Battle of the Neighbourhoods

Week 2

Finding a new business location

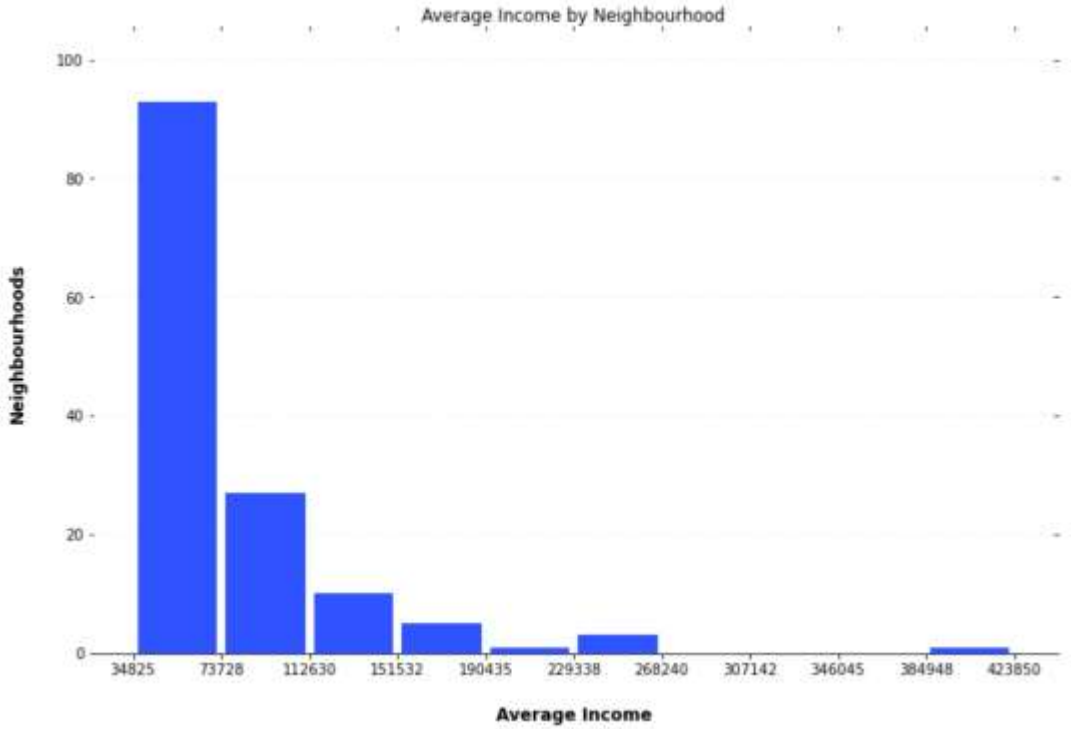
- Finding a perfect locations for a new business is critical
  - Determines who your clientele is
  - Determines what kind of traffic your business might generate
  - Can impact regulations and taxes
- Where is the perfect neighbourhood within Toronto to start a new nail salon business?

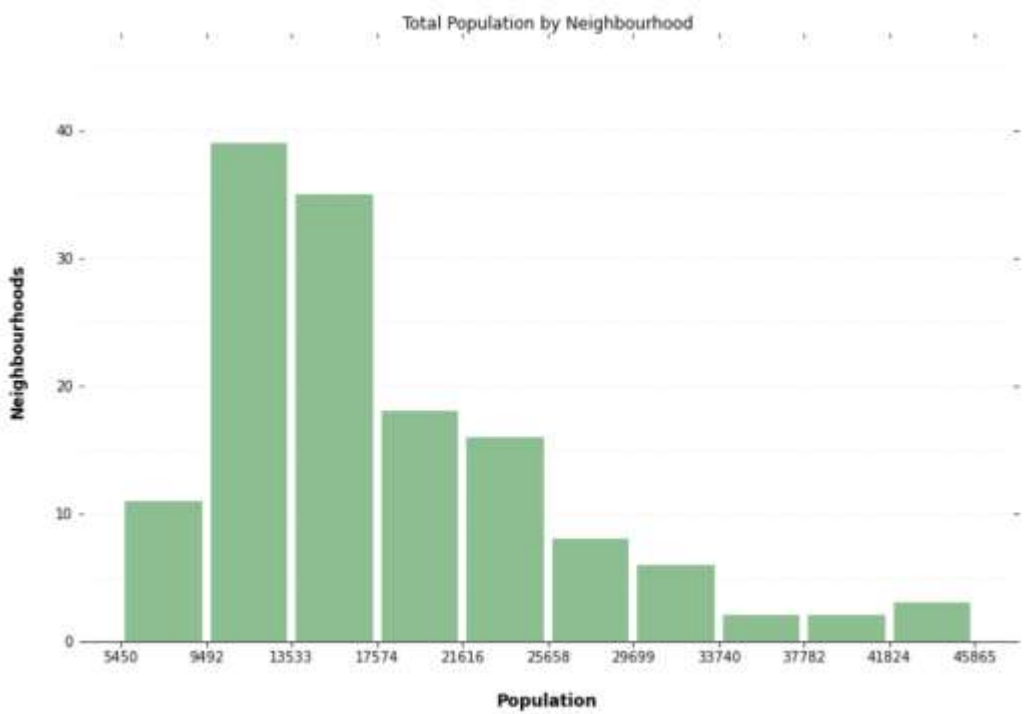
Data acquisition and cleaning

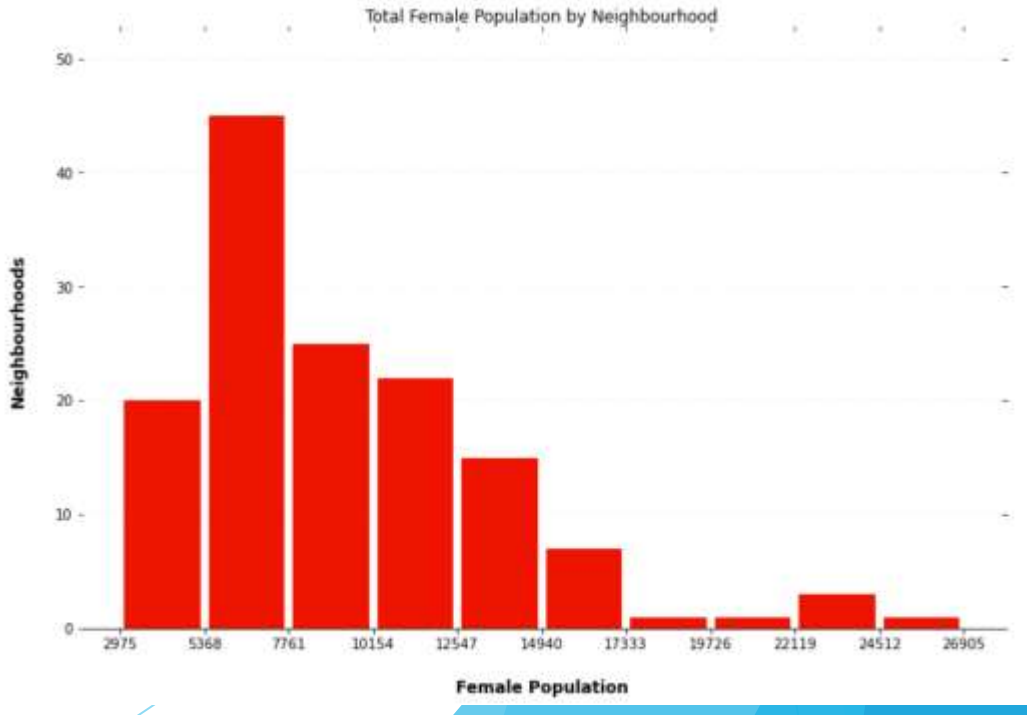
- Data pulled from two main sources
  - FourSquare – holds data on nearby venues given a certain location
  - Toronto Open Data – holds demographic and geographic information on all neighbourhoods within Toronto
- In total, 2020 venues were scraped from FourSquare for 140 Toronto neighbourhoods
- Unrelated columns were dropped and unfilled rows were managed
- Formed 2 datasets
  - One with data on all types of venues
  - One with data on only salon-type venues

Toronto Demographics

- Information on neighbourhood total population, female population and average income were graphed for exploratory analysis for use within the model



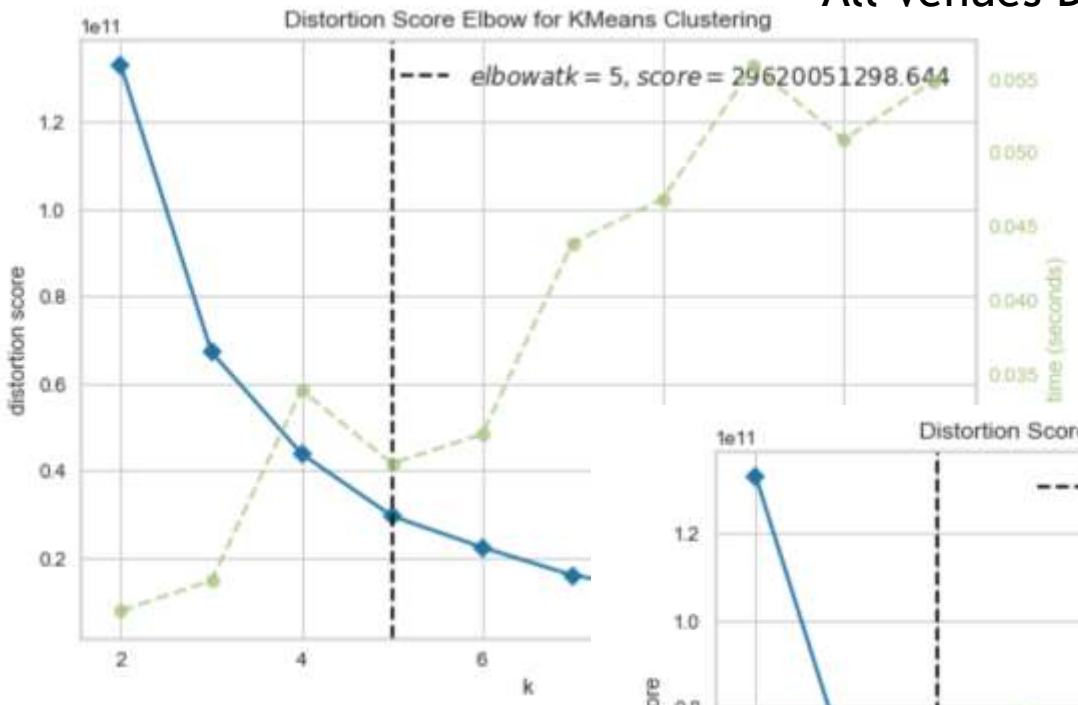




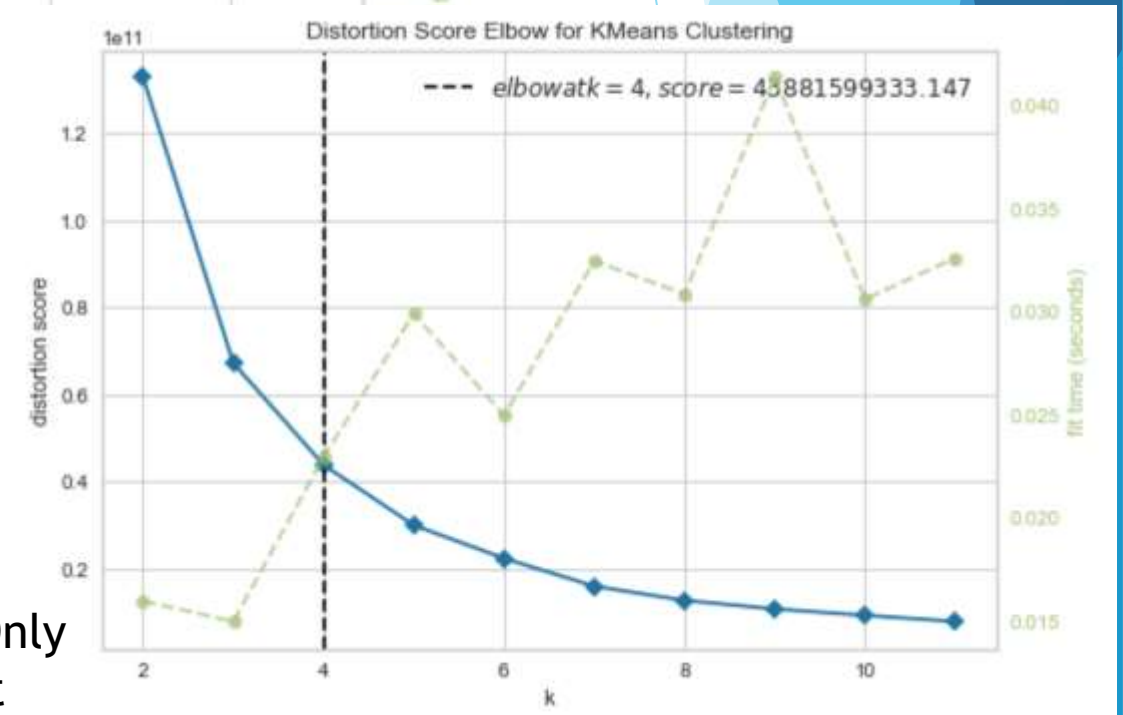
K Means – Finding a value for K

- A good value of K must be found in order to have meaningful data from the K Means model, as it determines how many clusters the model will try to form.
- The Elbow method can be used to determine the optimal value of K for both datasets
- K = 5 for both datasets

All Venues Dataset



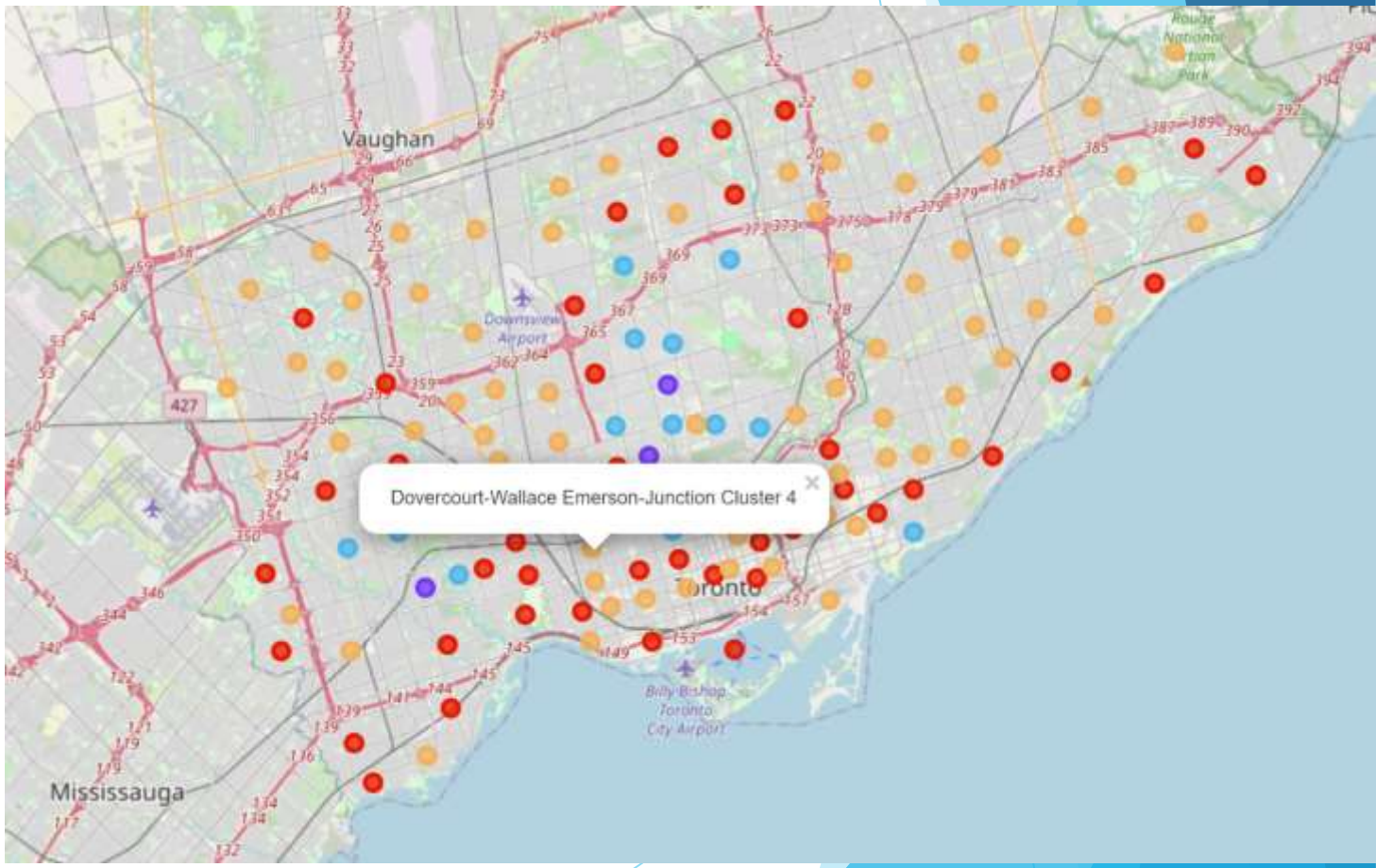
Salon Only Dataset



Mapping Results – All Venues

- Downtown Toronto and the outskirts are mostly made up of two clusters (red and orange on the map)
- The other three clusters (blue(2), purple(1), and light green(3)) forming a kind of ring in the suburbs surrounding downtown.

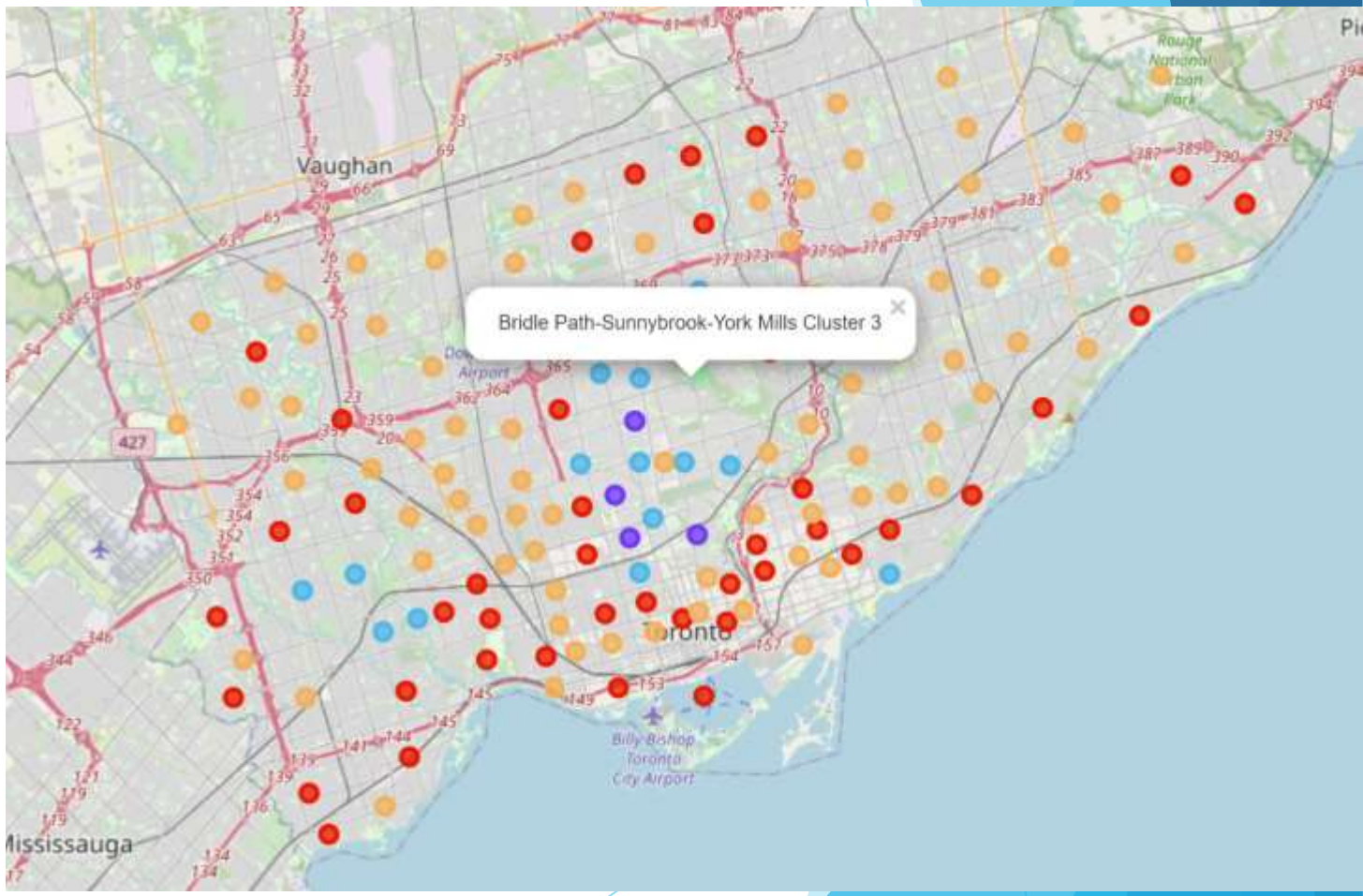
Colour	Cluster
Red	0
Purple	1
Blue	2
Light Green	3
Orange	4



Mapping Results – Salons Only

- Very similar to All Venues dataset
- Downtown Toronto and the outskirts are mostly made up of two clusters (red and orange on the map)
- The other three clusters (blue(2), purple(1), and light green(3)) forming a kind of ring in the suburbs surrounding downtown.

Colour	Cluster
Red	0
Purple	1
Blue	2
Light Green	3
Orange	4



Average Results

- The two clusters with the highest average income and lowest populations (1 and 3) have no nail salons or salon/barbershops.
  - There might be a reason why there are no similar businesses in any of these neighbourhoods.
- Cluster 2 has the highest rate of similar businesses, so neighbourhoods in that cluster can probably be dismissed, as the market might already be saturated.
- Though cluster 4 has a lower average income than cluster 0, it also has a higher average female population, and less competition, with a total lack of salon/barbershops, but enough nail salons to indicate that it is possible to run a successful business in those neighbourhoods.

All Venues Dataset

Cluster	Count	Population Avg	FemalePop Avg	AvgIncome	Nail Salon	Salon / Barbershop
0	43	14667	7644	79652	0.069767	0.023256
1	5	12973	6953	226033	0.000000	0.000000
2	14	15287	8112	135871	0.071429	0.071429
3	1	8210	4055	423850	0.000000	0.000000
4	77	20567	10597	57574	0.025974	0.000000

Salons Only Dataset

Cluster	Count	Population Avg	FemalePop Avg	AvgIncome	Nail Salon	Salon / Barbershop
0	43	14667	7644	79652	0.069767	0.023256
1	4	14018	7535	237051	0.000000	0.000000
2	15	14854	7880	138944	0.066667	0.066667
3	1	8210	4055	423850	0.000000	0.000000
4	77	20567	10597	57574	0.025974	0.000000

Conclusion

- From the clustering analysis done in this report, it is likely that a neighbourhood within Cluster 4 would be the most suitable location for a new nail salon business.
- This is because there are other nail salons in the cluster, implying that the business type is viable, but doesn't have the highest occurence of either nail salons or salons/barbershops, meaning that there will be less stiff competition for the new business.

- A next step would be to take the neighbourhoods of Cluster 4 and find more data on the area, maybe economic data about businesses, and perform another cluster analysis to break down the cluster even further.