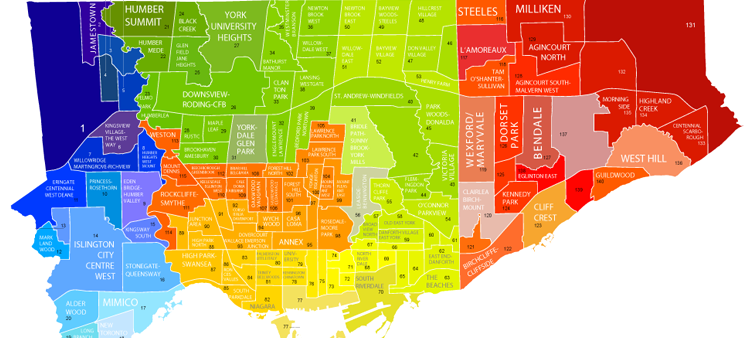
# **SHOULD WE MOVE TO WHICH PART OF TORONTO?**



**INTRODUCTION**

Let's imagine that you have a great new job that can boost up your career!

BUT... That new job is requiring you to move to Toronto, Canada - a new country, new culture, new everything for you - to start a new life. How can you ensure that you will be fine when living there? Do you able to pay the house price for your family, school fees for your children? What's the weather there? The crime rates? The social life?

This project would help you to make a better decision on choosing the best place that suits you and your family. It helps you by providing a lot of information about Toronto, including the neighborhoods, the most common venues in each of those places, the average household income, renovation fees, etc.

By leveraging K-mean clustering unsupervised machine learning algorithm, all the venues belong to 2 neighborhoods - which we will choose randomly - will be categorized and displayed on the map. From that, you can easily distinguish the differences between the 2 places, as well as the similarities between them, to support you giving the best decision!

**DATA SETS, API & METHODOLOGY**

2 data sets will be used to get more information about Toronto & its neighborhood

1. Foursquare API:   
   According to Wikipedia: *"Foursquare is a local search-and-discovery service mobile app which provides search results for its users. The app provides personalized recommendations of places to go to near a user's current location based on users' "previous browsing history, purchases, or check-in history""* We can leverage Foursquare's API to gather detailed information about a place, its venues, its photos & many more related tips.
2. NextHome's website:   
   <https://nexthome.ca/>   
   We can find apartments, homes or condos for rent or for sale in Canada with this website. It also provides other knowledge about a specific place and its related information, including: Average Household Income, Average Renovation Cost, Overall Age of Housing, Top 3 Popular Languages and many other valuable things.

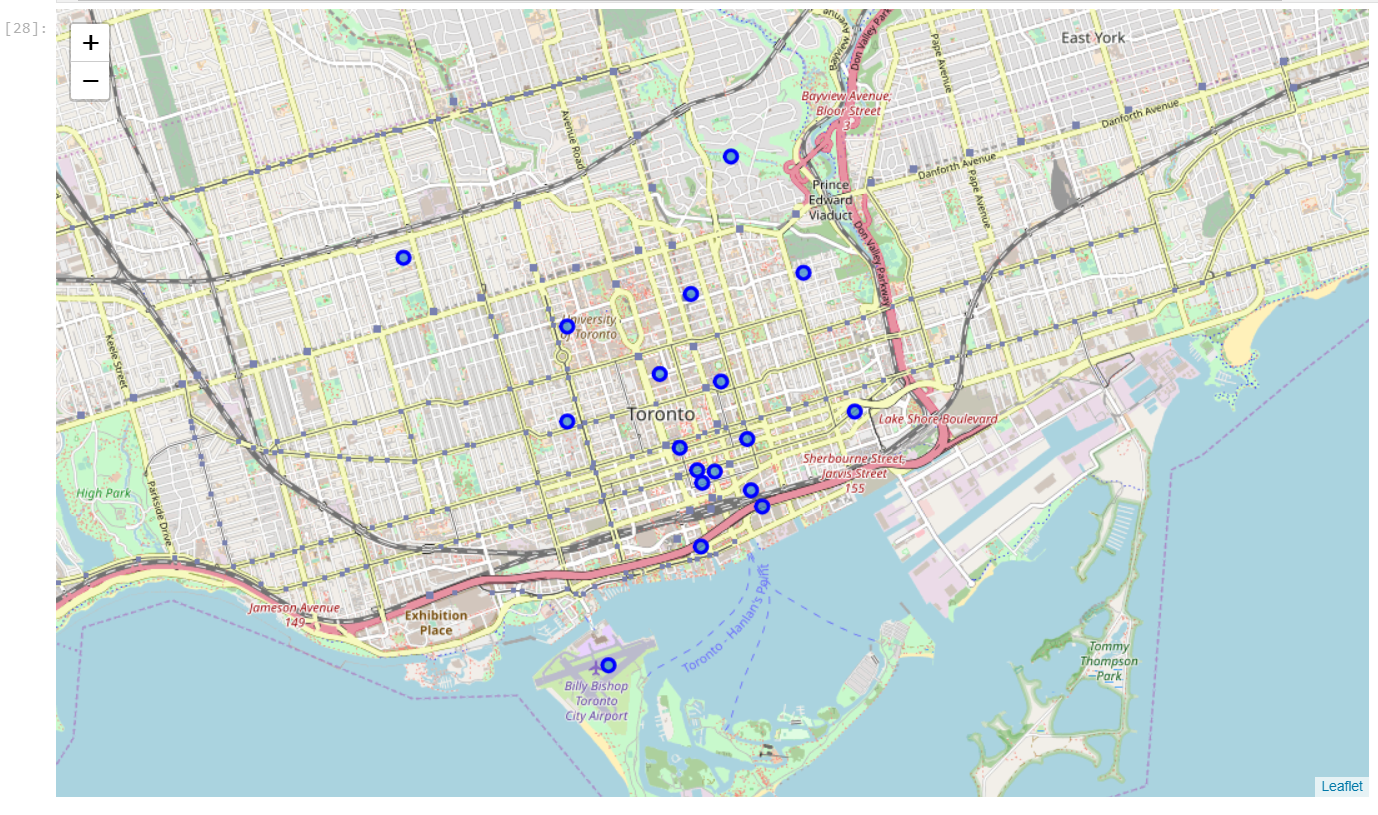
**How they will be used?**

* We need to register to Foursquare in order to use its API.
* Once we have the access keys, we can use them to retrieve information about various neighborhoods in Toronto.
* Then, build a map of Toronto with neighborhoods superimposed on top.
* Use K-Mean to clustering places into various clusters.
* Together with data from NextHome website, we then build a comprehensive information about places in Toronto.
* Visualize the data to give a better understanding
* And finally, make a conclusion!

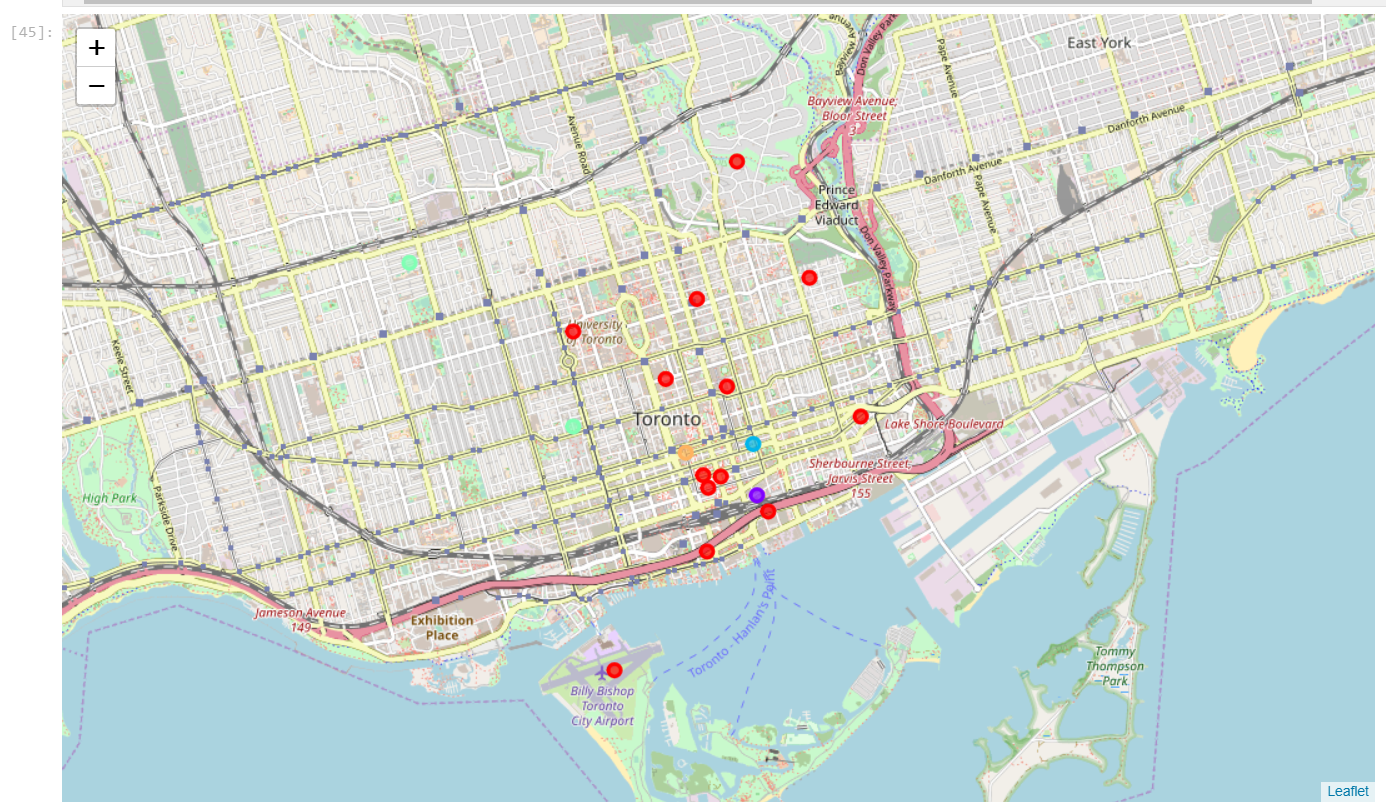
**API & LIBRARIES**

* **Foursquare API**: to gather information about any place in the world.
* **NumPy**: a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.
* **Pandas**: a library for data manipulation and analysis.
* **Geopy**: to locate the coordinates of addresses, cities, countries, and landmarks across the globe.
* **Requests**: a library to handle http requests
* **Matplotlib**: for embedding plots into applications.
* **Sklearn**: a machine learning library with various classification, regression and clustering algorithms.
* **Folium**: a map rendering library.

**RESULTS**

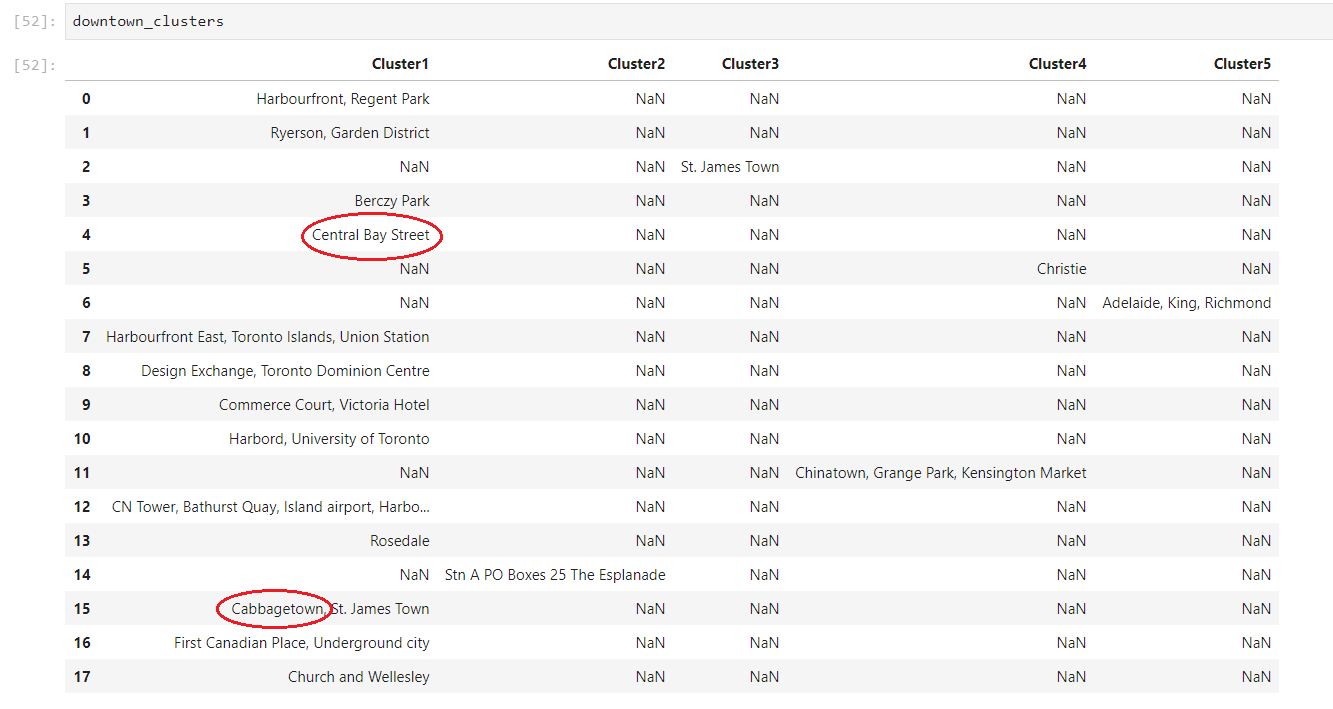
Downtown Toronto before clustering

Successfully add neighborhoods as markers on Downtown Toronto using data from Foursquare & Folium library

Downtown Toronto AFTER clustering

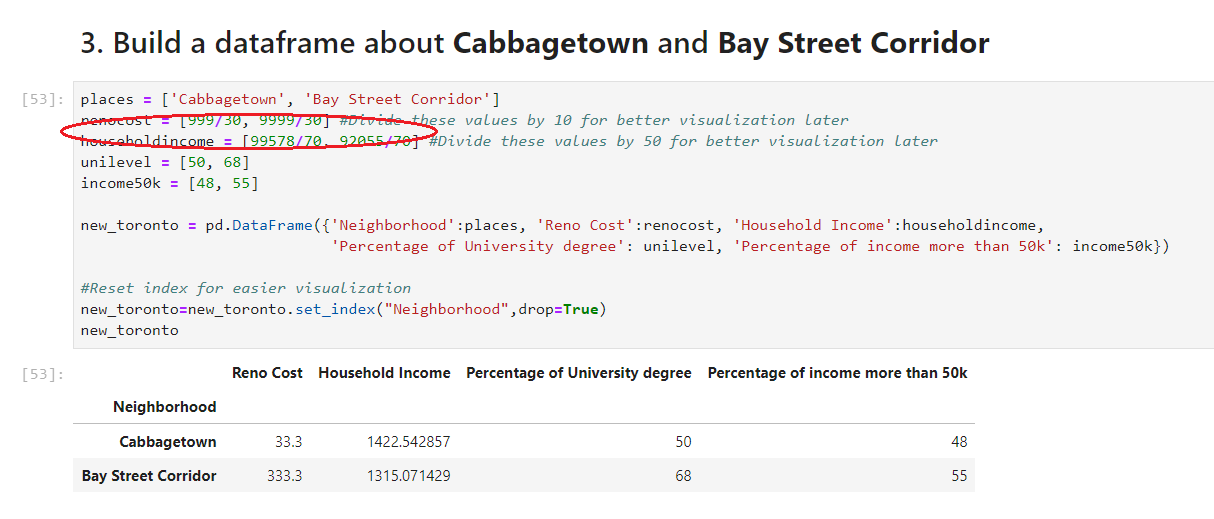
Clearly we can see that the neighborhoods have been divided into various clusters by using K-Mean algrorithm.

Now we can pick up 2 neighborhoods to explore more about them. But first, let consolidate our clusters for easier choosing neighborhoods.

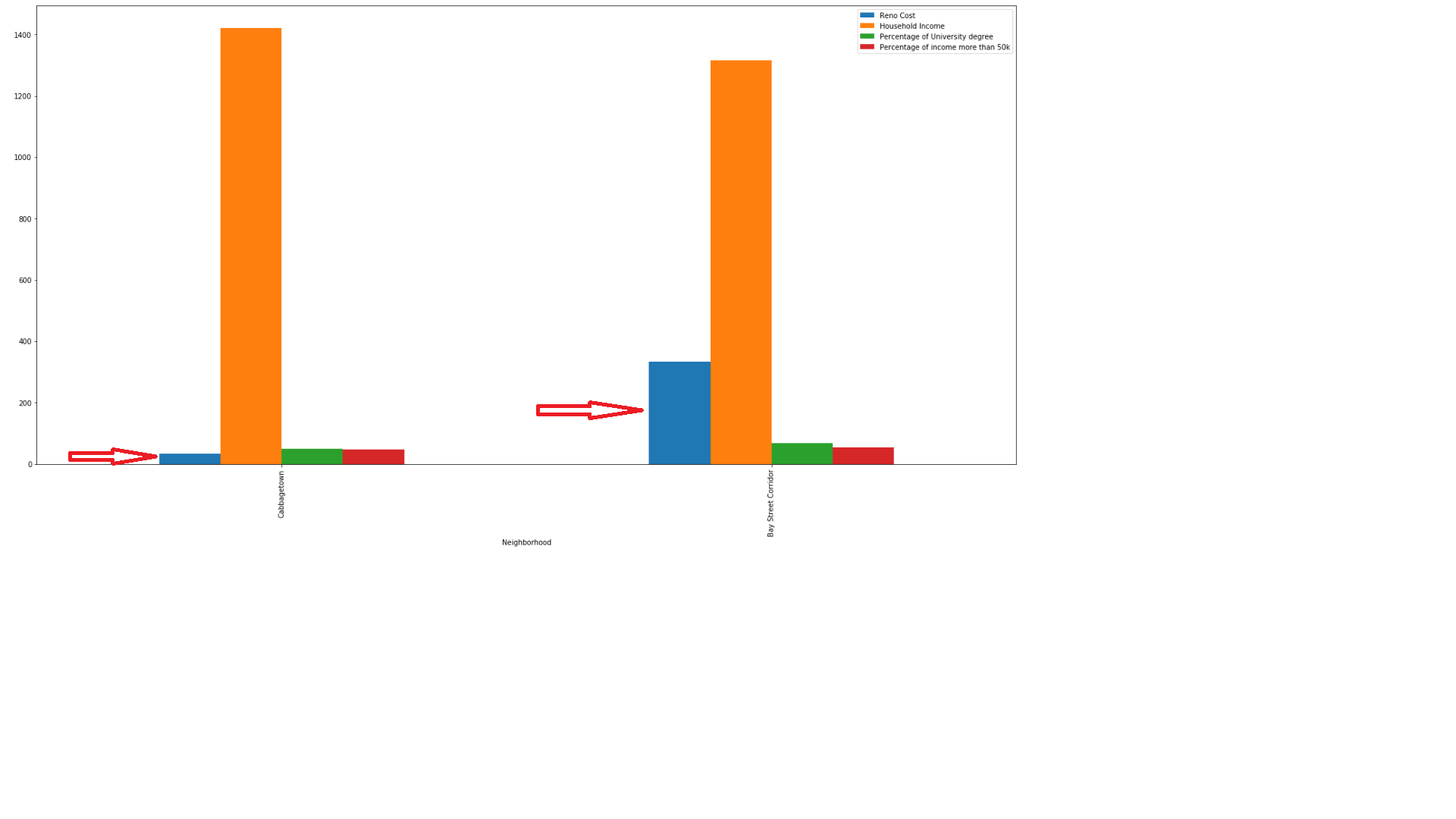


From now on, we will pick up 2 neighborhoods from the same cluster to explore more about them and to see whether they have anything different?

For a better visualization with other smaller values, we will divide the renovation cost by 30, divide the household income by 70 like below



Final chart!



From the above chart, we can clearly see that:

* + Renovation cost in Cabbagetown is significant lower than Bay Street Corridor
  + Household income in Cabbagetown is higher than Bay Street Corridor
  + Percentage of people who has graduated from University in Bay Street Corridor is slightly higher than Cabbagetown
  + Percentage of people who has income more than $50,000 in Bay Street Corridor is slightly higher than Cabbagetown.

**DISCUSSION**

We can see between the 2 neighborhoods, Cabbagetown and Bay Street Corridor, each of them has its own advantage. While Cabbagetown wins over 2 criteria, Bay Street also has its victory on Education and Middle-class income.

**CONCLUSTION**

This analysis concludes that Cabbagetown has the significant lower renovation cost and higher average income than Bay Street Corridor. Despite the percentage of people who has bachelor degree in Bay Street is slightly higher than Cabbagetown, as well as the percentage of population that belongs to the middle-class is higher too; living in Cabbagetown is definitely more than reasonable!   
**Hence, Cabbagetown wins over Bay Street Corridor for most of people!!!**

### THANK YOU!