

Introduction to Artificial Intelligence

# Exploring Best Places to Raise Children in Toronto

Group Members:

Rutvi Shah, Foram Gandhi, Kinjal Shah



Toronto is the largest city in Ontario, with a population of over 2.7 million people. With a high population that is growing day-by-day, the level of crime is bound to increase.

# Background and Business Plan

Parents strive to raise their children in a safe and accessible neighbourhood.

**Goal:** Help find the best neighbourhood to raise children



01

Schooling  
Options Available



02

Crime Rate  
(Break and Enter)



03

Children Population  
(ages 0-14)

# Data Explanation and Data Sources



- Neighbourhoods Profiles Dataset:
  - Used to get the population of children between the ages 0-14 in each neighbourhood
- Crime Rates Dataset:
  - Used to get the Break & Enter information for the year 2016
- Neighborhood Dataset:
  - Used to get the GeoJSON to format the folium map with proper formatting
- Foursquare Places API
  - Used to fetch venue information



# Methodology



1. Retrieve the number of schools in each neighbourhood
  - a. Make calls to the Foursquare Places API to fetch all the 'school' venues for each neighbourhood
  - b. Check the response for each neighborhood to see if it has matching search terms in the name
    - i. 'Elementary', 'primary', 'secondary', 'high', 'middle', 'preschool'
  - c. The number of schools is then written back to the .csv file


Agincourt North	Agincourt South-Malvern West	Alderwood	Annex	Banbury-Don Mills	Bathurst Manor	Bay Street Corridor	Bayview Village	Bayview Woods-Steeles	Bedfo
3,840	3,075	1,760	2,360	3,605	2,325	1,695	2,415	1,515	4,555
49	61	26	123	57	33	120	33	33	97
5	6	4	2	7	2	1	1	5	4

Show 10 per page

# Methodology Cont.



2. Obtain all the data from the .csv and format it into a data frame



	Pop. of Ages 0-14	Crime Rate of B&E	Num. of Schools
Agincourt North	3840.0	49.0	5.0
Agincourt South-Malvern West	3075.0	61.0	6.0
Alderwood	1760.0	26.0	4.0
Annex	2360.0	123.0	2.0
Banbury-Don Mills	3605.0	57.0	7.0
...	...	...	...
Wychwood	1860.0	29.0	7.0
Yonge-Eglinton	1800.0	19.0	7.0
Yonge-St.Clair	1210.0	12.0	7.0
York University Heights	4045.0	98.0	0.0
Yorkdale-Glen Park	1960.0	66.0	7.0

140 rows x 3 columns

# Methodology Cont.

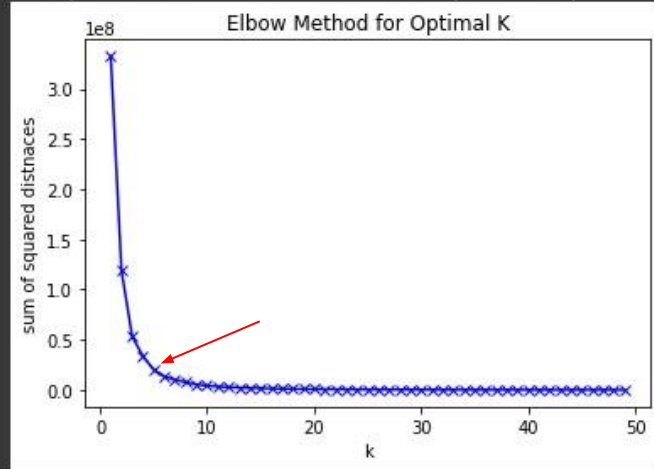


3. Find the optimal k value to perform the K-Means clustering

## Elbow Method

- i. Pick the 'elbow'

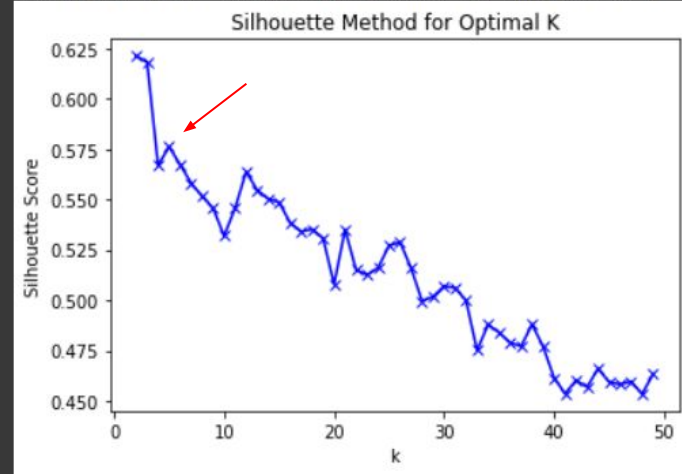
Text(0.5, 1.0, 'Elbow Method for Optimal K')



## Silhouette Method

- i. Pick the 'peak'

Text(0.5, 1.0, 'Silhouette Method for Optimal K')



# Optimal is $k = 5$

(Using Elbow Method)





# Methodology Cont.



4. Perform the K-Means clustering with  $k = 5$

	0	1
0	Agincourt North	4
1	Agincourt South-Malvern West	0
2	Alderwood	1
3	Annex	0
4	Banbury-Don Mills	4
...	...	...
135	Wychwood	1
136	Yonge-Eglinton	1
137	Yonge-St.Clair	1
138	York University Heights	4
139	Yorkdale-Glen Park	1

140 rows x 2 columns

```
Counter({1: 62, 0: 38, 4: 28, 2: 8, 3: 4})
```

**Column 0:** Neighbourhoods

**Column 1:** Cluster number

# Methodology Cont.



5. Calculated the scores for each neighbourhood and concatenated the 'Score' and 'Cluster' information to the data frame

$$\frac{[\text{Pop. of Ages 0-14}]}{[\text{Crime Rate of B\&E}]} \times [\text{Num. of Schools}]$$

	Pop. of Ages 0-14	Crime Rate of B&E	Num. of Schools	Score	Cluster
Agincourt North	3840.0	49.0	5.0	391.836735	4
Agincourt South-Malvern West	3075.0	61.0	6.0	302.459016	0
Alderwood	1760.0	26.0	4.0	270.769231	1
Annex	2360.0	123.0	2.0	38.373984	0
Banbury-Don Mills	3605.0	57.0	7.0	442.719298	4
...	...	...	...	...	...
Wychwood	1860.0	29.0	7.0	448.965517	1
Yonge-Eglinton	1800.0	19.0	7.0	663.157895	1
Yonge-St.Clair	1210.0	12.0	7.0	705.833333	1
York University Heights	4045.0	98.0	0.0	0.000000	4
Yorkdale-Glen Park	1960.0	66.0	7.0	207.878788	1

140 rows x 5 columns

# Methodology Cont.



6. Determine the best cluster according to mean score of all the neighbourhoods in the cluster

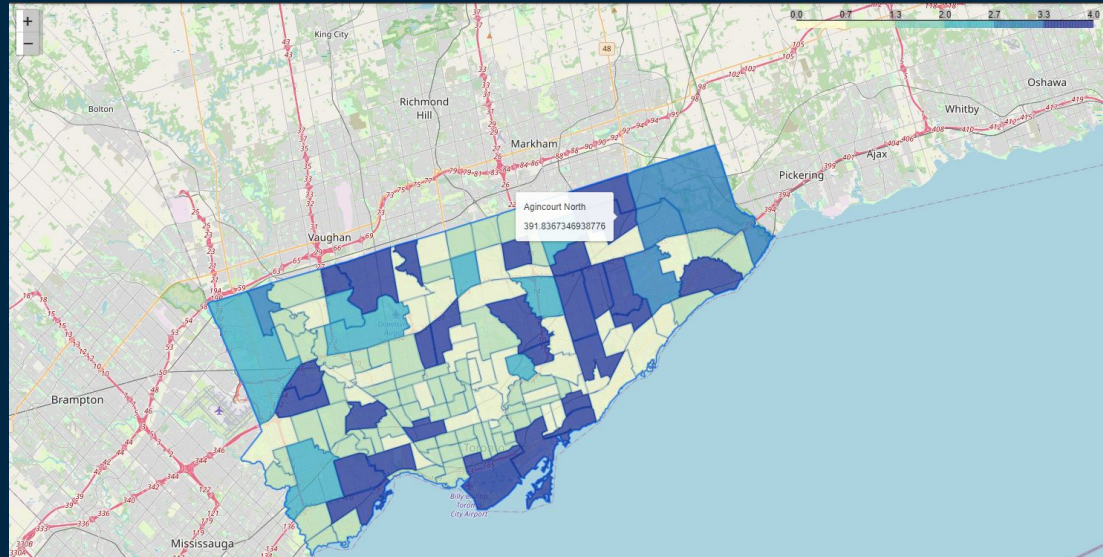
```
Cluster
0    320.653869
1    247.745093
2    903.085230
3    603.731439
4    264.266136
Name: Score, dtype: float64
```

2 is the best cluster for our goal as it has the highest mean score

# Methodology Cont.



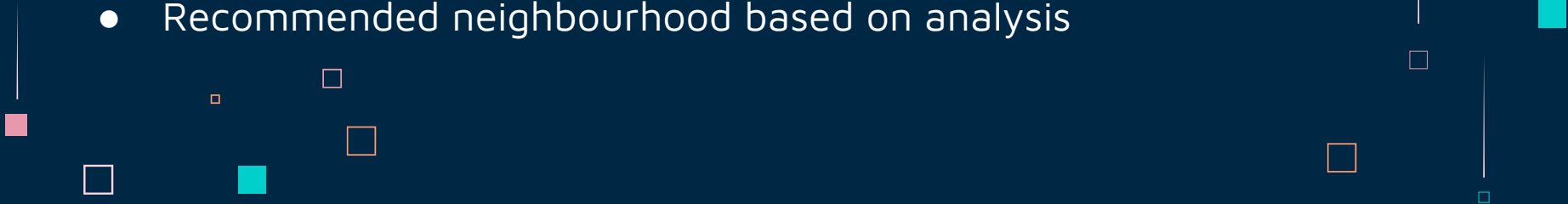
7. Use folium library and GeoJSON information to print neighbourhoods, then apply formatting to map so each neighbourhood is coloured according to the cluster they're a part of; labelled with their name and score



# Discussions and Conclusions



- Best neighbourhood for a family with children between the ages of 0 and 14 would be any neighbourhood in cluster 2
- Some neighbourhoods have a score of 0
- Correlation between crime rate and population of children
  - 3 factors - population, schooling, crime rate of B&E
  - Mapping data samples
  - Clustering with K-Means
  - Analyzing clusters
- Recommended neighbourhood based on analysis



DEMO

The background is a dark blue gradient. It is decorated with various geometric elements: small squares in teal, pink, and orange, and thin white vertical lines of varying lengths. These elements are scattered across the slide, creating a modern, minimalist aesthetic.

# Thank You!

CREDITS: This presentation template was created by [Slidesgo](#),  
including icons by [Flaticon](#), and infographics & images by [Freepik](#)