Identifying the Best Locations in the City of Toronto for a New Restaurant

MASOUD AL NAYARI

Agenda

- Introduction
- Aim
- Datasets
- Data Analysis
- Clustering
- Conclusion

Introduction

- Planning is an essential part of any business.

- What business to start depends on many factors like experience, passion, and available capital

- Restaurant business is one of the successful business

  - Number of people eating outside is continuously growing

Introduction (cont.)

- Factors to be considered for starting a new restaurant:
  - understanding the restaurant business
  - good-quality food
  - good customer service
  - Location
  - staff satisfaction
  - Marketing
  - smart menu

Aim

- identify and recommend the best locations in the city of Toronto to start a new restaurant

- Toronto is Ontario's capital city, Canada's largest municipality and the fourth largest city in North America.

- The city is home to a large immigrant population, and is a national and international hub for finance, communications and cultural life.

- Due to its diversity in culture and tourism spots, and its large inhabitants, choosing just the right location for a restaurant is a tricky task.

Datasets

▶ The project uses Foursquare location data with two more datasets

▶ First Dataset:

    ▶ Wikipedia web page that contains table of Toronto neighbourhoods and their postal codes

    ▶ https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

▶ Second Dataset

    ▶ dataset that contains latitudes and longitudes of Toronto Neighbourhoods (.csv file provided by IBM)

Datasets (cont.)

▶ A map showing Toronto's neighbourhoods

Data Analysis

- Wikipedia Web page scraping to get a dataframe
- Merging the two datasets into one dataframe

	Postal Code	Borough	Neighbourhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494

Data Analysis (cont.)

▶ Getting venues in all neighbourhoods

	Neighbourhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Parkwoods	43.753259	-79.329656	Brookbanks Park	43.751976	-79.332140	Park
1	Parkwoods	43.753259	-79.329656	Variety Store	43.751974	-79.333114	Food & Drink Shop
2	Victoria Village	43.725882	-79.315572	Victoria Village Arena	43.723481	-79.315635	Hockey Arena
3	Victoria Village	43.725882	-79.315572	Portugril	43.725819	-79.312785	Portuguese Restaurant
4	Victoria Village	43.725882	-79.315572	Tim Hortons	43.725517	-79.313103	Coffee Shop

▶ Applying data analysis to get the final dataframe:

	Neighbourhood	Non Restaurant	Restaurant	Count
0	Agincourt	0.800000	0.200000	0.05
1	Alderwood, Long Branch	1.000000	0.000000	0.07
2	Bathurst Manor, Wilson Heights, Downsview North	0.809524	0.190476	0.21
3	Bayview Village	0.500000	0.500000	0.04
4	Bedford Park, Lawrence Manor East	0.590909	0.409091	0.22

Clustering

- The purpose of clustering is to group similar neighbourhoods together

- The used dataset has three features: Restaurant, Non-restaurant venues ratio, and venues count

- Four clusters
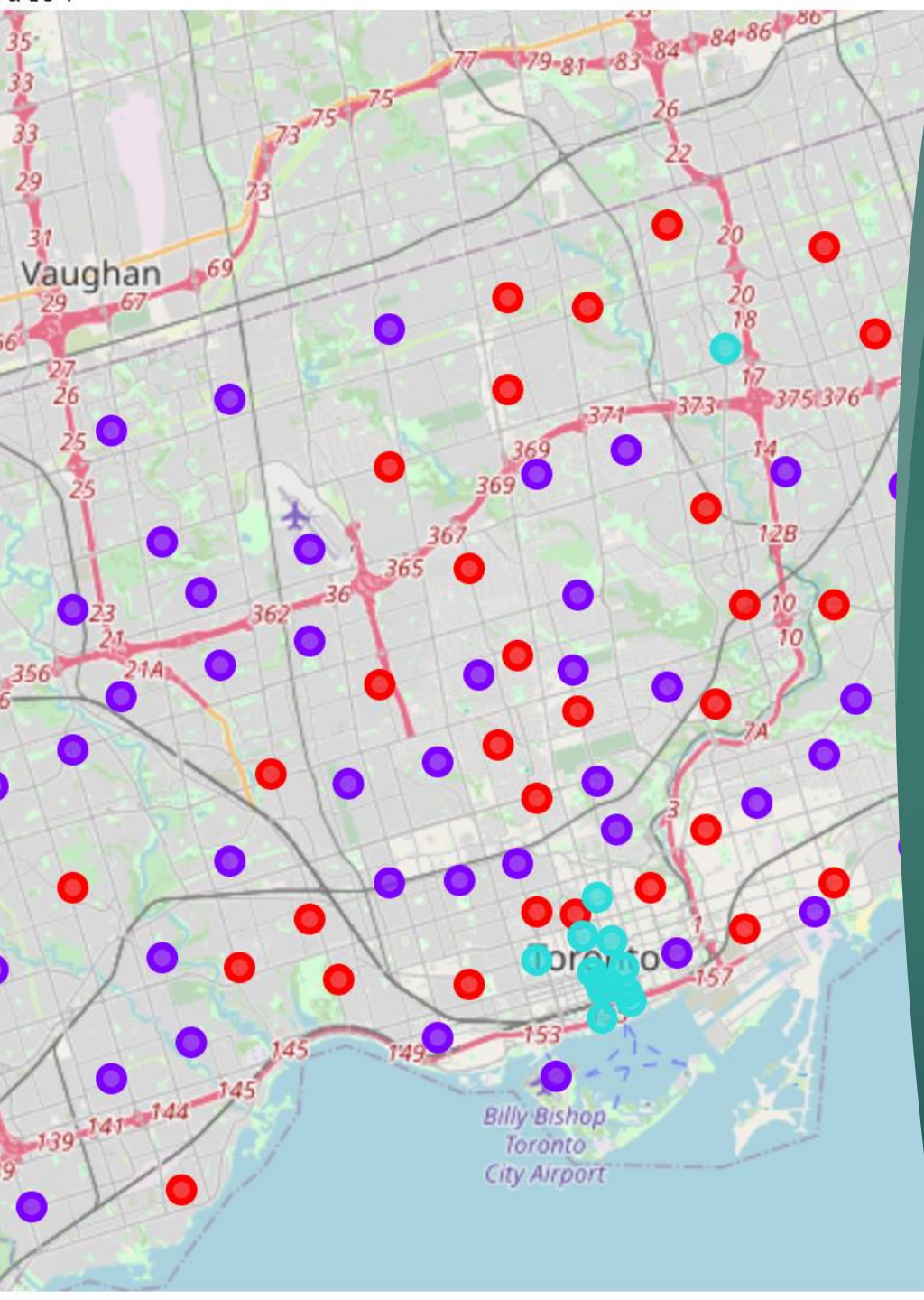
Clustering (cont.)

▶ Cluster 1

   ▶ the second largest group

   ▶ low number of venues

   ▶ moderate restaurants ratio

▶ Cluster 2

   ▶ Largest group

   ▶ low to moderate number of venues

   ▶ very low ratio of restaurants

(51, 8)

	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	Non Restaurant	Restaurant	Count
0	North York	Parkwoods	43.753259	-79.329656	1.0	1.000000	0.000000	0.02
2	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	1.0	0.931818	0.068182	0.44
3	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763	1.0	0.923077	0.076923	0.13
8	East York	Parkview Hill, Woodbine Gardens	43.706397	-79.309937	1.0	1.000000	0.000000	0.11
11	Etobicoke	West Deane Park, Princess Gardens, Martin Grov...	43.650943	-79.554724	1.0	1.000000	0.000000	0.01

	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	Non Restaurant	Restaurant	Count
1	North York	Victoria Village	43.725882	-79.315572	0.0	0.800000	0.200000	0.05
4	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	0.0	0.818182	0.181818	0.33
7	North York	Don Mills	43.745906	-79.352188	0.0	0.652174	0.347826	0.23
10	North York	Glencairn	43.709577	-79.445073	0.0	0.800000	0.200000	0.05
13	North York	Don Mills	43.725900	-79.340923	0.0	0.652174	0.347826	0.23

Clustering (cont.)

▶ Cluster 3

    ▶ **Best cluster**

    ▶ highest number of venues

    ▶ low restaurants ratio

▶ Cluster 4

    ▶ smallest group (2 members)

    ▶ Very low number of venues

    ▶ very high restaurants ratio

(13, 8)

	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	Non Restaurant	Restaurant	Count
9	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	2.0	0.790000	0.210000	1.00
15	Downtown Toronto	St. James Town	43.651494	-79.375418	2.0	0.717647	0.282353	0.85
20	Downtown Toronto	Berczy Park	43.644771	-79.373306	2.0	0.763636	0.236364	0.55
24	Downtown Toronto	Central Bay Street	43.657952	-79.387383	2.0	0.705882	0.294118	0.68
30	Downtown Toronto	Richmond, Adelaide, King	43.650571	-79.384568	2.0	0.770000	0.230000	1.00

(2, 8)

	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	Non Restaurant	Restaurant	Count
6	Scarborough	Malvern, Rouge	43.806686	-79.194353	3.0	0.0	1.0	0.01
65	Scarborough	Dorset Park, Wexford Heights, Scarborough Town...	43.757410	-79.273304	3.0	0.2	0.8	0.05

Conclusion

▶ The project successfully achieved the aim of identifying the best neighbourhood for starting a new restaurant.

▶ Cluster 3 has the best set of neighbourhoods to start a new restaurant and most of them are in a bustling borough (Downtown)

▶ **'Harbourfront East, Union Station, Toronto Islands'** neighbourhood is the ideal location for a new restaurant in Toronto.

   ▶ Buzzing area, more customers to feed

   ▶ Low number of running restaurants

Future Work

- Work can improved by including more features

- Venues ratings could be a good feature for better clustering

- Large number of highly-rated venues in a certain area makes it an ideal area for a new business

- Successful businesses draws more people

Thank You!