

Analyze Real Estate Investment Locations

Business Problem

- **Toronto as a dynamic metropolis with strong economy**
- **Potential to rent high standard apartments to wealthy customers**
- **Target group consists of private and commercial investors willing to buy real estate in Toronto**

Data Sources

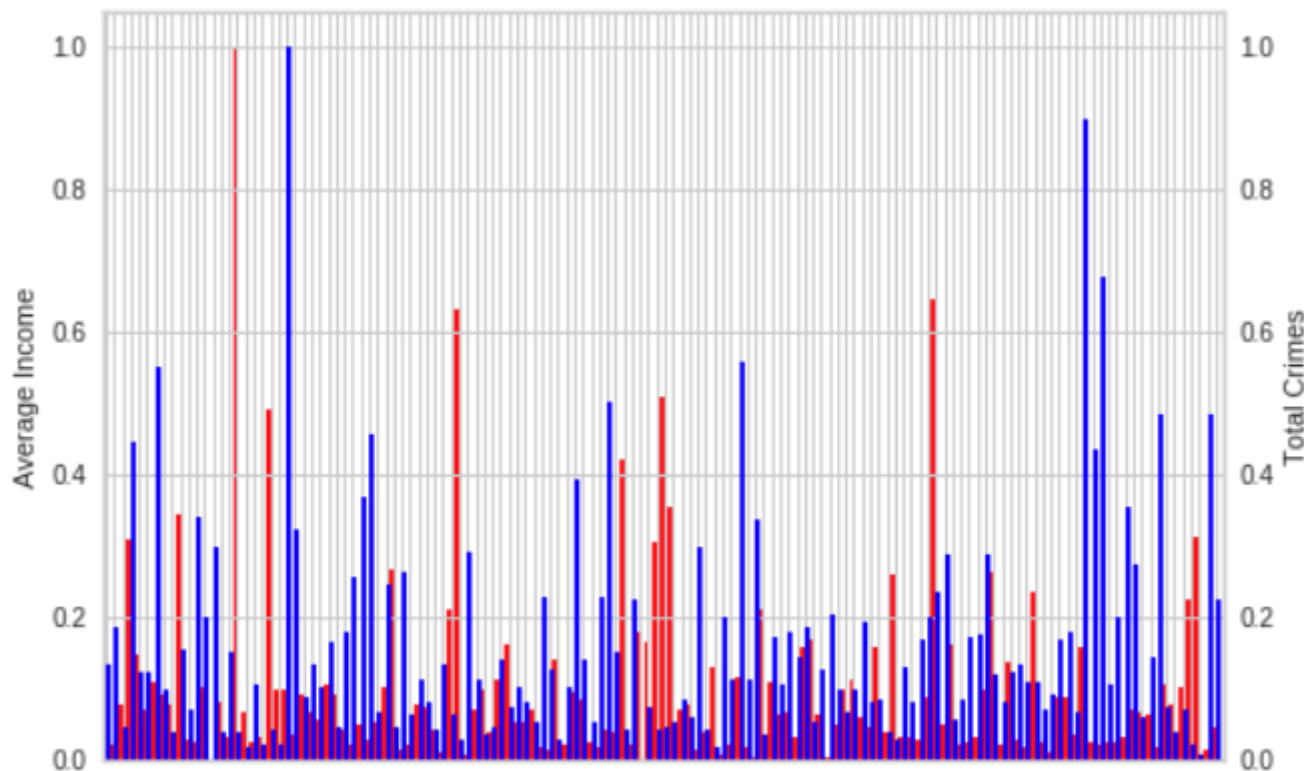
- **Toronto Crime Data: Major Crime Indicators (MCI) from the years 2014 to 2018**
<http://data.torontopolice.on.ca>
- **Demographic Data: Information about the neighborhoods in Toronto like average income or density**
<https://open.toronto.ca/dataset/neighbourhood-profiles>
- **Venue Data: Foursquare API**
<https://developer.foursquare.com>

Data Wrangling

- **Instead of single incidents use total number of crimes per neighborhood**
- **Unify neighborhood names in different datasets**
- **Drop unused data**
- **Merge datasets**
- **Add Latitudes and Longitudes (Nomatim)**
- **Fetch venue data for each neighborhood**

Exploratory Analysis I

Distribution of incomes and crime rates across neighborhoods



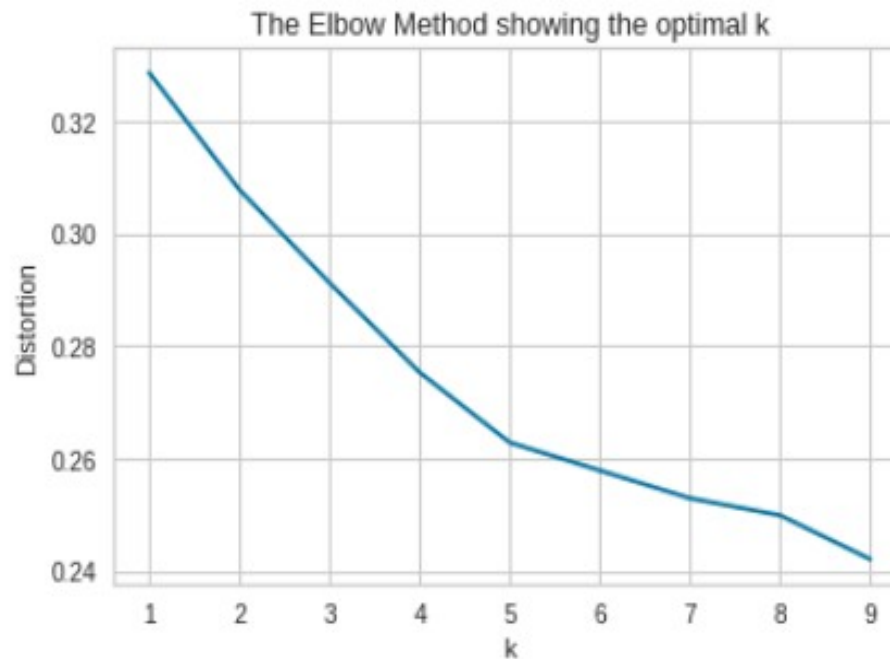
Exploratory Analysis II

Explore top venues of neighborhoods (example)

| | 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 | Agincourt North | Coffee Shop | Indian Restaurant | Ice Cream Shop | Chinese Restaurant | Bakery | Taco Place | Movie Theater | Juice Bar | Beer Store | Sporting Goods Shop |
| 1 | Agincourt South-Malvern West | Chinese Restaurant | Cantonese Restaurant | Restaurant | Park | Noodle House | Mediterranean Restaurant | Shopping Mall | Malay Restaurant | Breakfast Spot | Seafood Restaurant |
| 2 | Alderwood | Gas Station | Pharmacy | Convenience Store | Pizza Place | Park | Skating Rink | Gym | Storage Facility | Sandwich Place | Coffee Shop |
| 3 | Annex | Grocery Store | Beer Bar | Japanese Restaurant | Café | Bakery | Thai Restaurant | Italian Restaurant | Tea Room | Jewelry Store | Gift Shop |
| 4 | Banbury-Don Mills | Japanese Restaurant | Café | Coffee Shop | Restaurant | Gourmet Shop | Bakery | Bar | Supermarket | Movie Theater | Chocolate Shop |

K-Means Clustering

- **Segmentation of neighborhoods**
- **Start with elbow method**



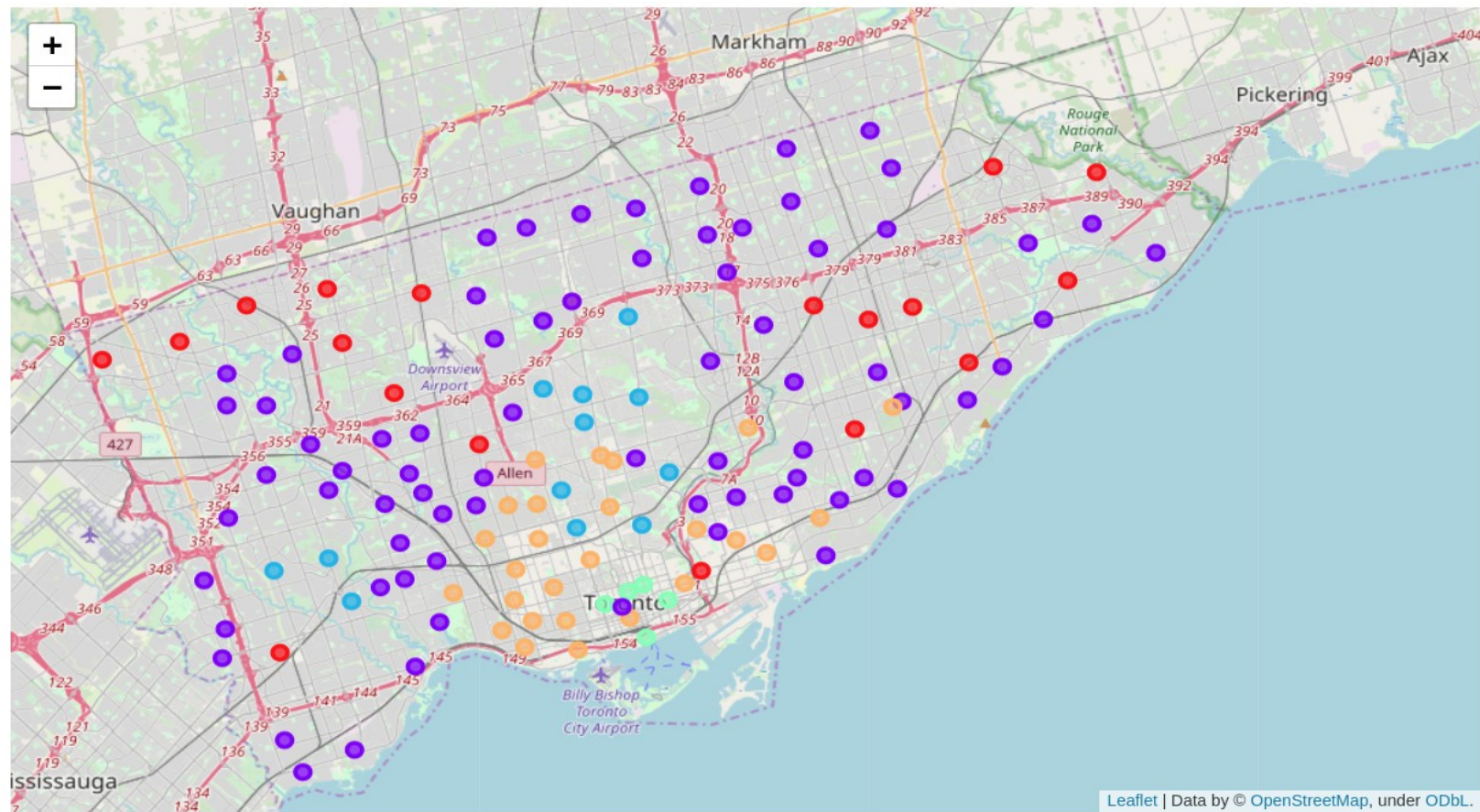
K-Means Clustering

Result of K-Means with K=5

- **Cluster 1:** High crime rate, low income, mid to low density and lot of lower priced food venues. Located mainly north-west and east of center.
- **Cluster 2:** Moderate crime rate, low income, mid to low density and mostly food and recreational venues (parks etc.). Spread across city excluding the center.
- **Cluster 3:** Low crime rate, high income, low density, more exclusive food and recreational venues (Sushi, Spa etc.) and some expensive stores. Located west and in the middle of town.
- **Cluster 4:** High crime rate, mid income, high density and a lot of cafés and coffee shops as well food venues. Located near the waterfront in the very center of the town.
- **Cluster 5:** Mixed crime rate and mixed income, high density. Mainly food venues. Located in the center.

Results

Clusters 3 and 5 seem to be promoting (orange and blue)



Conclusion / Future Directions

- **Two of five clusters seem to be promising for investments**
- **Some of the used data relatively old (2016)**
- **Data for real estate price trends missing**
- **Further analysis with more data needed**