



Exploring and Comparing Neighborhoods for Families Settling Down in Toronto

Introduction

- this project aims to explore and compare neighborhoods in Toronto for families who were forced to immigrate to Canada from different countries all over the world,
- choosing the right city to live in with a family and children is a difficult real-life problem.

Problem description

- a couple with a two sons want to immigrate to Toronto,
- their sons are 14 and 15 years old and should attend high school,
- they need to rent a house or apartment, which they can afford, and near a high school in a secure neighborhood,
- the neighborhood needs recreation places such as parks for everyday jogging and popular outdoor activities for families.

Data

- list of neighborhoods in Toronto from Wikipedia page
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- csv file with the latitude and the longitude coordinates of each neighborhood GeoSpatial Dataset https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDriverSkillsNetwork-DS0701EN-SkillsNetwork/labs_v1/Geospatial_Coordinates.csv
- not all neighborhoods (140) are in including e.g. Yonge-St.Clair - and check with <https://www.zipcodesonline.com/2020/06/postal-code-of-toronto-in-2020.html>

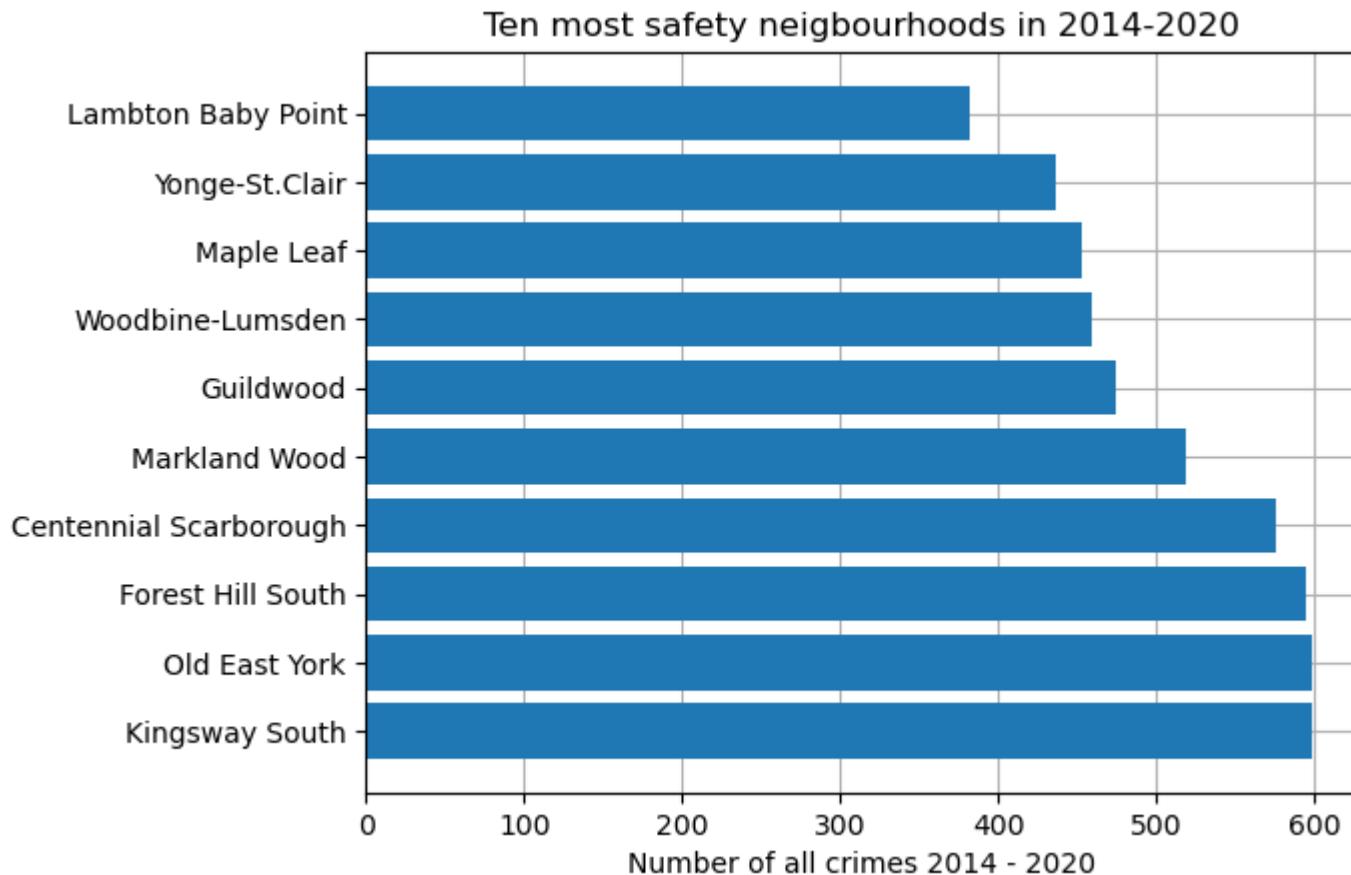
Data

- data about "Safety and crimes" we get from Toronto Police Service (Public Safety Data Portal) in *.csv file
https://data.torontopolice.on.ca/datasets/3a1a9c98146e470e94e814b0e3a3fbca_0/explore?location=43.718404%2C-79.378190%2C11.73&showTable=true
- data about prices of houses in Toronto in May 2021 from Zolo portal
<https://www.zolo.ca/toronto-real-estate/neighbourhoods> and in 2020 from <https://www.moneysense.ca/spend/real-estate/where-to-buy-real-estate-in-2020-city-of-toronto/>

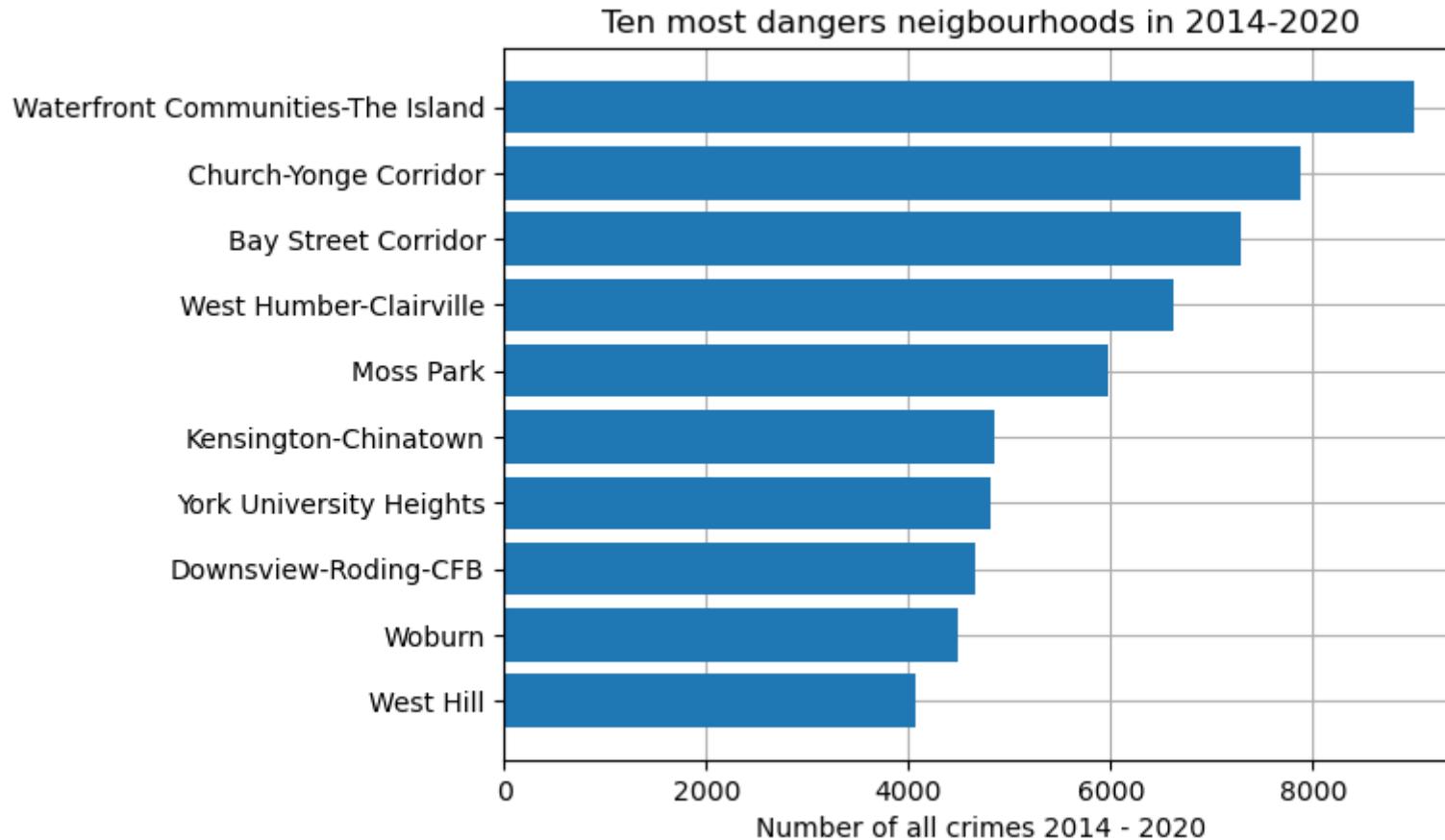
Data

- information about secondary schools in Toronto we download data from Fraser Institute for secondary schools in Toronto
<https://www.tdsb.on.ca/Find-your/School/Secondary>
- getting school addresses from Toronto District School Board and Toronto Catholic District School Board.
- using Foursquare API : "https://developer.foursquare.com/" to get all the venues in each neighborhood

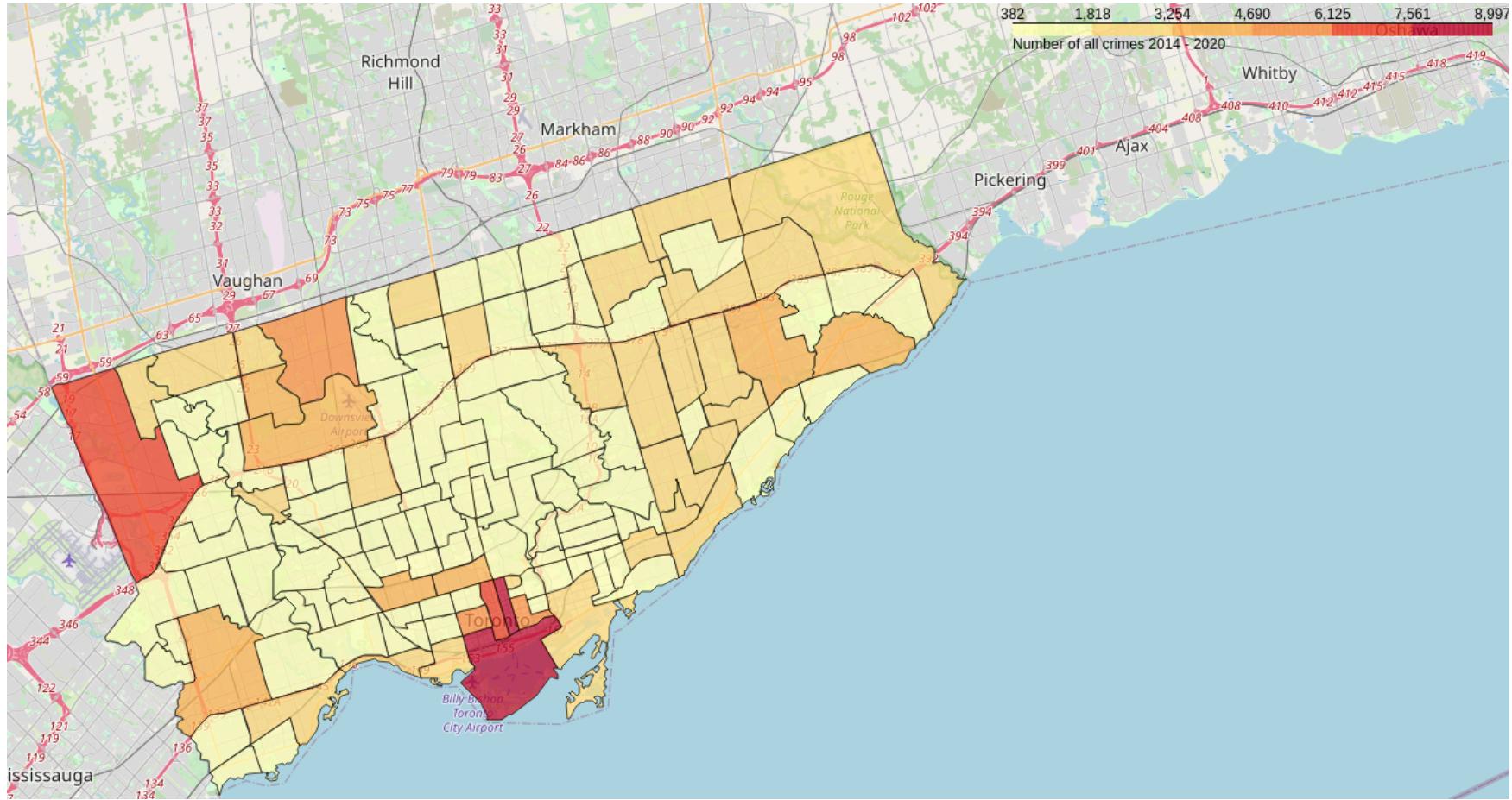
Analyzing safety neighborhoods in Toronto



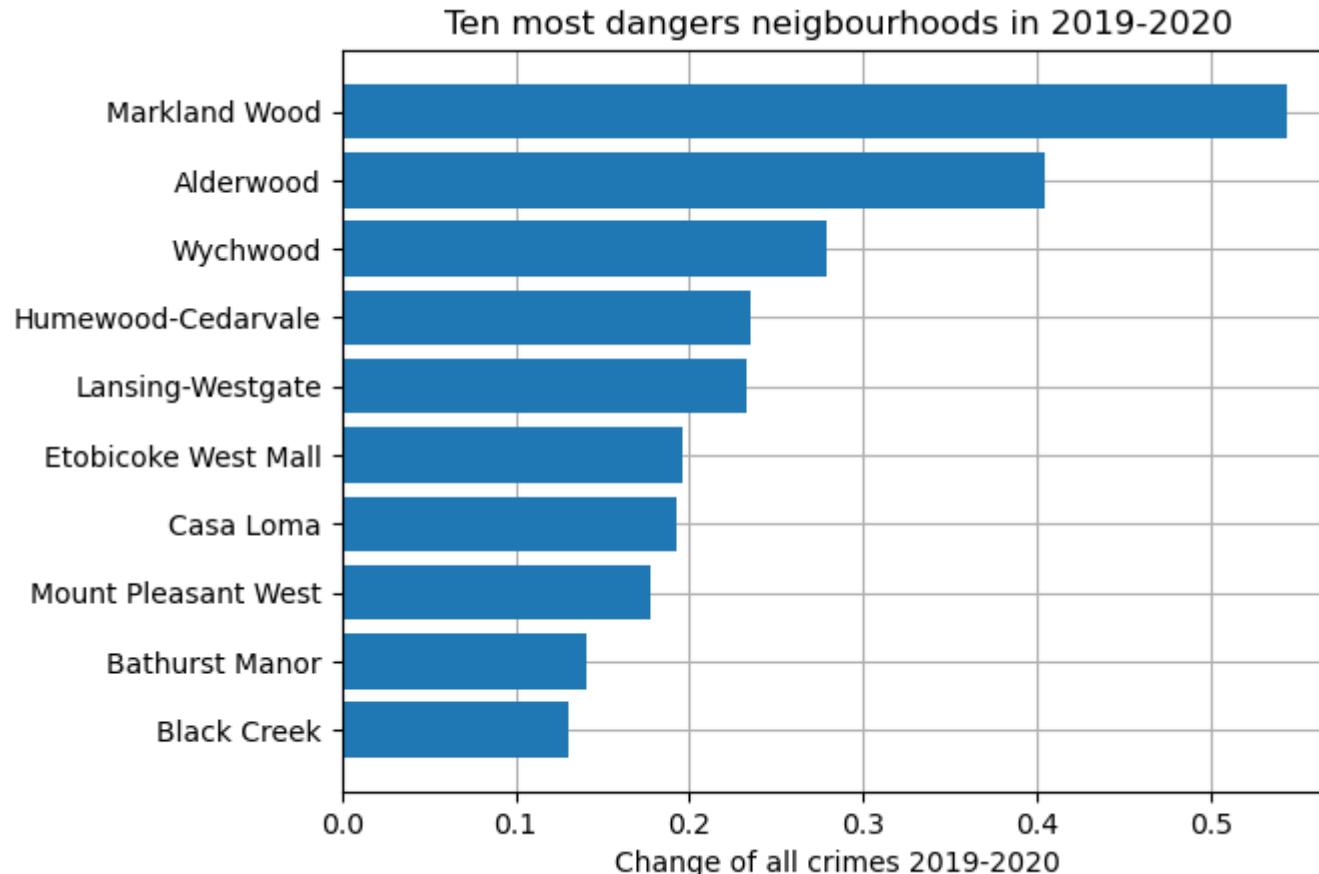
Analyzing safety neighborhoods in Toronto



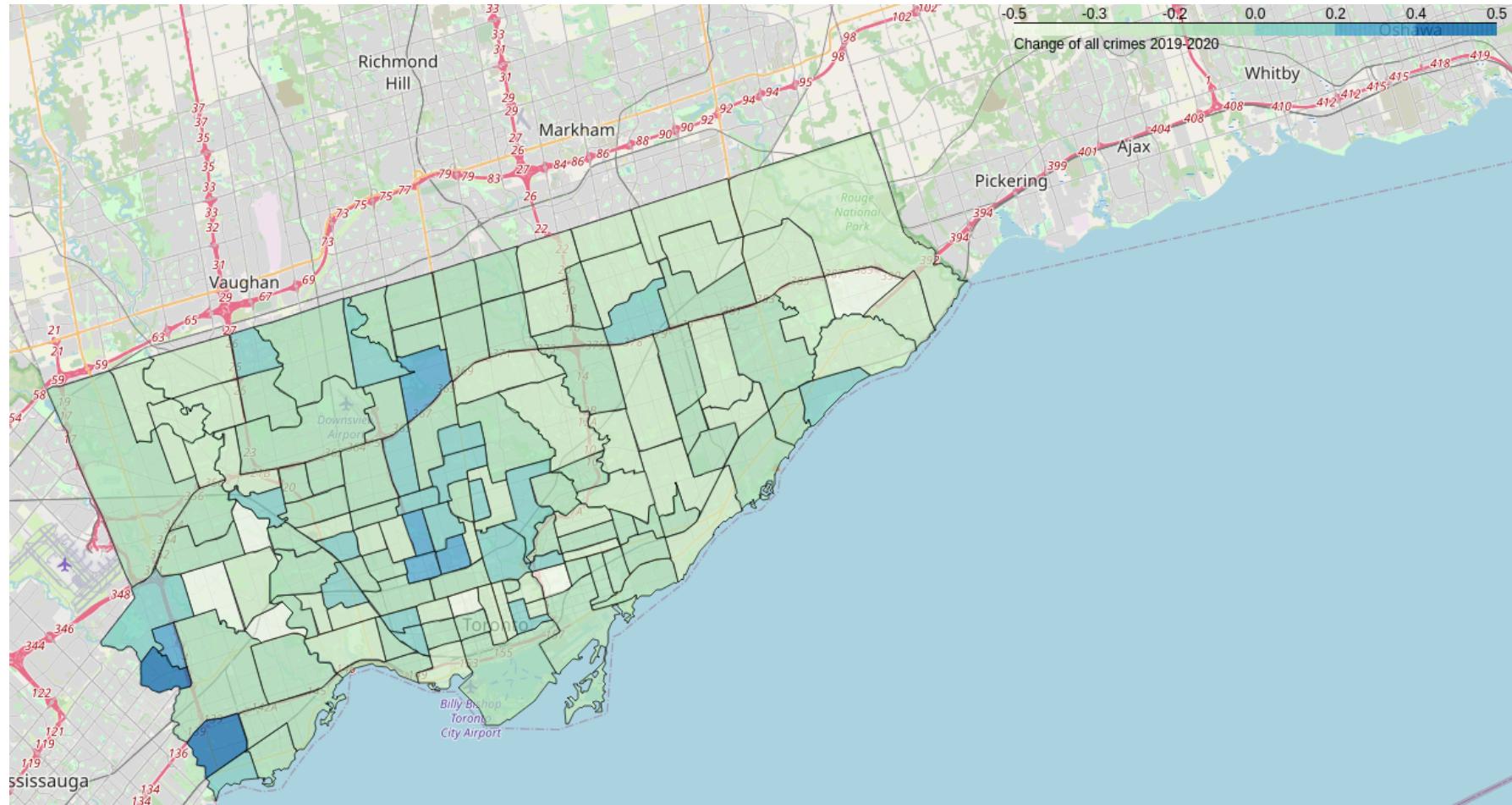
Analyzing safety neighborhoods in Toronto



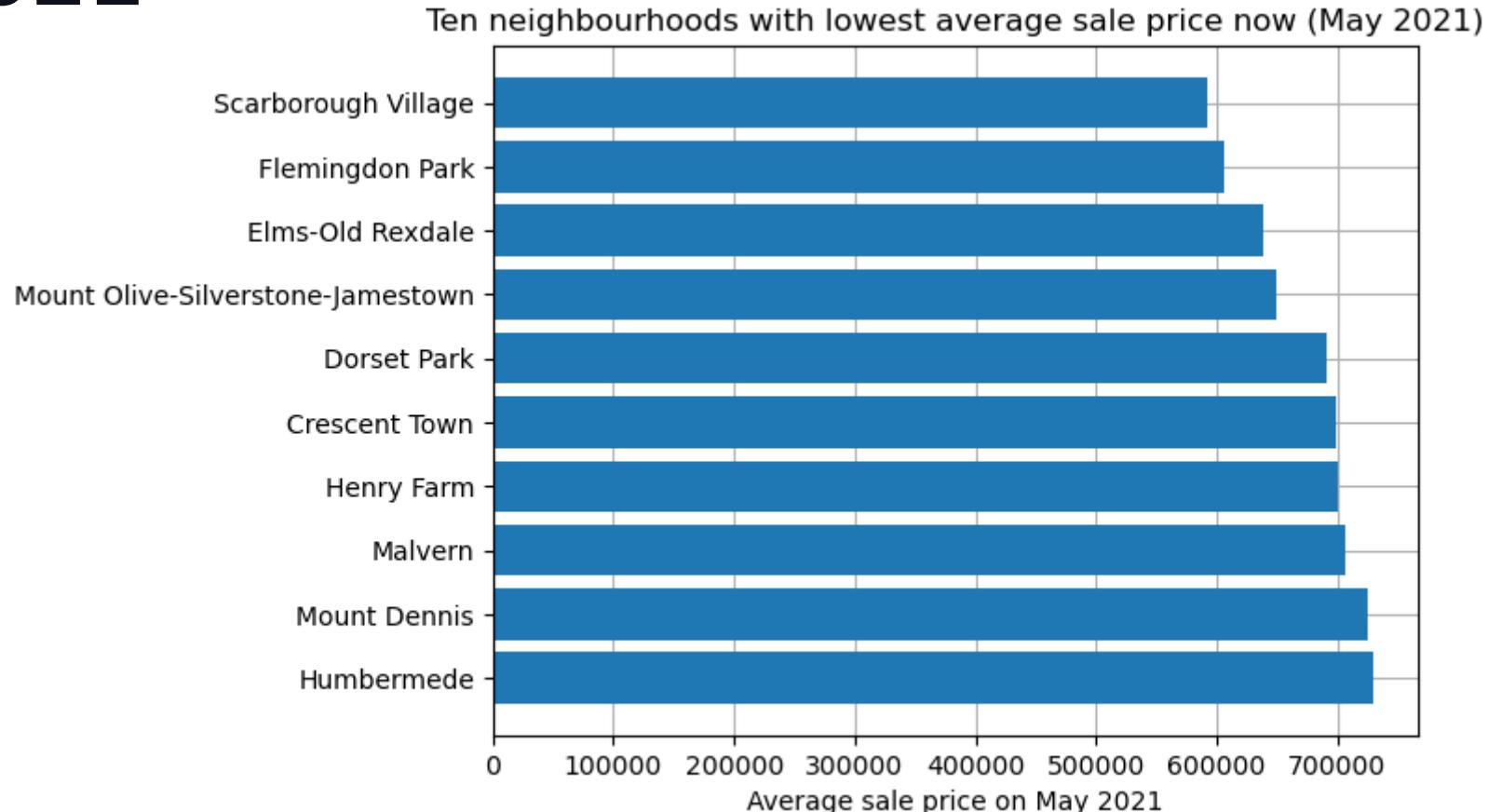
The percent of change in crime numbers in 2019-2020



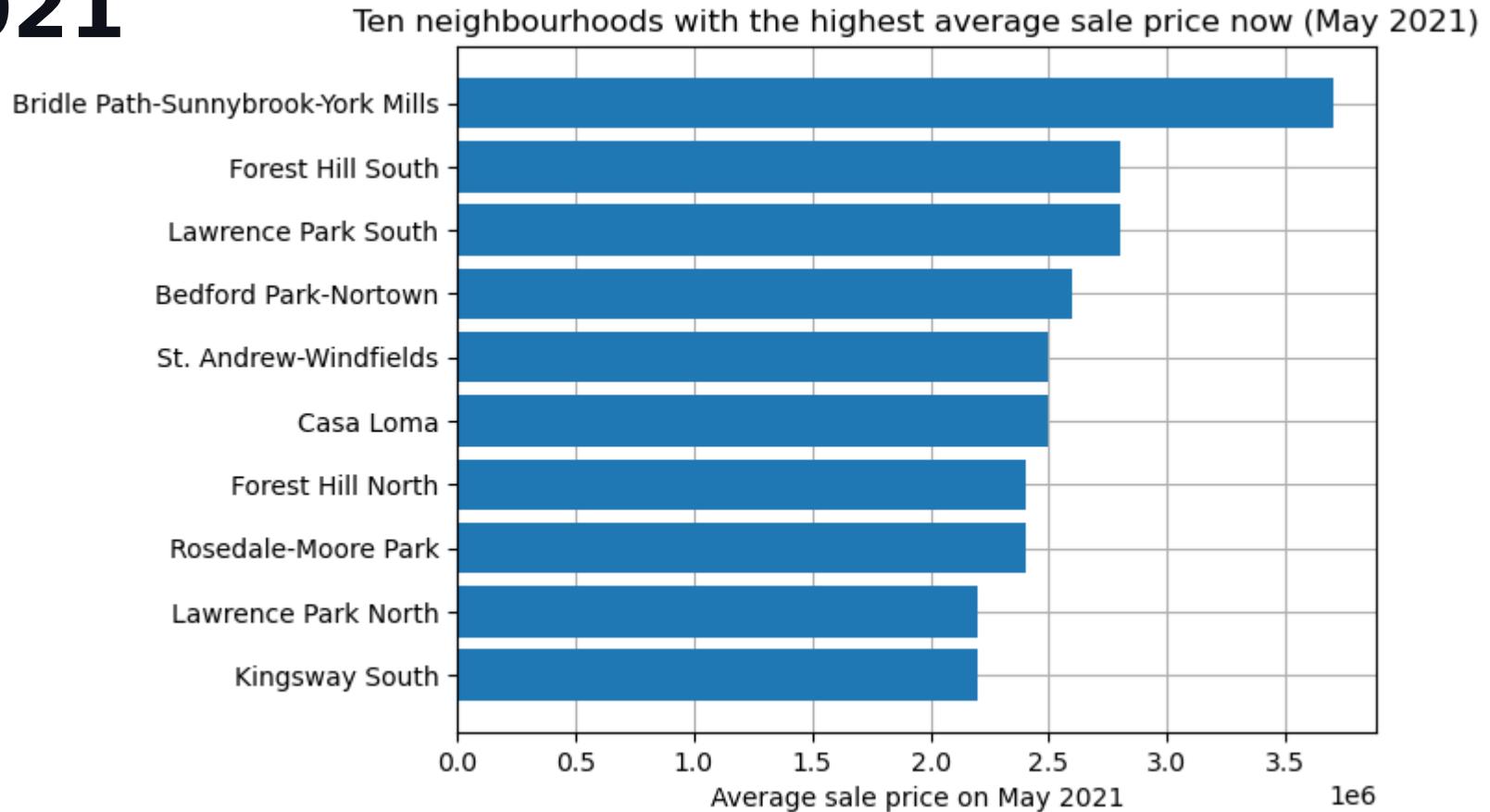
The percent of change in crime numbers in 2019-2020



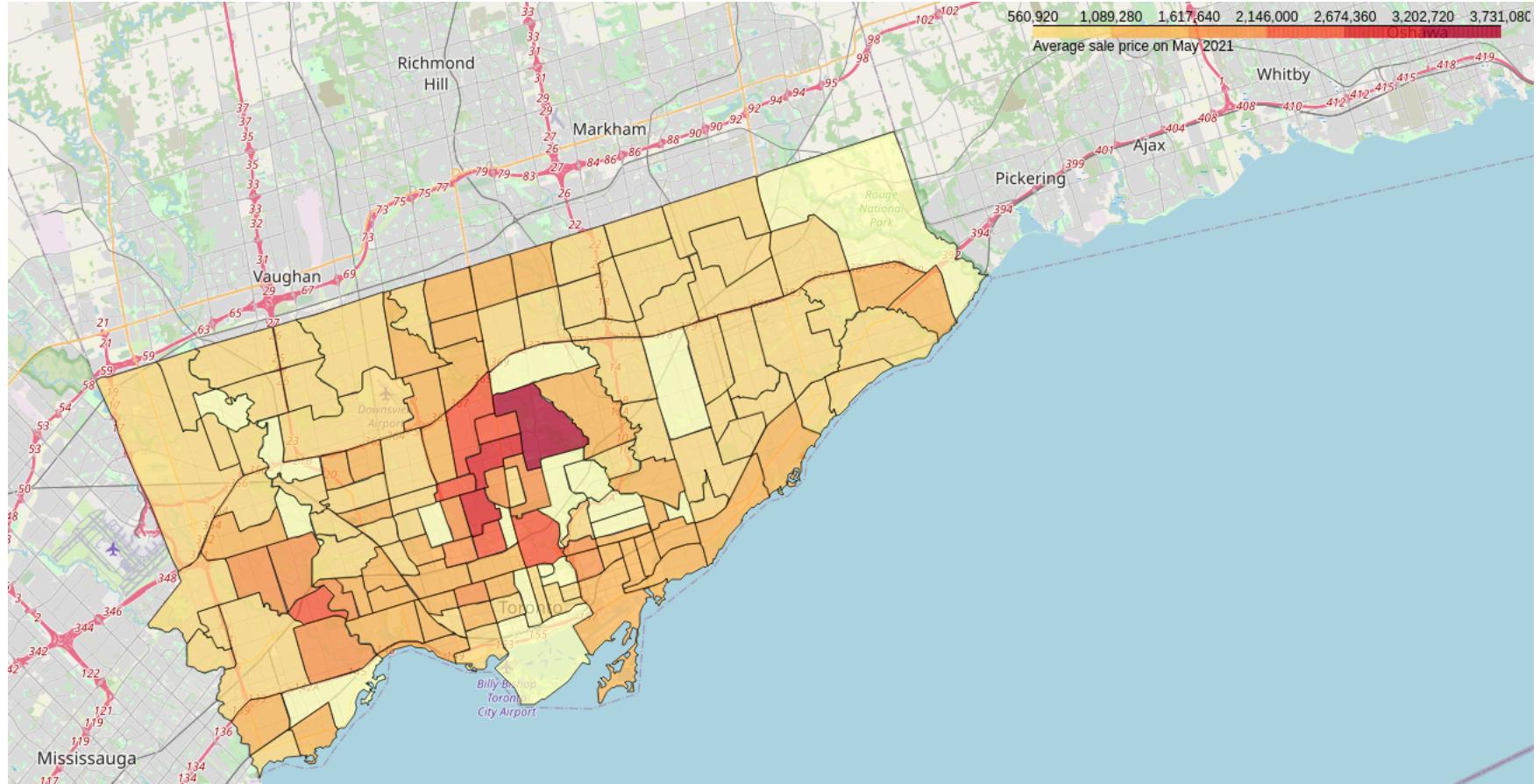
Analyzing house prices by neighborhoods in Toronto in May 2021



Analyzing house prices by neighborhoods in Toronto in May 2021

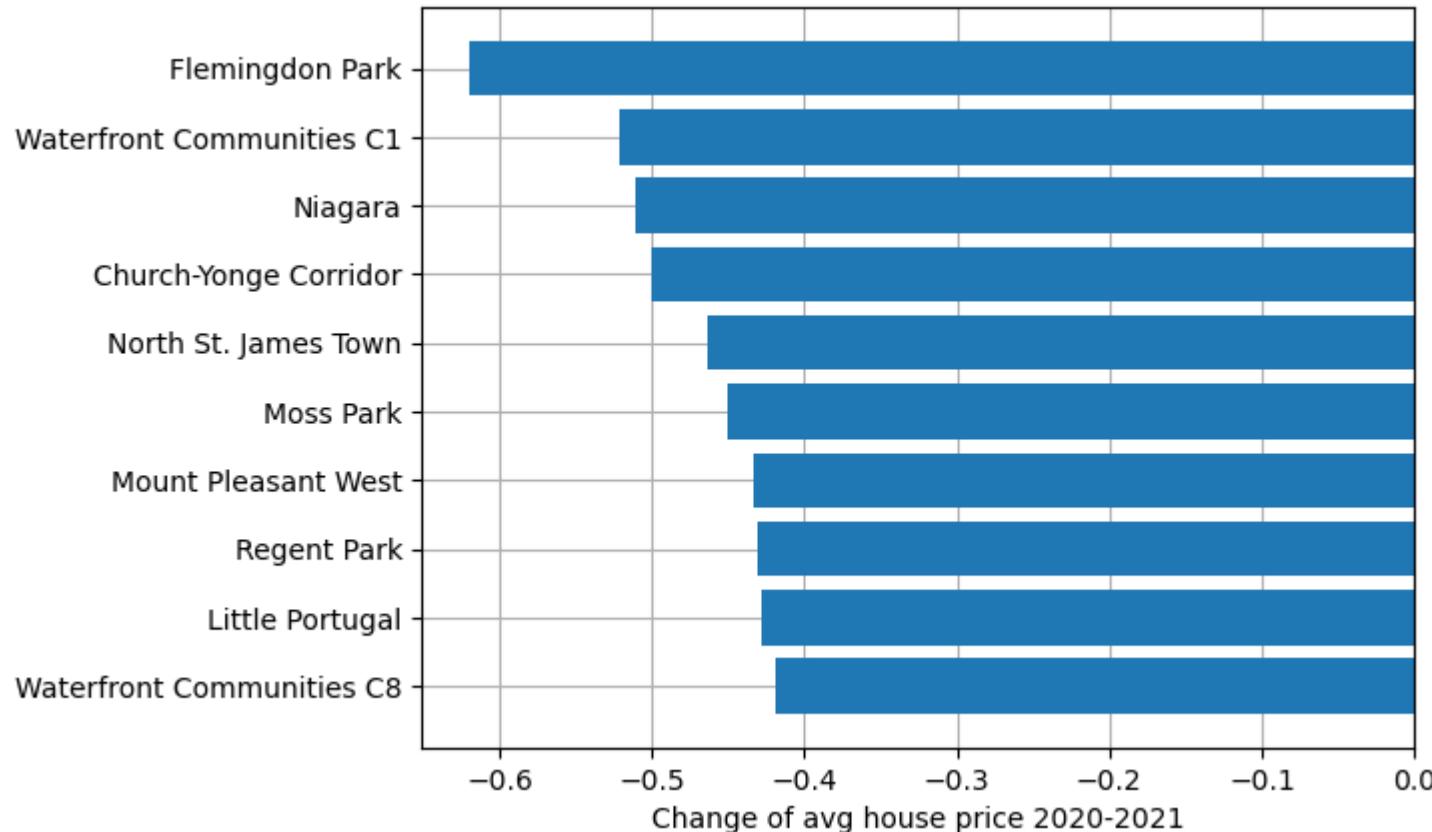


Analyzing house prices by neighborhoods in Toronto in May 2021

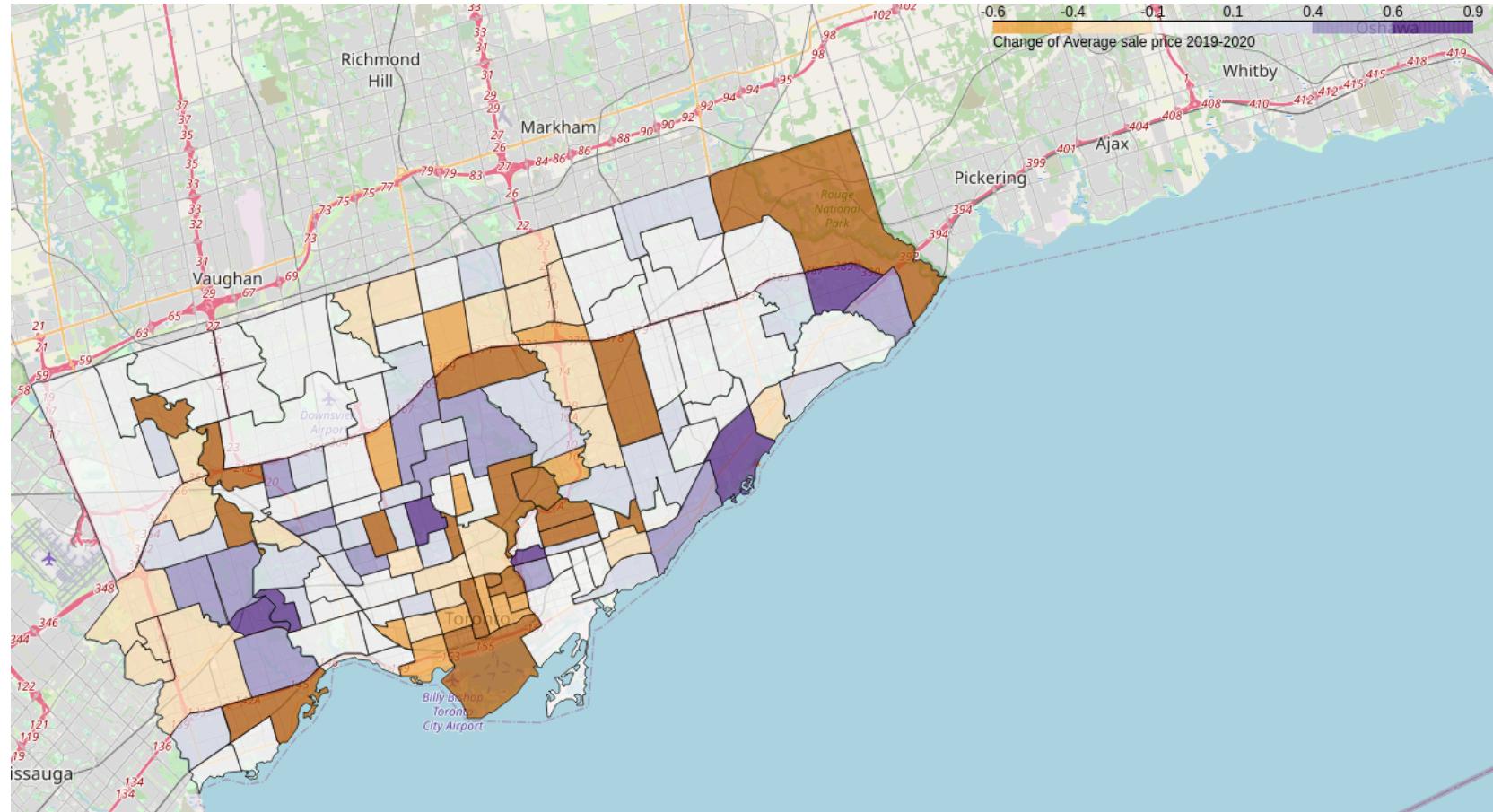


The percent of change of average house price from 2020 to May 2021

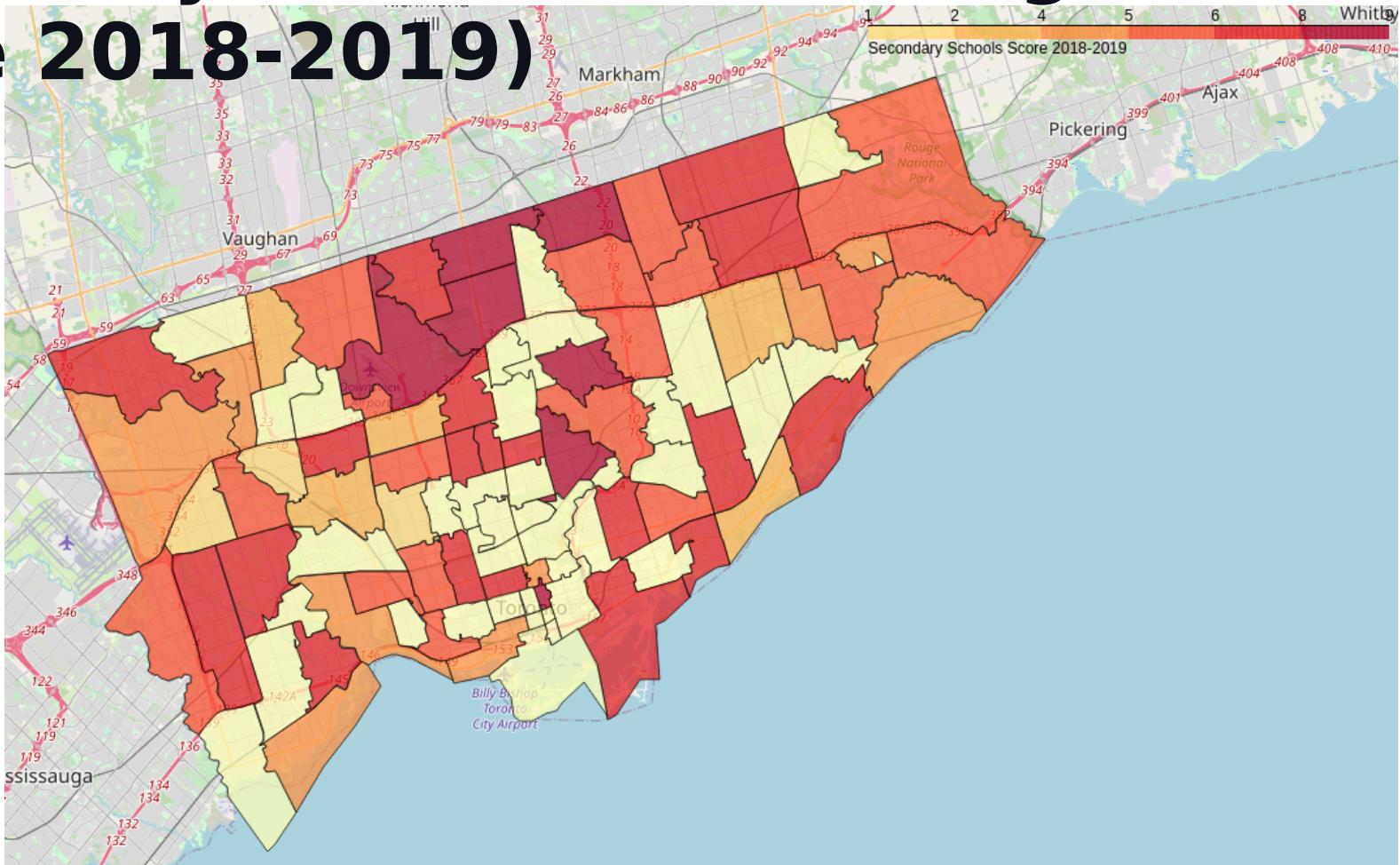
Ten neighbourhoods which have the lower average sale price 2020-2021



The percent of change of average house price from 2020 to May 2021



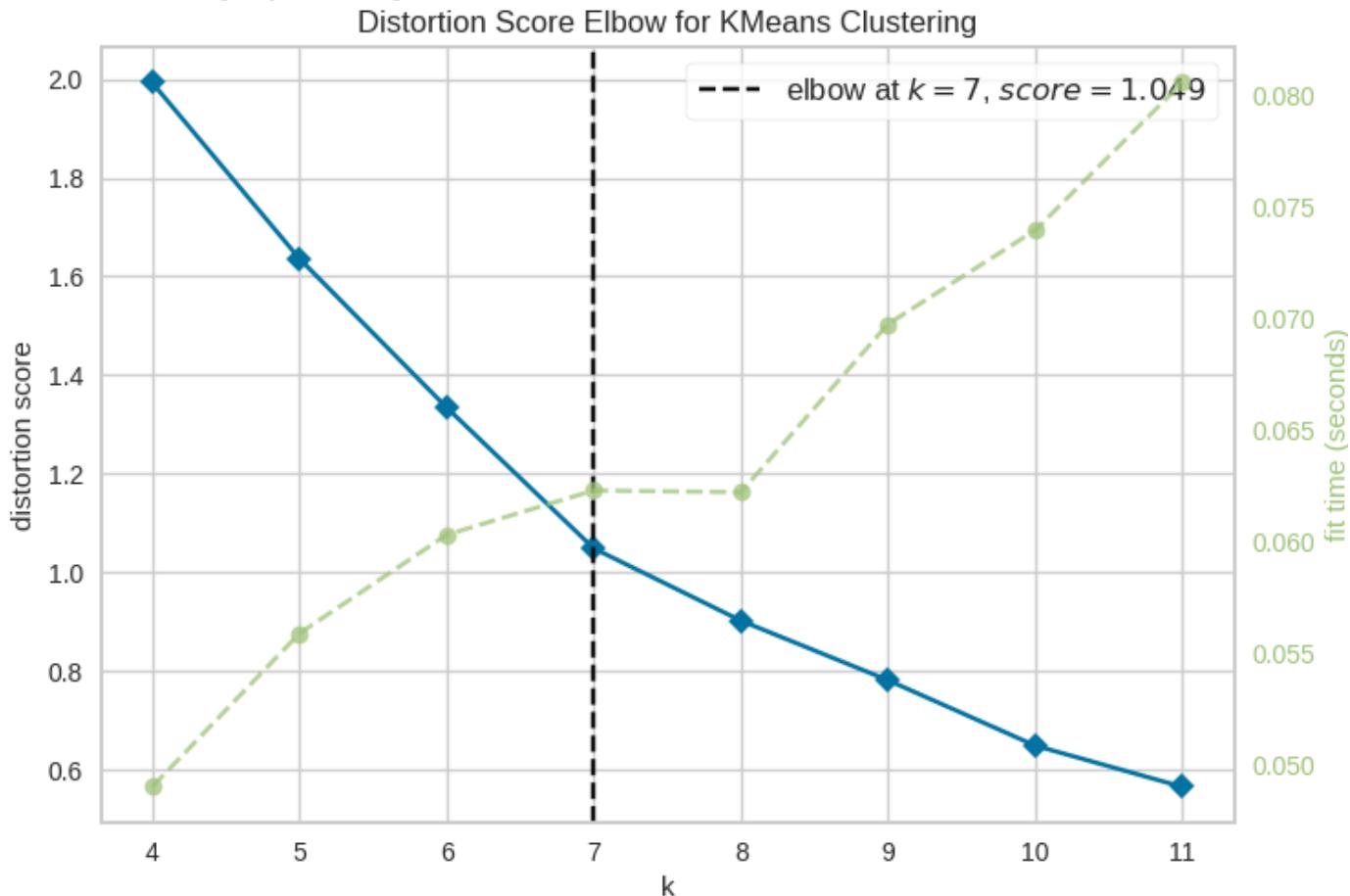
Map of neighborhoods with secondary schools (according to a score 2018-2019)



Clustering using only three columns 'Number of all crimes 2014 - 2020','School Score 2018-2019','Average sale price - now'.

	Postalcode	Borough	Neighborhood	Latitude	Longitude	Number of all crimes 2014 - 2020	School Score 2018-2019	Average sale price - now
0	M3A	North York	Parkwoods,Parkwoods,Parkwoods-Donald	43.753259	-79.329656	2014.0	6.0	1000000.0
1	M6A	North York	Lawrence Manor , Lawrence Heights,Briar Hill-B...	43.718518	-79.464763	1210.0	3.0	848000.0
2	M9A	Etobicoke	Islington Avenue,Edenbridge-Humber Valley,Humb...	43.667856	-79.532242	729.0	7.3	1800000.0
3	M1B	Scarborough	Malvern , Rouge,Agincourt South-Malvern West,M...	43.806686	-79.194353	3051.0	5.5	707000.0
4	M3B	North York	Don MillsNorth,Banbury-Don Mills,Don Mills	43.745906	-79.352188	1402.0	8.1	1500000.0
5	M5B	Downtown Toronto	Garden District,Church-Yonge Corridor,Garden D...	43.657162	-79.378937	7877.0	9.4	754000.0
6	M1C	Scarborough	Rouge Hill , Port Union , Highland Creek,Cente...	43.784535	-79.160497	2448.0	5.9	1300000.0
7	M3C	North York	Don MillsSouth,Flemington Park	43.725900	-79.340923	1375.0	5.9	606000.0
8	M4C	East York	Woodbine Heights,Danforth,Woodbine Heights,Woo...	43.695344	-79.318389	459.0	5.3	1100000.0

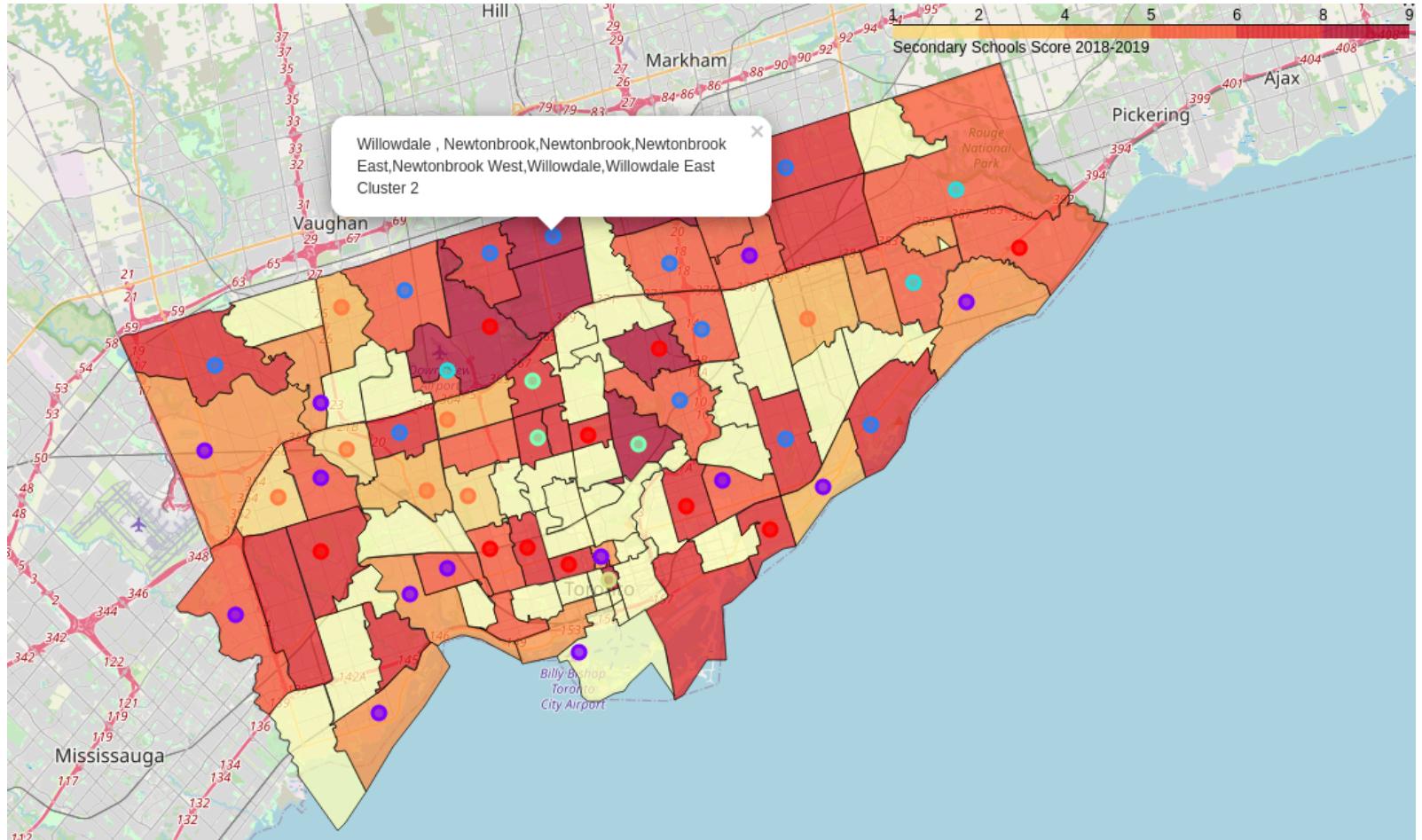
Using Elbow Method to get optimum k of the K-Means



Clusters

	Postalcode	Borough	Neighborhood	Latitude	Longitude	Number of all crimes 2014 - 2020	School Score 2018-2019	Average sale price - now	Cluster Labels
0	M3A	North York	Parkwoods,Parkwoods,Parkwoods-Donald...	43.753259	-79.329656	2014.0	6.0	1000000.0	2
1	M6A	North York	Lawrence Manor , Lawrence Heights,Briar Hill-B...	43.718518	-79.464763	1210.0	3.0	848000.0	6
2	M9A	Etobicoke	Islington Avenue,Edenbridge-Humber Valley,Humb...	43.667856	-79.532242	729.0	7.3	1800000.0	0
3	M1B	Scarborough	Malvern , Rouge,Ajincourt South-Malvern West,M...	43.806686	-79.194353	3051.0	5.5	707000.0	3
4	M3B	North York	Don MillsNorth,Banbury-Don Mills,Don Mills	43.745906	-79.352188	1402.0	8.1	1500000.0	0

Clustering 1



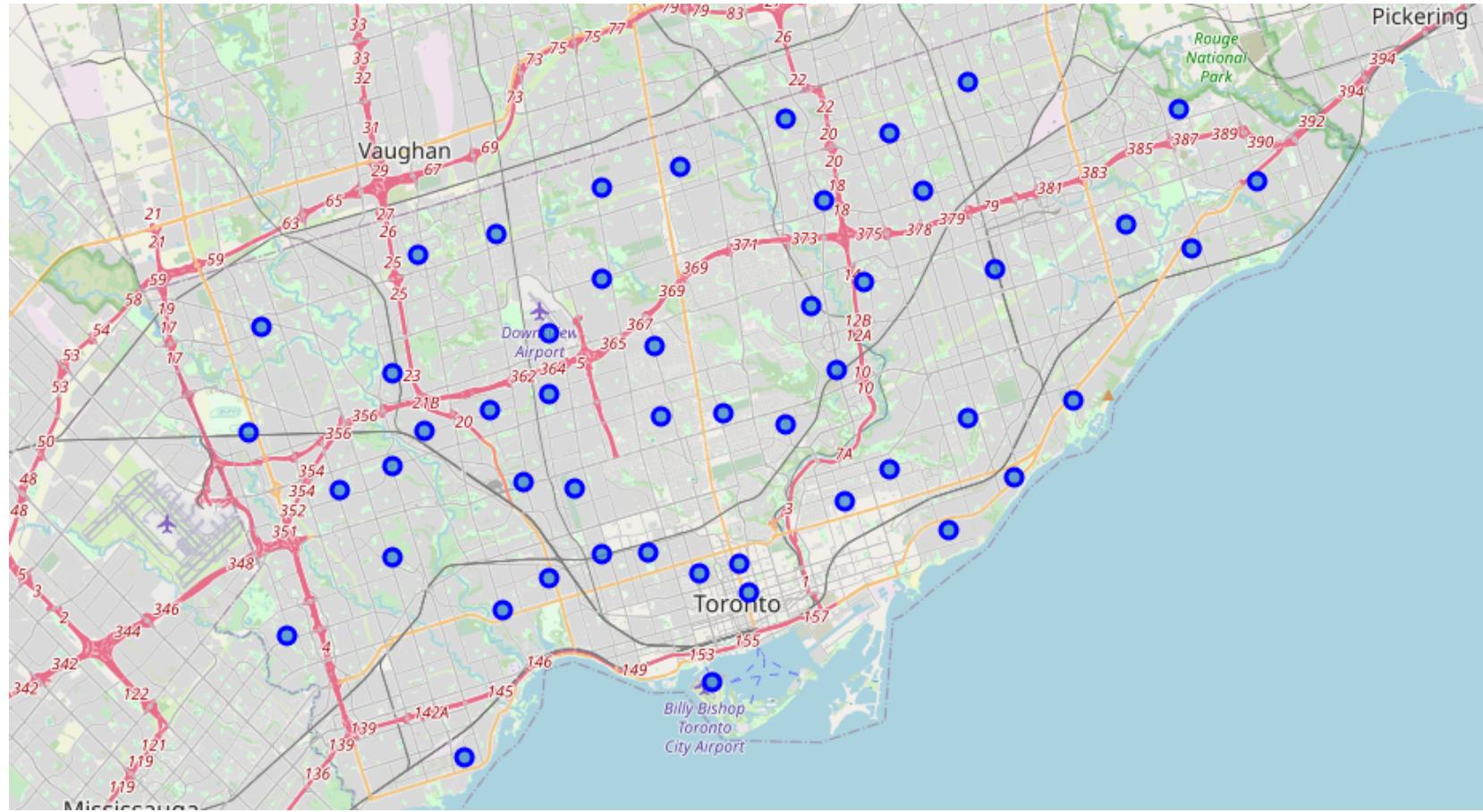
Cluster - examination

Cluster 2

```
[95]: toronto_merged.loc[toronto_merged['Cluster Labels'] == 2, toronto_merged.columns[[1]+[2] + list(range(5, toronto_merged.shape[1]))]]
```

	Borough	Neighborhood	Number of all crimes 2014 - 2020	School Score 2018-2019	Average sale price - now	Cluster Labels
0	North York	Parkwoods,Parkwoods,Parkwoods-Donald...	2014.0	6.0	1000000.0	2
7	North York	Don MillsSouth,Flemington Park	1375.0	5.9	606000.0	2
16	North York	Hillcrest Village,Bayview Woods-Steeles,Hillcr...	920.0	8.1	898000.0	2
19	North York	Fairview , Henry Farm , Oriole,Don Valley Vill...	703.0	6.3	1000000.0	2
20	North York	Northwood Park , York University,Northwood Par...	1764.0	6.0	805000.0	2
23	Scarborough	Golden Mile , Clairlea , Oakridge,Clairlea,Cla...	1958.0	7.8	892000.0	2
24	North York	North Park , Maple Leaf Park , Upwood Park,Map...	453.0	6.9	1200000.0	2
25	Scarborough	Cliffside , Cliffcrest , Scarborough Village W...	1459.0	7.1	592000.0	2
26	North York	Willowdale , Newtonbrook,Newtonbrook,Newtonbro...	1110.0	8.1	918000.0	2
38	North York	WillowdaleWest,Westminster-Branson,Willowdale ...	1098.0	7.2	1200000.0	2
43	Scarborough	Milliken , Agincourt North , Steeles East , L'...	2108.0	7.5	898000.0	2
46	Etobicoke	South Steeles , Silverstone , Humbergate , Jam...	920.0	6.9	898000.0	2
47	Scarborough	Steeles West , L'Amoreaux West,L'Amoreaux,L'Am...	920.0	6.2	898000.0	2

Venus from the Foursquare API

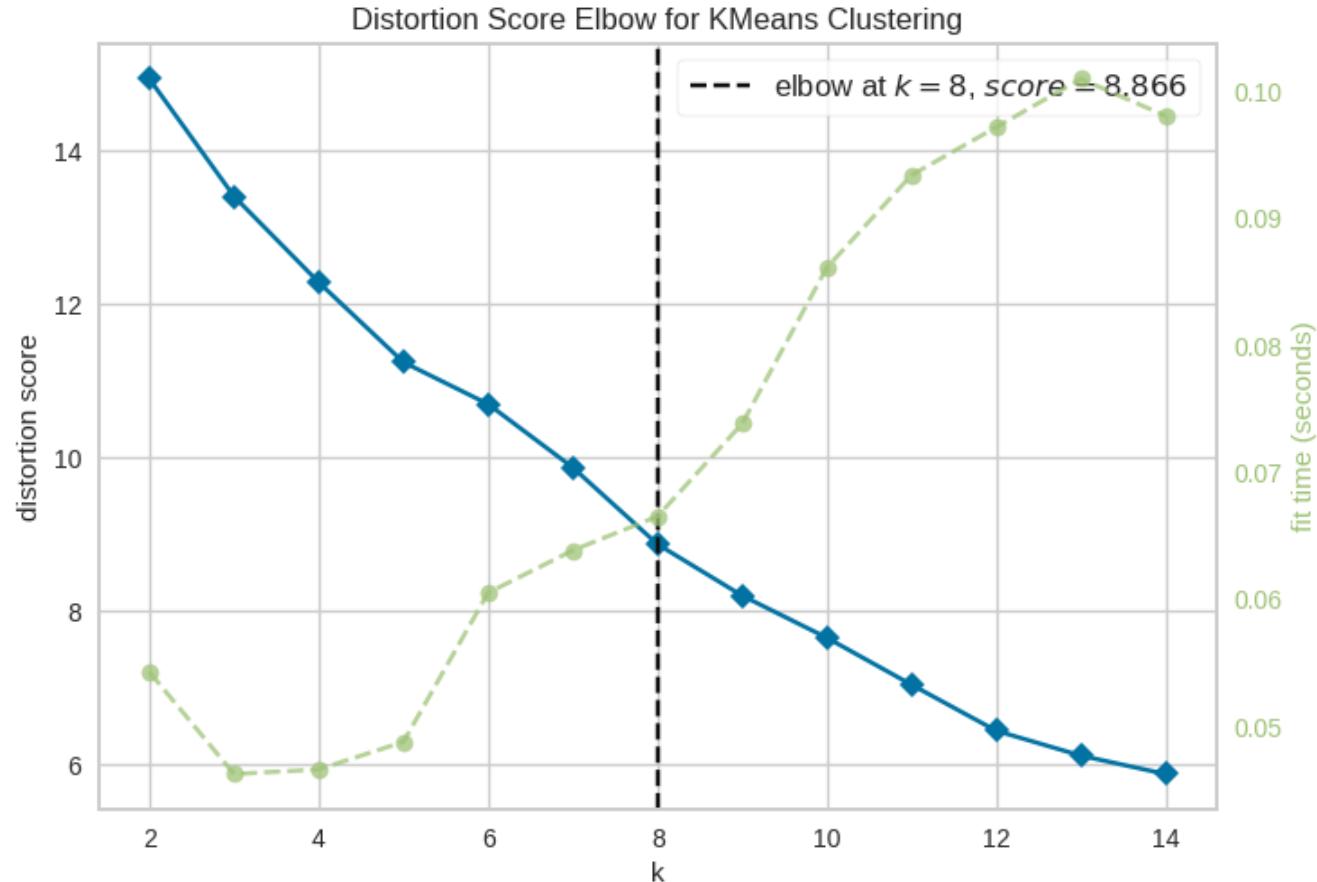


Venues from the Foursquare API , with the mean of the frequency of occurrence of each category:

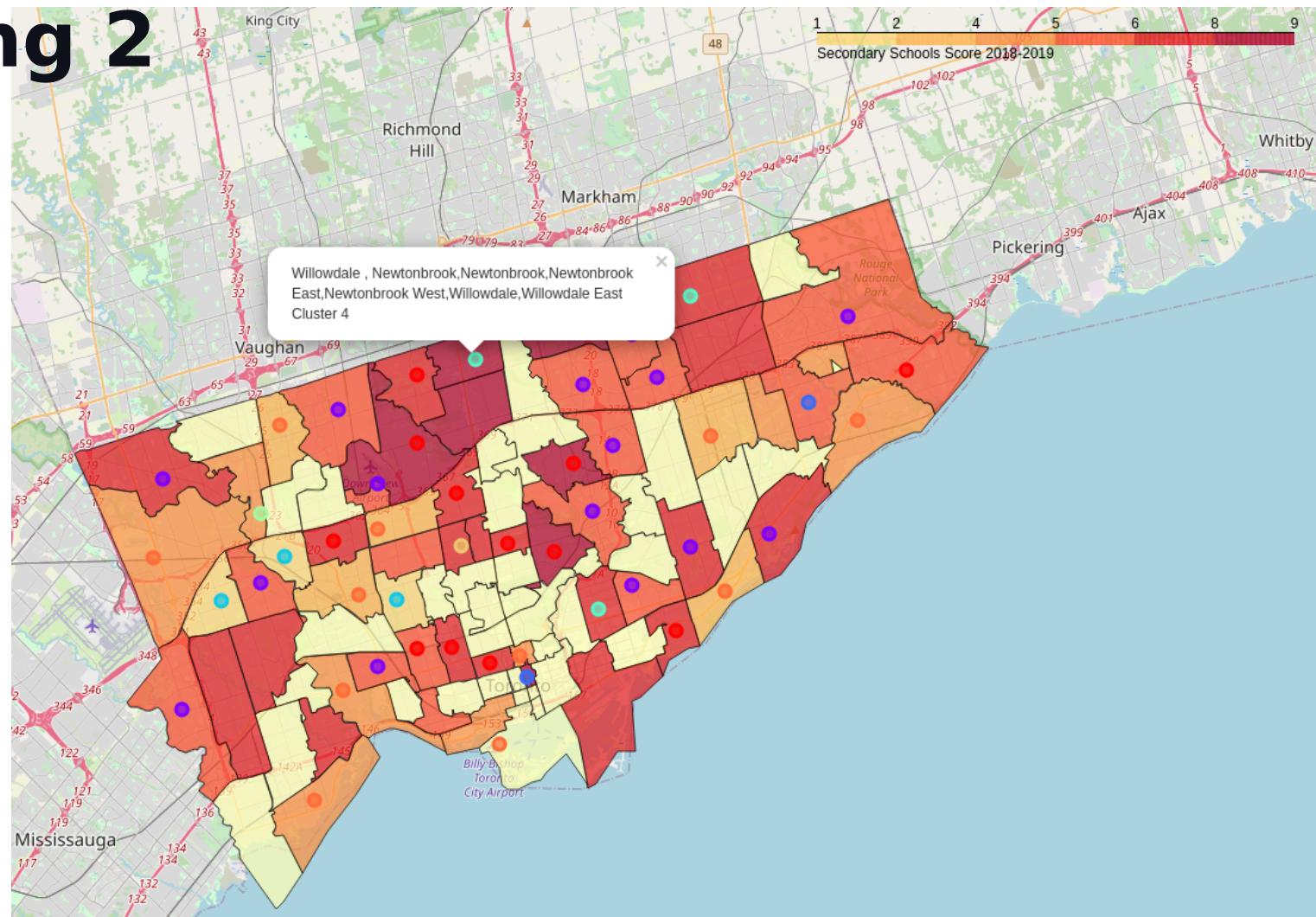
Clustering using only three columns 'Number of all crimes 2014 - 2020','School Score 2018-2019','Average sale price - now' + Venues

Sushi Restaurant	Tea Room	Thai Restaurant	Theater	Theme Restaurant	Toy / Game Store	Trail	Truck Stop	Vegetarian / Vegan Restaurant	Video Game Store	Video Store	Vietnamese Restaurant	Wine Bar	Wine Shop	Women's Store	Number of all crimes 2014 - 2020	School Score 2018-2019	Average sale price - now
0.058824	0.0	0.00	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.925550	0.897727	0.547809
0.040000	0.0	0.04	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.855103	0.704545	0.000000
0.000000	0.0	0.00	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.786791	0.318182	0.647410
0.000000	0.0	0.00	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.953169	0.397727	0.893426
0.000000	0.0	0.00	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.25	0.961041	0.011364	0.796813

Using Elbow Method to get optimum k of the K-Means



Clustering 2



The most optimum place to settle down in Toronto would be neighborhoods from cluster 5

Cluster 5

```
toronto_merged.loc[toronto_merged['Cluster Labels'] == 4, toronto_merged.columns[[1] + [2] + list(range(5, toronto_merged.shape[1]))]]
```

	Borough	Neighborhood	Number of all crimes 2014 - 2020	School Score 2018-2019	Average sale price - now	Cluster Labels	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
21	East York/East Toronto	The Danforth East, Blake-Jones, Broadview North...	739.0	7.3	1300000.0	4	Park	Metro Station	Convenience Store	Women's Store	Dance Studio	Discount Store	Diner	Dim Sum Restaurant
26	North York	Willowdale , Newtonbrook,Newtonbrook,Newtonbro...	1110.0	8.1	918000.0	4	Park	Women's Store	Curling Ice	Distribution Center	Discount Store	Diner	Dim Sum Restaurant	Dessert Shop
43	Scarborough	Milliken , Agincourt North , Steeles East , L...	2108.0	7.5	898000.0	4	Park	Playground	Intersection	Curling Ice	Discount Store	Diner	Dim Sum Restaurant	Dessert Shop

Discussion

- the most important observation - problem with getting all the needed information - sometimes there is no data or data are a few years old
- the second observation of the decision-making process on data - data are always changing
- e.g. venues from Foursquare API, a few hours - change the number of all venus in many categories and neighborhoods in Toronto - what changes our data

Conclusion

- data science and its tools helps people to choose a better place to settle down and live with children,
- data science simplifies looking for such things as a good school, secure neighborhood, price of a house, and place for spending outdoor time, which in older times seemed just as a "lottery"