FIND ME A HOME

CLUSTERING OF TORONTO'S NEIGHBORHOODS BASED ON CRIME RATE, LOCAL EMPLOYMENT, CHILDCARE SPACES AND NEARBY VENUES.

Analyzed by:

Nitish Bhardwaj

https://www.linkedin.com/in/nitishbhardwaj1912/

Business Problem

People migrating to an unknown location finds it difficult to find a new house as it requires time consuming and stressful research by them.

Their research is based on various aspects including Home Prices, Crime cases, Local Employment, Childcare spaces and nearby common venues in the neighbourhoods of Toronto.

People migrating to Toronto

People looking to move to another location/neighborhood in Toronto

Target Audience

Data

The geolocation, economics (home prices, local employment) and crime related details are fetched from Open Toronto Data website[1][2][3].
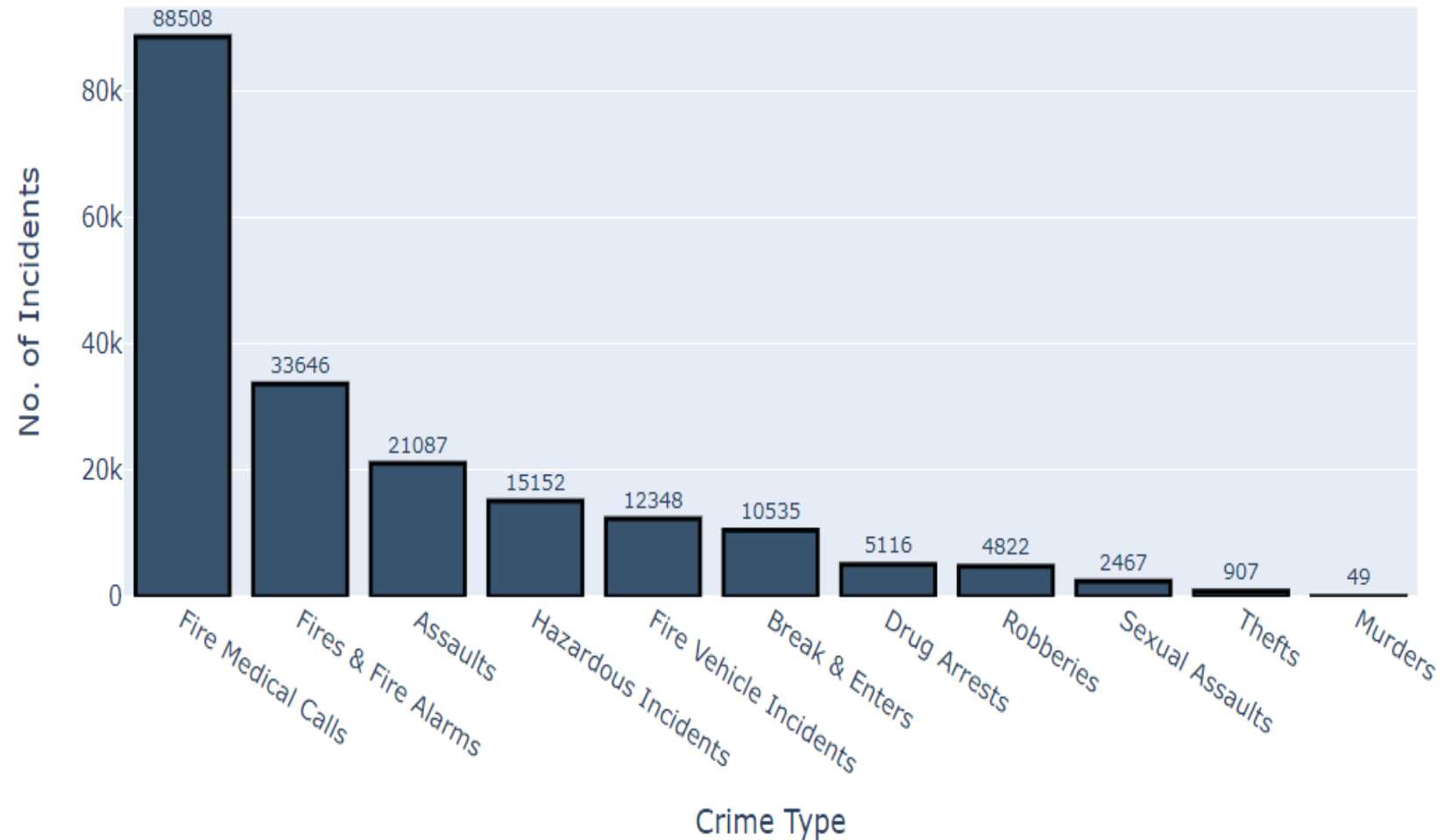
Based on the geolocation details, nearby venue details are fetched using FourSquare API[4].
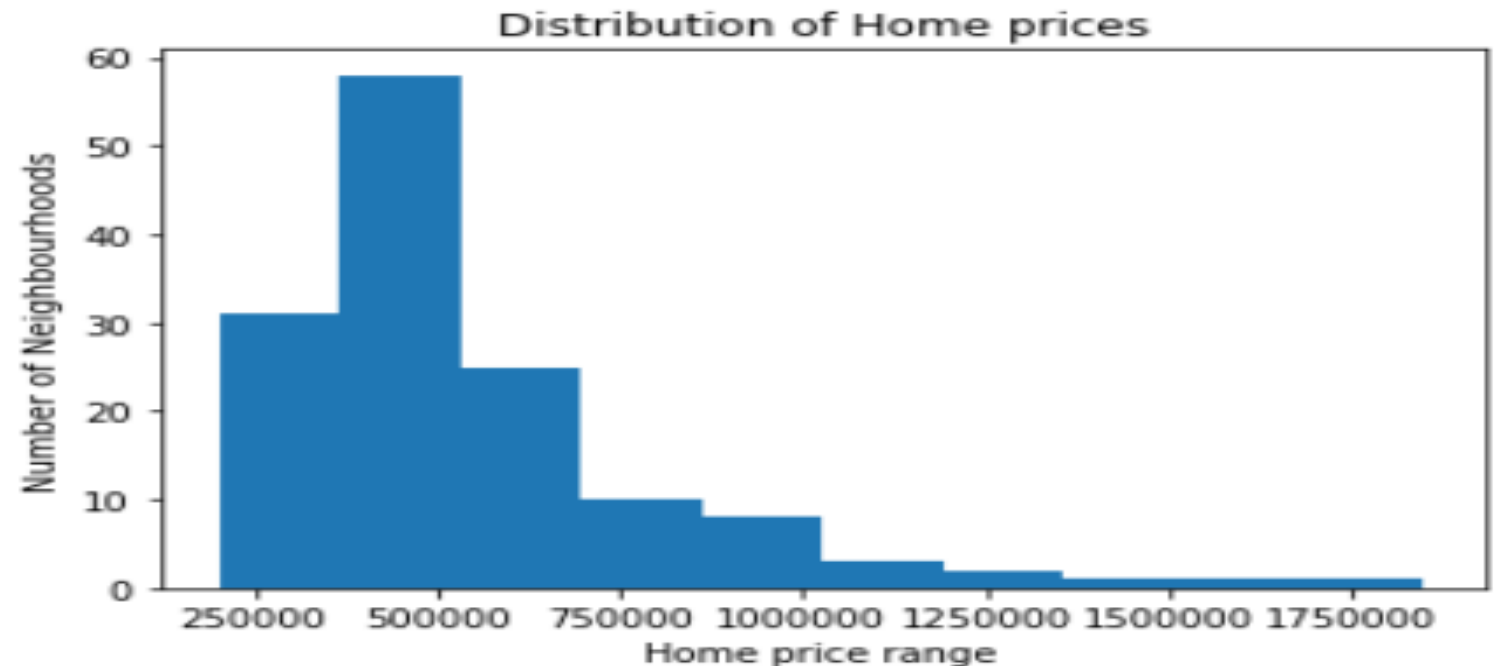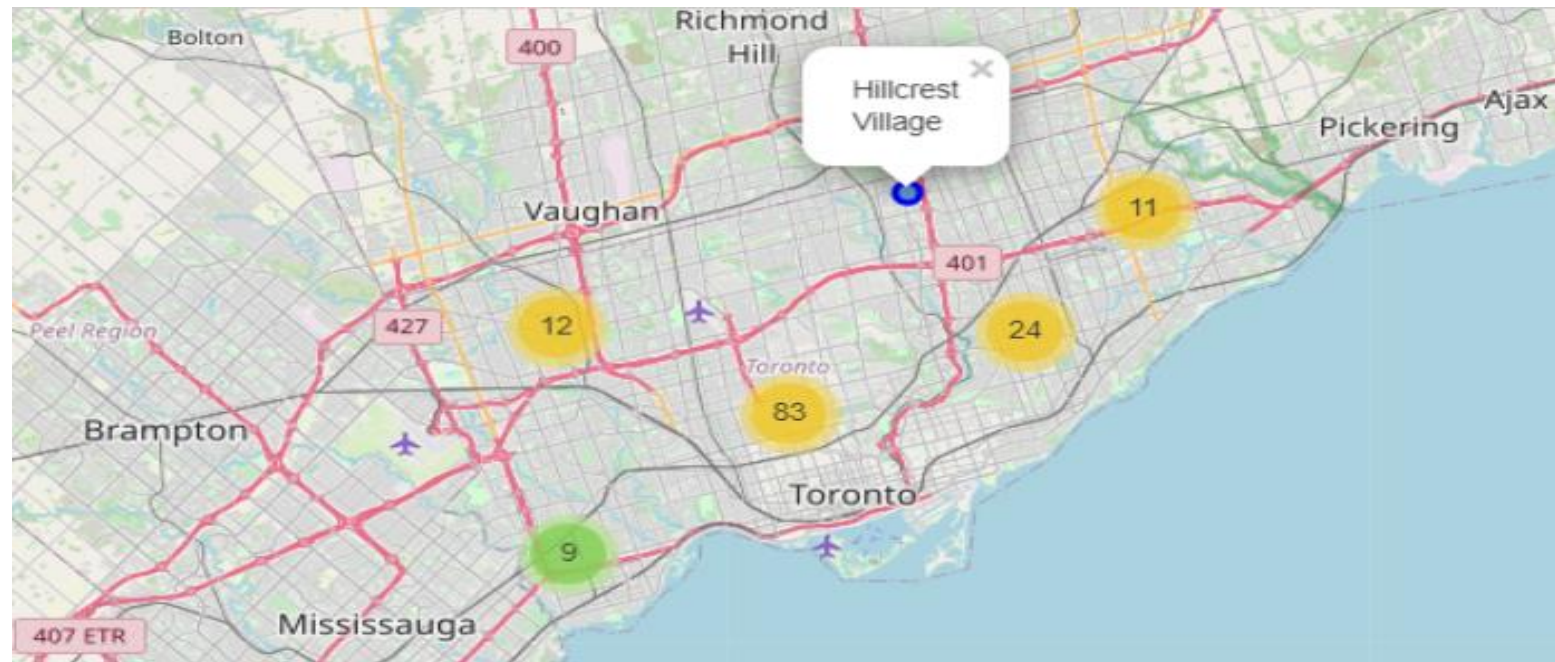
Exploratory Data Analysis(EDA)

1. No missing values are found in the data fetched from Open Toronto data.

2. High Fire medical calls and low murder cases are found in the dataset.
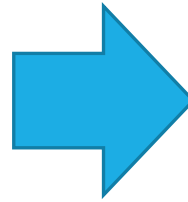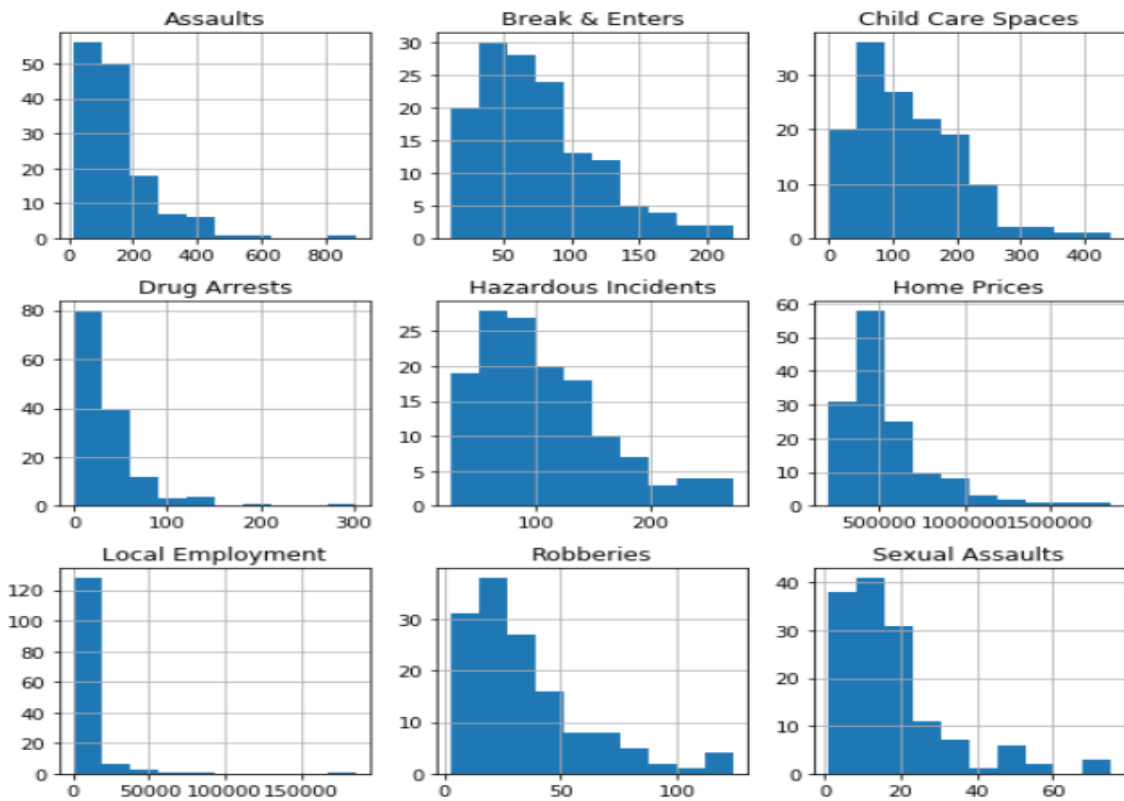

Crime frequency in Toronto

Exploratory Data Analysis (contd.)

1. Using Folium, geolocation details are validated by clustering based on location details.

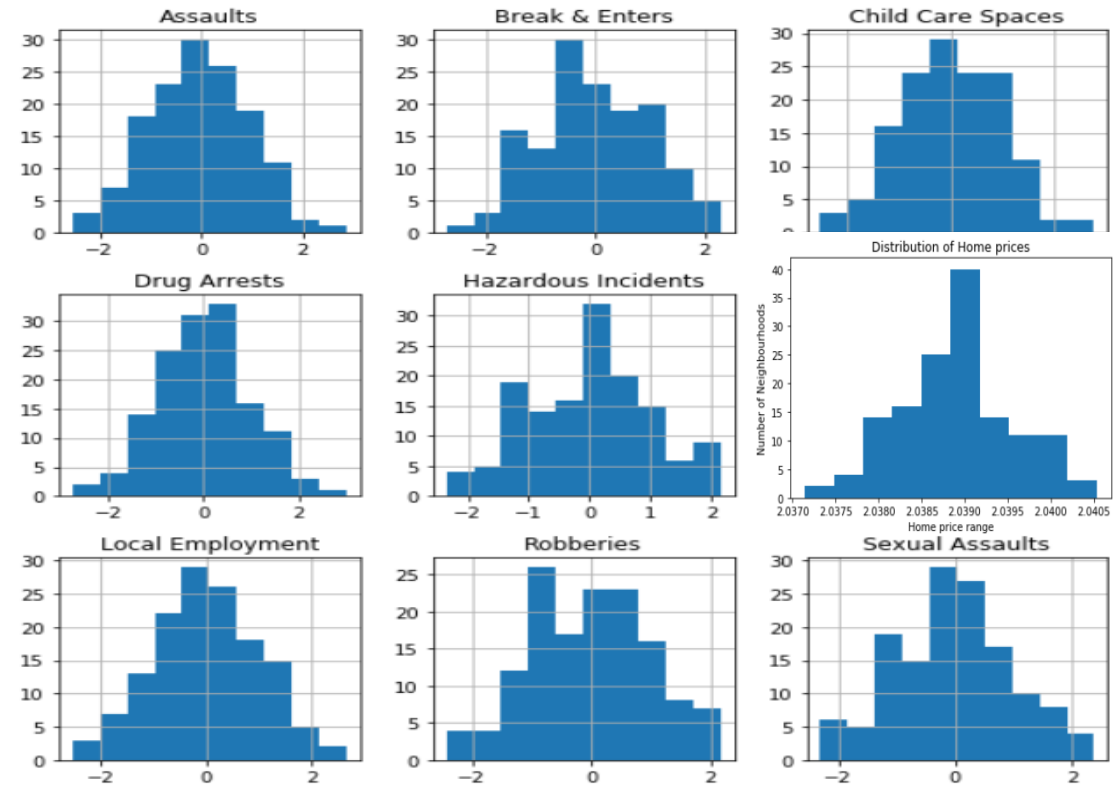2. Most of the neighborhoods in Toronto have a price range of 250K to 650K units.

Data Preparation

Raw Data:

Transformed Data:

Data Preparation (contd.)

Raw Venue Data:

	Neighbourhood	Venue	Venue Category
0	Wychwood	Wychwood Barns Farmers' Market	Farmers Market
1	Wychwood	Wychwood Barns	Event Space
2	Wychwood	Hillcrest Park	Park
3	Wychwood	Annabelle Pasta Bar	Italian Restaurant
4	Yonge-Eglinton	North Toronto Memorial Community Centre	Gym
...	...	...	...
2031	L'Amoreaux	Dollarama	Discount Store
2032	L'Amoreaux	Olympian Swimming	Gym Pool
2033	L'Amoreaux	Dumpling & Szechuan Cuisine (川流不息店)	Chinese Restaurant
2034	L'Amoreaux	Ocean Nails Spa	Nail Salon
2035	L'Amoreaux	Rogers Plus	Electronics Store

One-hot encoding:

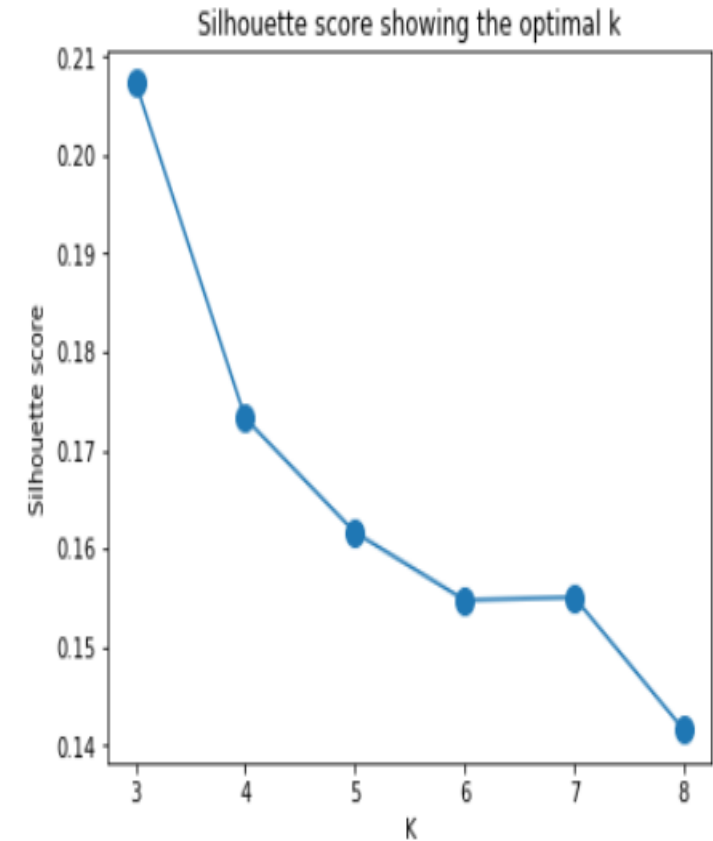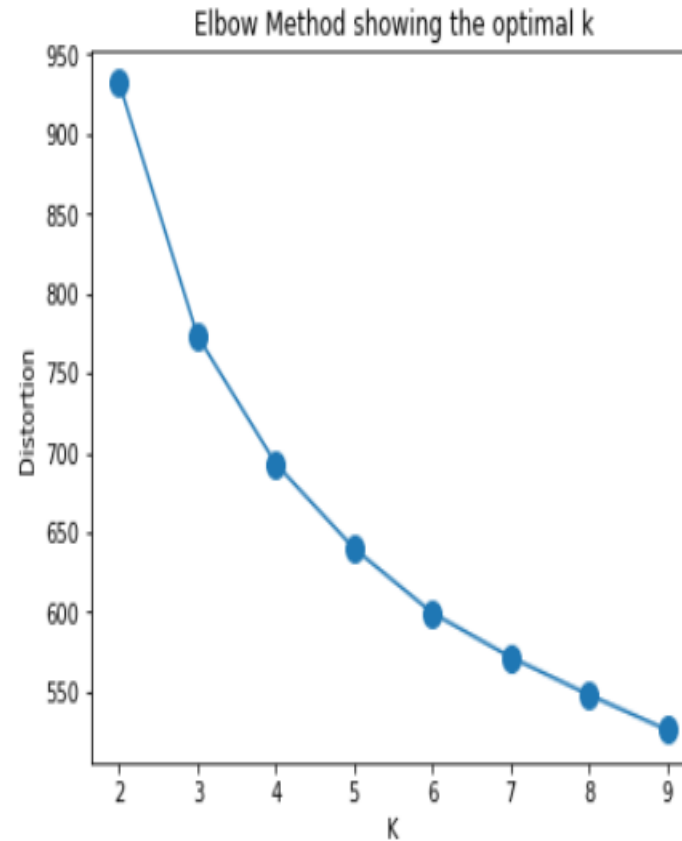| Neighbourhood	African Restaurant	American Restaurant	Amphitheater	Animal Shelter
Wychwood	0	0	0	0
Wychwood	0	0	0	0
Wychwood	0	0	0	0
Wychwood	0	0	0	0
Yonge-Eglinton	0	0	0	0

Re-formatted data by topmost venues

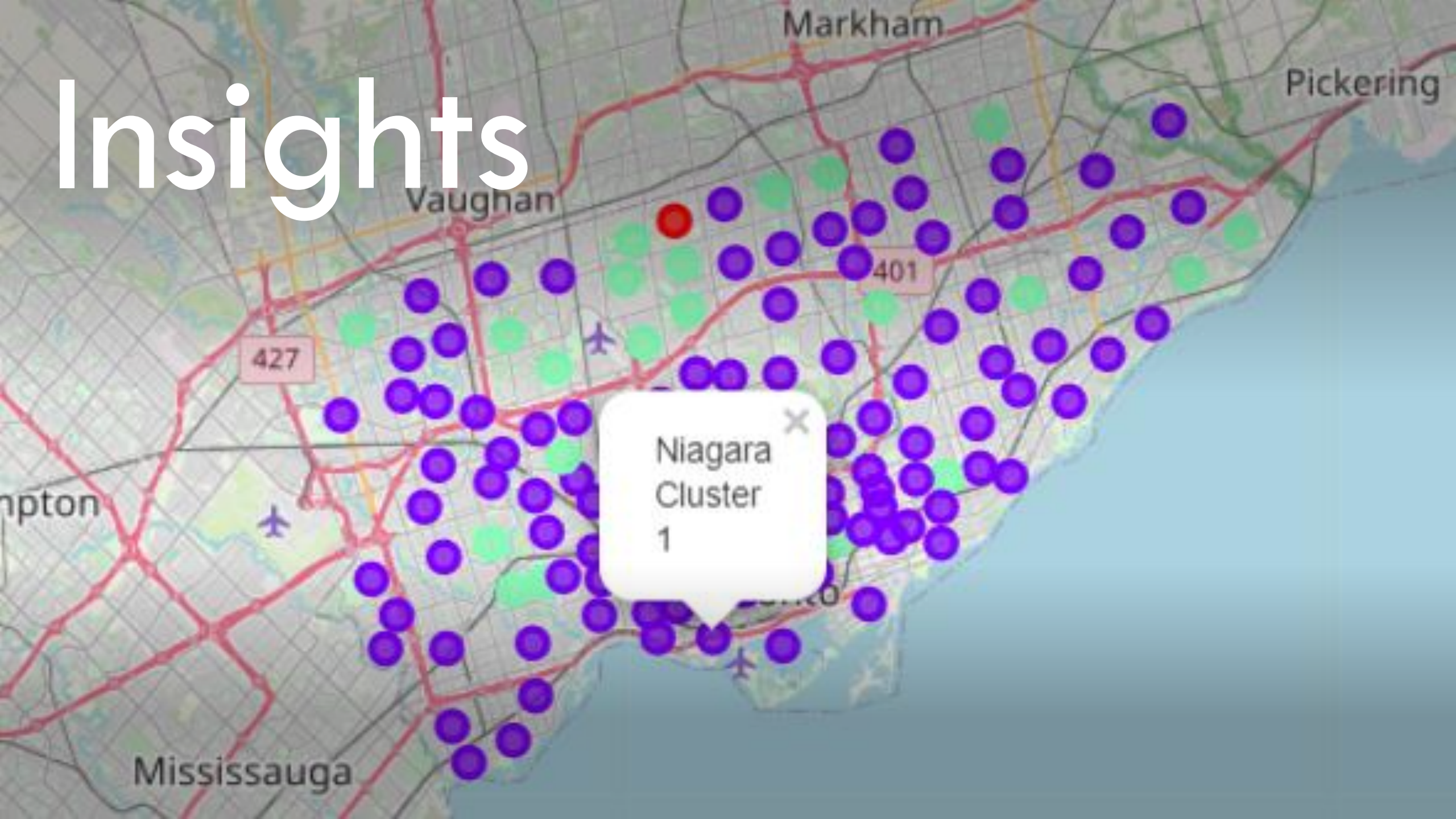| Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Wychwood	Convenience Store	Event Space	Park	Farmers Market	Electronics Store
Yonge-Eglinton	Coffee Shop	Fast Food Restaurant	Gym	Restaurant	Movie Theater
Yonge-St.Clair	Coffee Shop	Italian Restaurant	Restaurant	Sushi Restaurant	Bagel Shop
York University Heights	Bar	Japanese Restaurant	Miscellaneous Shop	Bank	Fast Food Restaurant
Yorkdale-Glen Park	Restaurant	Fast Food Restaurant	Furniture / Home Store	Fried Chicken Joint	Rental Car Location

Modeling and Evaluation

1. K-means clustering is performed on the data

2. Using Elbow method, it is found that the local optima can be found at K=3 or K=4.

3. Silhouette score, gave more confidence to select K=3 as the score is highest at that point.

4. Based on the analysis, three clusters are created.

Finding optimum K value for K-means by Elbow method and Silhouette score

Insights

Insights (contd.)

Cluster/ Segments	Cluster/Segment Details	Number of Neighbourhoods
Cluster 1	**House Price Range:** ~550K **Childcare Spaces:** Low to Medium **Crime Rate:** High. Common cases of Assaults, Break & Enter, and Hazardous Incidents. **Nearby Venues:** Less restaurants, photography studio and Farms, Farmers market **Local Employment:** Medium	1
Cluster 2	**House Price Range:** 250K to 600K **Childcare Spaces:** Medium to High. Some neighbourhoods don't have childcare spaces **Crime Rate:** Medium to High. Common cases of Assaults, Break & Enter. **Nearby Venues:** Restaunrants, coffee shops, pubs and bars **Local Employment:** Medium to High	114
Cluster 3	**House Price Range:** 250K to 800K **Childcare Spaces:** Medium to High (Each nighbourhood has a childcare space) **Crime Rate:** Low to High. Common cases of Assaults, Hazardous Incidents, and Thefts. **Nearby Venues:** Playgrounds or parks, restaurants, pubs/bars, gym and clothing store **Local Employment:** Medium to High	25

References

[1] "Open Data Dataset", *Open.toronto.ca*, 2020. [Online]. Available: https://open.toronto.ca/dataset/neighbourhoods/. [Accessed: 14- May- 2020].

[2] "Open Data Dataset", *Open.toronto.ca*, 2020. [Online]. Available: https://open.toronto.ca/dataset/wellbeing-toronto-economics/. [Accessed: 14- May- 2020].

[3] "Open Data Dataset", *Open.toronto.ca*, 2020. [Online]. Available: https://open.toronto.ca/dataset/wellbeing-toronto-safety/. [Accessed: 14- May- 2020].

[4] "Endpoints | Places API", Developer.foursquare.com, 2020. [Online]. Available: https://developer.foursquare.com/docs/places-api/endpoints/. [Accessed: 30- Apr- 2020].