

# **CAPSTONE PROJECT – BATTLE OF NEIGHBORHOODS**

**Detecting promising neighborhoods for  
the next boba storefront in Toronto**

Jiaqi Chen  
November 4, 2019

# MAKING EFFECTIVE DECISION ON THE LOCATION OF NEW STOREFRONTS IS CRUTIAL

- Growing boba market

The global bubble tea market size was valued at USD 1.89 Billion in 2018 and is projected to reach USD 3.49 Billion by the end of 2026.<sup>[1]</sup>

- Low cost of storefronts

- simple equipment
- low labour cost

- Fierce market competition

In 2019, at least 20 bubble tea brands have emerged on the Toronto Market

| Neighbourhood          |    |
|------------------------|----|
| Venue                  |    |
| Real Fruit Bubble Tea  | 20 |
| Tea Shop 168           | 15 |
| Sharetea               | 12 |
| Chatime                | 11 |
| Presotea               | 10 |
| CoCo Fresh Tea & Juice | 9  |
| Chatime Atealier       | 9  |
| Chatime 日出茶太           | 9  |
| The Alley              | 9  |
| ZenQ                   | 8  |

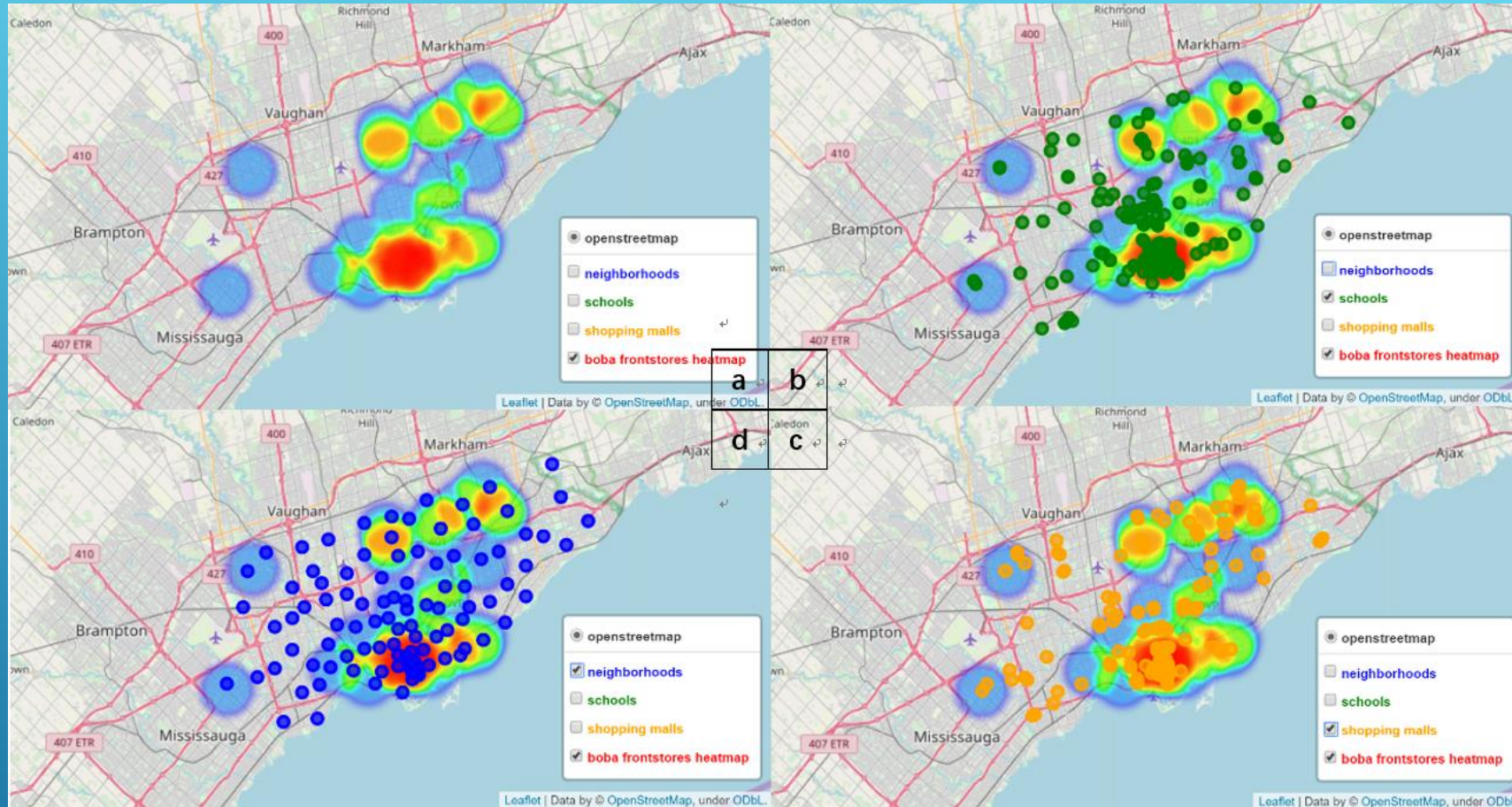
[1] <https://www.fortunebusinessinsights.com/industry-reports/bubble-tea-market-101564>

# DATA ACQUISITION AND CLEANSING

1. names of Toronto neighborhoods will be obtained from **Wikipedia** by using **BeautifulSoup**
2. coordinate of neighborhoods will be obtained from **cousera** in case of the disconnection of **geocoder**
3. number and coordinate of existing boba storefronts will be obtained using **Foursquare API**
4. number and coordinate of universities and colleges and their location in every neighborhood will be obtained using **Foursquare API**
5. number and coordinate of shopping places and their location in every neighborhood will be obtained using **Foursquare API**

*\* The exploring radius of above venues of each neighborhood is set to be 1km.*

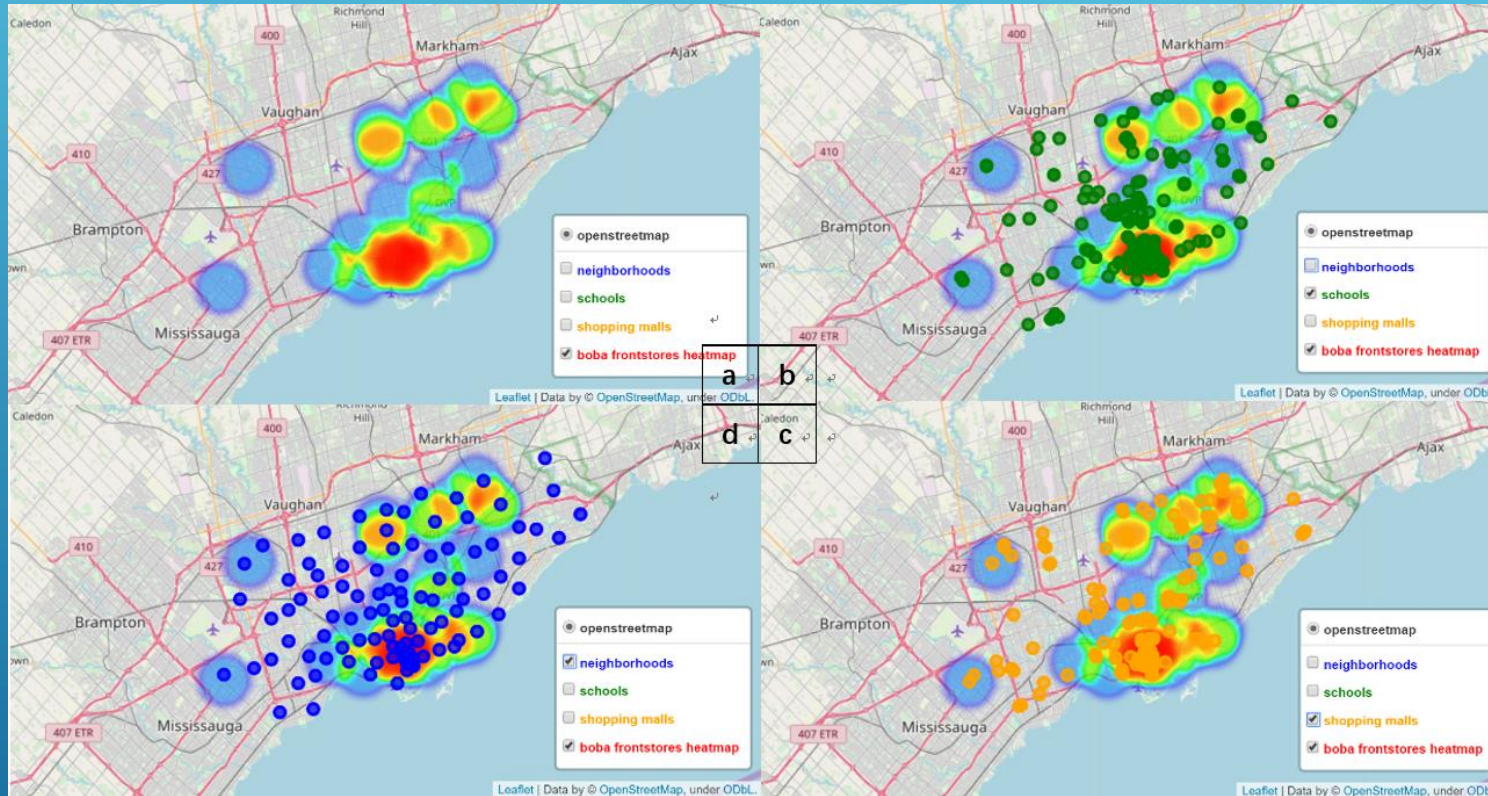
# DATA VISUALIZATION ON MAP



- a. heatmap of boba storefronts
- b. heatmap of boba storefronts and distribution of schools
- c. heatmap of boba storefronts and distribution of neighborhoods
- d. heatmap of boba storefronts and distribution of shopping places.



# DATA VISUALIZATION ON MAP

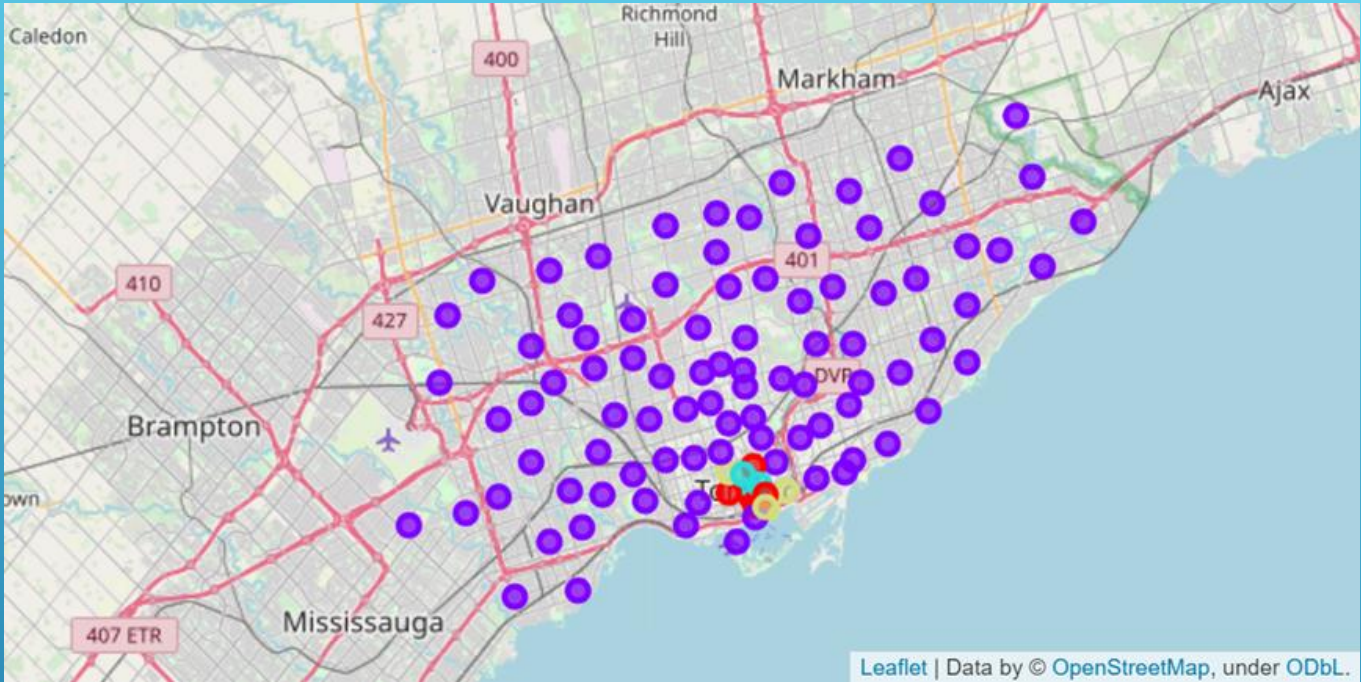


1. The distribution of bubble tea shops, schools and shopping malls are highly correlated.

2. Bubble tea shops mainly focus on old Toronto and North York.

3. Boba market competition is milder in the west and the northeast of Toronto than old Toronto.

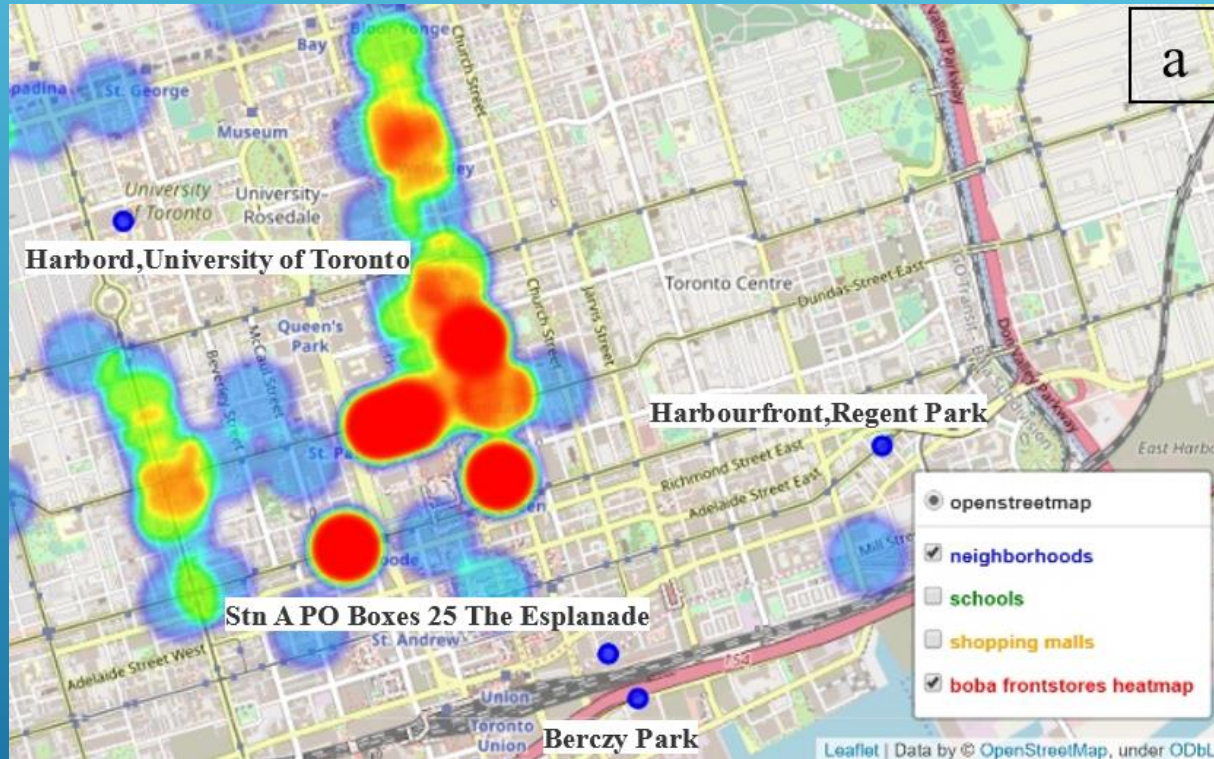
# DATA VISUALIZATION ON MAP



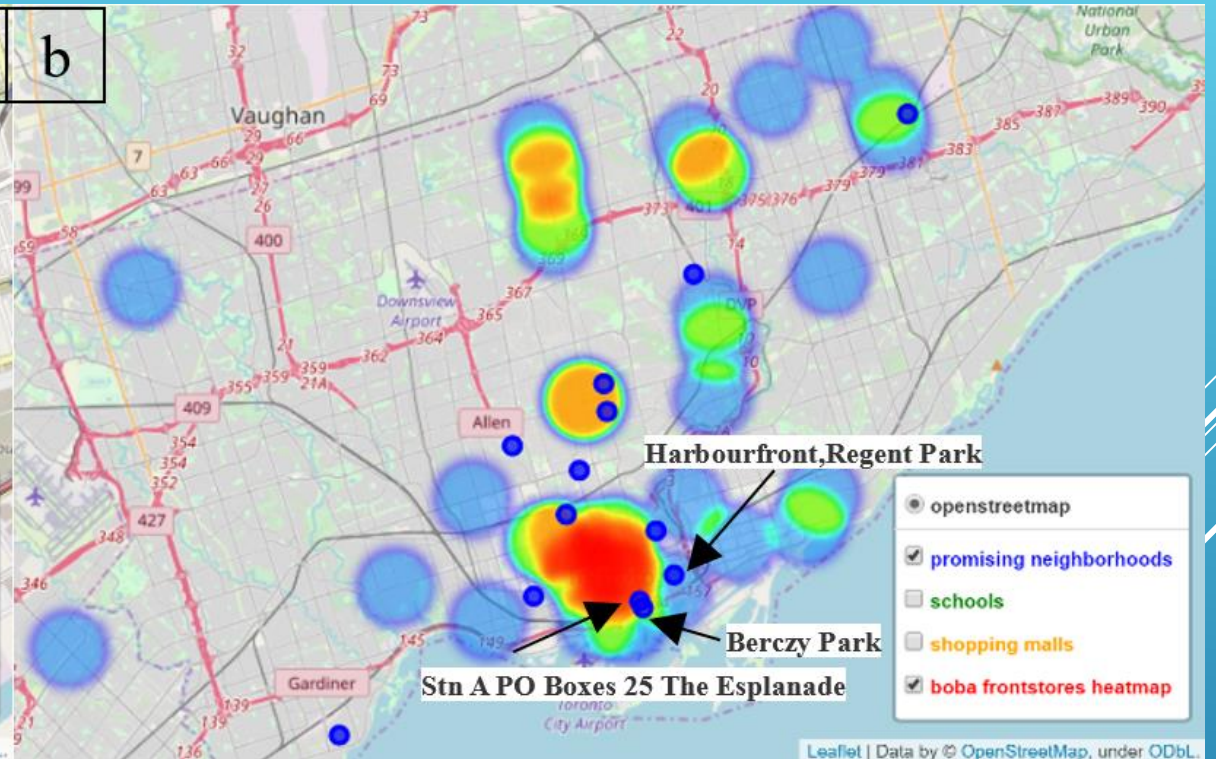
| Color - Cluster Number | Number of neighborhoods | Mean number of boba shops | Mean number of schools | Mean number of shopping places |
|------------------------|-------------------------|---------------------------|------------------------|--------------------------------|
| cluster0               | 7                       | 20.43                     | 24.14                  | 15.58                          |
| cluster1               | 88                      | 0.81                      | 1.55                   | 1.37                           |
| cluster2               | 3                       | 36.0                      | 35.67                  | 14.00                          |
| cluster3               | 4                       | 4.00                      | 19.75                  | 5.00                           |



# CONTRAST BETWEEN K-MEANS AND MANUAL SELECTION



a. map of cluster3



b. map of manual selection

Berczy Park, Harbourfront, Regent Park, Stn A PO Boxes 25 The Esplanade are shared by both approaches, and they all situate in the south of old Toronto near the seashore.

# CONTRAST BETWEEN K-MEANS AND MANUAL SELECTION

neighborhoods in cluster3

| Cluster Labels |   | Neighbourhood                   |
|----------------|---|---------------------------------|
| 0              | 3 | Berczy Park                     |
| 1              | 3 | Harbord,University of Toronto   |
| 2              | 3 | Harbourfront,Regent Park        |
| 3              | 3 | Stn A PO Boxes 25 The Esplanade |

neighborhoods in manual selection

| Cluster Labels |   | Neighbourhood                                     |
|----------------|---|---|
| 0              | 3 | Stn A PO Boxes 25 The Esplanade                   |
| 1              | 3 | Harbourfront,Regent Park                          |
| 2              | 3 | Berczy Park                                       |
| 3              | 1 | The Annex,North Midtown,Yorkville                 |
| 4              | 1 | Davisville  |
| 5              | 1 | Agincourt   |
| 6              | 1 | Davisville North                                  |
| 7              | 1 | Don Mills North                                   |
| 8              | 1 | Deer Park,Forest Hill SE,Rathnelly,South Hill,... |
| 9              | 1 | Little Portugal,Trinity                           |
| 10             | 1 | Cabbagetown,St. James Town                        |
| 11             | 1 | Humber Bay Shores,Mimico South,New Toronto        |
| 12             | 1 | Humewood-Cedarvale                                |



# CONCLUSIONS AND FUTURE DIRECTIONS

## CONCLUSIONS

- Used k-means to detect promising neighborhoods in Toronto for new storefronts
- Inspected the result of k-means by contrast with manual filtration

## FUTURE DIRECTIONS

***The ultimate goal of a boba shop owners is reducing cost and increasing profitability. There is improving room for this project***

- More accurate division of neighborhood boundary
- Consideration on more factors
  - rental costs
  - local consumption level
  - transportation
  - business hours