# Finding a Suitable Location to Open a Pastry Shop in Toronto

Reginald P. Chua

January 25, 2021

## 1. Introduction

Ontario has the strongest economy in Canada, contributing the highest GDP per capita in the country. Toronto, its capital, has one of the top food scenes in Ontario. As aspiring pastry chefs based in Toronto, we want to open our first Pastry shop here. Which neighborhoods in Toronto can we consider?

Our Pastry shop concept focuses on products that go well with either coffee or tea. While there will be a small area reserved for diners, we are considering take-aways to drive most purchases. The product selection, for now, will focus on pastry products only. As we are entering this industry and market, we want to take the opportunity to position ourselves in neighborhoods with coffee shops and tea houses that serve products which complement our pastries so that we can more easily tap our market. This can also set the foundation to expand our reach later.

## 2. Data Overview and Limitations

The analysis uses data gathered from and limited to the following sources:

- **Wikipedia**
  - An existing Wikipedia page contains information on postal codes, boroughs, and neighborhoods in Toronto. We will use web scraping to utilize the data.
  - https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

- **Geospatial Data**
  - A csv file on the web contains the coordinates of each postal code in Toronto, as provided by the IBM Data Science Capstone project.
  - https://cocl.us/Geospatial_data

- **Geopy library**
  - Neighborhood coordinates are supplied through this library.

- **FourSquare**
  - This API grants access to data about establishments within a neighborhood. We utilize the data on the categories of establishments in each neighborhood, focusing on competitor and complementary industries including the following establishments: Bakery, Café, Coffee Shop, Donut Shop, Pastry Shop, Tea Room

## 3. Target Audience

The information here may be of interest to the following groups of people, while being mindful of the context and data sources stated above:

a. Entrepreneurs in search of a brief analysis involving the pastry shop industry in Toronto.
b. Data Science students and learners
c. Junior Data Analysts or Junior Data Science Professionals

## 4. Methodology and Procedure

With the data described above, this study uses K-Means Clustering to group similar neighborhoods into clusters to help identify potential locations to open our first pastry shop. We carry out the following procedures:

1. Data acquisition and staging
    - Scrape data from Wikipedia. This contains data based on postal codes within Toronto.
    - Obtain longitudes and latitudes for each postal code using geospatial data.
    - Refine the data from postal code level down to neighborhood level.
    - Obtain the coordinates per neighborhood.
    - Blend FourSquare data into each neighborhood to have a look into the types of establishments within each neighborhood.

2. Data transformation
    - Perform one-hot encoding in preparation for clustering.
    - Calculate for the means for each establishment type in each neighborhood.

3. Cluster and examine
    - Use the K Elbow Visualizer to identify the best k to use for clustering.
    - Perform K-Means Clustering.
    - Study each cluster with the aid of visualization and by examining the data points.

## 5. Data Exploration and Analysis

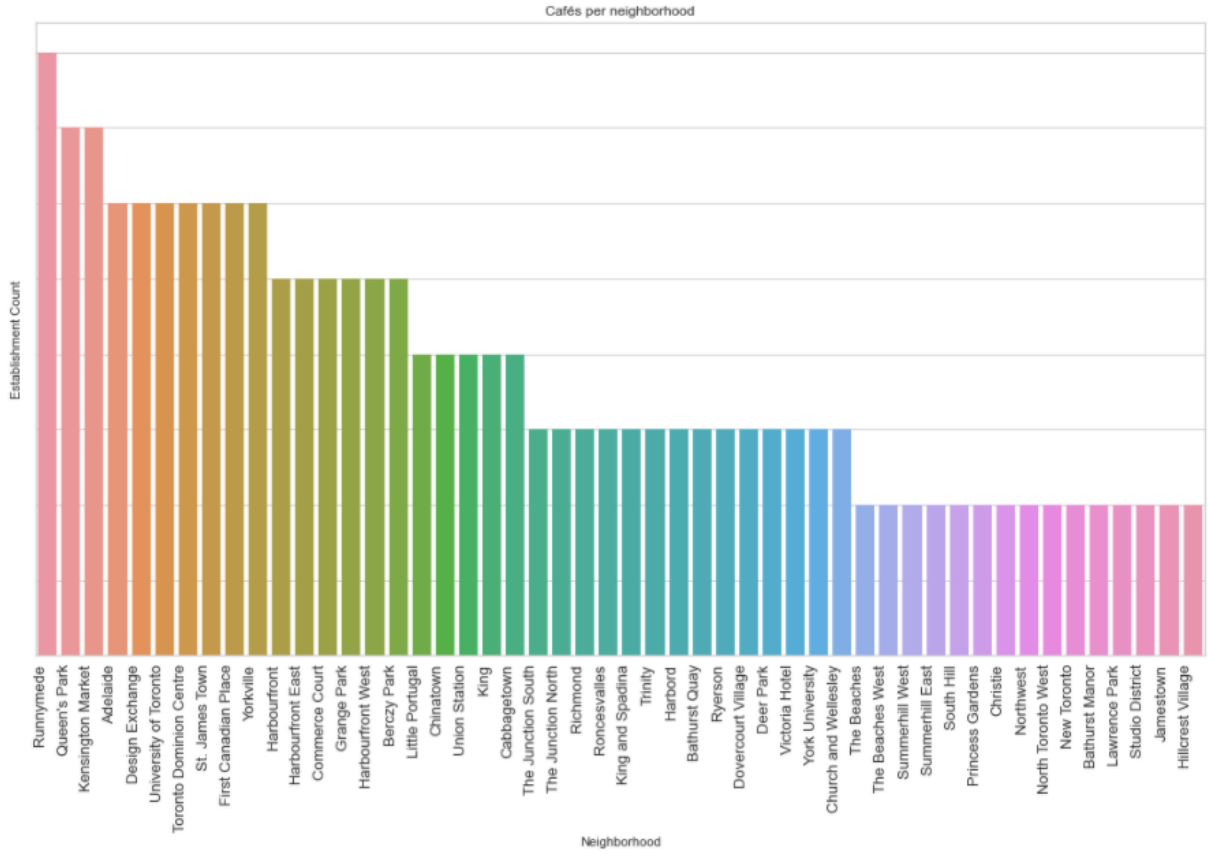
Inspecting the clusters (0-4), as we are looking to enter areas packed with Coffee Shops, Cafés, or Tea Rooms, as they complement our products, and with minimal competition, Cluster 3 is of interest.

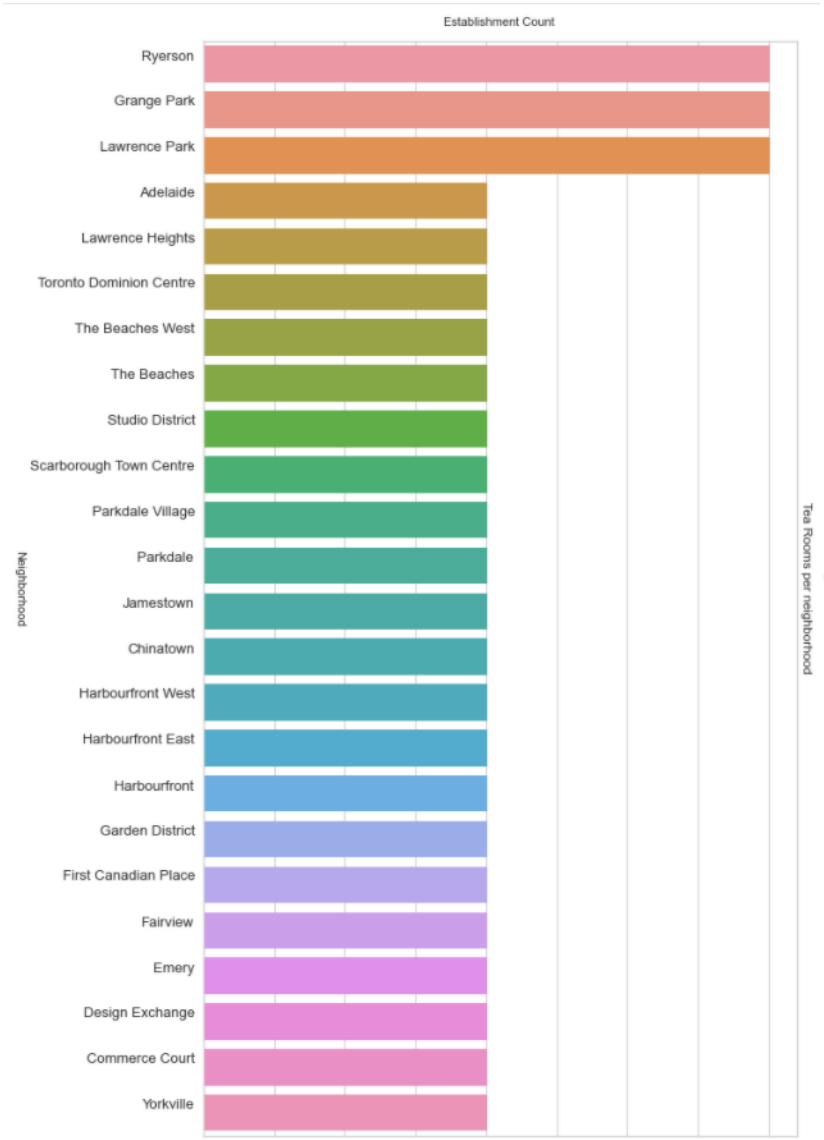
| | | Neighborhood | Venue |
|---|---|---|---|
| Cluster Labels | Venue Category | | |
| 0 | Bakery | 45 | 45 |
| | Café | 54 | 82 |
| | Coffee Shop | 74 | 60 |
| | Donut Shop | 5 | 4 |
| | Pastry Shop | 1 | 1 |
| | Tea Room | 13 | 9 |
| 1 | Bakery | 7 | 7 |
| | Café | 4 | 3 |
| | Coffee Shop | 9 | 3 |
| | Donut Shop | 3 | 2 |
| | Pastry Shop | 1 | 1 |
| | Tea Room | 4 | 2 |
| 2 | Bakery | 8 | 7 |
| | Coffee Shop | 1 | 1 |
| 3 | Bakery | 11 | 5 |
| | Café | 15 | 29 |
| | Coffee Shop | 36 | 25 |
| | Donut Shop | 1 | 1 |
| | Tea Room | 7 | 3 |
| 4 | Bakery | 1 | 1 |
| | Café | 4 | 11 |
| | Coffee Shop | 1 | 1 |

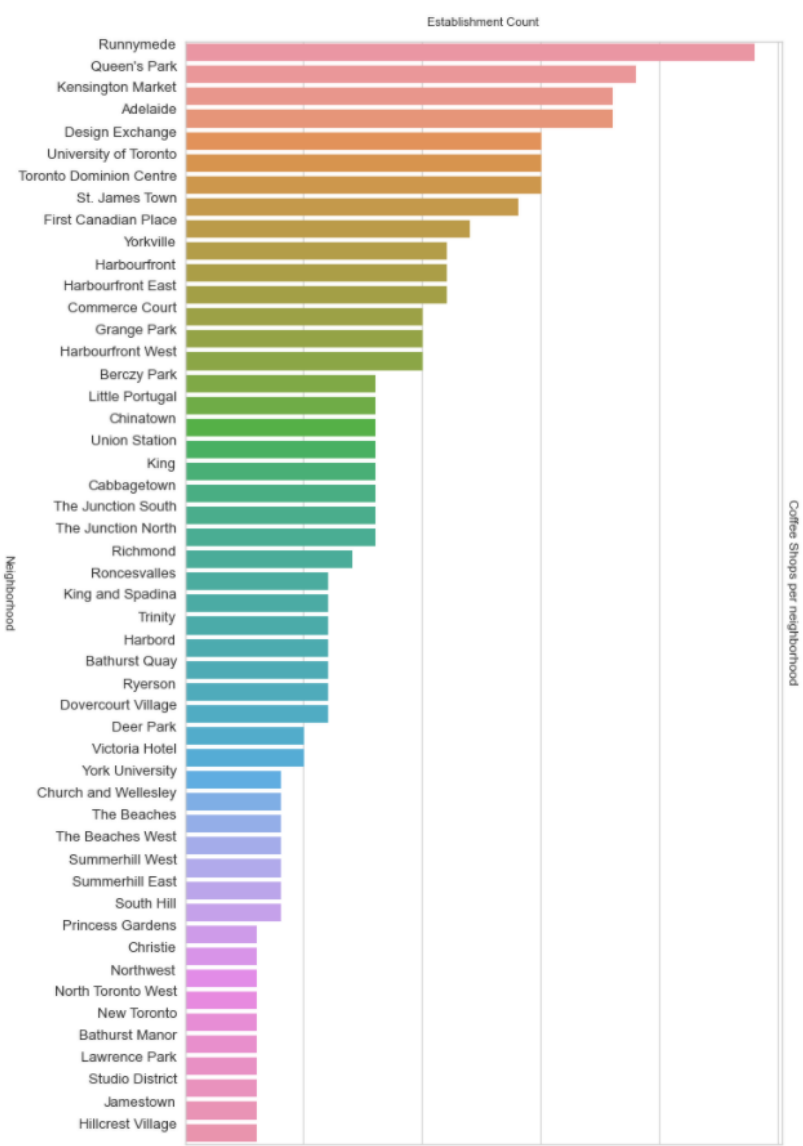
Within Cluster 3, we would like to prioritize neighborhoods that are densely populated with coffee shops, and with minimal competition: Oriole, Queen's Park, East Toronto, Mount Dennis, North Park, St Phillips, Martin Grove Gardens, Scarborough Village, Scarborough Village West, The Queensway East, Jamestown.

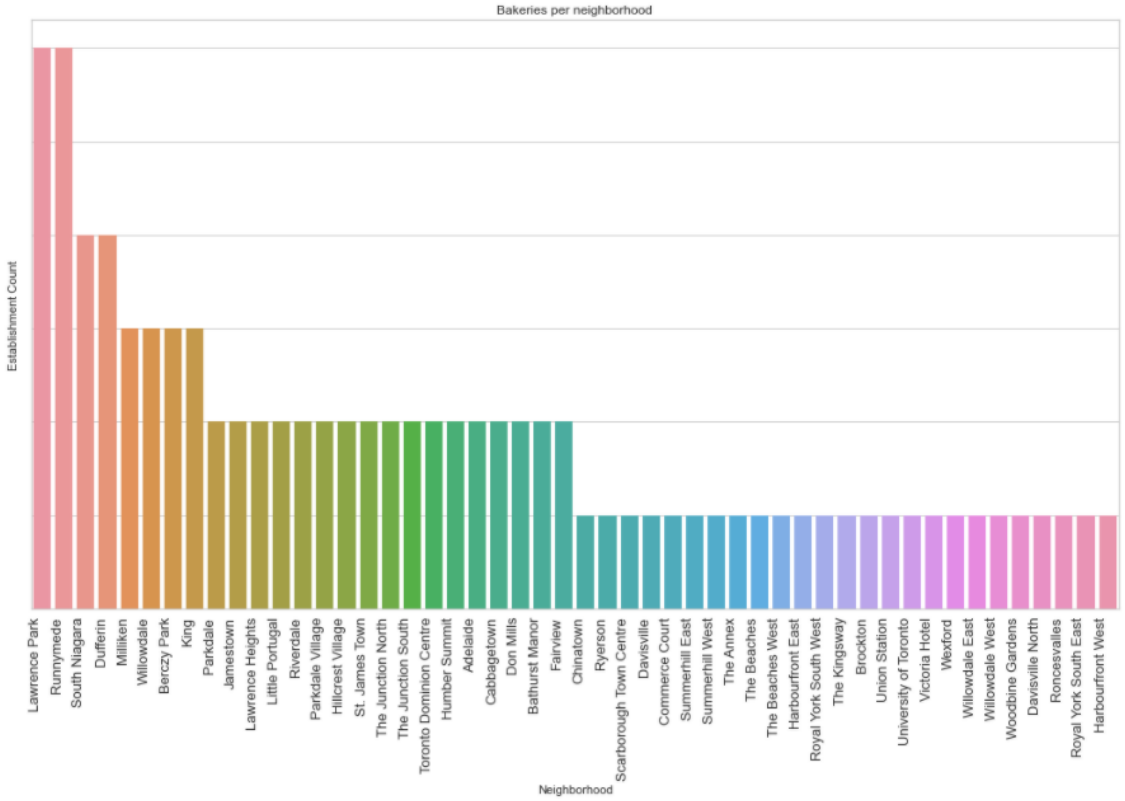
| | Neighbourhood | Cluster Labels | Bakery | Cafe | Coffee Shop | Donut Shop | Pastry Shop | Tea Room |
|---|---|---|---|---|---|---|---|---|
| 2 | Regent Park | 3 | 0.000000 | 0.000000 | 0.136364 | 0.000000 | 0.0 | 0.000000 |
| 3 | Harbourfront | 3 | 0.010000 | 0.050000 | 0.150000 | 0.000000 | 0.0 | 0.010000 |
| 6 | Queen's Park | 3 | 0.000000 | 0.092105 | 0.250000 | 0.013158 | 0.0 | 0.000000 |
| 31 | Morningside | 3 | 0.000000 | 0.083333 | 0.166667 | 0.000000 | 0.0 | 0.000000 |
| 37 | Central Bay Street | 3 | 0.000000 | 0.013699 | 0.136986 | 0.000000 | 0.0 | 0.000000 |
| 50 | Scarborough Village | 3 | 0.000000 | 0.000000 | 0.222222 | 0.000000 | 0.0 | 0.000000 |
| 51 | Fairview | 3 | 0.020408 | 0.020408 | 0.183673 | 0.000000 | 0.0 | 0.010204 |
| 53 | Oriole | 3 | 0.000000 | 0.000000 | 0.428571 | 0.000000 | 0.0 | 0.000000 |
| 55 | York University | 3 | 0.000000 | 0.142857 | 0.190476 | 0.000000 | 0.0 | 0.000000 |
| 56 | East Toronto | 3 | 0.000000 | 0.000000 | 0.250000 | 0.000000 | 0.0 | 0.000000 |
| 58 | Harbourfront East | 3 | 0.010000 | 0.050000 | 0.150000 | 0.000000 | 0.0 | 0.010000 |
| 59 | Union Station | 3 | 0.018182 | 0.072727 | 0.145455 | 0.000000 | 0.0 | 0.000000 |
| 71 | Design Exchange | 3 | 0.010000 | 0.060000 | 0.140000 | 0.000000 | 0.0 | 0.010000 |
| 78 | York Mills | 3 | 0.000000 | 0.000000 | 0.157895 | 0.000000 | 0.0 | 0.000000 |
| 83 | Victoria Hotel | 3 | 0.026316 | 0.078947 | 0.157895 | 0.000000 | 0.0 | 0.000000 |
| 84 | North Park | 3 | 0.000000 | 0.000000 | 0.250000 | 0.000000 | 0.0 | 0.000000 |
| 88 | Cliffside | 3 | 0.000000 | 0.000000 | 0.166667 | 0.000000 | 0.0 | 0.000000 |
| 90 | Scarborough Village West | 3 | 0.000000 | 0.000000 | 0.222222 | 0.000000 | 0.0 | 0.000000 |
| 91 | Willowdale | 3 | 0.020833 | 0.000000 | 0.166667 | 0.000000 | 0.0 | 0.000000 |
| 94 | Studio District | 3 | 0.000000 | 0.031746 | 0.174603 | 0.000000 | 0.0 | 0.015873 |
| 97 | Mount Dennis | 3 | 0.000000 | 0.000000 | 0.250000 | 0.000000 | 0.0 | 0.000000 |
| 101 | Cliffside West | 3 | 0.000000 | 0.000000 | 0.166667 | 0.000000 | 0.0 | 0.000000 |
| 102 | Willowdale | 3 | 0.020833 | 0.000000 | 0.166667 | 0.000000 | 0.0 | 0.000000 |
| 103 | Willowdale East | 3 | 0.020833 | 0.000000 | 0.166667 | 0.000000 | 0.0 | 0.000000 |
| 106 | Roselawn | 3 | 0.000000 | 0.076923 | 0.153846 | 0.000000 | 0.0 | 0.000000 |
| 113 | York Mills West | 3 | 0.000000 | 0.000000 | 0.157895 | 0.000000 | 0.0 | 0.000000 |
| 122 | Willowdale | 3 | 0.020833 | 0.000000 | 0.166667 | 0.000000 | 0.0 | 0.000000 |
| 123 | Willowdale West | 3 | 0.020833 | 0.000000 | 0.166667 | 0.000000 | 0.0 | 0.000000 |
| 131 | St. Phillips | 3 | 0.000000 | 0.000000 | 0.250000 | 0.000000 | 0.0 | 0.000000 |
| 132 | Martin Grove Gardens | 3 | 0.000000 | 0.000000 | 0.250000 | 0.000000 | 0.0 | 0.000000 |
| 151 | L'Amoreaux East | 3 | 0.000000 | 0.000000 | 0.166667 | 0.000000 | 0.0 | 0.000000 |
| 159 | Harbourfront West | 3 | 0.010000 | 0.050000 | 0.150000 | 0.000000 | 0.0 | 0.010000 |
| 160 | Bathurst Quay | 3 | 0.000000 | 0.120000 | 0.160000 | 0.000000 | 0.0 | 0.000000 |
| 166 | Jamestown | 3 | 0.020408 | 0.020408 | 0.183673 | 0.000000 | 0.0 | 0.010204 |
| 169 | L'Amoreaux West | 3 | 0.000000 | 0.000000 | 0.166667 | 0.000000 | 0.0 | 0.000000 |
| 171 | Alderwood | 3 | 0.000000 | 0.000000 | 0.142857 | 0.000000 | 0.0 | 0.000000 |
| 180 | Montgomery Road | 3 | 0.000000 | 0.000000 | 0.142857 | 0.000000 | 0.0 | 0.000000 |
| 188 | The Queensway East | 3 | 0.000000 | 0.000000 | 0.200000 | 0.000000 | 0.0 | 0.000000 |

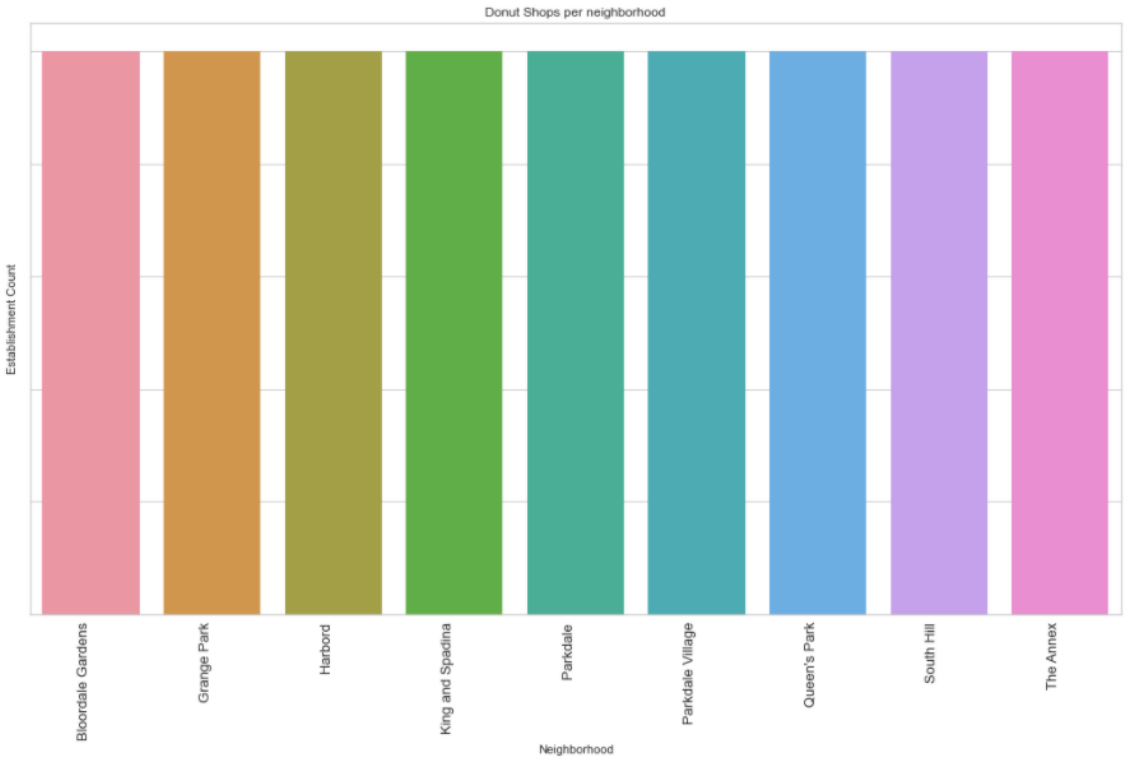
In the following graphs, we note that Runnymede (Cluster 0) has a high presence of Cafés, Coffee Shops, but also Bakeries.

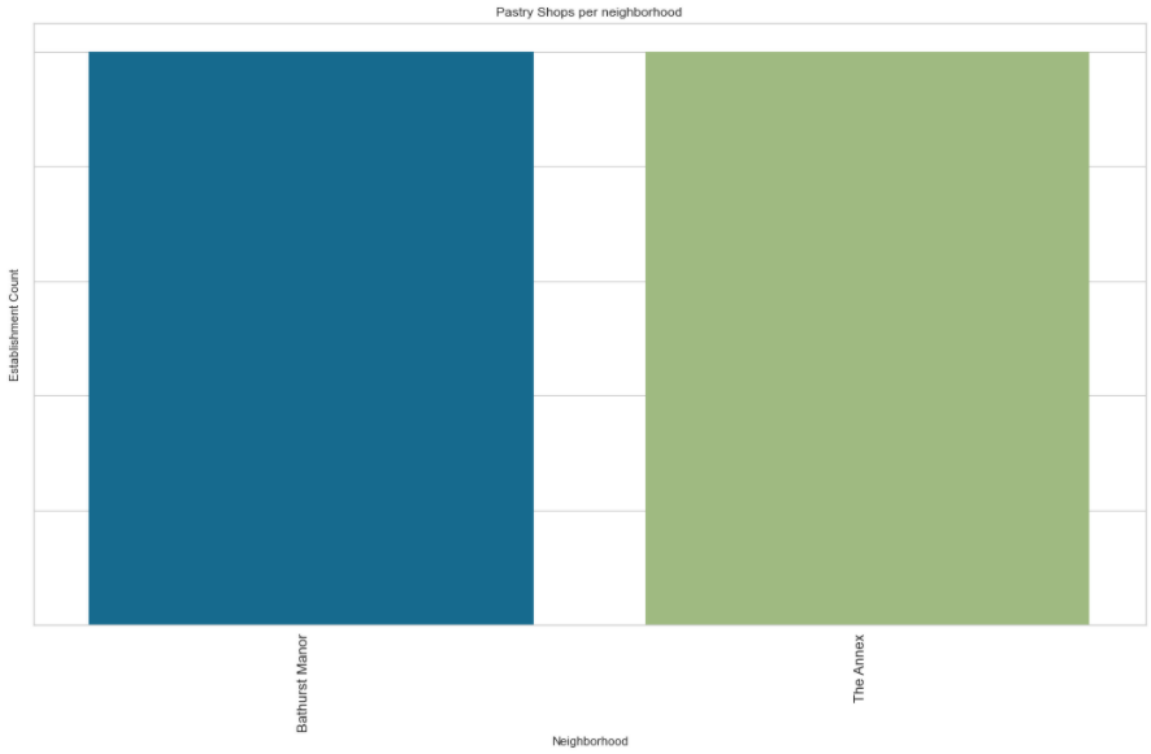



Establishment Count
Ryerson
Grange Park
Lawrence Park
Adelaide
Lawrence Heights
Toronto Dominion Centre
The Beaches West
The Beaches
Studio District
Scarborough Town Centre
Parkdale Village
Parkdale
Jamestown
Chinatown
Harbourfront West
Harbourfront East
Harbourfront
Garden District
First Canadian Place
Fairview
Emery
Design Exchange
Commerce Court
Yorkville
Neighborhood
Tea Rooms per neighborhood


Establishment Count
Runnymede
Queen's Park
Kensington Market
Adelaide
Design Exchange
University of Toronto
Toronto Dominion Centre
St. James Town
First Canadian Place
Yorkville
Harbourfront
Harbourfront East
Commerce Court
Grange Park
Harbourfront West
Berczy Park
Little Portugal
Chinatown
Union Station
King
Cabbagetown
The Junction South
The Junction North
Richmond
Roncesvalles
King and Spadina
Trinity
Harbord
Bathurst Quay
Ryerson
Dovercourt Village
Deer Park
Victoria Hotel
York University
Church and Wellesley
The Beaches
The Beaches West
Summerhill West
Summerhill East
South Hill
Princess Gardens
Christie
Northwest
North Toronto West
New Toronto
Bathurst Manor
Lawrence Park
Studio District
Jamestown
Hillcrest Village
Neighborhood
Coffee Shops per neighborhood


Bakeries per neighborhood
Establishment Count
Lawrence Park
Runnymede
South Niagara
Dufferin
Milliken
Willowdale
Berczy Park
King
Parkdale
Jamestown
Lawrence Heights
Little Portugal
Riverdale
Parkdale Village
Hillcrest Village
St. James Town
The Junction North
The Junction South
Toronto Dominion Centre
Humber Summit
Adelaide
Cabbagetown
Don Mills
Bathurst Manor
Fairview
Chinatown
Ryerson
Scarborough Town Centre
Davisville
Commerce Court
Summerhill East
Summerhill West
The Annex
The Beaches
The Beaches West
Harbourfront East
Royal York South West
The Kingsway
Brockton
Union Station
University of Toronto
Victoria Hotel
Wexford
Willowdale East
Willowdale West
Woodbine Gardens
Davisville North
Roncesvalles
Royal York South East
Harbourfront West
Neighborhood


Donut Shops per neighborhood
Establishment Count
Bloordale Gardens
Grange Park
Harbord
King and Spadina
Parkdale
Parkdale Village
Queen's Park
South Hill
The Annex
Neighborhood




Lastly, we note that the volume of donut shops and pastry shops is much lower than the other types of establishments in this study.

## 6. Recommendation

We are looking to enter areas packed with Coffee Shops, Cafés, or Tea Rooms, as they complement our products, where competition is minimal, as our brand is a new entrant in the market. Given this backdrop, Cluster 3 is of interest.

Within Cluster 3, we would like to prioritize neighborhoods that are densely populated with coffee shops and pose little competition. The neighborhoods are: Oriole, Queen's Park, East Toronto, Mount Dennis, North Park, St Phillips, Martin Grove Gardens, Scarborough Village, Scarborough Village West, The Queensway East, Jamestown.

We note that Runnymede (Cluster 0) has a high presence of Cafés, Coffee Shops, but also Bakeries. While this neighborhood is of a lesser priority as of this analysis, one would be prudent to familiarize one’s self with the scene at this neighborhood for two reasons.

First, competition may come from establishments here seeking to expand elsewhere. As such, we should take steps to prevent or challenge such moves, should they choose to expand to our locations of interest.

Second, it would be prudent to earmark this location for further expansion later when the brand is more established and able to operate in a more competitive environment.

Moreover, as there is a notably low volume of donut shops and pastry shops, additional confirmation is required, for example via a visual inspection of the recommended neighborhoods, to validate the situation as the data suggests.

For further analysis, as the accuracy of the models has room for improvement, consider augmenting the study with more data from other sources, such as store ratings, a richer list of establishments and locations, foot traffic or population density.