CAPSTONE PROJECT

THE BATTLE OF NEIGHBORHOODS – BY AMIT BENDRE

INTRODUCTION & PROBLEM STATEMENT

- In this project, we study two important cities Manhattan & Toronto. The objective of the study is to segment areas of Manhattan & Toronto into most common places using Foursquare API.
- Using segmentation and clustering, we aim to find out similarities and dissimilarities of both the cities.
- The target audience of this project are tourists who wish to finalize their travel plan based on the findings.

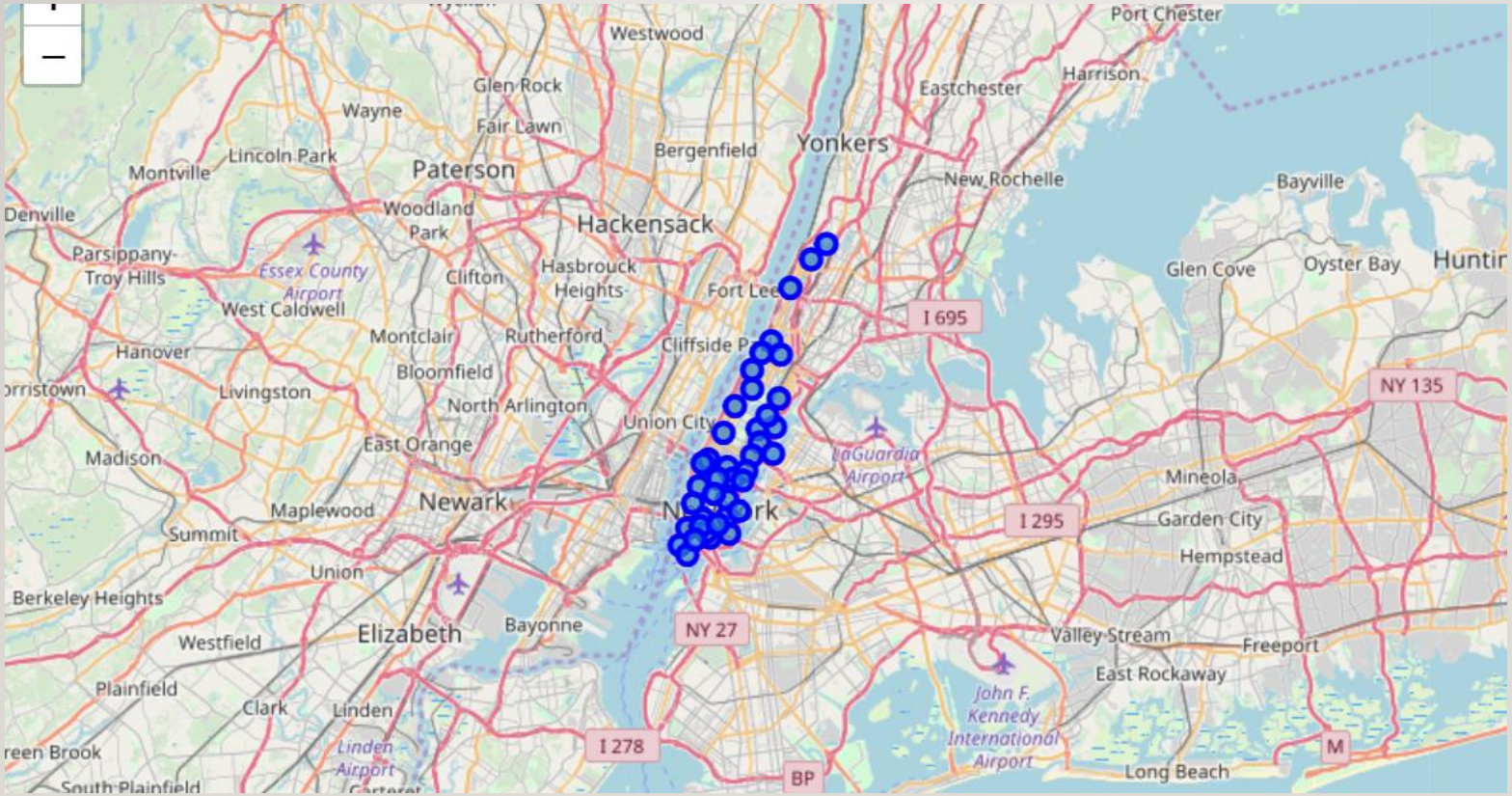
DATA

- We will use the Foursquare API to explore neighborhoods in New York/Manhattan. We will use the explore function to get the most common venue categories in each neighborhood, and then use this feature to group the neighborhoods into clusters.
- For Manhattan dataset exists for free on the web. Link to the dataset: https://geo.nyu.edu/catalog/nyu_2451_34572
- For the Toronto neighborhood data, a Wikipedia page exists that has all the information we need to explore and cluster the neighborhoods in Toronto. You will be required to scrape the Wikipedia page and wrangle the data.
- Dataset link for Canada/Toronto: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

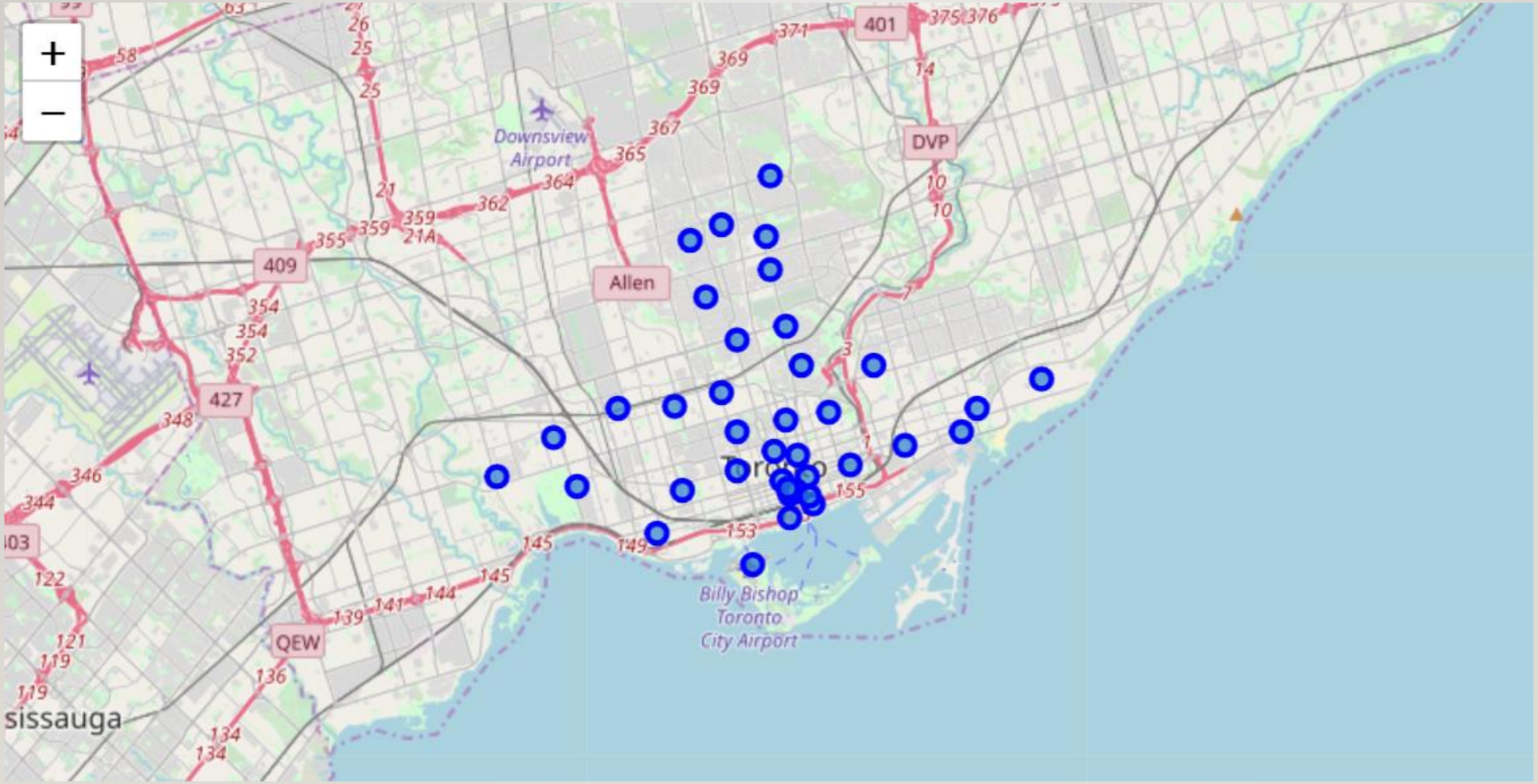
METHODOLOGY

- We have used k-means Clustering to segment the neighborhoods.
- With k-means Clustering we aim to group neighborhoods together based on similar characteristic.
- Once we get Cluster groups of both Manhattan & Toronto, we do compare and contract to see how these two cities are identical and different.

MANHATTAN MAP



TORONTO MAP


401
DVP
409
427
QEW
Allen
Downsview Airport
Billy Bishop Toronto City Airport
sissauga

CLUSTER DETAILS

Manhattan Cluster:

	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
37	Stuyvesant Town	Bar	Boat or Ferry	Park	Playground	Gas Station	Pet Service	Farmers Market	Basketball Court	Baseball Field	Cocktail Bar

Toronto Cluster:

	Borough	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	9th Most Common Venue	10th Most Common Venue
4	Central Toronto	4	Bus Line	Park	Swim School	Dim Sum Restaurant	Women's Store	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store

RESULT & CONCLUSION

- For Manhattan with k-mean = 3, Discount Bar, Boat & Park are the top 3 common venues. For Toronto with k-mean = 5, Bus line, Park and Swim School are the top 3 common venues.
- As you can see, the clustering is completely different, which mean, if tourist have to carefully select their city of visit as per preferences as identified in the clusters.