Coursera Capstone Project – Battle of the Neighborhoods

To explore and analyze the venues in the neighborhoods of Philadelphia and Toronto using FourSquare venue data

1. Introduction

1.1 Background

When visiting or moving to a new city there can be difficulties when trying to choose or discover restaurants, stores and other local venues to explore in the neighborhoods. This can be daunting for both a user and business perspective when venturing forth into new territory.

When exploring new neighborhoods, there is a difficult task of analyzing various forms of criteria that can drastic impact your decision making. Such factors include, distance, price tier, ratings, photos and tips which is a lot of criteria too look through.

The Philadelphia area has many universities and colleges make it a top study destination, as the city has evolved into an educational and economic hub, with an estimated gross metropolitan product of $490 billion in 2019. This makes the twelve neighborhoods of Philadelphia a hot spot for new residents, travelers and businesses. The twelve neighborhoods of Philadelphia are Center City, South Philadelphia, Southwest Philadelphia, West Philadelphia, Lower North Philadelphia, Upper North Philadelphia, Bridesburg-Kensington-Port Richmond, Roxborough-Manayunk, Germantown-Chestnut Hill, Olney-Oak Lane, Near Northeast Philadelphia and Far Northeast Philadelphia.

Toronto is an international center for business, finance, arts and culture in the provincial capital of Ontario. Toronto is also an education and economic hub with diversified strengths in technology, design, food services, education and many more. This makes the ten boroughs a hot spot for new residents, travelers and businesses much like Philadelphia. The ten boroughs that make up Toronto are North York, Downtown Toronto, Etobicoke, Scarborough, East York, York, East Toronto, West Toronto, Central Toronto, Mississauga.

1.2 Problem

With the utilization of Foursquare venue data we can utilize this to explore and compare the city of Philadelphia and Toronto. This project will focus on how this data can be used to find new insights required for business opportunities in certain neighborhoods, cluster and segment venues to provide for a better user experience. We can then also compare and distinguish which city venues are better.

By segmenting and clustering venues based on our Foursquare venue data, we can compare and distinguish venue similarities to observe which neighborhoods would be good business ventures. Also, we will utilize factors such as user likes, ratings, tips, photos and distance to venues to seek if there are correlations that exist.

1.3 Interest

Interested parties might include business start-ups, stakeholders, and existing owners. Business start-ups might be looking for areas that lack certain categorical venues to further enhance a neighborhoods overall experience. Stakeholders would also be interested in this and also as a means to see how their invested interests are performing as a whole comparatively within city neighborhoods and to another city. Existing business will also benefit to see how the venues likes, tips and photos will bring effectiveness on their price tie and ratings.

2. Data Acquisition and Cleaning

2.1 Data Sources

Our two primary sources of data will be location information provided from data scrapped from Wikipedia on the neighborhood information for Philadelphia and Toronto plugged into python's geocoder library. Foursquare will provide our venue details to help explore, segment and analyze the neighborhoods of Philadelphia and Toronto. Below are the links of the Wikipedia sources for location data that was utilized in conjunction with Python's geocoder:

Philadelphia: https://en.wikipedia.org/wiki/List_of_Philadelphia_neighborhoods

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

2.2 Data Cleaning

After scraping the data from Wikipedia there was a lot of effort put forth in the cleaning and organizing of the data. Each set of data was cleaned to provide our location information. Before obtaining location data, the number of neighborhoods per borough were extracted for Frequency Distribution analysis.(Figures 1.1 & 1.2) The unique Boroughs' location information was obtained using the geocoder library to figure out the latitude and longitude for each of the boroughs. (Figure 2.1 & 2.2).

Figure 1.1 – Philadelphia Neighborhood Frequency Distribution by Borough

	Neighborhood	%-overall
South Philadelphia	29	16.02%
West Philadelphia	29	16.02%
Center City	21	11.6%
Near Northeast Philadelphia	19	10.5%
Far Northeast Philadelphia	16	8.84%
Southwest Philadelphia	15	8.29%
Lower North Philadelphia	12	6.63%
Germantown-Chestnut Hill	11	6.08%
Olney-Oak Lane	11	6.08%
Upper North Philadelphia	6	3.31%
Bridesburg-Kensington-Port Richmond	6	3.31%
Roxborough-Manayunk	6	3.31%

Figure 1.2 – Toronto Neighborhood Frequency Distribution by Borough

	Neighborhood	%-overall
North York	24	23.3%
Downtown Toronto	19	18.45%
Scarborough	17	16.5%
Etobicoke	12	11.65%
Central Toronto	9	8.74%
West Toronto	6	5.83%
East York	5	4.85%
East Toronto	5	4.85%
York	5	4.85%
Mississauga	1	0.97%

Figure 2.1 – Philadelphia Borough Localization Data

	Borough	Latitude	Longitude
0	Center City	39.952544	-75.165219
1	South Philadelphia	39.964110	-75.161050
2	Southwest Philadelphia	39.910040	-75.186370
3	West Philadelphia	40.053132	-75.028511
4	Lower North Philadelphia	39.964158	-75.198802
5	Upper North Philadelphia	40.059110	-75.052180
6	Bridesburg-Kensington-Port Richmond	39.980900	-75.099600
7	Roxborough-Manayunk	40.037990	-75.223080
8	Germantown-Chestnut Hill	40.078489	-75.211934
9	Olney-Oak Lane	40.041130	-75.124050
10	Near Northeast Philadelphia	40.092800	-74.987030
11	Far Northeast Philadelphia	40.092800	-74.987030

Figure 2.2 – Toronto Borough Localization Data

	Borough	Latitude	Longitude
0	North York	43.768260	-79.412630
1	Downtown Toronto	43.658200	-79.368320
2	Etobicoke	43.644360	-79.567130
3	Scarborough	43.772200	-79.256660
4	East York	43.691800	-79.327030
5	York	43.692080	-79.478630
6	East Toronto	43.659030	-79.349010
7	West Toronto	43.664712	-79.346346
8	Central Toronto	43.609727	-79.492844
9	Mississauga	43.587260	-79.644940

Next, we are able to append our venue information to our dataframes. We query foursquare to find our top 25 venues in each borough, making sure to return back information on each. This query focuses on obtaining the location information for each of the venues in the borough and their associated foursquare ID.(Figure 3.1 & 3.2)

Figure 3.1 – Philadelphia Borough Venue Localization & Categorization

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue ID	Distance	Venue Latitude	Venue Longitude	Venue Category
0	North York	43.76826	-79.41263	The Keg	5a35b4443abcaf37eb1a0d88	191	43.766579	-79.412131	Steakhouse
1	North York	43.76826	-79.41263	Konjiki Ramen	5a02789d0a464d3112a58785	144	43.766998	-79.412222	Ramen Restaurant
2	North York	43.76826	-79.41263	Toronto Centre for the Arts	4ad4c062f964a520c3f720e3	255	43.766228	-79.414115	Theater
3	North York	43.76826	-79.41263	Loblaws	4ae257cff964a520758d21e3	66	43.768722	-79.412101	Grocery Store
4	North York	43.76826	-79.41263	Satay Sate	57f92db0498ee70159702002	179	43.766690	-79.412100	Indonesian Restaurant

Figure 3.2 – Toronto Borough Venue Localization & Categorization

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue ID	Distance	Venue Latitude	Venue Longitude	Venue Category
0	North York	43.76826	-79.41263	The Keg	5a35b4443abcaf37eb1a0d88	191	43.766579	-79.412131	Steakhouse
1	North York	43.76826	-79.41263	Konjiki Ramen	5a02789d0a464d3112a58785	144	43.766998	-79.412222	Ramen Restaurant
2	North York	43.76826	-79.41263	Toronto Centre for the Arts	4ad4c062f964a520c3f720e3	255	43.766228	-79.414115	Theater
3	North York	43.76826	-79.41263	Loblaws	4ae257cff964a520758d21e3	66	43.768722	-79.412101	Grocery Store
4	North York	43.76826	-79.41263	Satay Sate	57f92db0498ee70159702002	179	43.766690	-79.412100	Indonesian Restaurant

Finally, we utilize the foursquare one last time to obtain each venues respective details for user likes, ratings, price tier, photo count and tips. This information is compiled in to one final dataframe that will be utilized for the remainder of the project to address the questions and problems we have. (Figure 4.1 & 4.2)

Figure 4.1 – Finalized Philadelphia Venue Localization & Information Dataset

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue ID	Distance	Venue Latitude	Venue Longitude	Venue Category	like_counts	rating	photo_count	reasons_count	Tips count
Center City	39.9525435	-75.165219	Dilworth Park	4bde0d566198c9b6c5cc12ff	49	39.9527718	-75.164723	Park	311	9.1	668	1	23
Center City	39.9525435	-75.165219	Philadelphia Film Center	47bdd66df964a520da4d1fe3	195	39.950835	-75.164683	Movie Theater	88	8.9	177	1	20
Center City	39.9525435	-75.165219	City Hall Courtyard	4f29e8e3e4b02f0aff55b2a7	138	39.9524842	-75.163592	Plaza	58	8.9	214	1	3
Center City	39.9525435	-75.165219	Del Frisco's Double Eagle Steak House	4ab2ac0bf964a520d66b20e3	177	39.9509564	-75.165459	Steakhouse	275	8.8	309	1	85
Center City	39.9525435	-75.165219	The Ritz-Carlton, Philadelphia	4a68db6ff964a52023cb1fe3	152	39.9514456	-75.164149	Hotel	199	8.8	599	1	63

Figure 4.2 – Finalized Toronto Venue Localization & Information Dataset

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue ID	Distance	Venue Latitude	Venue Longitude	Venue Category	like_counts	rating	photo_count	reasons_count	Tips count
North York	43.76826	-79.41263	The Keg	5a35b4443abcaf37eb1a0d88	191	43.7665789	-79.412131	Steakhouse	25	8.5	7	0	3
North York	43.76826	-79.41263	Konjiki Ramen	5a02789d0a464d3112a58785	144	43.7669977	-79.412222	Ramen Restaurant	39	8.3	68	1	8
North York	43.76826	-79.41263	Toronto Centre for the Arts	4ad4c062f964a520c3f720e3	255	43.7662283	-79.414115	Theater	46	8.1	145	1	13
North York	43.76826	-79.41263	Loblaws	4ae257cff964a520758d21e3	66	43.768722	-79.412101	Grocery Store	90	7.8	94	1	13
North York	43.76826	-79.41263	Satay Sate	57f92db0498ee70159702002	179	43.76669	-79.4121	Indonesian Restaurant	8	7.8	4	0	4

2.3 Data Limitations

Currently, the two limiting factors in a more robust data set is the fact that we have to limit the API calls for premium data, as a foursquare personal account restricts the premium call daily limit to 500. Thus, we look at the boroughs overall and only the top 25 venues in each borough. We are also limited by the amount of available data provided by foursquare at each of these venues, as they can be missing important information pertaining to our regression analyses. As such, this is more a modeling of what is possible with our the queriable data and could be expanded upon when limitations are lifted.

3. Methods

Initially, an overall frequency table will be generated with percentages to distinguish the two important characteristics. First how many neighborhoods make up each borough and the overall percentage that borough contributes to the cities total neighborhood count. We can utilize this to understand if there are any significant differences between our cities.

The information extraction from Foursquare's premium calls will allow us to utilize the information on user likes, tips, photo counts, rating and distance to see if there are any correlations that exist between these variables.

We will use the data to obtain information regarding the top 25 venue categories in each borough to observe the frequency distribution of what venue categories the borough is lacking. This will provide insight to potential new areas to recommend business start-ups looking to create new venues or add to a low category. We will utilize those that are one or less for our dataset.

Finally, we can utilize a k-means algorithm to cluster and segment our venues based on tips, photo counts, ratings, user likes and even distance. This can greatly enhance the user experience when selecting venues.

4. Analysis

4.1 Distribution Frequencies

4.1.a Distribution Frequencies – Neighborhoods & Boroughs

Figure 5.1 – Toronto & Philadelphia Neighborhood Breakdown

	Neighborhood	%-overall
North York	24	23.3%
Downtown Toronto	19	18.45%
Scarborough	17	16.5%
Etobicoke	12	11.65%
Central Toronto	9	8.74%
West Toronto	6	5.83%
East York	5	4.85%
East Toronto	5	4.85%
York	5	4.85%
Mississauga	1	0.97%

	Neighborhood	%-overall
South Philadelphia	29	16.02%
West Philadelphia	29	16.02%
Center City	21	11.6%
Near Northeast Philadelphia	19	10.5%
Far Northeast Philadelphia	16	8.84%
Southwest Philadelphia	15	8.29%
Lower North Philadelphia	12	6.63%
Olney-Oak Lane	11	6.08%
Germantown-Chestnut Hill	11	6.08%
Roxborough-Manayunk	6	3.31%
Upper North Philadelphia	6	3.31%
Bridesburg-Kensington-Port Richmond	6	3.31%

Figure 5.2 – Toronto and Philadelphia Overall Breakdown

	Toronto	Philadelphia
Borough Count	10	12
Total Neighborhoods	103	181
Average Neighborhood per Borough	10.3	8.58
Top 3 Neighborhoods Total(% Overall)	60(58.25%)	79(43.65%)

```
Philadelphia's overall neighborhood count is 43.09% higher than Toronto's
```

4.1.b Distribution Frequencies – Venue Details per Borough

Figure 6.1 – Toronto Venue Details Distribution Frequency

Neighborhood	Venue-s	User Likes	like %	Avg-Rating	Total Photos	Total Tips
East Toronto	25	1225	26.45%	8.19	1326	496
Scarborough	38	944	20.38%	6.68	928	278
North York	26	900	19.43%	7.38	1270	260
Mississauga	22	633	13.67%	6.99	1036	220
West Toronto	23	425	9.18%	7.95	418	231
Downtown Toronto	15	322	6.95%	7.63	528	152
East York	11	75	1.62%	6.65	85	40
Etobicoke	5	57	1.23%	6.72	38	16
Central Toronto	5	47	1.01%	6.94	38	11
York	4	3	0.06%	6.25	2	5

Figure 6.2 – Philadelphia Venue Distribution Frequency

Neighborhood	Venue-s	User Likes	like %	Avg-Rating	Total Photos	Total Tips
Center City	25	3602	58.0%	8.64	6891	893
South Philadelphia	25	1003	16.15%	7.47	1397	474
Germantown-Chestnut Hill	16	388	6.25%	7.38	461	195
Southwest Philadelphia	8	304	4.9%	6.64	98	117
Bridesburg-Kensington-Port Richmond	10	261	4.2%	7.64	152	72
Upper North Philadelphia	10	220	3.54%	7.37	112	60
Roxborough-Manayunk	8	176	2.83%	7.10	107	44
Olney-Oak Lane	5	92	1.48%	7.30	102	23
Lower North Philadelphia	8	72	1.16%	7.11	48	27
Far Northeast Philadelphia	2	34	0.55%	5.60	42	12
Near Northeast Philadelphia	2	34	0.55%	5.60	42	12
West Philadelphia	3	24	0.39%	6.77	32	6

4.2 Venue Category Distribution Analysis

Figure 7.1 – Toronto Venue Category Distribution

Neighborhood	Central Toronto	Downtown Toronto	East Toronto	East York	Etobicoke	Mississauga	North York	Scarborough	West Toronto	York	Total
Venue Category											
Total	5	15	25	11	5	22	26	38	23	4	174
Coffee Shop	0	0	3	5	4	6	3	8	1	0	30
Sandwich Place	0	1	0	4	0	0	0	8	0	4	17
Clothing Store	0	0	0	0	0	3	0	6	0	0	9
Café	0	1	2	0	0	2	2	0	0	0	7
Bakery	0	0	2	0	0	0	1	0	2	0	5
Vietnamese Restaurant	0	0	0	0	0	0	0	0	5	0	5
Bar	0	0	2	0	0	0	0	1	2	0	5
Yoga Studio	0	1	0	0	0	1	0	0	1	0	3
Grocery Store	1	0	0	0	0	0	1	1	0	0	3
Restaurant	0	1	1	0	0	0	1	0	0	0	3

Figure 7.2 – Philadelphia Venue Category Distribution

Neighborhood	Bridesburg-Kensington-Port Richmond	Center City	Far Northeast Philadelphia	Germantown-Chestnut Hill	Lower North Philadelphia	Near Northeast Philadelphia	Olney-Oak Lane	Roxborough-Manayunk	South Philadelphia	Southwest Philadelphia	Upper North Philadelphia	West Philadelphia	Total
Venue Category													
Total	10	25	2	16	8	2	5	8	25	8	10	3	122
Convenience Store	3	0	0	0	0	0	0	3	0	0	3	1	10
Sandwich Place	2	0	0	2	0	0	0	0	3	1	0	0	8
Pizza Place	0	1	0	1	0	0	1	0	4	0	0	0	7
Pharmacy	0	0	0	0	0	0	2	2	1	0	1	0	6
Coffee Shop	2	1	0	2	0	0	0	0	0	1	0	0	6
Sporting Goods Shop	0	0	2	0	0	2	0	0	0	0	0	0	4
Bar	1	0	0	1	0	0	0	0	1	0	1	0	4
Italian Restaurant	0	0	0	1	0	0	0	0	1	1	1	0	4
Donut Shop	0	0	0	0	0	0	0	0	0	2	0	2	4
Hotel	0	1	0	0	0	0	0	0	1	1	0	0	3

4.3 Regression Analysis

Figure 8.1 – Overall Like Counts vs Ratings

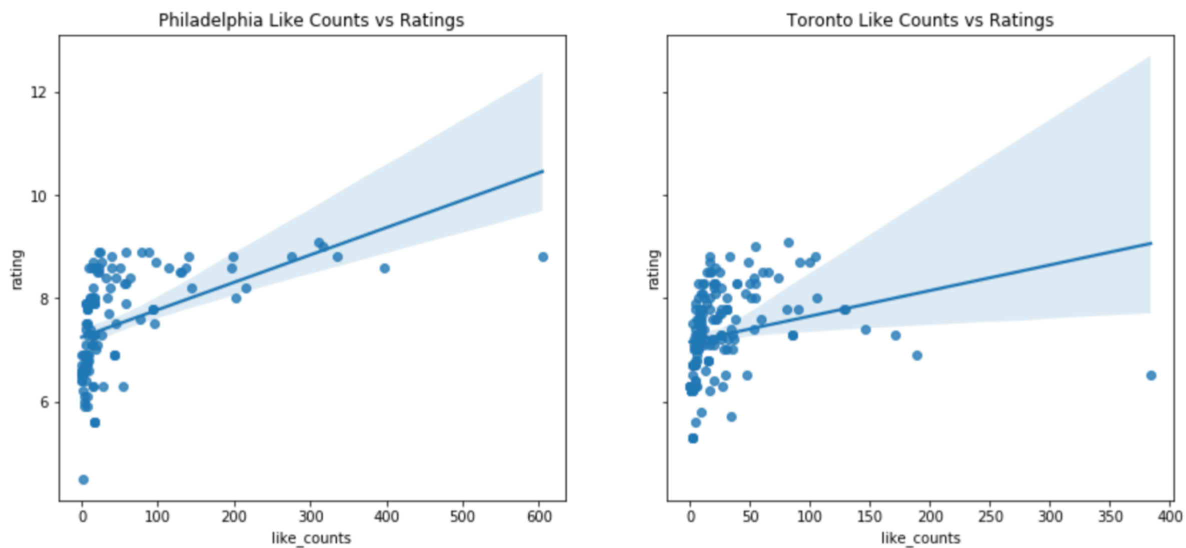


Figure 8.2 – Like Counts at 100 and 30 vs Ratings

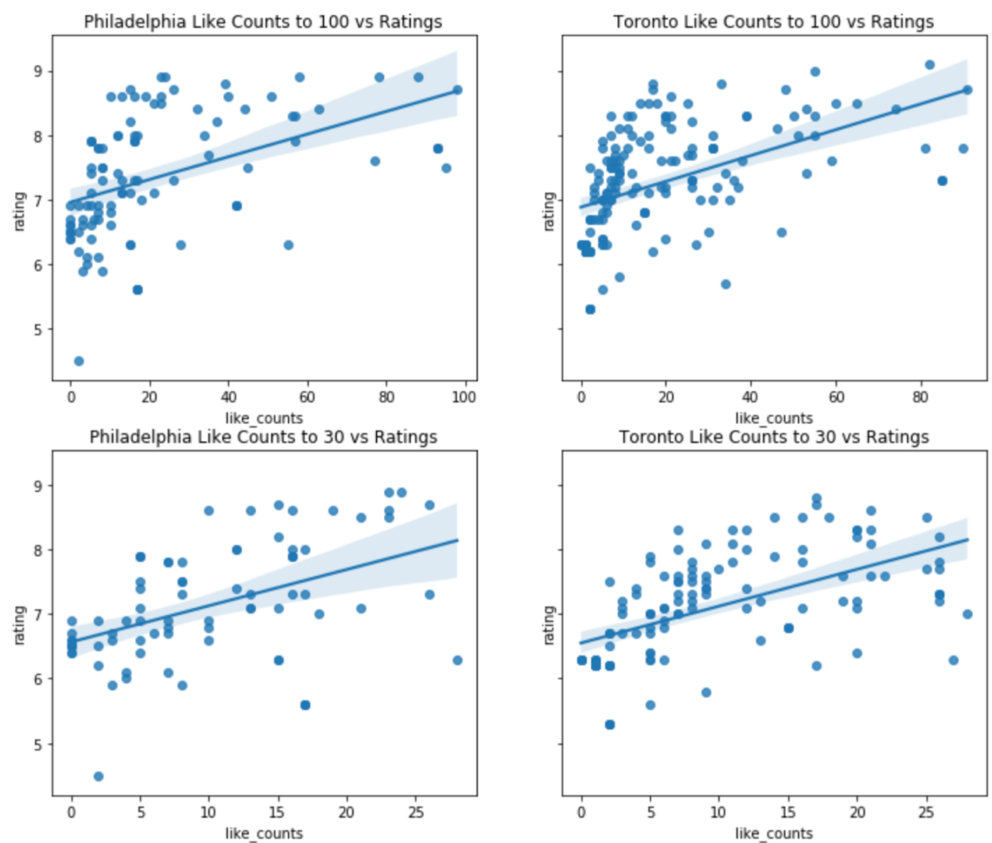


Figure 8.3 – Overall Distance vs Ratings

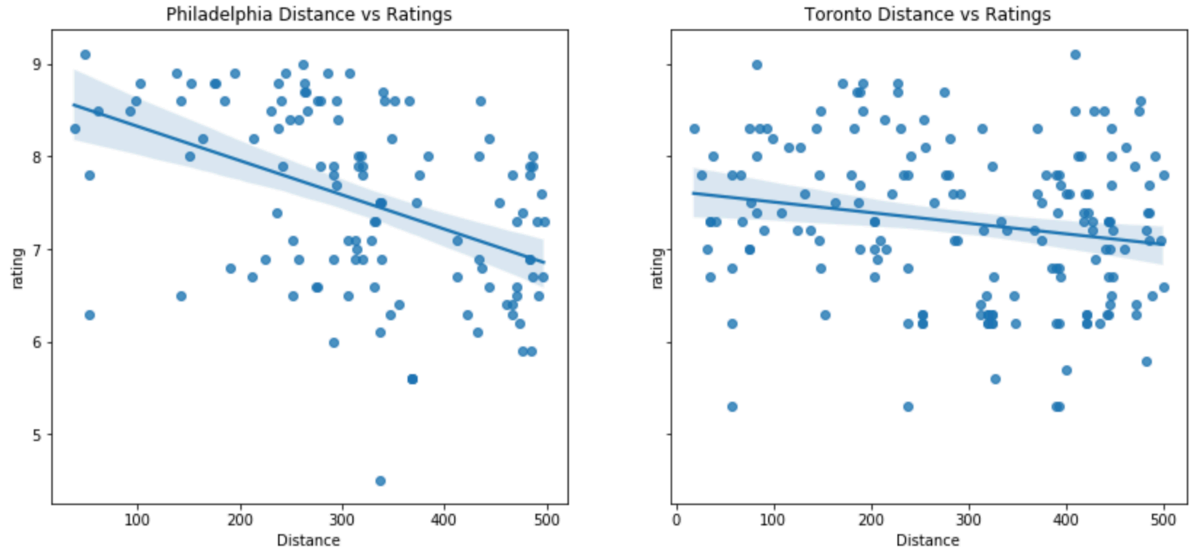


Figure 8.4 – Overall Photo Counts vs Ratings

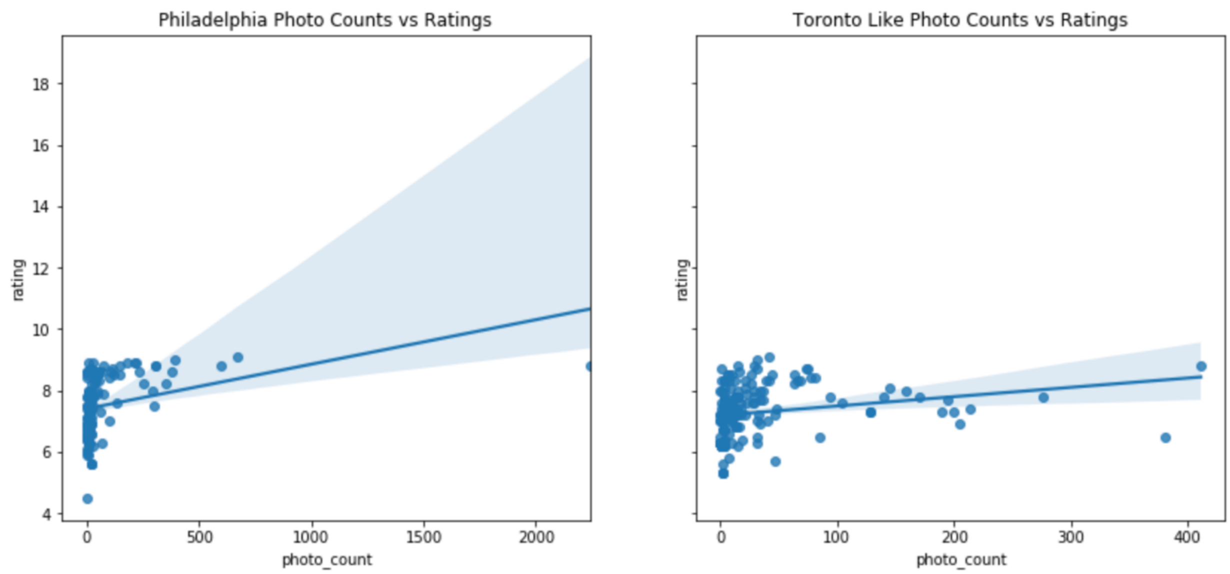


Figure 8.5 – Photo Counts at 200 and 50 vs Ratings

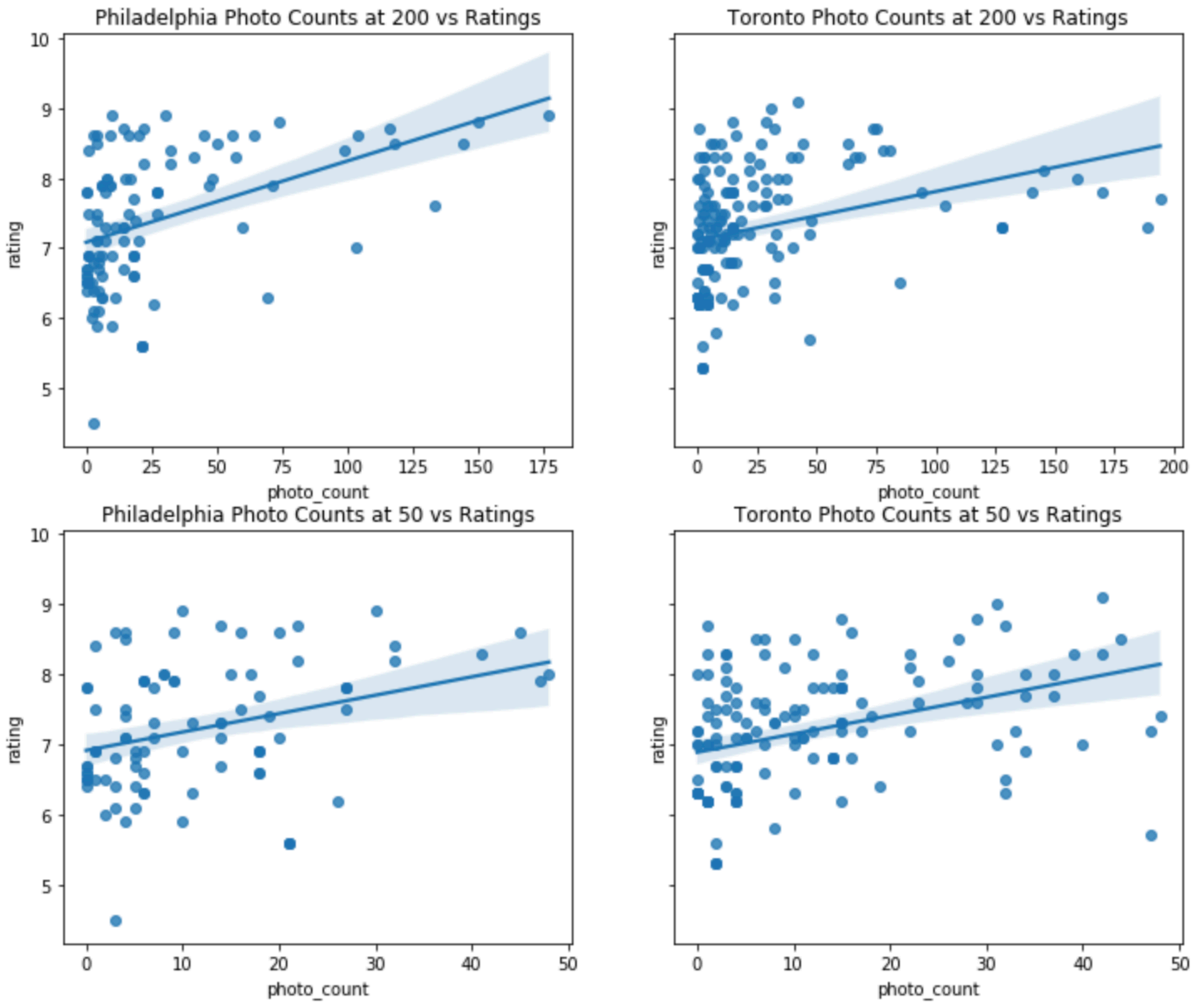
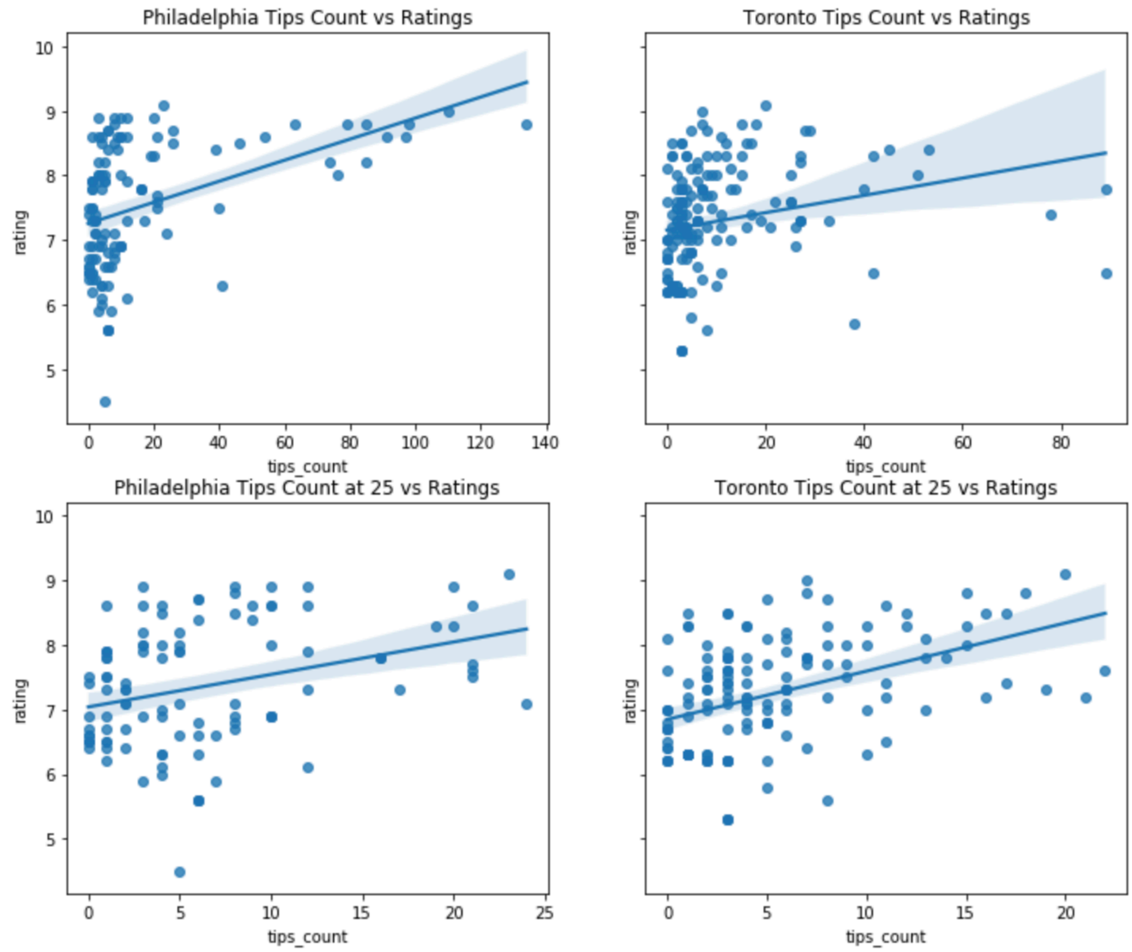


Figure 8.6 – Overall Tip Counts vs Ratings & Tip Counts at 25 vs Ratings



4.4 K-Means Clustering

Figure 9.1 – Elbow Method: Number of Clusters determination for Philadelphia

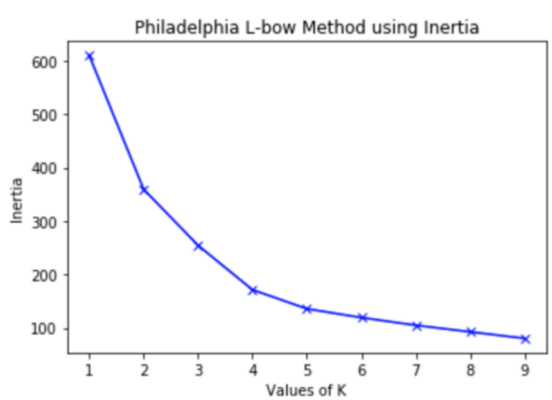


Figure 9.2 – Elbow Method: Number of Clusters determination for Toronto

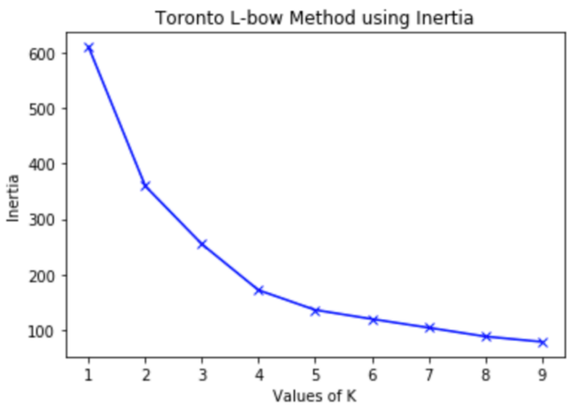


Figure 9.3 – Philadelphia Mean Values for Cluster Centroids

Labels	Neighborhood Latitude	Neighborhood Longitude	Distance	Venue Latitude	Venue Longitude	like_counts	rating	photo_count	reasons_count	tips_count
0	39.984504	-75.168295	257.373134	39.984397	-75.168568	32.985075	7.683582	36.761194	0.343284	10.343284
1	39.952543	-75.165219	175.000000	39.954123	-75.165303	605.000000	8.800000	2243.000000	1.000000	134.000000
2	39.955698	-75.164082	226.090909	39.954291	-75.163304	229.090909	8.509091	371.545455	1.000000	70.545455
3	40.027419	-75.124175	442.906977	40.027678	-75.124426	20.348837	6.969767	16.069767	0.162791	7.720930

Figure 9.4 – Toronto Mean Values for Cluster Centroids

Labels	Neighborhood Latitude	Neighborhood Longitude	Distance	Venue Latitude	Venue Longitude	like_counts	rating	photo_count	reasons_count	tips_count
0	43.683038	-79.434537	93.028571	43.683081	-79.434464	26.542857	7.542857	25.657143	0.371429	8.342857
1	43.692782	-79.379355	430.702703	43.693332	-79.378964	19.878378	7.179730	22.554054	0.175676	7.567568
2	43.696153	-79.396270	144.714286	43.696761	-79.395907	156.857143	7.557143	262.857143	1.000000	52.285714
3	43.697905	-79.373153	260.396552	43.698073	-79.372790	19.534483	7.236207	21.758621	0.224138	8.465517

Figure 9.5 – Cities Clusters on Like Counts vs Distance

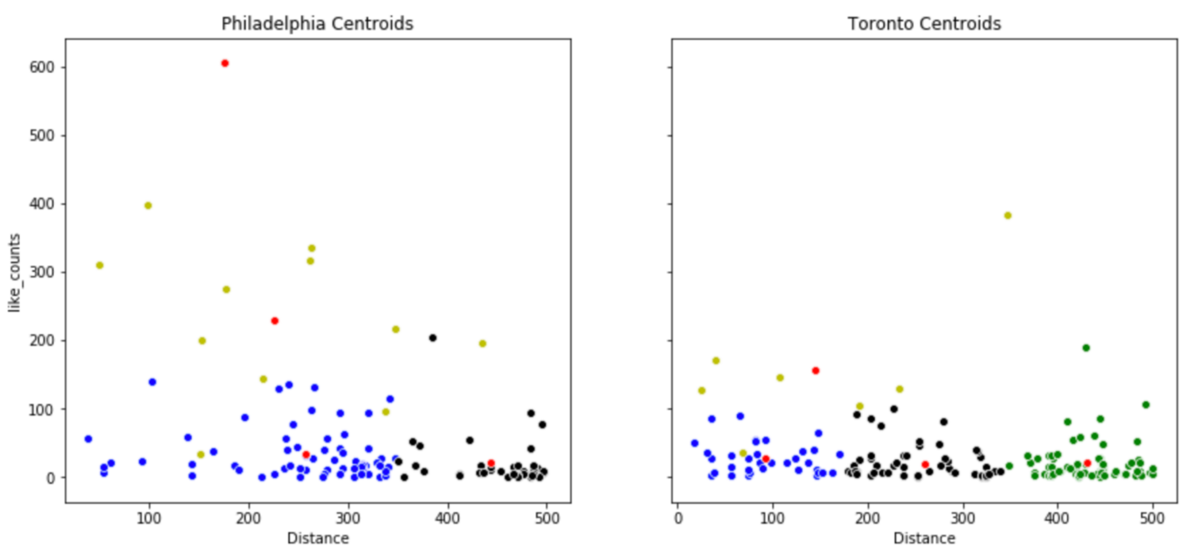


Figure 9.6 – Cities Clusters on Rating vs Distance

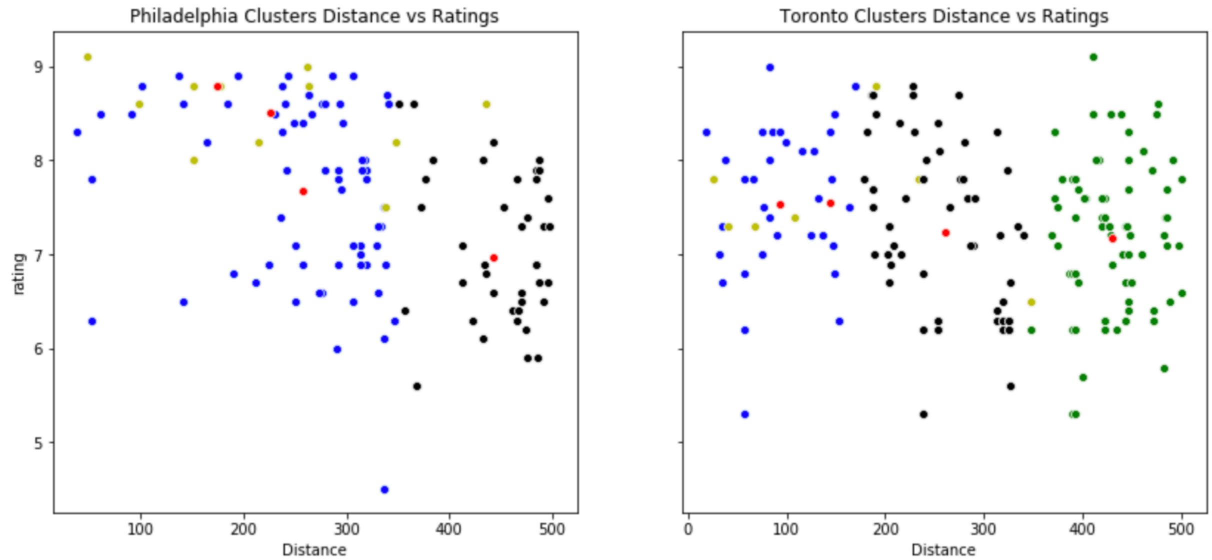


Figure 9.7 – Cities Clusters on Photo Count vs Distance

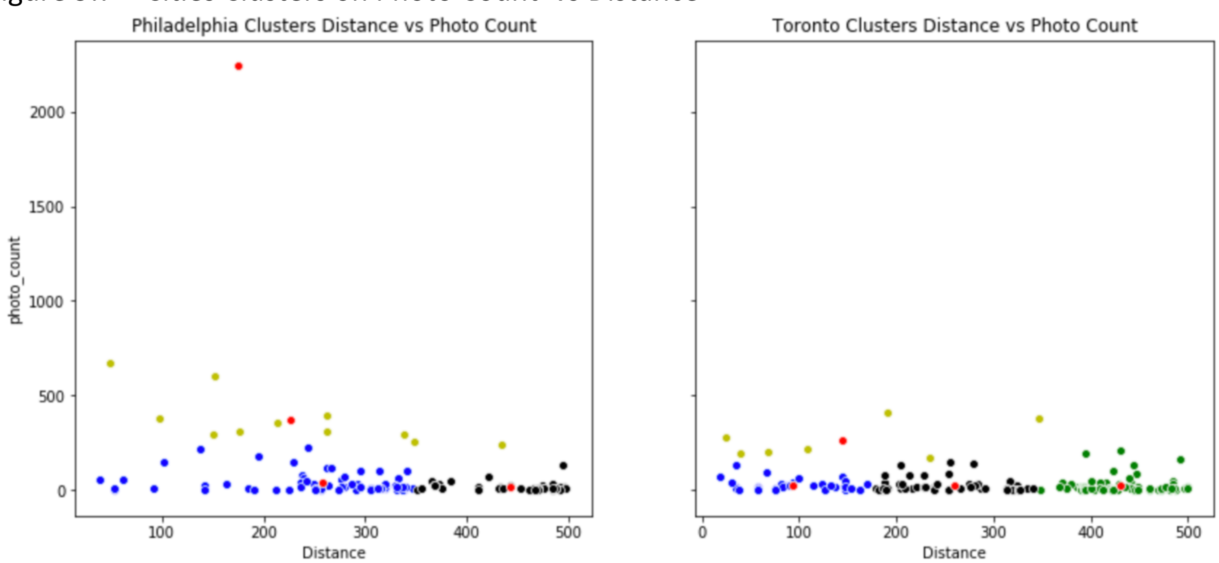


Figure 9.8 – Cities Clusters on Tip Count vs Distance

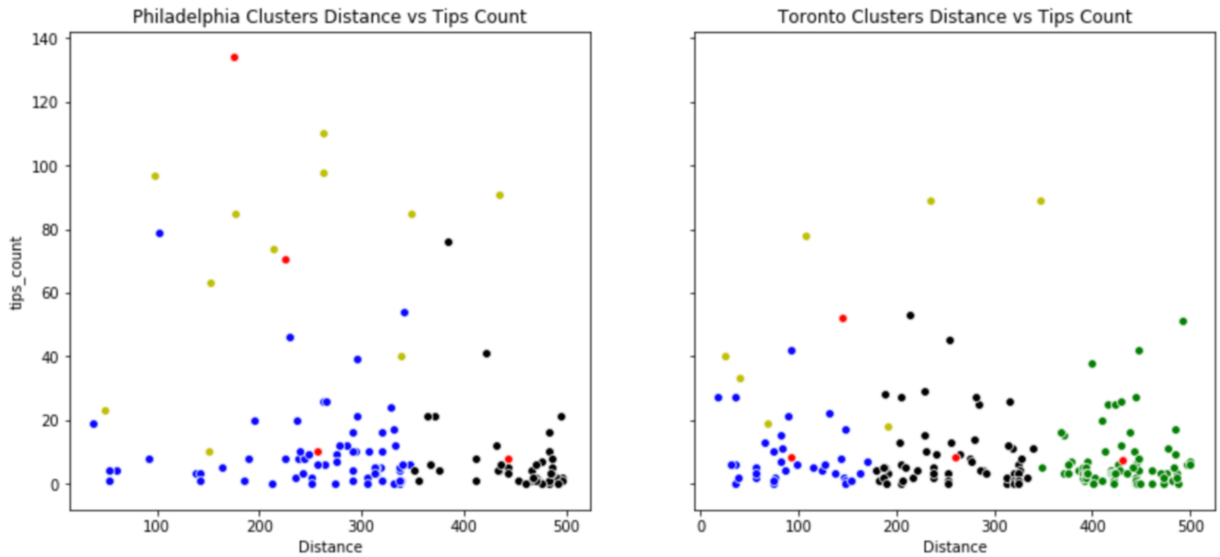


Figure 9.9 – Philadelphia Before Clustering

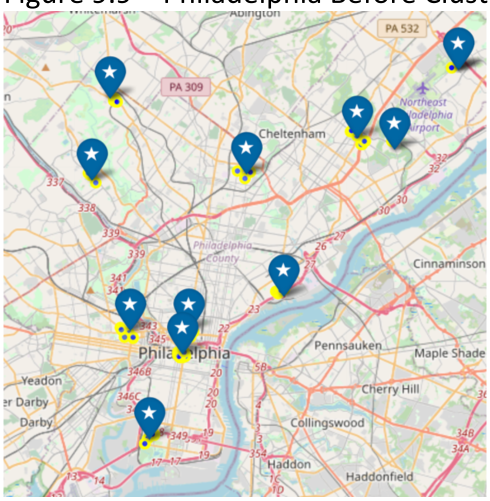


Figure 9.10 – Philadelphia Clustered

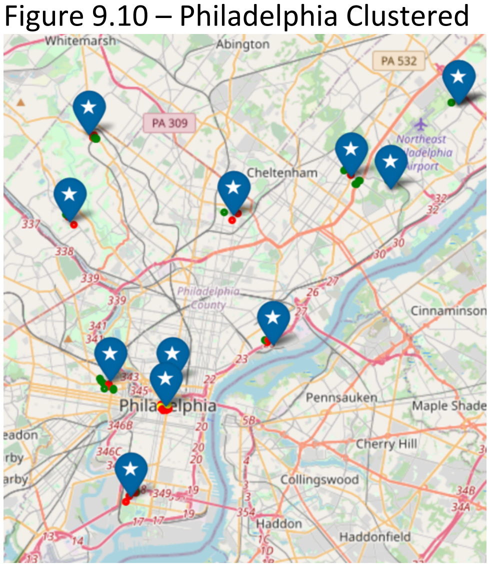


Figure 9.11 – Toronto Before Clustering

Figure 9.12 – Toronto Clustered

5. Results

Philadelphia's overall neighborhood count is 43.09% higher than Toronto's Neighborhoods. The top three boroughs in Toronto that make up 58.25% of the City are North Yok, Downtown Toronto and Scarborough. The top three boroughs in Philadelphia that make up 43.65% of the City are South Philadelphia, West Philadelphia and Center City. Toronto has an average of 10.3 neighborhoods per borough whereas Philadelphia has an average 8.58. (Figures 5.1 and 5.2)

Toronto has 26.45% of it's overall likes from venues within East Toronto, 20.38% in Scarborough, and 19.43% in North York. Whereas Philadelphia has 58% of it's overall likes located in Center City and 16.15% in South Philadelphia. Philadelphia has much more drastic separation of ratings for venues. (Figure 6.1 and 6.2)

The top 10 venues in each location were extracted and accounted for in each of the Boroughs for their respective cities. Toronto's top 10 included coffee shops, sandwich places, clothing stores, café's, bakeries, Vietnamese restaurants, bar, yoga studio, grocery store and restaurants. (Figure 7.1) Philadelphia's top 10 included Convenience Store, Sandwich Place. Pizza Place, Pharmacy, Coffee Shop, Sporting Goods Shop, Bar, Italian Restaurant, Donut Shop, and Hotels. (Figure 7.2)

Overall like counts, photo counts, and tip counts with ratings showed a positive correlation for both Toronto and Philadelphia. (Figures 8.1, 8.4 & 8.6) This persisted when observing a subset of these factors where the majority of the counts existed.(Figure 8.2, 8.5 & 8.6) Whereas the distance showed a negative correlation with ratings for both Philadelphia and Toronto. (Figure 8.3)

Our elbow point results showed that we should select a cluster range from 3 to 4 to have the best results. (Figure 9.1 & 9.2) Our mean values for the centroid points where then determined from the mean values of our k-means cluster with a n=4. (Figure 9.3 & 9.4) Next we observed the clusters values for likes, photos, tips and ratings by their distance to distinguish potential groupings or categorizations of clusters. (Figure 9.5, 9.6, 9.7 & 9.8). Finally, we visualized the venues in each borough both before and after clustering for Philadelphia (Figure 9.9 & 9.10) and Toronto (Figure 9.11 & 9.12).

6. Discussion

Our analysis of the venues at the boroughs within Philadelphia and Toronto observed the significance of venue likes, photo count and tips. This will provide useful information for venue owners to mangers and can be driving factors when trying to promote or establish new business practices within these areas. Segmentation of these venues based on these significant parameters will improve the user experience at these venues as well.

The descriptive analysis of the venue categories demonstrated the top categories in each of the boroughs and how they are distributed. These findings can provide a clear indicator to recommend business owners and startups where popular locations may be to start new business. This can be where there are limited to no venues in a borough that has potential for growth.

When comparing the two cities we see that Philadelphia has more neighborhoods within its limits than Toronto. However, 58% of the likes consist from Center City alone, whereas Toronto has a more even distribution of likes between East Toronto, Scarborough, and North York. Center City and East Toronto are the major locations of each Philadelphia and Toronto, respectively, as they boast ratings above an 8 and have venues with the majority of the likes, photos, tips and ratings.

We demonstrated in the regression analyses that ratings are positively correlated with likes, photo count and tips for each venue within Philadelphia and Toronto. We also observed ratings are negatively correlated with distances. This information is crucial when exploring ways to improve upon existing or new businesses within the area. It will also factor in where a good location for might be for these businesses to start.

The clustering of our venues in the boroughs gave us insight for Philadelphia and Toronto that most likes, ratings, photo count and tip counts were not affected by the distance on cluster label 2, as they had a tendency to have their values range higher than the rest despite of distance. The remains the same for Toronto as well, meaning these might be the hotspots for activity in their Boroughs and can provide means for people to evaluate and visit these venues.

7. Conclusions

In this project, we explored the effectiveness of distance, photos, tips, ratings and likes. We compared the venues, boroughs and neighborhoods of both Cities and revealed a strong positive correlation to ratings when observing likes, photos and tips for both cities. We also observed a small negative correlation with ratings with distance for Toronto and an even greater negative correlation for Philadelphia. While there are some limitations to this project, as larger data from Foursquare cannot be pulled without higher access and not all data pulled contains information for our venues that’s required for the analysis. Overall, we have established a working model that can only improve evaluation of venues and boroughs for potential locations for start-ups, business ventures and even travel hot spots as more information is made available.

Thank you!