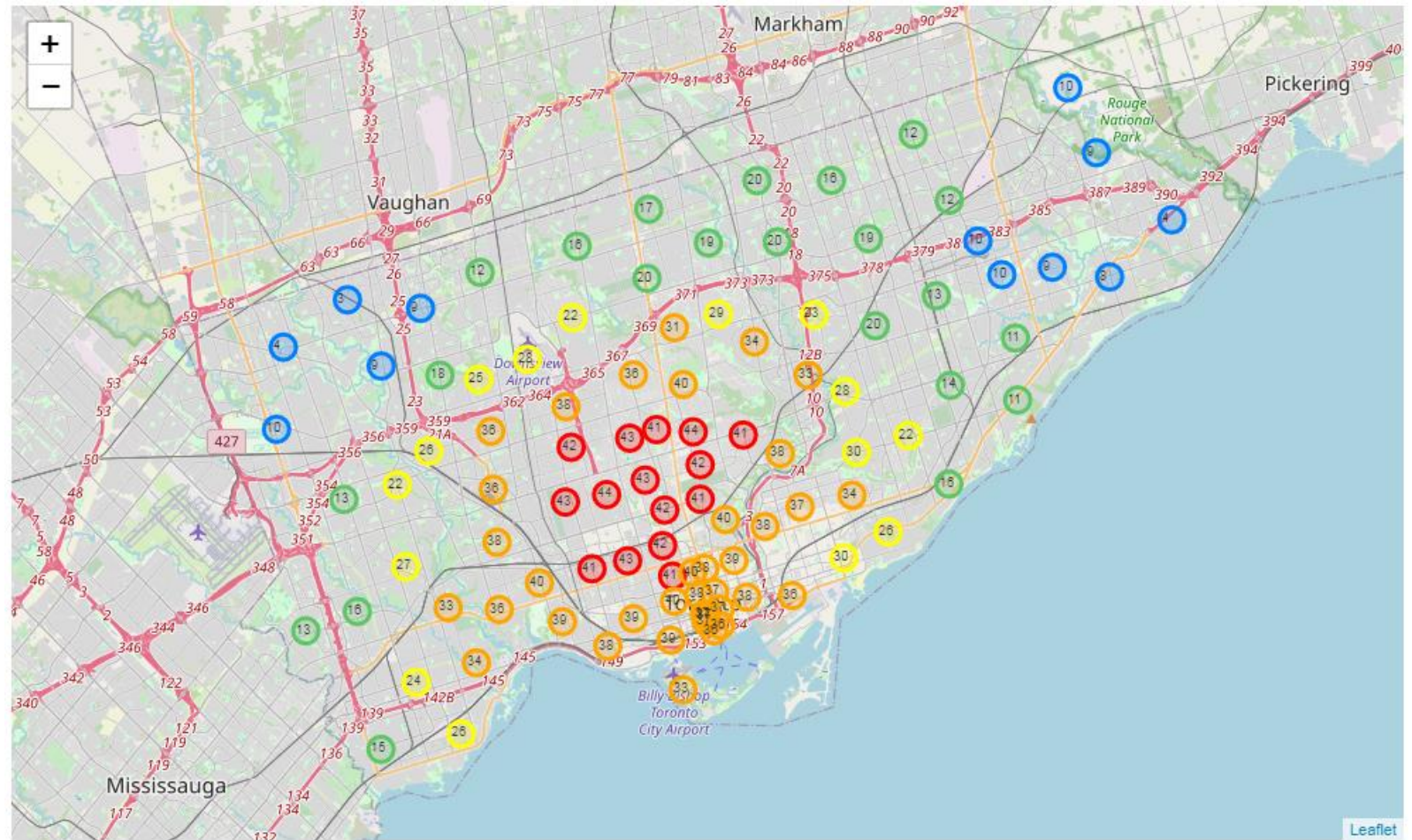
	PostalCode	Borough	Latitude	Longitude	VenueName	VenueCategory	VenueLat	VenueLong
0	M1B	Scarborough	43.8113	-79.193	West Hill Animal Clinic	Veterinarian	43.8113	-79.138002
1	M1B	Scarborough	43.8113	-79.193	Markham Road Animal Hospital	Veterinarian	43.8113	-79.229159
2	M1B	Scarborough	43.8113	-79.193	Whites Road Animal Hospital	Veterinarian	43.8113	-79.122314
3	M1B	Scarborough	43.8113	-79.193	Ashcott Veterinary Clinic	Veterinarian	43.8113	-79.290771
4	M1B	Scarborough	43.8113	-79.193	Markham Veterinary Clinic	Veterinarian	43.8113	-79.238510

- Venue count by FSA:

	PostalCode	VenueCount	Latitude	Longitude
0	M4P	44	43.7135	-79.3887
1	M6C	44	43.6915	-79.4307
2	M5N	43	43.7113	-79.4195
3	M6G	43	43.6683	-79.4205
4	M5P	43	43.6966	-79.4120

Results:

- Venues clustered by FSA and superimposed on the map of Toronto:



Results:

- Population and venue count by FSA:

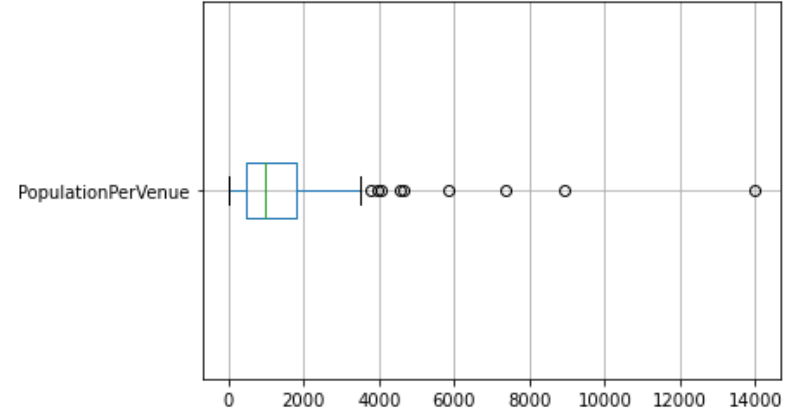
	PostalCode	VenueCount	Latitude	Longitude	Population
0	M4P	44	43.7135	-79.3887	20039.0
1	M6C	44	43.6915	-79.4307	24596.0
2	M5N	43	43.7113	-79.4195	16610.0
3	M6G	43	43.6683	-79.4205	32086.0
4	M5P	43	43.6966	-79.4120	19423.0

- Population and venue count by FSA, now with added columns: population per venue ratio, and the population per venue ratio normalized using 'min-max' normalization method:

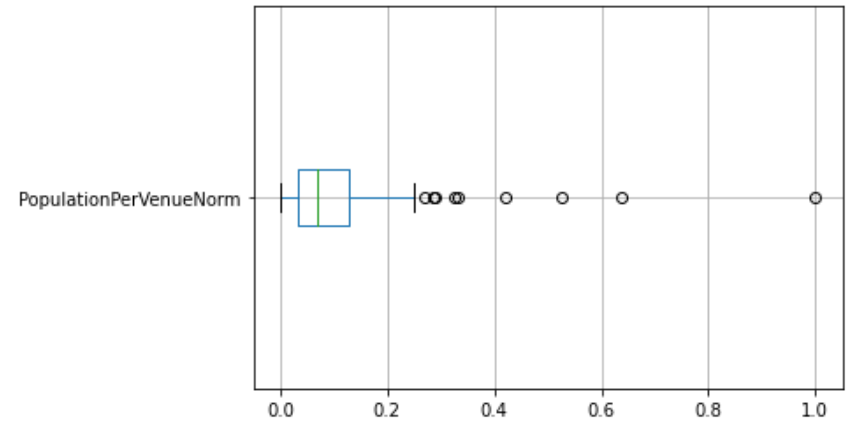
	PostalCode	VenueCount	Latitude	Longitude	Population	PopulationPerVenue	PopulationPerVenueNorm
0	M4P	44	43.7135	-79.3887	20039	455.431818	0.032555
1	M6C	44	43.6915	-79.4307	24596	559.000000	0.039958
2	M5N	43	43.7113	-79.4195	16610	386.279070	0.027612
3	M6G	43	43.6683	-79.4205	32086	746.186047	0.053338
4	M5P	43	43.6966	-79.4120	19423	451.697674	0.032288

Results:

- “PopulationPerVenue” column visualized in the boxplot:

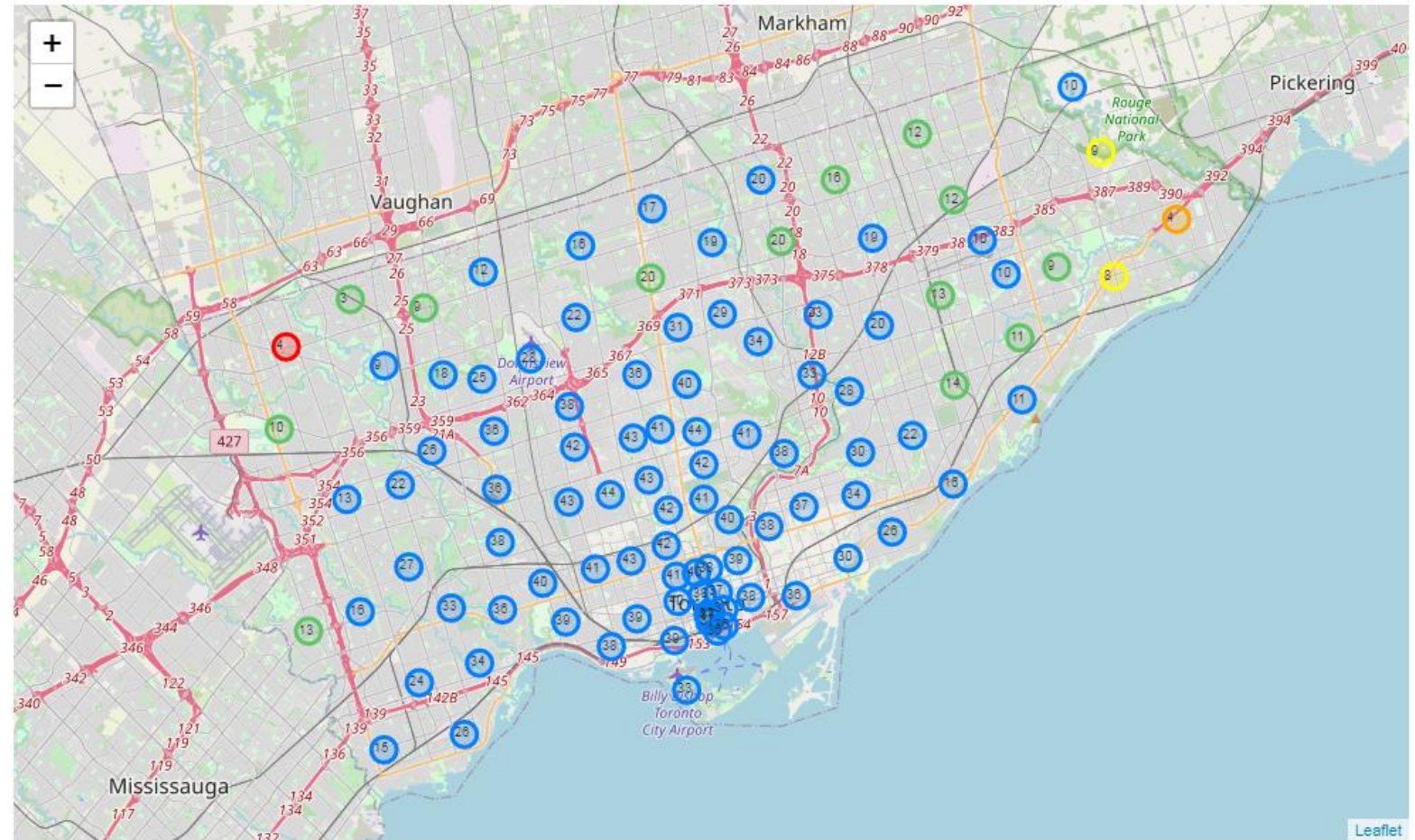


- “PopulationPerVenueNorm” column (normalized alues) visualized in the boxplot:



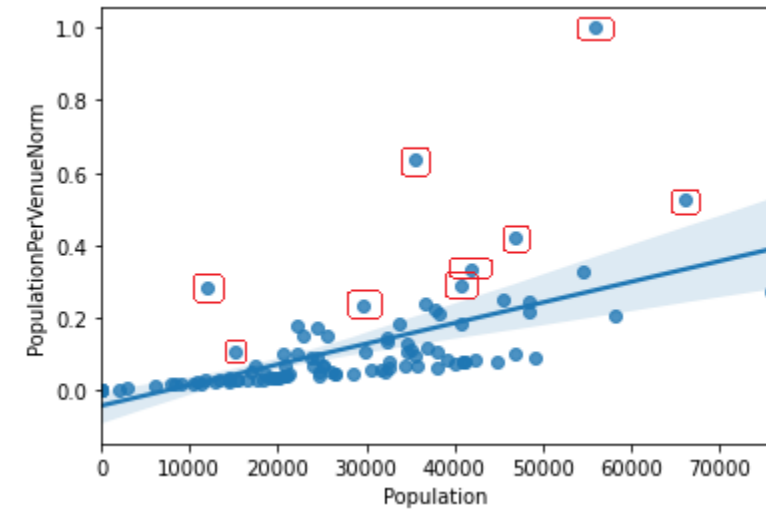
Results:

- Population / Venues ratio clustered by FSA and superimposed on the map of Toronto; colour coding is based on normalized values:



Results:

- Scatter plot of the Linear Regression calculated on "Population" and "PopulationPerVenueNorm":



- Final dataframe sorted by the 'PopulationPerVenue' column in descending order shows the optimal candidate:

	PostalCode	VenueCount	Latitude	Longitude	Population	PopulationPerVenue	PopulationPerVenueNorm
100	M9V	4	43.7432	-79.5876	55959	13989.750000	1.000000
99	M1C	4	43.7878	-79.1564	35626	8906.500000	0.636645
96	M1B	9	43.8113	-79.1930	66108	7345.333333	0.525051
98	M1E	8	43.7678	-79.1866	46943	5867.875000	0.419441
97	M3N	9	43.7568	-79.5210	41958	4662.000000	0.333244
87	M1V	12	43.8177	-79.2819	54680	4556.666667	0.325715
91	M9W	10	43.7144	-79.5909	40684	4068.400000	0.290813
101	M9L	3	43.7598	-79.5565	11950	3983.333333	0.284732
68	M2N	20	43.7673	-79.4111	75897	3794.850000	0.271259
82	M1P	13	43.7612	-79.2707	45571	3505.461538	0.250574
81	M1K	14	43.7298	-79.2639	48434	3459.571429	0.247293

Discussion:

- Existing distribution of veterinary clinics across Toronto is more or less even
- There are a few areas where the ratio suggests that the large local population may benefit from more veterinary clinics in the area:
 1. FSA 'M9V' in Etobicoke, located in the north-western part of the city has the population of 55,959 and features only 4 veterinary clinics: 13,989.75 ratio.
 2. FSA 'M1C' in Scarborough, with 4 clinics serving the population of 35,626 (8906.5 ratio)
 3. FSA 'M1B' in Scarborough, with 9 clinics serving the population of 66,108 (7,345.33 ratio).

Conclusion:

- The analysis was based on venue category, geographical location, population density by FSA and population/venue ratio
- Further analysis could include additional parameters:
 - income
 - education
 - age
 - timeline

Thank you!