

# Finding the Right Neighborhood:

## Toronto and the Twin Cities

By: Alexander Stetzer

# Introduction

- Goal:
  - Use Foursquare venue data to compare the neighborhoods of Toronto and the Twin Cities
  - Help potential homebuyers in finding similar neighborhoods to move to based on venue data.
- Interested Parties:
  - Potential homebuyers or renters looking to find a new place within one of the other two cities
  - Real Estate agencies looking for neighborhoods to help suggest new places for clients

# Data Sources

- Data comes from the cities wikipedia pages
  - [Minneapolis](#)
  - [St. Paul](#)
  - [Toronto](#)
- Location data for Toronto from postal code geospatial csv
- Location data for the Twin Cities comes from Google Geocoder API
- Venue Data from Foursquare API

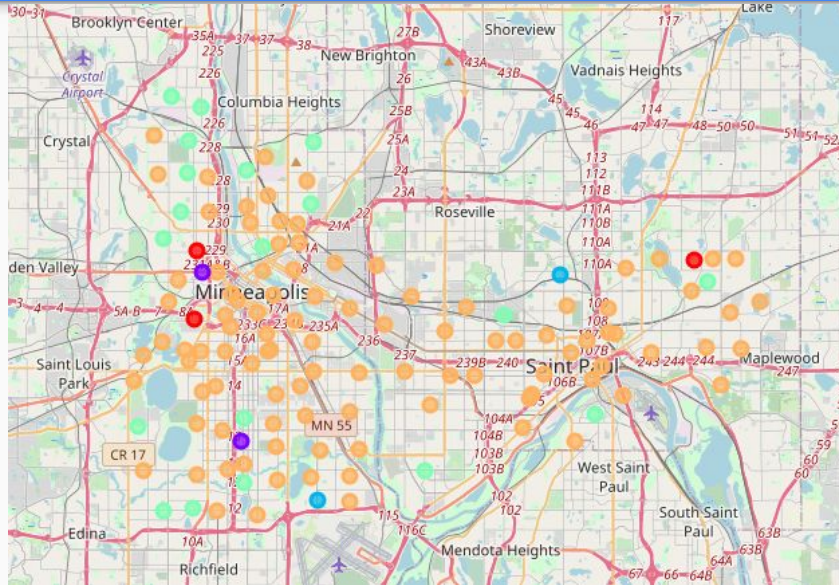
# Data Cleaning

- One Hot Encoding on the venue data
  - Creates a DataFrame with rows of venues and columns of venue categories
- Data Grouped by Neighborhood
  - Allows for ease of comparison between neighborhoods
- Mean taken of each venue category
  - Normalizes the data for each neighborhood

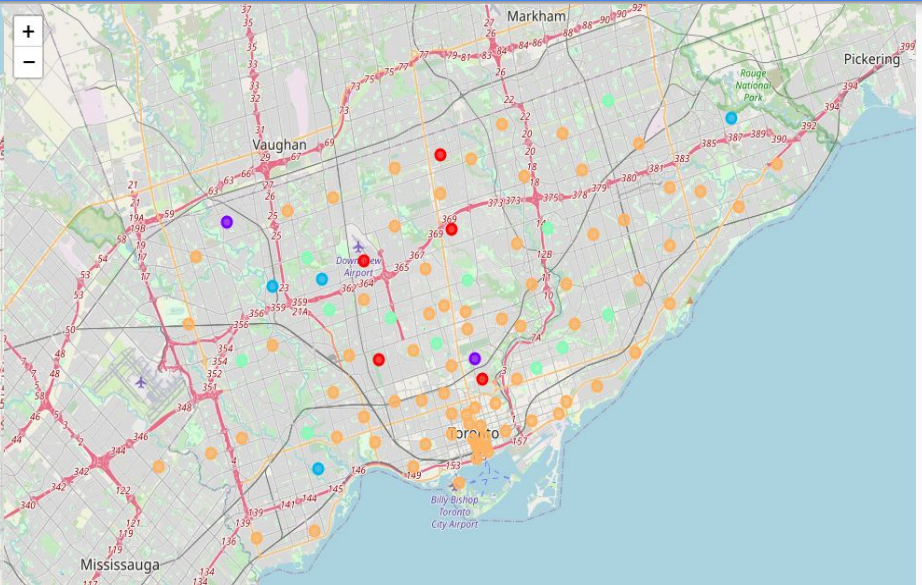
# Methods

- KMeans Clustering
  - Five clusters selected
  - Clusters used as first step of similarity testing
- Cosine Similarity
  - $\cos(\theta) = (A \cdot B) / (\|A\| * \|B\|)$
  - Allows for comparison based on similarities of each venue category
- Final Printables
  - Find\_neigh function allows for user to input home neighborhood
  - Function returns top five similar neighborhoods and top ten venues of each neighborhood

# Clustering Results



Twin Cities, Minnesota, United States



Toronto, Ontario, Canada

# Cosine Similarity

## Cluster 0:

	City 1	Neighborhood 1	City 2	Neighborhood 2	Cosine Sim
0	Minneapolis	Lowry Hill	Toronto	Willowdale, Newtonbrook	1.000000
1	Minneapolis	Lowry Hill	Toronto	York Mills West	0.894427
2	Minneapolis	Lowry Hill	Toronto	Rosedale	0.816497
3	Toronto	Caledonia-Fairbanks	Minneapolis	Lowry Hill	0.816497
4	Minneapolis	Near North	Toronto	Willowdale, Newtonbrook	0.816497

## Cluster 2:

	City 1	Neighborhood 1	City 2	Neighborhood 2	Cosine Sim
0	Toronto	Humberlea, Emery	St. Paul	North of Maryland	0.816497
1	St. Paul	North of Maryland	Toronto	Old Mill South, King's Mill Park, Sunnylea, Hu...	0.577350
2	Toronto	Humberlea, Emery	Minneapolis	Wenonah	0.577350
3	St. Paul	North of Maryland	Minneapolis	Wenonah	0.471405
4	Toronto	DownsviewCentral	St. Paul	North of Maryland	0.471405

## Cluster 4:

	City 1	Neighborhood 1	City 2	Neighborhood 2	Cosine Sim
0	Minneapolis	Standish	Toronto	Woburn	0.894427
1	Toronto	Harbourfront East, Union Station, Toronto Islands	Minneapolis	Standish	0.809524
2	Toronto	Richmond, Adelaide, King	Minneapolis	Standish	0.776750
3	Minneapolis	East Harriet	Toronto	Roselawn	0.750000
4	Toronto	Ontario Provincial Government	Minneapolis	Standish	0.744208

## Cluster 1:

	City 1	Neighborhood 1	City 2	Neighborhood 2	Cosine Sim
0	Toronto	Humber Summit	Minneapolis	Regina	0.57735
1	Toronto	Moore Park, Summerhill East	Minneapolis	Regina	0.57735
2	Toronto	Humber Summit	Minneapolis	Sumner-Glenwood	0.50000
3	Toronto	Moore Park, Summerhill East	Minneapolis	Sumner-Glenwood	0.50000

## Cluster 3:

	City 1	Neighborhood 1	City 2	Neighborhood 2	Cosine Sim
0	Minneapolis	Folwell	St. Paul	Highland Park	0.816497
1	Minneapolis	Linden Hills	Toronto	Milliken, Agincourt North, Steeles East, L'Amo...	0.654654
2	Toronto	DownsviewWest	Minneapolis	Willard Hay	0.566947
3	Minneapolis	Lind-Bohanon	St. Paul	Riverview	0.516398
4	Toronto	Glencairn	Minneapolis	Linden Hills	0.507093

# Final Printables for Near North, Minneapolis

- These are the DataFrames the Find\_neigh function returns

	City 1	Neighborhood 1	City 2	Neighborhood 2	Cosine Sim
0	Minneapolis	Near North	Toronto	Willowdale, Newtonbrook	0.816497
1	Minneapolis	Near North	Toronto	York Mills West	0.730297
2	Minneapolis	Near North	Toronto	Caledonia-Fairbanks	0.666667
3	Minneapolis	Near North	Toronto	Rosedale	0.666667
4	Minneapolis	Near North	Toronto	DownsviewEast	0.577350

- Top:
  - Top 5 Similar Neighborhoods
- Bottom:
  - Top Ten Most Common Venues

	Cluster Labels	City	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	0	Minneapolis	Near North	Park	Miscellaneous Shop	Wine Bar	Yoga Studio	Donut Shop	Drugstore	Eastern European Restaurant	Electronics Store	Elementary School	English Restaurant
1	0	Toronto	Willowdale, Newtonbrook	Park	Yoga Studio	Ethiopian Restaurant	Donut Shop	Drugstore	Eastern European Restaurant	Electronics Store	Elementary School	English Restaurant	Escape Room
2	0	Toronto	York Mills West	Park	Convenience Store	Yoga Studio	Donut Shop	Drugstore	Eastern European Restaurant	Electronics Store	Elementary School	English Restaurant	Escape Room
3	0	Toronto	Caledonia-Fairbanks	Park	Women's Store	Pool	Yoga Studio	Doner Restaurant	Donut Shop	Drugstore	Eastern European Restaurant	Electronics Store	Elementary School
4	0	Toronto	Rosedale	Park	Trail	Playground	Fish & Chips Shop	English Restaurant	Doner Restaurant	Fish Market	Donut Shop	Drugstore	Eastern European Restaurant
5	0	Toronto	DownsviewEast	Airport	Park	Yoga Studio	Ethiopian Restaurant	Donut Shop	Drugstore	Eastern European Restaurant	Electronics Store	Elementary School	English Restaurant



# Discussion

- Interesting observations
  - Downtown areas typically in similar cluster
    - Overlap of venues
    - Similar venue audience
  - Two Neighborhoods of 1.0 Cosine Similarity
    - Lowry Hill, Minneapolis
    - Willowdale, Newtonbrook, Toronto
  - Find\_neigh function found very similar neighborhoods

# Conclusion

- Foursquare data analyzed and similar neighborhoods found for potential homebuyers and other interested parties
- Improvements:
  - Add Housing Costs
  - Crime Rates
  - Venue Quality
  - Other Key Neighborhood Data