

Final Project of Applied Data Science Capstone - Find a new living area in Toronto

Coursera - Applied Data Science Capstone
Capstone Project - The Battle of Neighborhoods

2021.05.05

Chien-Hsi Chen

Content

- ▶ Introduction
- ▶ Data for this project
- ▶ Methods
- ▶ Results
- ▶ Discussion and Conclusion

Introduction

- ▶ Many purposes to move, new job, study, retirement
- ▶ Moving is hard, especially to a strange place or foreign countries
- ▶ Information of the place to move is everywhere but also nowhere
- ▶ Quality of life (QoL) is the key to decide where to live
 - ▶ Many factors
- ▶ A quick and thorough way to gather information to help decide where would be a suitable residential area is required

Scenario for this project

- ▶ David, a software engineer
 - ▶ Home office and only sometimes to the office for meetings.
 - ▶ The distance between his home and office is not important
 - ▶ The quality of life is the thing he cares the most
- ▶ Requirements for the new home
 - ▶ Restaurants right down the street within 200 meters
 - ▶ Fresh coffee from the coffee shops
 - ▶ Work out in a gym after the long working time
 - ▶ Avoid any areas that may cause allergies, such as parks, flower stores, pet stores, or something similar

Data for this project

- ▶ Borough data of Toronto from Wikipedia
- ▶ Location data from Foursquare

Methods

- Retrieve necessary location data from Wiki and Foursquare

	PostalCode	Borough	Neighborhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park, Harbourfront
3	M6A	North York	Lawrence Manor, Lawrence Heights
4	M7A	Queen's Park	Ontario Provincial Government

Clean up data

- ▶ Retrieve location data from Foursquare
- ▶ Categorized the shops into bigger categories

	Neighborhood	All restaurants	All sports	All coffee	All allergies	All bars	Accessories Store	Adult Boutique	Art Gallery	Arts & Crafts Store	Auto Workshop	BBQ Joint	Bank	Bike Shop	Bookstore	Boutique	Bubble Tea Shop	Burrito Place	Cheese Shop	Chocolate Shop	Clot S
0	Alderwood, Long Branch	0.333333	0.0	0.333333	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	Bathurst Manor, Wilson Heights, Downsview North	0.714286	0.0	0.142857	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	Bedford Park, Lawrence Manor East	0.666667	0.0	0.111111	0.0	0.111111	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	Birch Cliff, Cliffside West	0.000000	0.0	1.000000	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	Brockton, Parkdale Village	0.000000	0.0	0.500000	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

2021-05-05

7

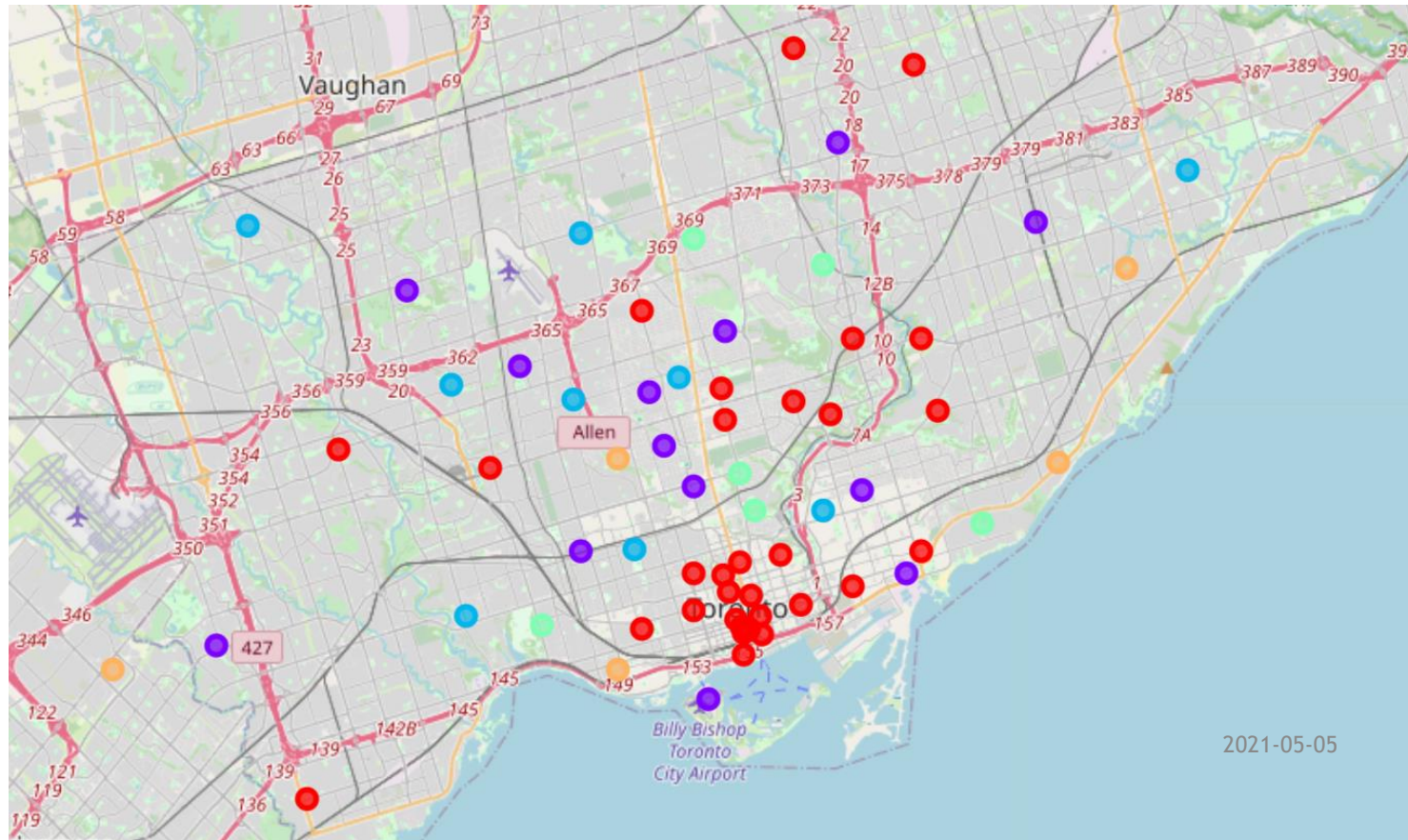
Perform analysis, clustering, and visualization

- K-means clustering performed
- Sorted the dataframe with the top 5 venues

Cluster Labels		Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	0	Alderwood, Long Branch	All restaurants	All coffee	Pharmacy	Wings Joint	Organic Grocery
1	2	Bathurst Manor, Wilson Heights, Downsview North	All restaurants	All coffee	Deli / Bodega	Jewelry Store	Opera House
2	0	Bedford Park, Lawrence Manor East	All restaurants	All coffee	All bars	Cupcake Shop	Kids Store
3	4	Birch Cliff, Cliffside West	All coffee	All restaurants	Jewelry Store	Opera House	Music Venue
4	4	Brockton, Parkdale Village, Exhibition Place	All coffee	Playground	All restaurants	Jewelry Store	Opera House

Perform analysis, clustering, and visualization

- The distribution of all the clustered neighborhoods differentiated by color



2021-05-05

Results

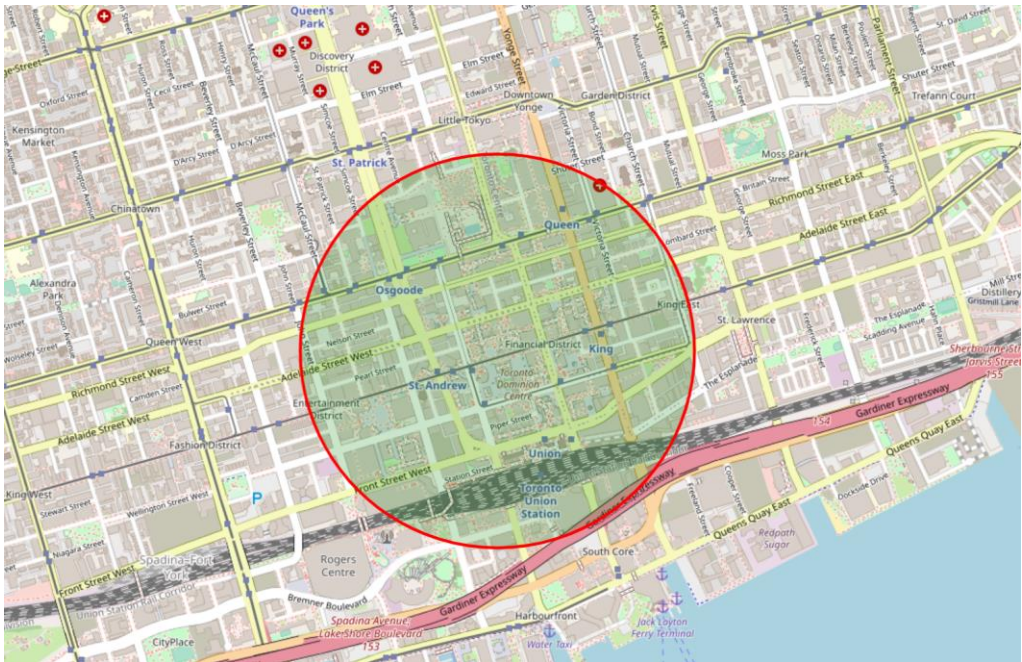
- We narrow down from 100 neighborhoods to around 30, based on the requirement of the client

	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
70	Etobicoke	0	All restaurants	All coffee	Jewelry Store	Opera House	Music Venue
79	Central Toronto	0	All restaurants	All coffee	Dessert Shop	Toy / Game Store	Bookstore
80	Downtown Toronto	0	All sports	All coffee	BBQ Joint	All restaurants	Trail
84	Downtown Toronto	0	All restaurants	All coffee	All bars	Arts & Crafts Store	Dessert Shop
90	Scarborough	0	All restaurants	Bank	Pharmacy	All sports	All coffee
92	Downtown Toronto Stn A	0	All restaurants	All bars	All coffee	All allergies	Hotel
93	Etobicoke	0	All restaurants	All coffee	Pharmacy	Wings Joint	Organic Grocery
96	Downtown Toronto	0	All restaurants	All coffee	All allergies	All bars	General Entertainment
97	Downtown Toronto	0	All restaurants	All coffee	All sports	Deli / Bodega	All bars
99	Downtown Toronto	0	All restaurants	Burrito Place	Martial Arts School	All allergies	All bars

Results

- The neighborhood with postal code “M5X” should be the suitable choice and is shown on the map

PostalCode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
97	MSX	Downtown Toronto First Canadian Place, Underground city	43.648429	-79.38228	0	All restaurants	All coffee	All sports	Deli / Bodega	All bars



Discussion and Conclusion

- ▶ The criteria for getting the end results in this project
 - ▶ Top common venue: the restaurants
 - ▶ 2nd and 3rd common venue: coffee shops and sports related facilities
 - ▶ Exclusion criteria: venues that could cause allergy
- ▶ The most suitable neighborhoods for David to live in Toronto
 - ▶ First Canadian Place or
 - ▶ Underground city, which are both inside Downtown Toronto.

Future work

- ▶ A more thorough investigation of the price of apartments located in these two neighborhoods should be conducted
- ▶ Further information required from the client
 - ▶ Size of the apartment
 - ▶ Room numbers
 - ▶ Interior design
 - ▶ And more...