

# Battle of the Neighborhoods

## Finding the best Neighborhood in Toronto

### Table of Contents

1. Introduction
2. Target Audience
3. Data Overview
4. Methodology
5. Discussion
6. Conclusion

## 1. Introduction:

This project is a part of IBM Data Science Professional Certificate which aims to teach students on how to define a business problem, acquire the data, and analyze that data using various visualization, statistical tools, and machine learning tools to solve problem.

The objective of this project is to use Foursquare location data and regional clustering of venue information to determine what might be the 'best' neighborhood in Toronto for nature-lover and pet-owner who leads a very active lifestyle.

### Background:

Toronto is the capital city of the Canadian province of Ontario. With a recorded population of 2,731,571 in 2016, it is the most populous city in Canada and the fourth most populous city in North America.

Toronto is all about the latest and greatest in culture, design, dining, and entertainment. But everyone needs a break from big city bustle, and that's where green spaces come in. A diverse eastern city known for its distinctive neighborhoods, Toronto has more than 1,500 parks to explore.

### Problem:

If I am to move to Toronto, which neighborhood with the best park and avenues would be the most suitable for me?

I am, and many other, are nature lovers and pet owners, would love to find a neighborhood with great amenities and other types of venues that suit my active, on-the-go lifestyle with my 2 dogs. I'd love to find a location where the richness of nature are preserved while also includes venues such as:

- Parks, as #1 priority
- Coffee shop or cafe

- Restaurants of various cuisines
- Nearby vet and pet services
- Gym

**Problem statement:**

What is/are the best neighborhood to live in Toronto for me as a nature lover/ pet-parent with an active lifestyle?

## 2. Target Audience

This information provided by this report would be useful for people, especially nature-lovers, pet-owners, with active, on-the-go lifestyle who are interested in comparing different neighborhood of the city of Toronto and are interested in finding the best neighborhood to live in.

## 3. Data Overview

**Source 1:** For the Toronto neighborhood data, a Wikipedia page exists that has all the information which we need to explore and cluster the neighborhoods in Toronto. We will scrape the Wikipedia page and wrangle the data, clean it, and then read it into a pandas dataframe.

After cleansing the data, the next step was to analyze it. We then created a map using Folium and colour-coded each Neighborhood depending on what Borough it was located in.

This will give me:

- Postal code
- Borough
- Neighborhood

**Source 2:** Geographical location data using geocoder package

Using the Latitude and Longitude collected from the Geocoder package, we merged the two tables together based on Postal Code to get the local venue within a 500-meter radius

This will give me:

- Postal code
- Latitude
- Longitude

**Source 3:** Venue data using Foursquare

Next, we used the Foursquare API to get a list of all the venues in Toronto which included gyms, parks, schools, café shops, restaurants etc. We then merged the Foursquare Venue data with the Neighborhood data which then gave us the nearest venue for each of the neighborhoods.

This will give me:

- Venue id
- Venue distance
- Venue name
- Venue category
- Venue latitude
- Venue longitude

And to compare the different parks, I will be using:

- Likes
- Rating
- Tips

## 4. Methodology

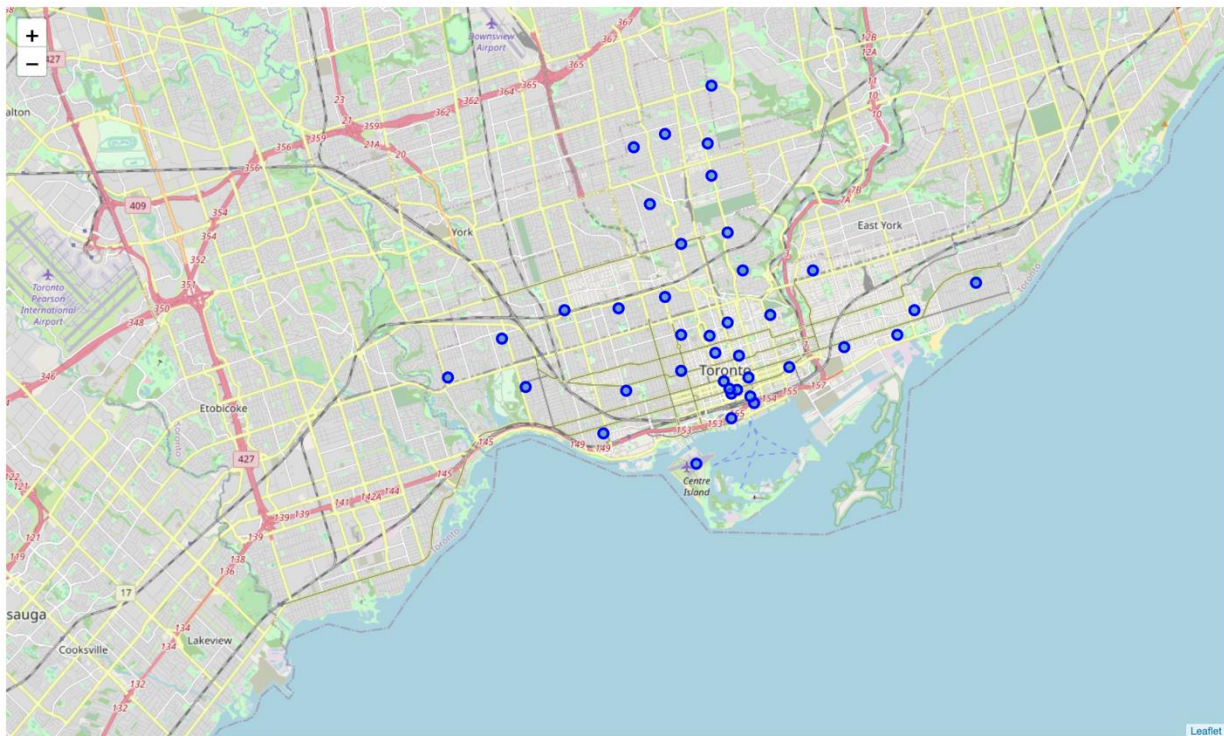
- Data will be collected from [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M), scraped with Beautiful Soup, and cleaned and processed into a Pandas dataframe.

	Postalcode	Borough	Neighborhood
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
5	M1J	Scarborough	Scarborough Village
6	M1K	Scarborough	Kennedy Park, Ionview, East Birchmount Park
7	M1L	Scarborough	Golden Mile, Clairlea, Oakridge
8	M1M	Scarborough	Cliffside, Cliffcrest, Scarborough Village West
9	M1N	Scarborough	Birch Cliff, Cliffside West

- Using the latitude and longitude collected from the Geocoder package, we merged the two tables together based on Postalcode.

	Postalcode	Borough	Neighborhood	Latitude	Longitude
0	M4E	East Toronto	The Beaches	43.676357	-79.293031
1	M4K	East Toronto	The Danforth West, Riverdale	43.679557	-79.352188
2	M4L	East Toronto	India Bazaar, The Beaches West	43.668999	-79.315572
3	M4M	East Toronto	Studio District	43.659526	-79.340923
4	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790
5	M4P	Central Toronto	Davisville North	43.712751	-79.390197
6	M4R	Central Toronto	North Toronto West, Lawrence Park	43.715383	-79.405678
7	M4S	Central Toronto	Davisville	43.704324	-79.388790
8	M4T	Central Toronto	Moore Park, Summerhill East	43.689574	-79.383160
9	M4V	Central Toronto	Summerhill West, Rathnelly, South Hill, Forest...	43.686412	-79.400049

- Using python Folium library, we get to visualize geographic details of Toronto and its neighborhoods.



- Using FourSquare API, we locate and get a list of all the venues in Toronto which included gyms, parks, schools, café shops, restaurants etc. We will explore the neighborhoods and segment them. We set the LIMIT parameter to 100, which would limit the number of venues returned by the Foursquare API and the radius of 500 meter.



	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	The Beaches	43.676357	-79.293031	Glen Manor Ravine	43.676821	-79.293942	Trail
1	The Beaches	43.676357	-79.293031	The Big Carrot Natural Food Market	43.678879	-79.297734	Health Food Store
2	The Beaches	43.676357	-79.293031	Grover Pub and Grub	43.679181	-79.297215	Pub
3	The Beaches	43.676357	-79.293031	Upper Beaches	43.680563	-79.292869	Neighborhood
4	The Danforth West, Riverdale	43.679557	-79.352188	Pantheon	43.677621	-79.351434	Greek Restaurant
5	The Danforth West, Riverdale	43.679557	-79.352188	MenEssentials	43.677820	-79.351265	Cosmetics Shop
6	The Danforth West, Riverdale	43.679557	-79.352188	Dolce Gelato	43.677773	-79.351187	Ice Cream Shop
7	The Danforth West, Riverdale	43.679557	-79.352188	Mezes	43.677962	-79.350196	Greek Restaurant
8	The Danforth West, Riverdale	43.679557	-79.352188	Cafe Fiorentina	43.677743	-79.350115	Italian Restaurant
9	The Danforth West, Riverdale	43.679557	-79.352188	La Diperie	43.677702	-79.352265	Ice Cream Shop

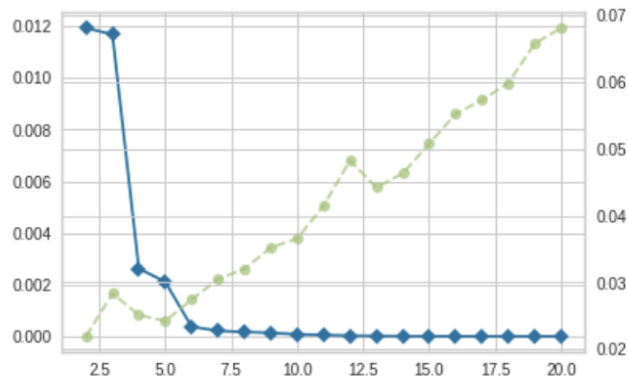
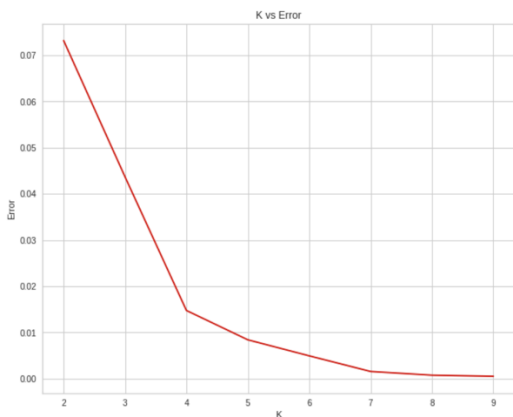
- Using One hot encoding, we analyze the data by transforming Categorical Data into Numerical Data for Machine Learning algorithms. For each of the neighborhoods, individual venues were turned into the frequency at how many of those venues were located in each neighborhood.



(1171, 215)

	Neighborhoods	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	Antique Shop	Aquarium	Art Gallery	Art Museum	Arts & Crafts Store	Asian Restaurant	Athletics & Sports
0	The Beaches	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	The Beaches	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	The Beaches	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	The Beaches	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	The Danforth West, Riverdale	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	The Danforth West, Riverdale	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	The Danforth West, Riverdale	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	The Danforth West, Riverdale	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	The Danforth West, Riverdale	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	The Danforth West, Riverdale	0	0	0	0	0	0	0	0	0	0	0	0	0	0

- Using k-means, with optimal k = 5, we cluster the neighbourhoods based on 'park'.

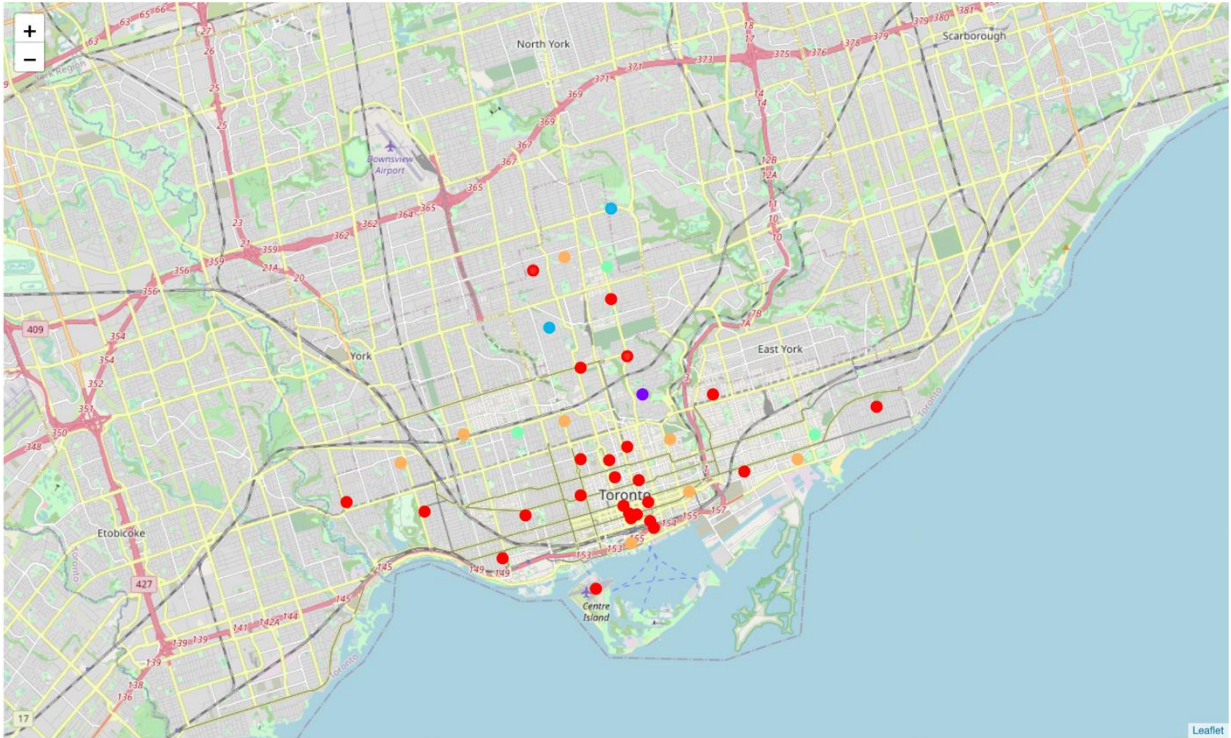


	Neighborhood	Park	Cluster Labels
0	Berczy Park	0.020000	0
1	Brockton, Parkdale Village, Exhibition Place	0.000000	0
2	Business reply mail Processing Centre, South C...	0.062500	4
3	CN Tower, King and Spadina, Railway Lands, Har...	0.000000	0
4	Central Bay Street	0.020000	0
5	Christie	0.125000	3
6	Church and Wellesley	0.020000	0
7	Commerce Court, Victoria Hotel	0.020000	0
8	Davisville	0.030303	0
9	Davisville North	0.111111	3

- We then merged the Foursquare Venue data with our clustered park data which then gave us the nearest venue for each of the Neighborhoods based on 'park'.

	Neighborhood	Park	Cluster Labels	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Berczy Park	0.02	0	43.644771	-79.373306	The Keg Steakhouse + Bar - Esplanade	43.646712	-79.374768	Restaurant
0	Berczy Park	0.02	0	43.644771	-79.373306	LCBO	43.642944	-79.372440	Liquor Store
0	Berczy Park	0.02	0	43.644771	-79.373306	Fresh On Front	43.647815	-79.374453	Vegetarian / Vegan Restaurant
0	Berczy Park	0.02	0	43.644771	-79.373306	Goose Island Brewhouse	43.647329	-79.373541	Beer Bar
0	Berczy Park	0.02	0	43.644771	-79.373306	Meridian Hall	43.646292	-79.376022	Concert Hall
0	Berczy Park	0.02	0	43.644771	-79.373306	Berczy Park	43.648048	-79.375172	Park
0	Berczy Park	0.02	0	43.644771	-79.373306	Hockey Hall Of Fame (Hockey Hall of Fame)	43.646974	-79.377323	Museum
0	Berczy Park	0.02	0	43.644771	-79.373306	St. Lawrence Market (South Building)	43.648743	-79.371597	Farmers Market
0	Berczy Park	0.02	0	43.644771	-79.373306	D.W. Alexander	43.648333	-79.373826	Cocktail Bar
0	Berczy Park	0.02	0	43.644771	-79.373306	Dog Fountain	43.647998	-79.375361	Fountain

- Using python Folium library again, we get to visualize geographic details of Toronto and its neighborhoods in clusters based on 'park' frequency.



- Using FourSquare API, we obtain Likes, Rating, and Tips from all of the parks in Toronto.

	Borough	Neighborhood	ID	Name	Likes	Rating	Tips
0	East Toronto	The Beaches	4bae1ed7f964a520ba863be3	Kew Gardens	33	9.0	12
1	East Toronto	The Beaches	4cb32a6539458cfaaf5a18a0	Kew Gardens Play Park	8	7.6	2
2	East Toronto	The Danforth West, Riverdale	4bf079c06f8aa59336f2c13c	Withrow Park	46	9.1	11
3	East Toronto	The Danforth West, Riverdale	4ba5a1e2f964a5202b1839e3	Riverdale Park East	132	9.4	28
4	East Toronto	India Bazaar, The Beaches West	4baa029cf964a52085433ae3	Woodbine Park	87	8.0	17
5	East Toronto	India Bazaar, The Beaches West	4b5f9a3df964a52003c429e3	Greenwood Park	32	8.2	11
6	East Toronto	India Bazaar, The Beaches West	4bae1ed7f964a520ba863be3	Kew Gardens	33	9.0	12
7	East Toronto	India Bazaar, The Beaches West	4ad4c05ef964a520bef620e3	Ashbridge's Bay Park	45	7.9	14
8	East Toronto	India Bazaar, The Beaches West	4afef402f964a520103222e3	Monarch Park	26	8.6	9
9	East Toronto	Studio District	4ddfbaa185035f3a44e8df6	Underpass Park	27	8.8	2

- To compare all the parks, we calculate for average rating for parks in Toronto.

	Neighborhood	Average Rating
17	Little Portugal, Trinity	9.300000
33	The Danforth West, Riverdale	9.250000
13	Harbourfront East, Union Station, Toronto Islands	9.033333
5	Christie	9.000000
20	Parkdale, Roncesvalles	8.950000
0	Berczy Park	8.933333
28	Stn A PO Boxes	8.933333
22	Regent Park, Harbourfront	8.833333
34	Toronto Dominion Centre, Design Exchange	8.800000
29	Studio District	8.800000
11	First Canadian Place, Underground city	8.800000
14	High Park, The Junction South	8.800000
3	CN Tower, King and Spadina, Railway Lands, Har...	8.711111
31	The Annex, North Midtown, Yorkville	8.700000
27	St. James Town, Cabbagetown	8.700000
21	Queen's Park, Ontario Provincial Government	8.650000
35	University of Toronto, Harbord	8.650000
9	Davisville North	8.633333
8	Davisville	8.633333
4	Central Bay Street	8.600000
6	Church and Wellesley	8.600000
24	Roselawn	8.600000
19	North Toronto West, Lawrence Park	8.566667
16	Lawrence Park	8.550000
7	Commerce Court, Victoria Hotel	8.500000
26	St. James Town	8.500000
23	Rosedale	8.500000
15	India Bazaar, The Beaches West	8.340000
32	The Beaches	8.300000
10	Dufferin, Dovercourt Village	8.175000
12	Forest Hill North & West, Forest Hill Road Park	8.150000
30	Summerhill West, Rathnelly, South Hill, Forest...	8.133333
1	Brockton, Parkdale Village, Exhibition Place	8.100000



## 5. Result and Discussion

### Cluster 0:

Cluster 1 has the lowest number of park (frequency = 0 - 0.030). This will not be in consideration.

### Cluster 1:

Cluster 1 of Rosedale has the highest number of park (frequency = 0.5) however the most common venues only include playground, park, and trail. This will not be in consideration.

### Cluster 2:

Cluster 2 of Forest Hill North & West, Forest Hill Road Park, and Lawrence Park has the second highest number of park (frequency = 0.25 - 0.33) however its venues only include bus, swim, sushi, trail, jewelry, and park. This will not be in consideration.

### Cluster 3:

Cluster 3 has the third highest number of park (frequency = 0.11 - 0.125) with many venues which I require such as gym, sport store, many variety of restaurants and cuisines, bonus of dog run and pet stores. Two neighborhoods that stand out are:

- Christie, 1 park, located in Downtown Toronto.
- Davisville North, 3 parks, located in Central Toronto.
- India Bazaar and The Beaches West neighborhoods, 5 parks, located in East Toronto.

This looks to me that India Bazaar and The Beaches West neighborhoods is the winner with the highest number of park. To be more throughout with the analysis, I will compare ratings of these parks.

### Cluster 4:

Cluster 4 has the forth highest number of park (frequency = 0.04 - 0.06) with just as many venues such as Cluster 3.

- Dufferin, Dovercourt Village, 4 parks, West Toronto
- High Park, The Junction South, 2 parks, West Toronto
- Business reply mail Processing Centre has 3 parks, East Toronto
- North Toronto West, Lawrence Park, 3 parks, Central Toronto
- St. James Town, Cabbagetown, 4 parks, Downtown Toronto
- Regent Park, Harbourfront, 3 parks, Downtown Toronto
- Harbourfront East, Union Station, Toronto Islands, 3 parks, Downtown Toronto

I will look Cluster 3 and Cluster 4 in depth since they are very similar. I will look at rating for neighborhood with high number of parks and compare their ratings. The neighborhoods that I will be comparing are:

- India Bazaar and The Beaches West neighborhoods, 5 parks
- Dufferin, Dovercourt Village, 4 parks
- St. James Town, Cabbagetown, 4 parks

## 6. Conclusion

Solely based on the number of parks per neighborhood and the type of venues it has, and the average rating, I create a list of options below:

- St. James Town, Cabbagetown, 4 parks - average rating = 8.4  
Venues: pet store, pizza place, general entertainment, jewelry store, butcher, bakery, café, Italian restaurant, Farmers Market, snack place, deli/ bodega, Caribbean restaurant, Tawainese restaurant, gift shop, pub, Thai restaurant, coffee shop, bank, liquor store, beer store, playground, sandwich place, pizza place, Chinese restaurant, pharmacy, Japanese restaurant, diner, clothing store, Indian restaurant
- India Bazaar, The Beaches West - average rating = 8.34  
Venues: liquor store, brewery, Italian restaurant, pub, pet store, steakhouse, burrito place, movie theater, sandwich place, pizza place, coffee shop, sushi restaurant, fast food restaurant, ice-cream shop, gym, fish & chip shop
- Dufferin, Dovercourt Village, 4 parks - average rating = 8.1 Venues: music venue, grocery store, bar, middle eastern restaurant, brewery, pharmacy, bank, supermarket, bakery

These neighborhood would be great options for me and anyone who desires a location filled with good amenities and venues such as gym, sport store, many variety of restaurants and cuisines, and pet stores and pet-related services, etc. that cater to the need of a pet-owner with very active lifestyle.

The best option is St. James Town, Cabbagetown, 4 parks with an average rating of 8.4. It has the highest number of venues and seems to be very active and busy.

My personal choice is India Bazaar, The Beaches West, 5 parks with an average rating of 8.34. I chose this because they have the sandy beaches and vibrant boardwalk give me opportunity to socialize and interact with plenty of surrounding neighborhoods.

However, more analyses should be done to compare other factors such as rent, crime rate, school quality, lifestyle, noise level, walking distance etc. to know for sure, but this is too big for the scope of this assignment.