“THE BATTLE OF NEIGHBORHOODS” IBM Data science specialization capstone project

Preparer: T.L. Lau / Date: 2019-07-08

Report on Meat and Cooking Ingredient Shop Opportunities in North York, Ontario

**Table of Contents:**

* Introduction
* Data
* Methodology
* Result
* Discussion
* Conclusion
* References
* Acknowledgement
* Appendix

**Introduction:**

North York is an administrative division in Toronto, Ontario. It is located directly north of Old Toronto between Etobicoke to the west and Scarborough to the east. To take advantage of her high-density residences, rapid transit, and high multiculturality, Mr Nakamura is looking for opportunity to invest in North York. In particular, he is considering opening a meat and cooking ingredient shop, and the relevant warehouse in North York, whereby he will be importing from Japan and other Asian countries.

In this regard, Mr Nakamura has sought our advice in providing information regarding restaurants and food joints in North York, Ontario. In particular, he would like to know where the most concentrated areas for restaurants and food joints in North York are, i.e. where his potential customers will be mostly located.

In this report, we set out our analysis and recommendation to Mr Nakamura, as well as our basis in data science and machine learning methodology.

**Data:**

We have attempted to obtain geo-locational information (i.e. longitudes and latitudes) of North York, as well as of its neighborhoods in Ontario. In additional, we obtained data about restaurants and food joints in North York and its neighborhoods. With the above data, we performed segmenting and clustering using K-means algorithm and from such observation, formed our recommendations to Mr Nakamura.

To obtain geo-location information and data regarding restaurants and food joints in North York and its neighborhoods, the following sources were sought:

* <https://en.wikipedia.org>

Postal codes of 211 boroughs in Ontario, Canada were obtained, which were used to combine with the geo-locational information for further data preparation.

* <https://ibm.ent.box.com>

Geo-locational information, i.e. longitude and latitude, of Ontario, Canada were obtained. These geo-locational information were combined with their postal codes as mentioned above, and were filtered for North York.

* <https://www.foursquare.com>

Data restaurants and food joints information were obtained from FourSquare using their API, based on the geo-location information of North York. Both “basic” and “advanced” venue information from within 1,000km of the respective neighbourhood centers provided to the FourSquare API were sought, limiting to 500 resultant returns. These information include:

* + Venue
  + Venue Summary
  + Venue Category
  + Distance

**Methodology:**

The above information were formed into a database, together with a list of “venue category of interest”, which composed of all restaurants and food joints only. The “venue category of interest” has been agreed with Mr Nakamura (*Appendix A)*. The database was fed to the K-means algorithm for cluster ranking and generated the result in this report.

**Result:**

By utilizing the K-means clustering algorithm, the highest ranking cluster group based on the “venue category of interest” is the borough **Willowdale South** in North York.

**Discussion:**

By applying the K-means clustering algorithm, we had set the K size to 7, i.e. 7 clusters were generated from the 38 boroughs. These 38 boroughs composed of 8184 venues of 156 unique categories obtained from FourSquare. Based on the list of “venue category of interest” agreed with Mr Nakamura, we have trimmed the unique categories to 44, and fed such to the K-means clustering algorithm. The K-means algorithm returned a cluster result as in *Appendix B*.

Further data regarding the location of ports, traffic and road conditions, residential density, etc can be obtained to do a more thorough analysis and study.

**Conclusion:**

Based on the result, we would recommend Mr Nakamura to open the meat and cooking ingredient shop in Willowdale South, North York, Ontario.

**References:**

* [*https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of\_Canada:\_M*](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
* [*https://ibm.ent.box.com/shared/static/9afzr83pps4pwf2smjjcf1y5mvgb18rr.csv*](https://ibm.ent.box.com/shared/static/9afzr83pps4pwf2smjjcf1y5mvgb18rr.csv)
* [*https://api.foursquare.com*](https://api.foursquare.com)

**Acknowledgement:**

* **Coursera**
* **IBM**

*Appendix A*

*List of Venue Category of Interest*

1. Caribbean Restaurant
2. Café
3. Fast Food Restaurant
4. Fish & Chips Shop
5. Food & Drink Shop
6. Pizza Place
7. Chinese Restaurant
8. Portuguese Restaurant
9. Vietnamese Restaurant
10. Sushi Restaurant
11. Burger Joint
12. Fried Chicken Joint
13. Greek Restaurant
14. Restaurant
15. Seafood Restaurant
16. Sandwich Place
17. Korean Restaurant
18. Italian Restaurant
19. Japanese Restaurant
20. Asian Restaurant
21. Salad Place
22. Thai Restaurant
23. Diner
24. Breakfast Spot
25. Cafeteria
26. Latin American Restaurant
27. Mediterranean Restaurant
28. Middle Eastern Restaurant
29. Dim Sum Restaurant
30. American Restaurant
31. Falafel Restaurant
32. Turkish Restaurant
33. Empanada Restaurant
34. Hot Dog Joint
35. Indian Restaurant
36. Ramen Restaurant
37. Comfort Food Restaurant
38. Wings Joint
39. Steakhouse
40. Indonesian Restaurant
41. Burrito Place
42. Snack Place
43. French Restaurant
44. Eastern European Restaurant

*Appendix B*

*K-Means Clustering Result*

|  |  |  |
| --- | --- | --- |
| **Cluster**  **Ranking** | **Number of**  **Boroughs** | **Name of**  **Boroughs in North York, Ontario** |
| 1 | 2 | Willowdale South |
| 2 | 6 | Bedford Park  Don Mills South  Flemingdon Park  Lawrence Heights  Lawrence Manor  Lawrence Manor East |
| 3 | 6 | Don Mills North  Fairvie  Henry Farm  Newtonbrook  Oriol  Willowdale |
| 4 | 3 | Glencair  Northwood Park  York University |
| 5 | 6 | Bathurst Manor  CFB Toronto  Downsview East  Downsview North  Downsview Northwest  Wilson Heights |
| 6 | 9 | Bayview Village  Downsview  Downsview Central  Hillcrest Village  Humber Summit  North Par  Parkwoods  Upwood Par  York Mills West |
| 7 | 7 | Downsview West  Emery  Humberlea  Silver Hill  Victoria Village  Willowdale West  York Mills |
|  | **38** |  |