Leveraging Ontology to Build Smart City -- Venue Location Analysis

1. Introduction

Location-based data analytics have been facilitating effective commercial marketing and decision making. It is of critical importance for decision-makers to assess venue patterns when establishing a new facility (e.g., hotel, coffee shop). For example, the location-based data would help if an investor is looking to open a new coffee shop.

Extant literature has been devoted to unveiling venue characteristic analysis and building machine learning models, such as mono-centric models, agglomeration models, and multi-dimension models (Yang et al., 2005). Specifically, prior research has investigated to index venue characteristics to improve decision quality.

However, limited research has explored neighborhoods' similarity in term of venue patterns to facilitate location selection. In this study, we are interested in illustrating how to build the neighbors' similarity matrix regarding venues features to facilitate location selection. We will take Etobicoke as an example.

2. Data

We will use the city of Toronto, specifically the Etobicoke Borough as an example. The structure neighborhood data with coordinates are not readily available. Therefore, we build a crawler to collect and structure the neighborhood information from the Wikipedia webpage [1]. Specifically, we collected information on postal code, borough, and neighborhood (Figure 1).

[1] https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

	Postal Code	Borough	Neighbourhood
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Regent Park, Harbourfront
5	M6A	North York	Lawrence Manor, Lawrence Heights
6	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government
...	...	...	...
160	M8X	Etobicoke	The Kingsway, Montgomery Road, Old Mill North
165	M4Y	Downtown Toronto	Church and Wellesley
168	M7Y	East Toronto	Business reply mail Processing Centre, South C...
169	M8Y	Etobicoke	Old Mill South, King's Mill Park, Sunnylea, Hu...
178	M8Z	Etobicoke	Mimico NW, The Queensway West, South of Bloor,...

103 rows × 3 columns

Figure 1. Neighborhood Data

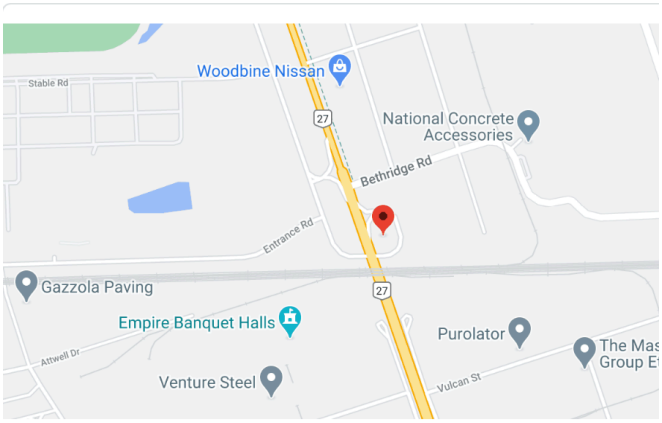
Meanwhile, we collect the coordinates of these neighborhood via Geocoder (Figure 2).

	Postal Code	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476
...	...	...	...
98	M9N	43.706876	-79.518188
99	M9P	43.696319	-79.532242
100	M9R	43.688905	-79.554724
101	M9V	43.739416	-79.588437
102	M9W	43.706748	-79.594054

103 rows × 3 columns

Figure 2. Neighborhood Coordinates

We collect all venue-relevant information via Foursquare API, with the latitude and the longitude coordinates of venues in each neighborhood (Figure 3).



Map for 43.7067,-79.59

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	West Deane Park, Princess Gardens, Martin Grov...	43.650943	-79.554724	Majestic Tees	43.649774	-79.555437	Print Shop
1	Eringate, Bloordale Gardens, Old Burnhamthorpe...	43.643515	-79.577201	LCBO	43.642099	-79.576592	Liquor Store
2	Eringate, Bloordale Gardens, Old Burnhamthorpe...	43.643515	-79.577201	Starbucks	43.641312	-79.576924	Coffee Shop
3	Eringate, Bloordale Gardens, Old Burnhamthorpe...	43.643515	-79.577201	The Beer Store	43.641313	-79.576925	Beer Store
4	Eringate, Bloordale Gardens, Old Burnhamthorpe...	43.643515	-79.577201	Shoppers Drug Mart	43.641312	-79.576924	Pharmacy
5	Eringate, Bloordale Gardens, Old Burnhamthorpe...	43.643515	-79.577201	Pizza Hut	43.641845	-79.576556	Pizza Place
6	Eringate, Bloordale Gardens, Old Burnhamthorpe...	43.643515	-79.577201	Cafe Sympatico	43.641820	-79.576721	Café
7	Eringate, Bloordale Gardens, Old Burnhamthorpe...	43.643515	-79.577201	Burnhamthorpe Mall	43.641741	-79.576891	Shopping Plaza
8	Eringate, Bloordale Gardens, Old Burnhamthorpe...	43.643515	-79.577201	Hasty Market	43.641871	-79.576370	Convenience Store
9	Westmount	43.696319	-79.532242	Starbucks	43.696338	-79.533398	Coffee Shop
10	Westmount	43.696319	-79.532242	Mayflower Chinese Food	43.692753	-79.531566	Chinese Restaurant

Figure3. Venues Dataset

3. Method

In this study, we aim to retrieve venues information and build a similarity matrix the neighborhoods in terms of venues. We follow a three-step procedure:

Step1—We map the neighborhoods to explore the neighborhoods; Step2—We conduct a venue similarity analysis via building the venue array and calculate the similarly matrix; and Step3—We statistically compared the similarities between neighborhoods and suggest potential places for new facility construction.

4. Findings

Through visualizing the neighborhoods, we can observe the neighborhoods distribution. There are large neighborhoods (larger bubble) and smaller neighborhoods (smaller bubble).

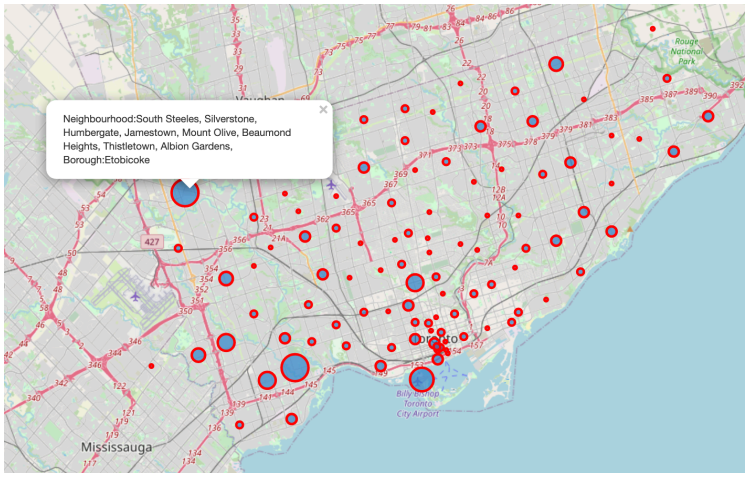


Figure. 4 Neighborhoods Map

In particular, we would construct a venue knowledge matrix for the Etobicoke area to see how the neighborhoods are different in terms of facilitate intensity.

Neighborhood	Bakery	Bar	Baseball Field	Beer Store	Burger Joint	Bus Line	Café	Chinese Restaurant	Coffee Shop	Convenience Store	Discount Store	Fast Food Restaurant	Fried Chicken Joint	Grocery Store	Gym	Hardware Store	Kids Store	Liquor Store	Park	Pharmacy	Pizza Place	Pool	Print Shop	Pub	Rental Car Location	San
Alderwood, Long Branch	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1	1	0	0	1	0	
Eringate, Bloordale Gardens, Old Burnhamthorpe, Markland Wood	0	0	0	1	0	0	1	0	1	1	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	
Kingsview Village, St. Phillips, Martin Grove Gardens, Richview Gardens	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	
Mimico NW, The Queensway West, South of Bloor, Kingsway Park South West, Royal York South West	1	0	0	0	1	0	0	0	0	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	
New Toronto, Mimico South, Humber Bay Shores	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Figure. 5 Neighborhoods Venue Samples

From this Venue knowledge matrix, we can see that there are each neighborhood has its own characteristic in relation to venues. For example, Alderwood, Long Branch has coffee shops while Eringate, Bloordale Gardens, Old Burnhamthorpe, Markland Wood has both coffee shops and beer stores.

We calculate the different distributions of venues in cosine similarity and present the results in Figure 6. from this similarity matrix, we observe that Neighborhood Alderwood, Long Branch and Neighborhood Westmount has venues similarity over than 0.6, while Neighborhood Alderwood, Long Branch and Neighborhood Kingsview Village, St. Phillips, Martin Grove Gardens, Richview Gardens has very limited similarity.

	0	1	2	3	...	7	8	9	10
0	1.000000	0.433013	0.0	0.109109	...	0.433013	0.0	0.0	0.612372
1	0.433013	1.000000	0.0	0.094491	...	0.375000	0.0	0.0	0.353553
2	0.000000	0.000000	1.0	0.000000	...	0.000000	0.0	0.0	0.000000
3	0.109109	0.094491	0.0	1.000000	...	0.188982	0.0	0.0	0.133631
4	0.408248	0.353553	0.0	0.000000	...	0.000000	0.0	0.0	0.500000
5	0.000000	0.000000	0.0	0.000000	...	0.000000	0.0	0.0	0.000000
6	0.000000	0.000000	0.0	0.000000	...	0.000000	0.0	0.0	0.000000
7	0.433013	0.375000	0.0	0.188982	...	1.000000	0.0	0.0	0.530330
8	0.000000	0.000000	0.0	0.000000	...	0.000000	1.0	0.0	0.000000
9	0.000000	0.000000	0.0	0.000000	...	0.000000	0.0	1.0	0.000000
10	0.612372	0.353553	0.0	0.133631	...	0.530330	0.0	0.0	1.000000

Figure. 6 Neighborhoods Similarity.

5. Discussion

Our findings reveal that the neighborhoods in Etobicoke have large neighborhoods and small neighborhoods. These neighborhoods have unique features of venues in terms of location and quantity. Specifically, some neighborhoods are similar in venue setting while others are not. Our findings also suggest that the matrix can help decision-makers to analyze the locations for new facilities.

6. Conclusion

This study makes two major contributions. First, it leverages three data sources to construct a venue coordinates dataset to analyze neighborhood similarity. Second, it demonstrates a method to construct a venue similarity matrix for decision-makers to analyze venues in neighborhoods. Further research direction includes ontologizing the venue and coordinating information to establish a smart city knowledge map of neighborhoods.

7. Reference:

Yang, Y., Tang, J., Luo, H., & Law, R. (2015). Hotel location evaluation: A combination of machine learning tools and web GIS. *International Journal of Hospitality Management*, *47*, 14-24.