Recommendations for Opening New Restaurant in Toronto, Canada

Vikramjit Singh Rathee

Ph.D. in Chemical & Biomolecular Engineering (University of Notre Dame, USA)

Website: https://vikramjitsinghrathee.com/

Libraries Utilized and Project Notebook Link

- Numpy, Pandas, Math – Data analysis
- Matplotlib – Data visualization
- Beautiful Soup – Scraping webpage data
- Geopy – Obtaining geospatial coordinates (latitude and longitude)
- Folium – Visualization of Toronto map and marking important areas
- Sklearn – KMeans unsupervised clustering algorithm
- Requests – Sending organic, grass-fed HTTP/1.1 requests

Complete Project Notebook link

Methodology – Explore Important Factors & Provide Recommendations

- Multiple factors determine which locations are ideal for opening a new restaurant
  - Interested in specialized cuisine or fine dinning restaurant
  - Preferably 3 star and above

- We discuss important factors and motivation in sections and then provide final recommendation
  - Section – 1: Best Borough and Top Recommended Venues in Toronto
  - Section – 2: Property Prices in Toronto
  - Section – 3: Crime and Average Family Income in Toronto
  - Section – 4: Comparison with Competing Restaurants in Toronto
  - Section – 5: Final Recommendations

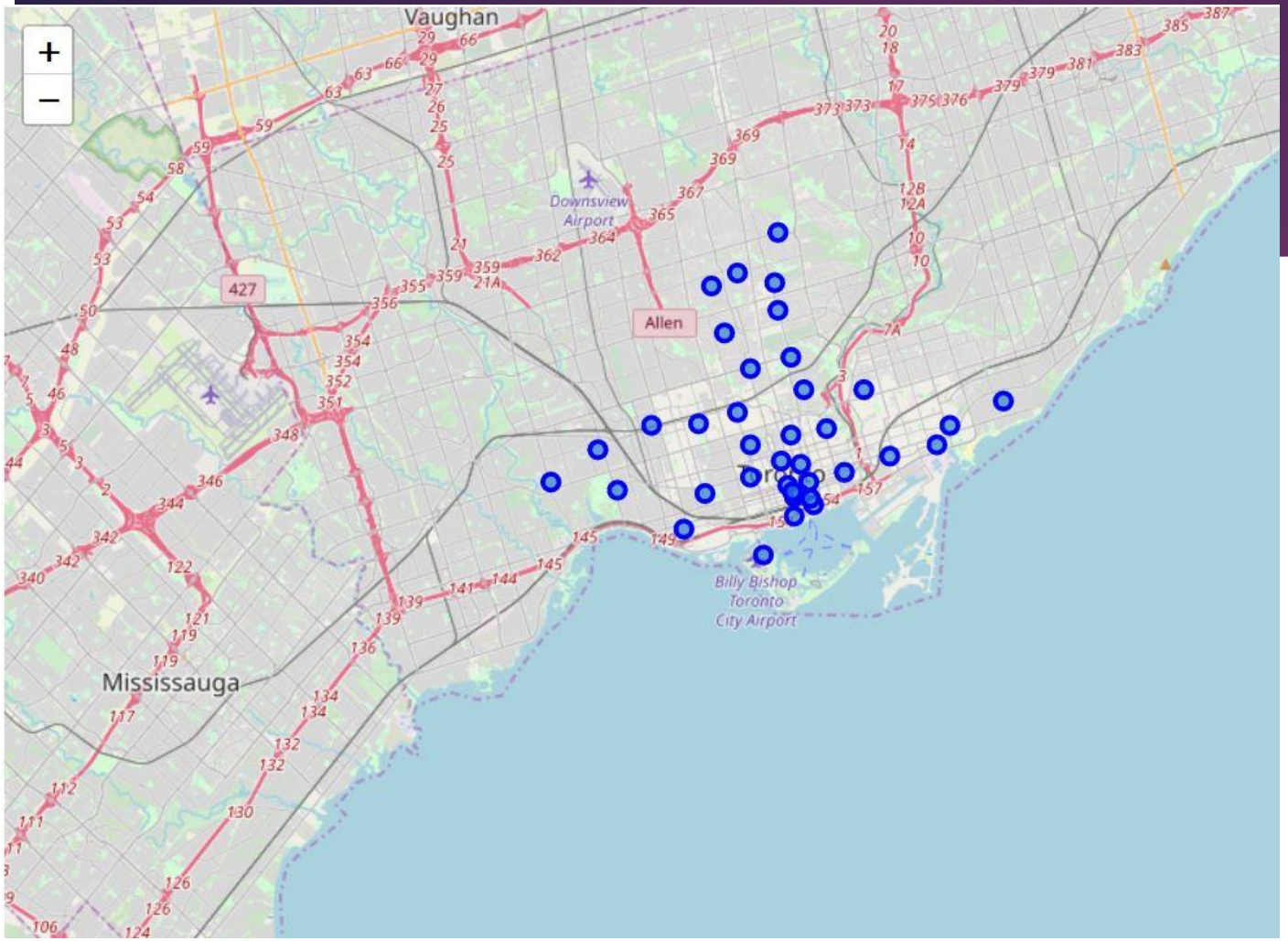
Best Borough and Top Recommended Venues in Toronto

Section - 1

Motivation

- Find the best area in Toronto with regards to density of top recommended venues for tourists
  - Map these top recommended areas as well top 5 venues and top 3 restaurants
- **Why?** - Identification of areas that are in proximity to top recommended venues in Toronto would be useful; Opening new restaurant in these areas increase the chances of tourists visiting the restaurant

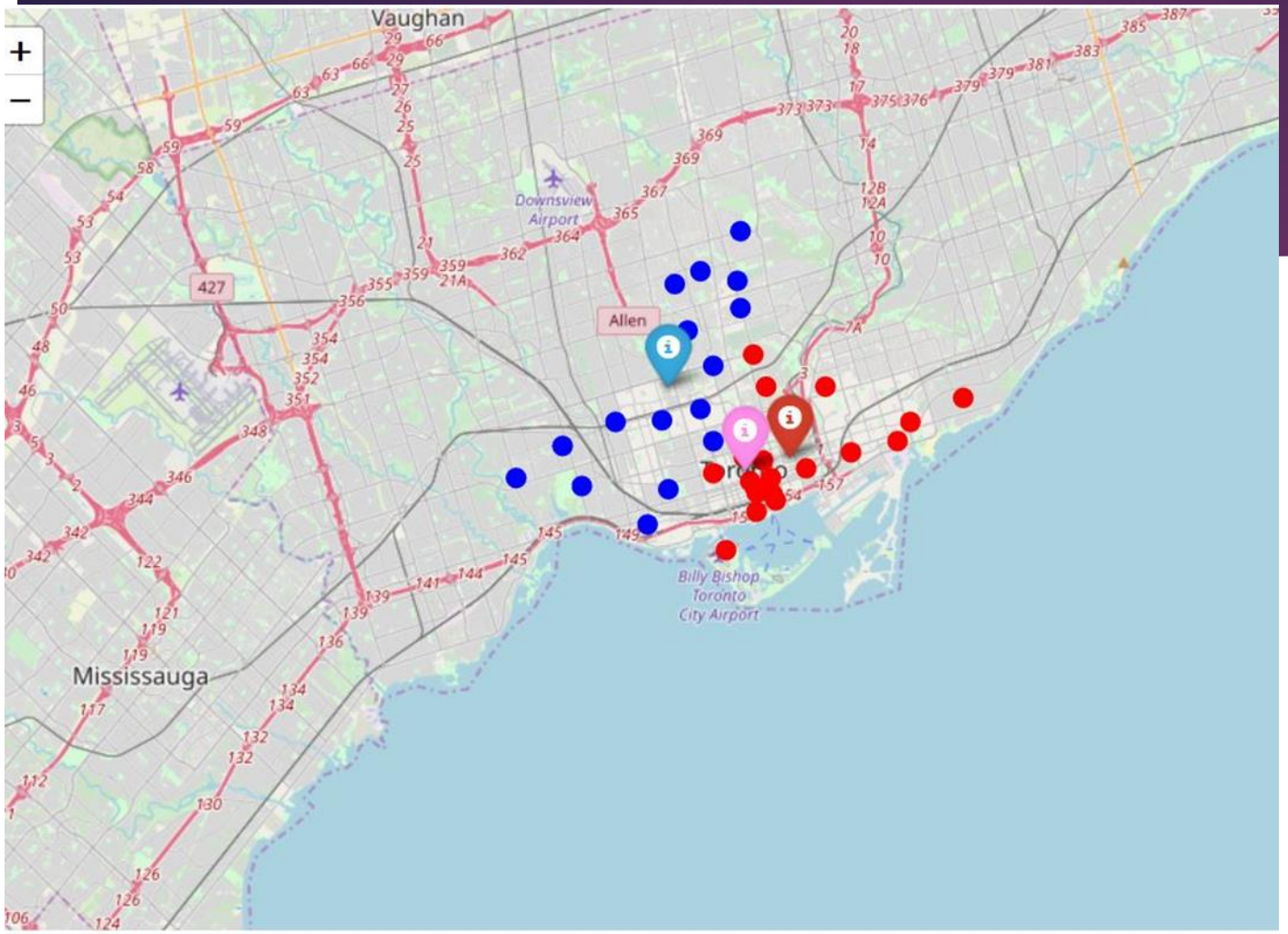
Distribution of Postal Codes



- Data obtained by scraping the Wikipedia page
  - Postal codes, boroughs and neighborhoods in Toronto
- Blue points depict geographical coordinates of the postal codes on the Toronto map

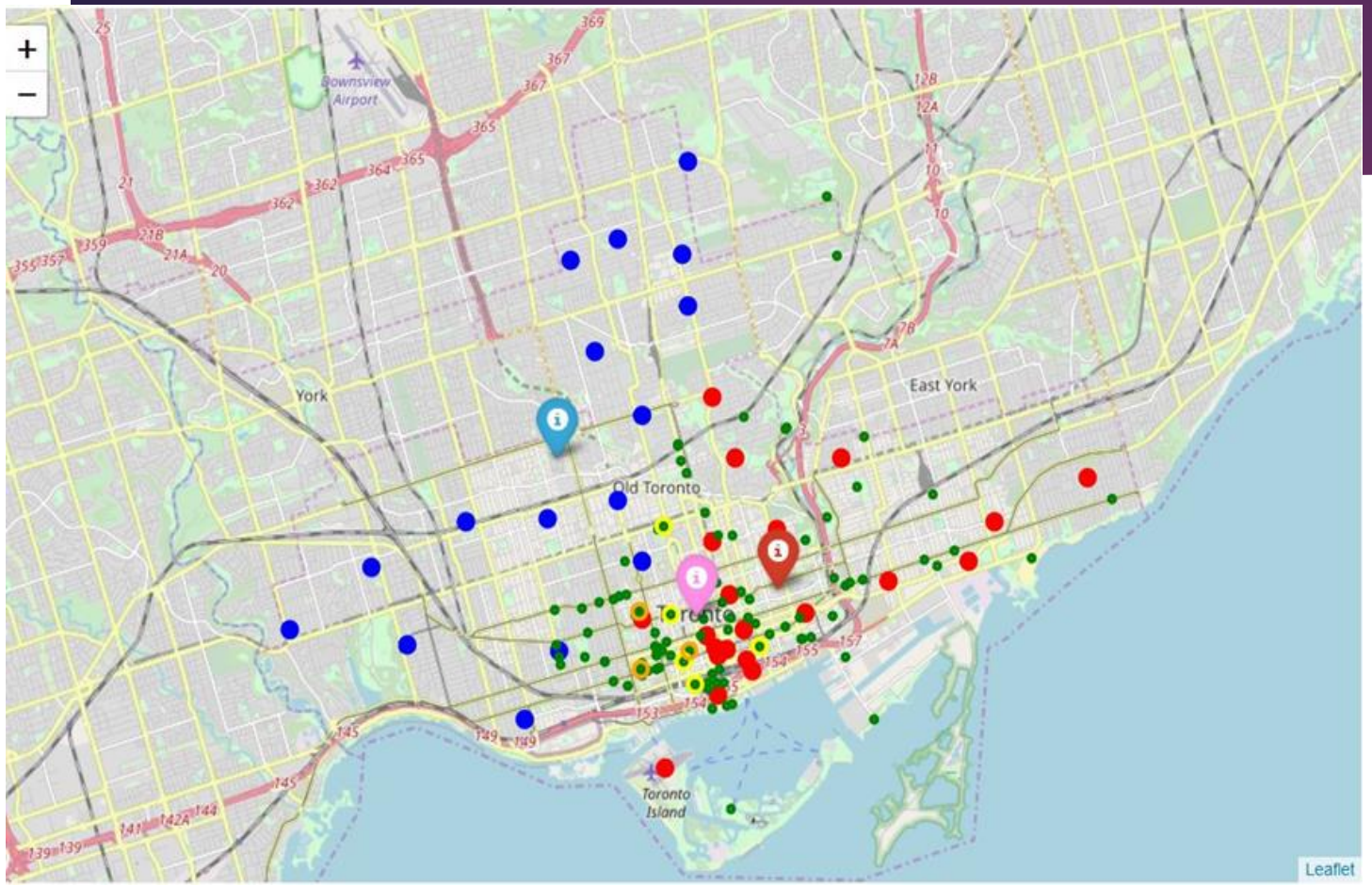
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

Clustering Postal Codes

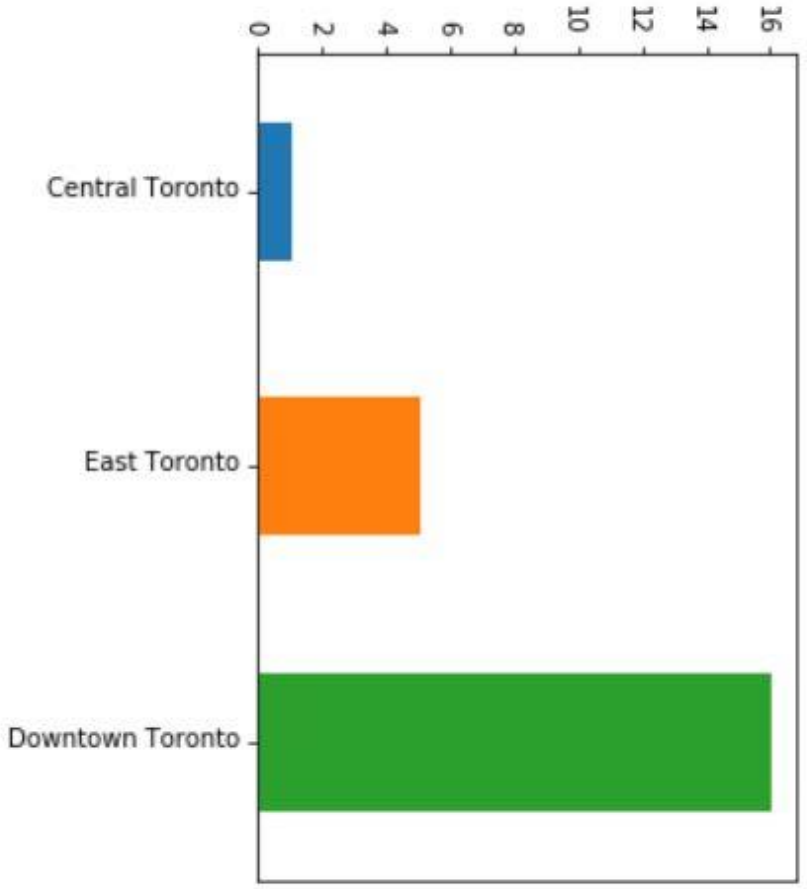


- Cluster the postal codes into 2 groups
  - KMeans clustering (unsupervised machine learning algorithm)
  - Red and blue points indicate postal codes belonging to Cluster 0 and Cluster 1, respectively.
  - Pink, red and blue pinpoint marker represent the Toronto center, Cluster 0 and Cluster 1 center coordinates, respectively.

Top Recommended Venues



- Extract top recommended venues based on number of ‘likes’.
  - Utilized Foursquare API
  - Green circles - Top recommended venues
  - Yellow hollow rings - Top 5 recommended venues; Orange hollow rings - Top 3 restaurants
- Top recommended venues are largely grouped in Cluster 0 (red circle markers) and few in Cluster 1



Bar plot depicts the distribution of top recommended venues in the boroughs that reside in Cluster 0 (red circle markers)

Summary - Top venues and best borough

- Downtown Toronto Borough has most recommended venues for the visiting tourists
- Downtown Toronto is in Cluster 0 (red circle markers)

- Most influx of tourists in Downtown
- Advisable to open restaurant that is close to or in this borough

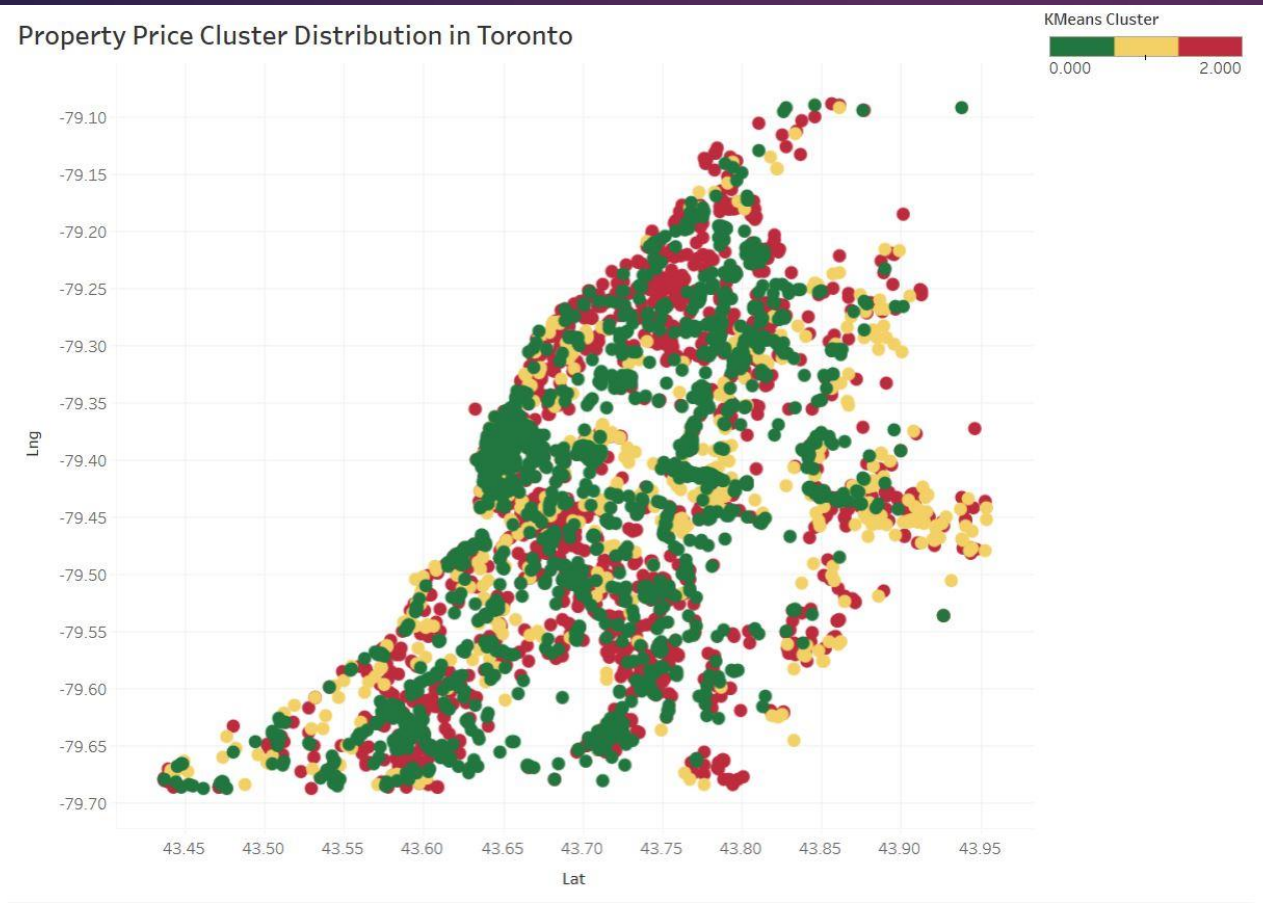
Property Prices in Toronto

Section - 2

Motivation

- Useful to cluster properties in Toronto based on their prices
  - Mapping on geographical coordinates will help with visualization
- **Why?** – Clustering properties based on their prices can segregate low, medium and high priced properties!
  - Will provide info about the property pricing that is to be expected in the preferable areas

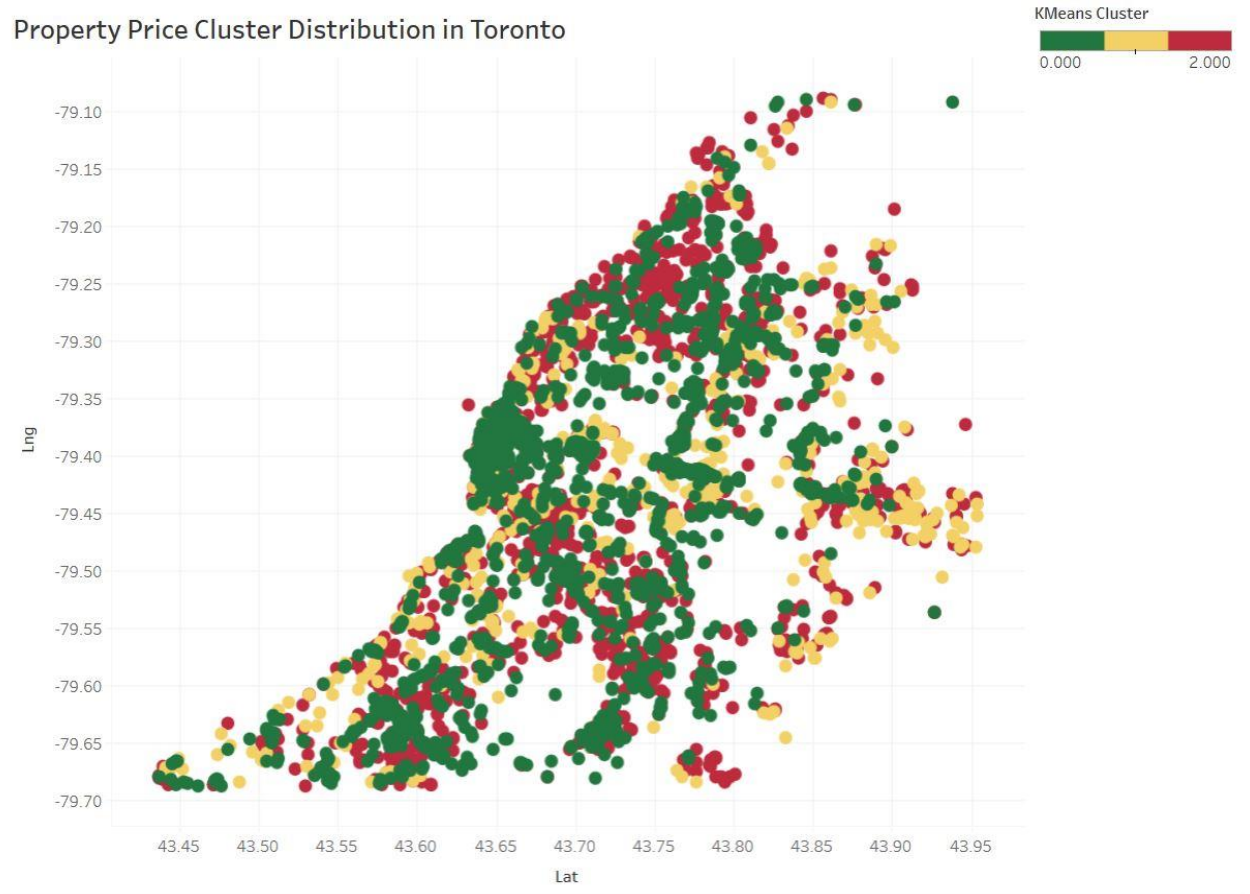
Property Price Clustering



- Cluster the property location and prices into 3 groups using KMeans algorithm
  - Low, Medium and High prices
  - 3D clustering, dimensions of each point are Latitude, Longitude, Price

Cluster	Cluster group #	Cluster center value
Low price	0	~ $ 348342
Medium price	2	~ $ 721802
High price	1	~ $ 1309962

https://www.kaggle.com/mnabaee/ontarioproperties



Summary - Property Pricing

Properties of low to high prices are distributed throughout Toronto and are not localized to particular areas

This is good, because this gives us more options regarding location.

Other factors must be considered to narrow down options

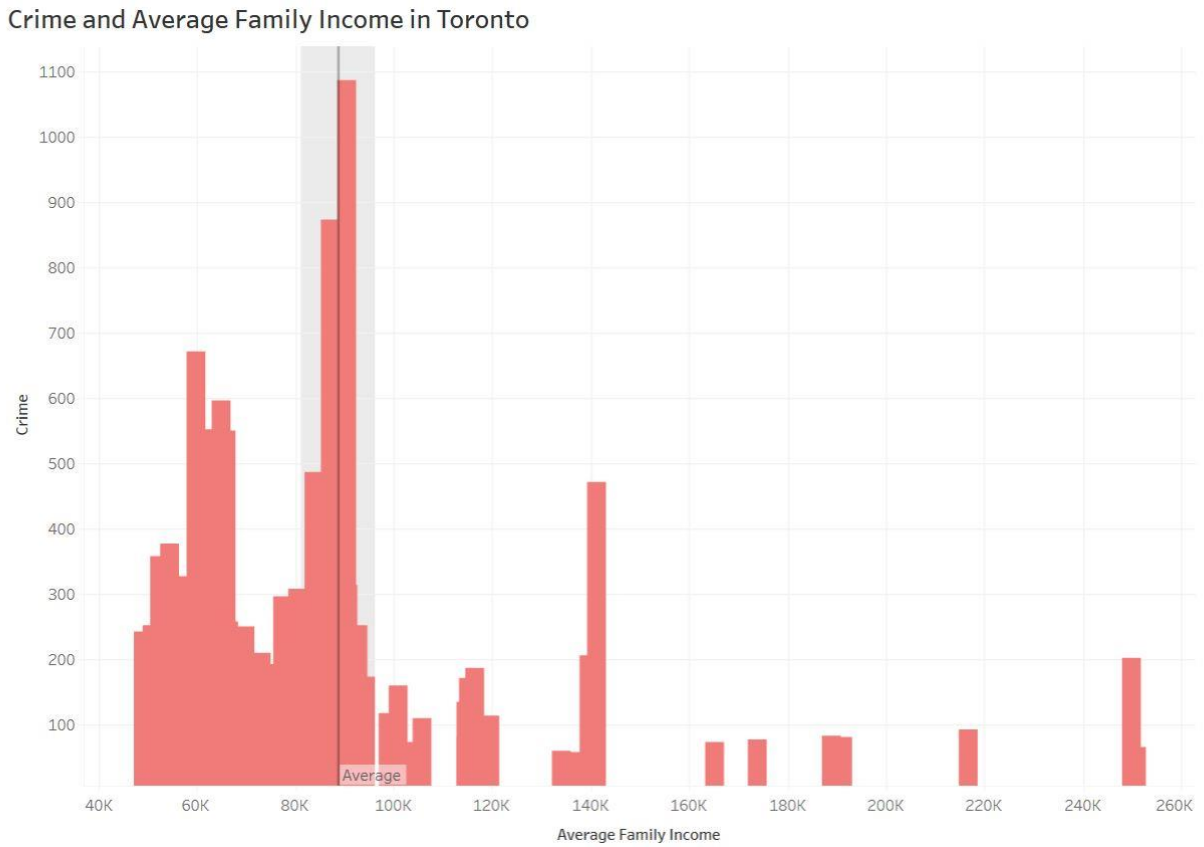
Crime and Average Family Income in Toronto

Section - 3

Motivation

- Explore crime, Average family income - preferable areas in Toronto for opening restaurant
- **Low crime areas are essential**
  - Safe and relaxing environment.
  - Crime here is combined metric consisting of - Murders, Thefts, Vehicle Thefts, Robberies, Break & Enters, Sexual Assaults, Assaults, Drug Arrests cases.
- **Areas with high 'Average Family Income'** - more likely to have access to leisure income to spend on eating at restaurants.
- **Proximity to tourists**

Crime and Average Family Income



Narrowing option - extract areas with areas with 'Average Family Income' > $96309

- **'Average Family Income' > 95% Confidence Interval (average) = $ 96309, shown in gray color, seems to be the best fit for opening restaurant.**
  - These are high income areas - more leisure income to spend on restaurants
  - Areas have less crime incidences

https://www.toronto.ca/city-government/data-research-maps/neighbourhoods-communities/wellbeing-toronto/

Preferable Locations Based on Crime and Average family Income

	Neighbourhood	Neighbourhood Id	With Bachelor Degree or Higher	Average Family Income	Healthy Food Index	Pop 25-45 years	Crime	Lat	Lng
2	Edenbridge-Humber Valley	9.0	5190.0	119581.0	46.41	3550.0	114.0	43.671478	-79.516731
21	Banbury-Don Mills	42.0	10850.0	115239.0	21.31	6410.0	171.0	43.734804	-79.357243
40	The Beaches	63.0	9235.0	139757.0	31.53	5870.0	207.0	43.671024	-79.296712
59	High Park-Swansea	87.0	10665.0	116465.0	46.53	8085.0	187.0	43.644940	-79.478313
61	Runnymede-Bloor West Village	89.0	4165.0	114798.0	45.64	2940.0	82.0	43.651778	-79.475923
65	Annex	95.0	16590.0	141111.0	38.85	11485.0	471.0	43.670338	-79.407117
67	Yonge-St.Clair	97.0	6775.0	173751.0	32.65	4320.0	78.0	43.688210	-79.394004
69	Yonge-Eglinton	100.0	6255.0	140907.0	31.95	4225.0	144.0	43.706748	-79.398327
71	Lawrence Park South	103.0	7880.0	216754.0	37.53	3390.0	93.0	43.729199	-79.403252
74	Humewood-Cedarvale	106.0	6060.0	105770.0	32.85	5045.0	110.0	43.690248	-79.422097
88	Centennial Scarborough	133.0	4055.0	102278.0	9.35	2935.0	74.0	43.787491	-79.150768
89	Highland Creek	134.0	3055.0	98857.0	0.63	2820.0	118.0	43.790117	-79.173334
95	Rosedale-Moore Park	98.0	11425.0	249884.0	32.20	4895.0	203.0	43.690388	-79.383297
96	Forest Hill South	101.0	5575.0	251035.0	18.49	2745.0	66.0	43.693559	-79.413902

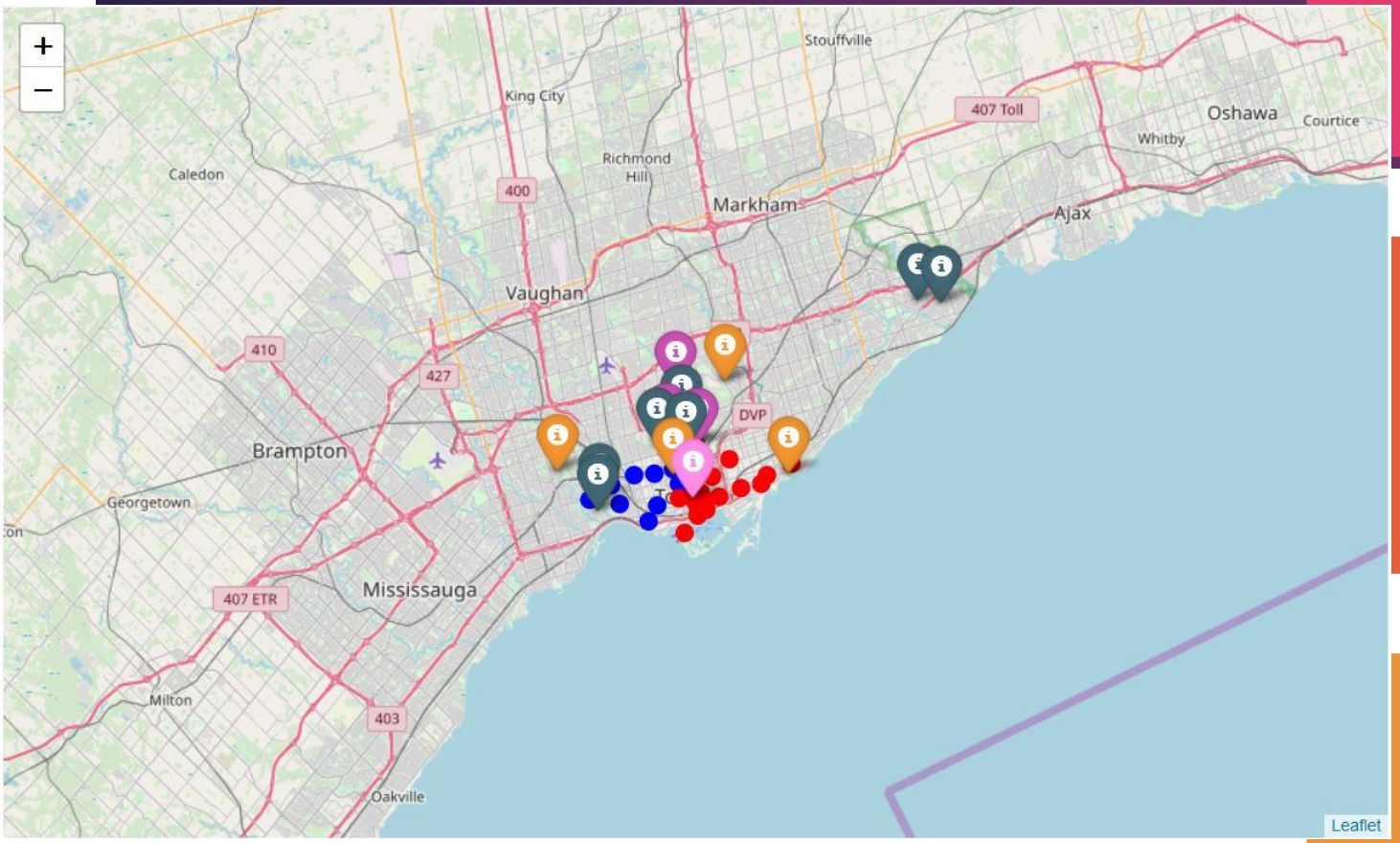
- Useful to assign property pricing cluster to each of these places
  - Property pricing expected for opening restaurant in these areas
- Approximating property prices in these areas - Find the nearest property in the property pricing Cluster and then assign that Cluster group.

Preferable Locations Based on Crime and Average family Income

- - Low price cluster (~$350K)
- - Medium price cluster (~$700K)
- - High price cluster (~$1.3M)

4 areas in Cluster 0 (Low price cluster)
3 areas in Cluster 1 (High price cluster)
7 areas in Cluster 2 (Medium price cluster)

Summary - Preferable Locations



Options narrowed down to 14 preferred areas in Toronto for opening a new restaurant

- Low crime incidences
- High income families

Proximity to Tourists would be useful for sustained customer flux.

4 areas in Cluster 0 (Low price cluster)

3 areas in Cluster 1 (High price cluster)

7 areas in Cluster 2 (Medium price cluster)

Comparison with Competing Restaurants in Toronto

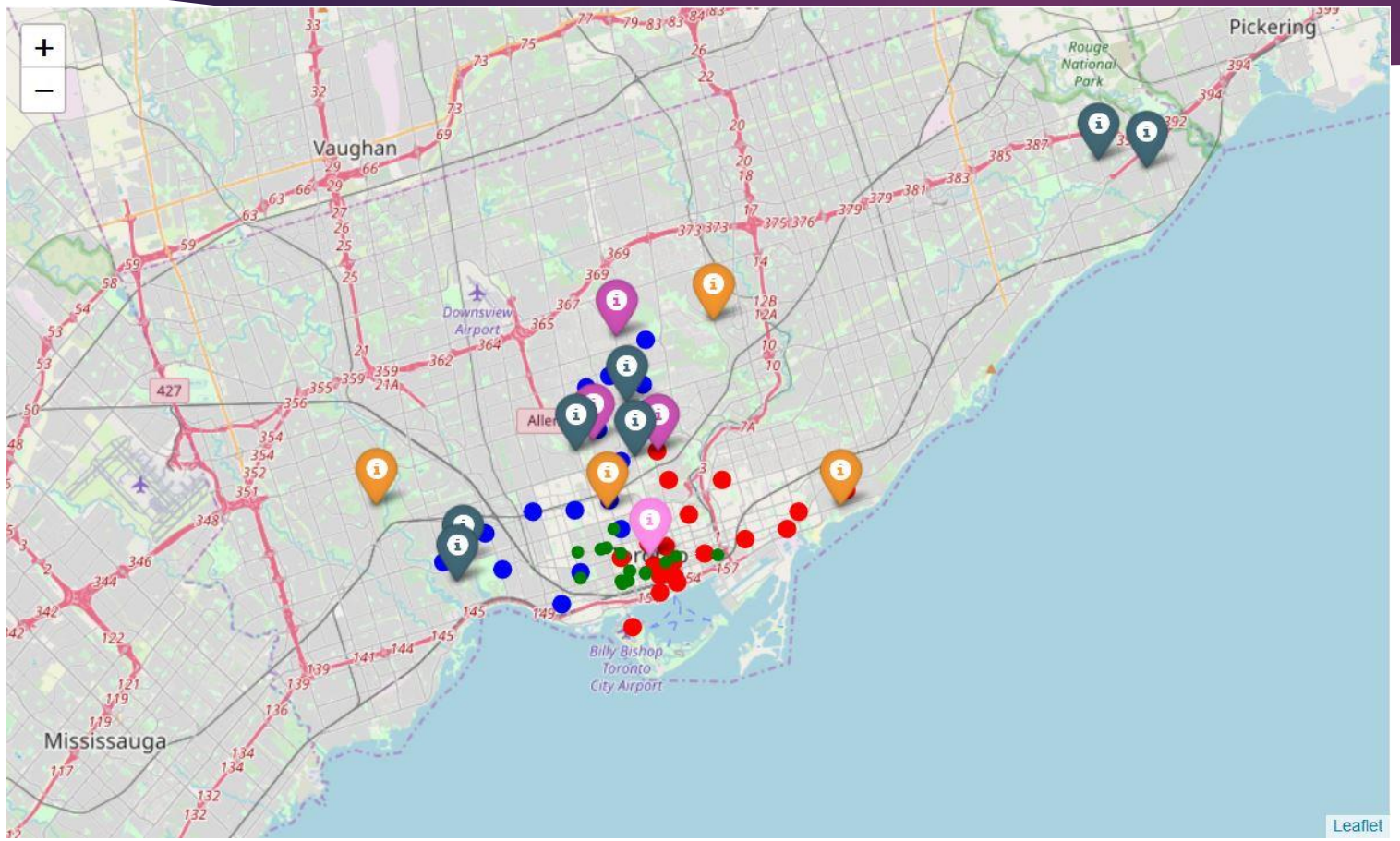
Section - 4

Motivation

- Explore top recommended restaurants in Toronto
  - These are competitions!
  - Locations – Marking them on the Toronto map
  - Cuisine categories and distribution
  - Popularity – number of ‘likes’

- **Why?** – Will provide info about the cuisine to offer in the new restaurant and locations to avoid that will be too competitive due to presence of other established restaurants.

Location of Competing Restaurants



- Green circle markers here depicts only the restaurants that were in the top recommended venues in Toronto.
  - Grouped around red circle markers - Largely spread in the Downtown area.

Distribution of Competing Restaurants

categories	count	mean	min	max
Restaurant	3.0	64.0	40.0	78.0
Italian Restaurant	2.0	290.5	174.0	407.0
Mediterranean Restaurant	2.0	145.5	49.0	242.0
Mexican Restaurant	2.0	295.0	292.0	298.0
Tapas Restaurant	2.0	277.0	261.0	293.0
Thai Restaurant	2.0	369.5	71.0	668.0
American Restaurant	1.0	249.0	249.0	249.0
French Restaurant	1.0	74.0	74.0	74.0
New American Restaurant	1.0	179.0	179.0	179.0
Ramen Restaurant	1.0	176.0	176.0	176.0
Seafood Restaurant	1.0	177.0	177.0	177.0
Theme Restaurant	1.0	27.0	27.0	27.0

- Count – number of restaurants of that category. For e.g. there are 2 top recommended Italian restaurants
- mean, min and max - number of likes for the restaurants in each category
- **Thai and Italian restaurants are the favorites!**
  - **max ‘likes’ - 668 and 407 respectively**

Summary – Competing top recommended restaurants

categories	count	mean	min	max
Restaurant	3.0	64.0	40.0	78.0
Italian Restaurant	2.0	290.5	174.0	407.0
Mediterranean Restaurant	2.0	145.5	49.0	242.0
Mexican Restaurant	2.0	295.0	292.0	298.0
Tapas Restaurant	2.0	277.0	261.0	293.0
Thai Restaurant	2.0	369.5	71.0	668.0
American Restaurant	1.0	249.0	249.0	249.0
French Restaurant	1.0	74.0	74.0	74.0
New American Restaurant	1.0	179.0	179.0	179.0
Ramen Restaurant	1.0	176.0	176.0	176.0
Seafood Restaurant	1.0	177.0	177.0	177.0
Theme Restaurant	1.0	27.0	27.0	27.0

- Lot of different cuisines offered.
- Highly liked Thai and Italian restaurants
- Multiple restaurant options for each cuising

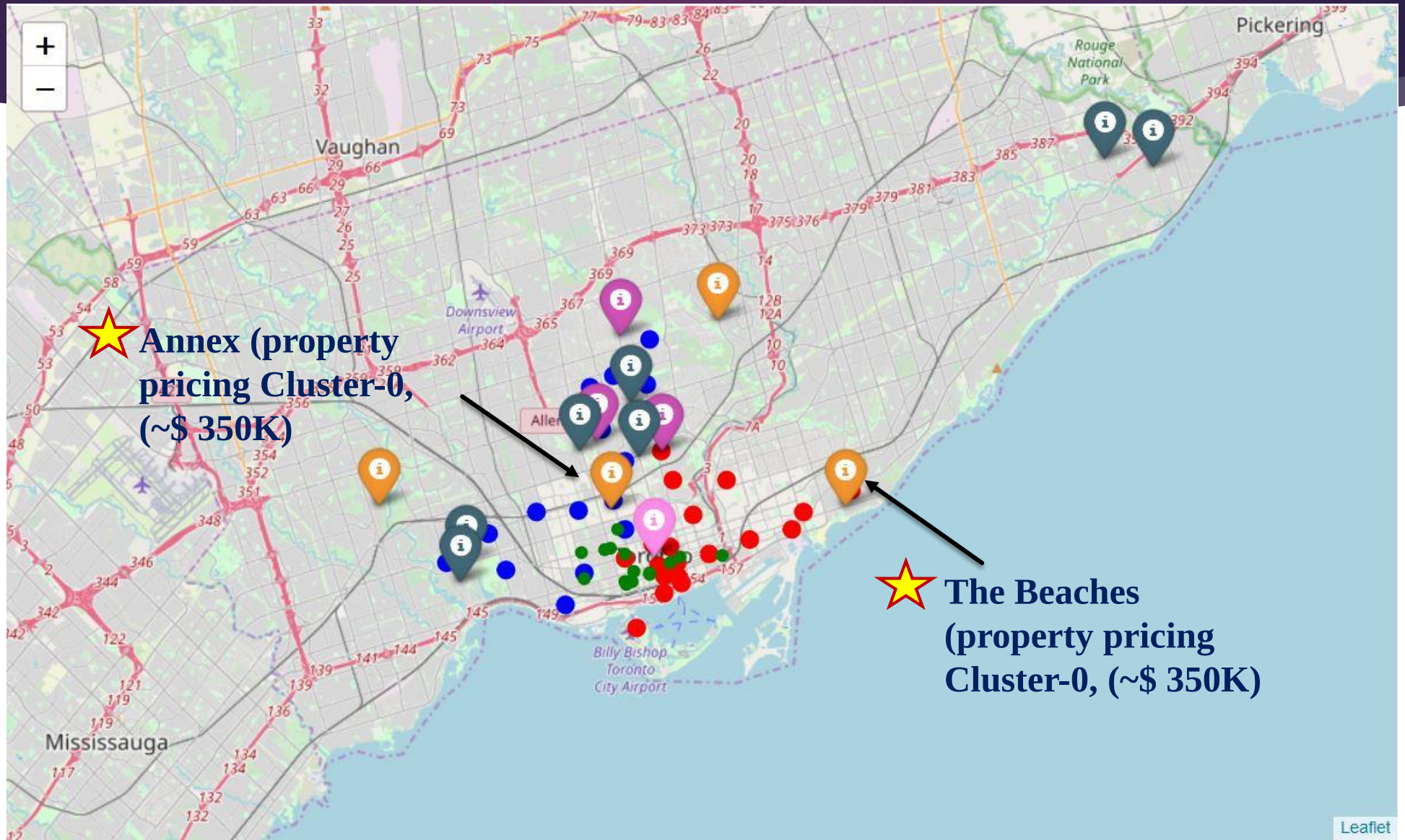
Tough competition!

Would be useful to offer cuisine which competing restaurants don’t offer

Final Recommendations

Section - 5

Recommendations for Location



Noteworthy Reasons for Final Recommendations

- Low investment for buying/renting restaurant property.
- Lot of choices as indicated in Section – 2.
- Low crime incidences.
- Local demographic has leisure income to spend on eating out in restaurants.
- Both areas reside close to the Cluster (red circle markers) that has the top recommended venues in Toronto (Section - 1). Hence, plenty of tourist customer influx is expected .
- An unexpected insight – competing restaurants are tightly grouped (green circle markers) but recommended areas are further away from these restaurants. Therefore, less competition since the customer wouldn’t have multiple choices for restaurants in these areas.

Extra Suggestions - Cuisine

- **New restaurant could offer cuisine not offered by competing restaurants.**
  - Indian or Chinese Cuisine.
- **'The Beaches' is by the ocean - beneficial location due to the view of the ocean**
  - Seafood cuisine restaurant a possibility
- **French or Theme restaurants are also viable**
  - Only 1 restaurant for each and aren’t very popular either
- **It is advised not to open Italian or Thai restaurants**
  - Popular Thai and Italian restaurants with high amount of likes already exist.
  - Multiple recommended restaurants for these cuisines are present in the Toronto area.