Real Estate Neighborhood Recommendations: City of Toronto

By Sergio Conde

Problem

- Customer & Product Segmentation
- Being able to identify homes or neighborhoods that specific customer segments might want. The better you know your available products or where to find them fast as possible, the best you can help customers and fastrack sales

Solution

- Data!
- By collecting data from cities and their neighborhoods, real estate companies can create better informed product recommendations for their clients
- In this report, I will focus on the city of Toronto and attempt to identify neighborhoods that tailor to specific real estate customers

Data

- Sources of data:
  - FourSquare
    - Acquiring venue data for each neighborhood by using their free API
    - Thing to consider: there is a limit to how many venues I can receive at once in free version
  - Wikipedia
    - Scraping tables to obtain all the neighborhoods in Toronto

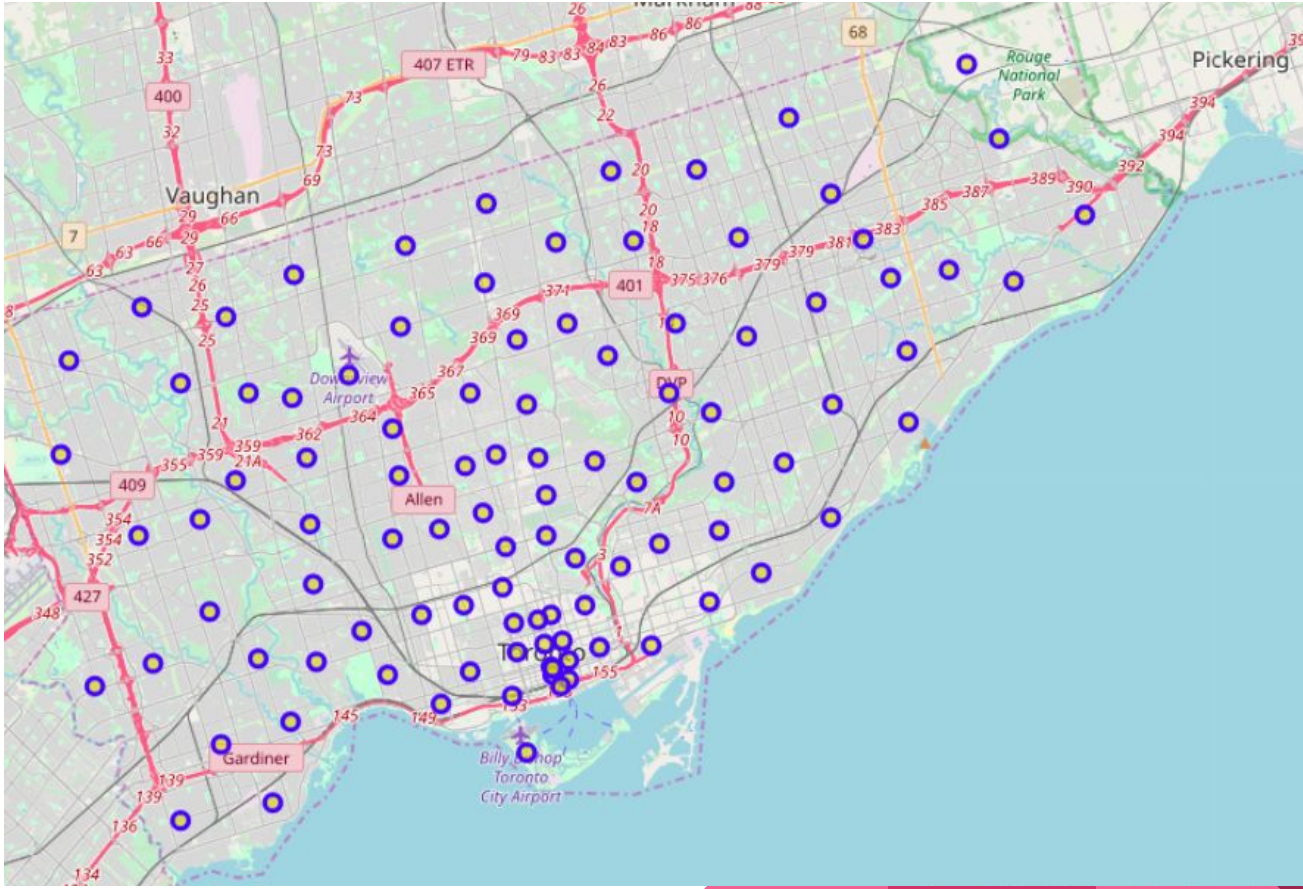
Methodology

1. Get Toronto neighborhood postcodes from Wikipedia
2. Use geolocation python library to get latitude and longitude coordinates of neighborhoods from their postcodes (necessary for FourSquare API input)
3. Feed coordinates into FourSquare API to get all the venues in each neighborhood
4. Analyze data & make some visuals
5. Use K-means clustering machine learning algorithm to group neighborhoods into four distinct segments.
6. Identify potential customers for those clustered neighborhoods

Data Analysis

- Higher neighborhood concentration around city downtown

Toronto Neighborhood Map



Data Analysis

- About 61% of neighborhoods have less than 50 venues
- Keep in mind FourSquare venue amount limit is low

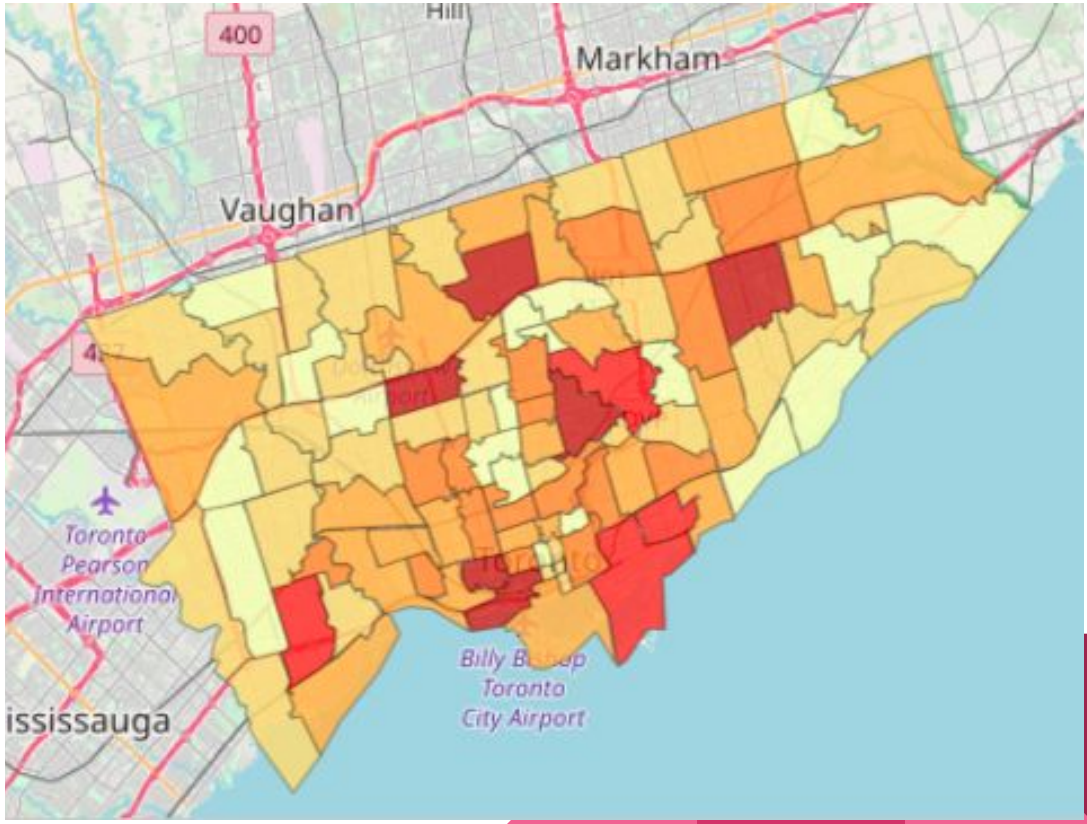
Top 10 Neighborhoods by Venue Count

	Neighbourhood	Venue Count
0	Dorset Park, Scarborough Town Centre, Wexford ...	118
1	CN Tower, Bathurst Quay, Island airport, Harbo...	114
2	Willowdale South	104
3	Little Portugal, Trinity	102
4	Lawrence Heights, Lawrence Manor	101
5	Leaside	101
6	Flemingdon Park, Don Mills South	87
7	The Beaches West, India Bazaar	86
8	Studio District	86
9	Kingsway Park South West, Mimico NW, The Queen...	81
10	Humewood-Cedarvale	78

Data Analysis

- Higher overall venue activity around downtown area
- Venue busy neighborhoods have calmer neighborhoods around them
- Super busy neighborhoods (dark red= 99+ venues) are usually not next to each other

Venue Distribution by Neighborhood
(the darker the more venues)



Data Analysis

- 338 unique venue categories
- Parks are very common, way to go Toronto!
- Coffee shops, cafes, and pizzas take the stage as well
- One of the least common: nudist beach
    - Surprised they got one though!

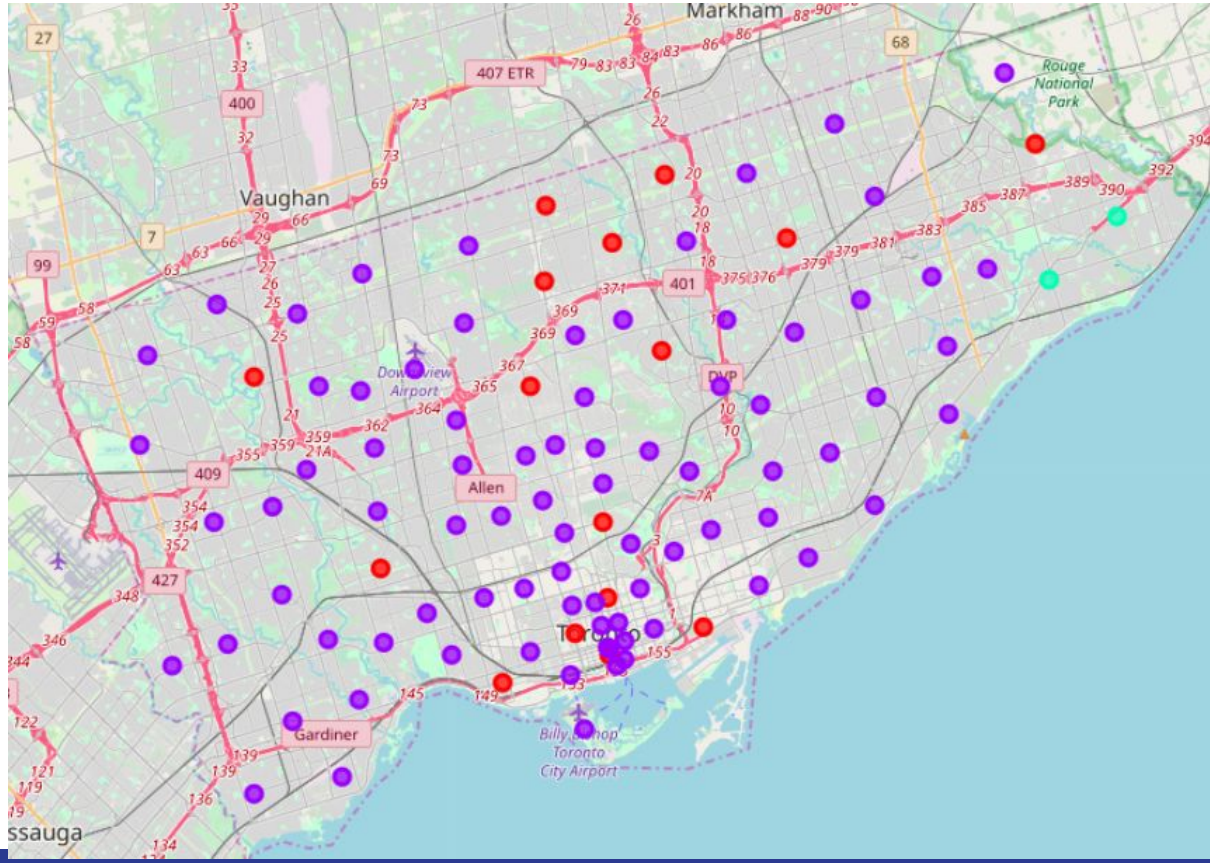
Most Common Venues

Coffee Shop	362
Pizza Place	156
Café	148
Park	142
Sandwich Place	118
Fast Food Restaurant	112
Bakery	106
Italian Restaurant	104
Restaurant	94
Pharmacy	93
Grocery Store	93
Bank	84
Japanese Restaurant	65
Sushi Restaurant	64
Chinese Restaurant	64
Bar	62
Burger Joint	52
Gym	49
Pub	48
Breakfast Spot	48

Results

- Used K-Means clustering algorithm with k=3 clusters
- Purple: tourist vibe
- Turquoise: small town atmosphere
- Red: residential/tourist atmosphere

Clusters Map
(each color = one cluster)



Results

- Cluster 0 (Purple):
  - Tourist friendly neighborhoods with frequent hotels, bars, coffee shops, greater food diversity, and entertainment
  - Perfect for customers looking for airbnb, hotel, other tourist related properties
  - It could also be perfect for people looking for a fast paced lifestyle: young singles

Results

- Clusters 1 (Turquoise):
  - Small town vibe neighborhoods. They have essentials such as supermarkets, pharmacies, and food places. Nothing crazy here. Perfect for customers looking for homes with a small town feel

Results

- Cluster 2 (Red):
  - Lots of food places, entertainment here and there but not as common as cluster 0
  - Frequent grocery stores, and pharmacies. This is a mix of small town and tourist friendly neighborhoods
  - Perfect for customers looking for a calm, yet fun nearby neighborhoods. Perfect for families
  - Also relevant to airbnb/real estate investors looking to provide local/residential type vacations to customers due to its semi-tourist status!

Conclusion

- Data can help businesses find insights into type of markets, products, and customers, as shown in my neighborhood analysis of the city of Toronto
- Using FourSquare, geolocation libraries, and publicly available data of the city, I was able to find insights on the types of venues available per neighborhood in the city
- Taking venues into account, 3 neighborhood clusters were extracted from the data bia K-Means Clustering:
    - Tourist Neighborhoods
    - Small Town Neighborhoods
    - Semi-Tourist Residential Neighborhoods

Conclusion

- With this data, real estate companies in Toronto can rapidly pinpoint best fit neighborhoods based on their client needs! It gives them an idea on where to first look for their customers dream home
- Looking forward:
  - This data can also be used to help potential business owners locate the best neighborhoods to set up shop based on competition (types of venues nearby) or foot traffic (amount of venues)
  - The data can be enhanced with better venue location services, as the FourSquare API free version is very limiting to the amount of venues that can be obtained

Thank You!