

ANYI FU

EXPLORE NEIGHBOURHOODS IN DOWNTOWN TORONTO

Creating an area map and adding neighbours as marks on it is helpful to see the distribution of the local business and/or other buildings, and segmenting and classifying those neighbourhoods are useful for analyzing local business as well.

Target: people who want to open a new restaurant in downtown Toronto area

Abstract & Problem

DATA SOURCES

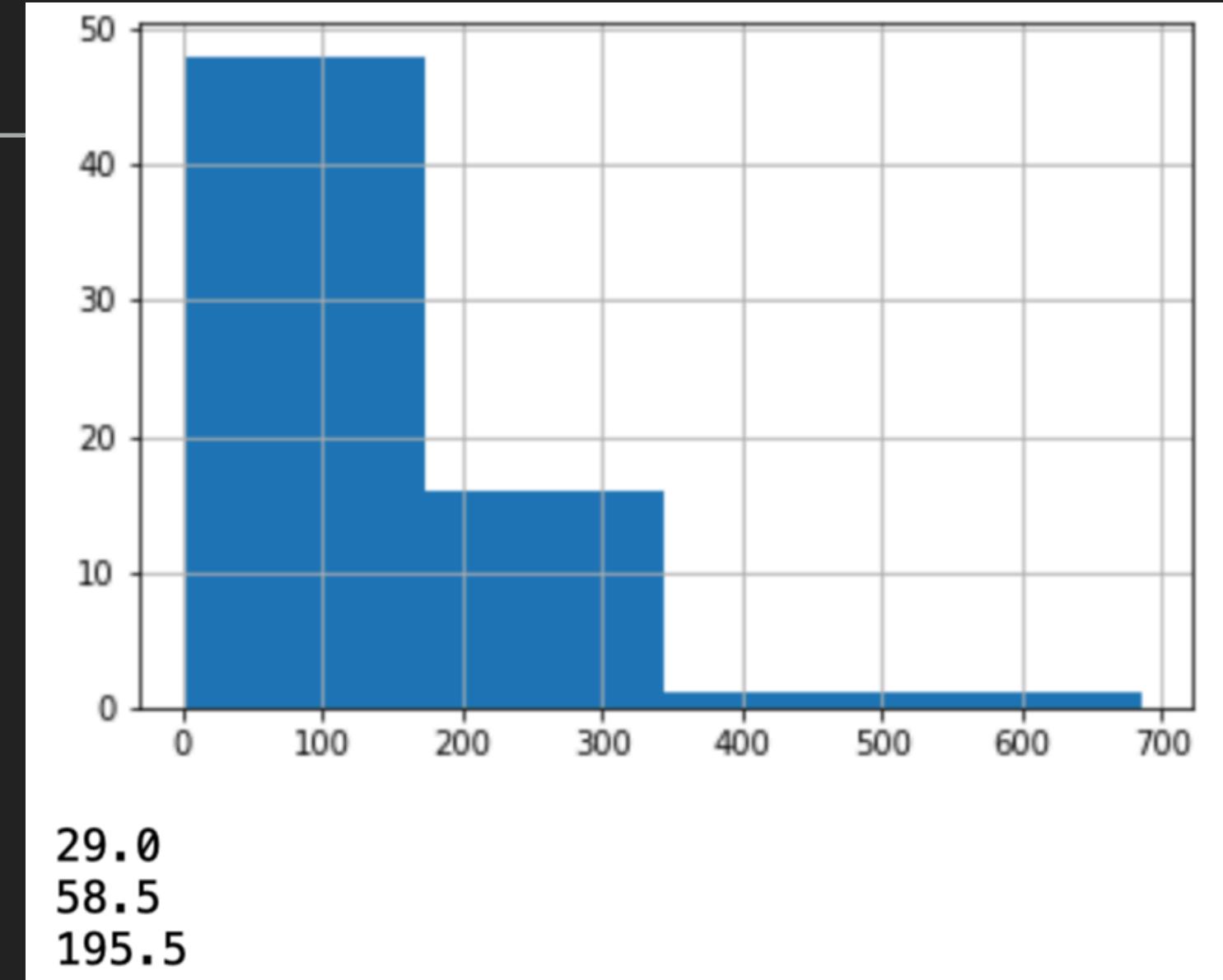
- ▶ Postal codes, borough, and neighbourhoods information from Wikipedia
["https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M"](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
- ▶ CSV file with latitude and longitude values
["https://cocl.us/Geospatial_data"](https://cocl.us/Geospatial_data)
- ▶ Venue name, ID, location, category, and count of likes form Foursquare API
["https://developer.foursquare.com/developer/"](https://developer.foursquare.com/developer/)

DATA CLEANING

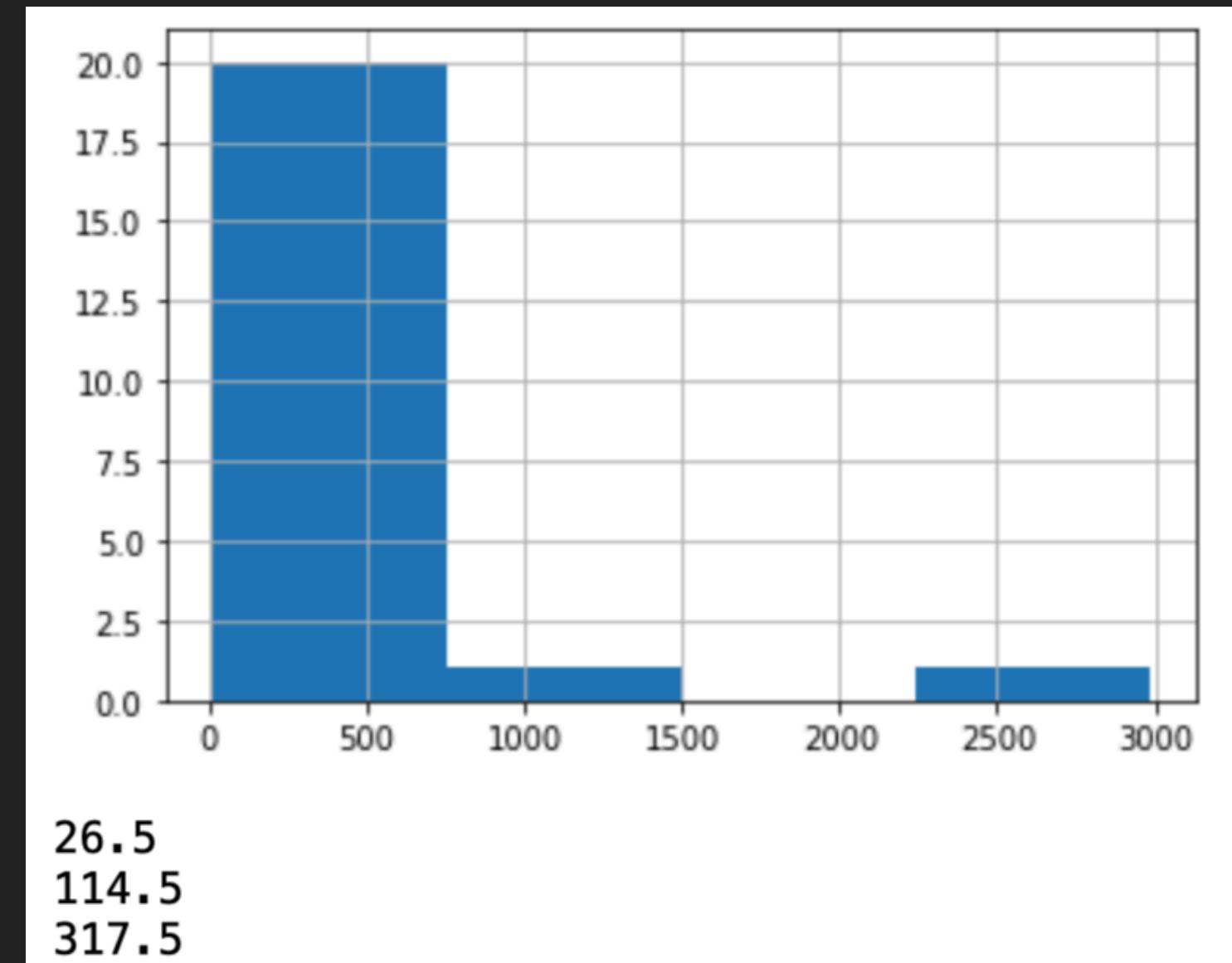
- ▶ Deal with the “Not assigned” values
- ▶ Get latitude and longitude values for each post code
- ▶ Get a clean data-frame

RE-CATEGORIZATION & PREPARATIONS

- ▶ Pull the detailed data from Foursquare API
- ▶ Deal with the ambiguous and messy categories, and re-categorize data by creating new categories but with smaller amount.
- ▶ Split the data, create two sub-data-frames: Restaurants data and Tourism data
- ▶ Visualizing base on histograms and find the 25th, 50th, and 75th percentiles for each data-frame.



Histogram and percentiles - restaurant



Histogram and percentiles - tourism

ONE-HOT ENCODING

- ▶ Used on categorical data
- ▶ Transfer categorical data to numerical data
- ▶ In the project: apply to “categories” and “total likes categories”

| | Name | other_tour | shopping | above avg | below avg | great | poor |
|---|--------------------------|------------|----------|-----------|-----------|-------|------|
| 0 | Downtown Toronto | 1 | 0 | 1 | 0 | 0 | 0 |
| 1 | Nathan Phillips Square | 1 | 0 | 0 | 0 | 1 | 0 |
| 6 | LUSH | 0 | 1 | 0 | 0 | 0 | 1 |
| 7 | Textile Museum of Canada | 1 | 0 | 0 | 0 | 0 | 1 |
| 8 | CF Toronto Eaton Centre | 0 | 1 | 0 | 0 | 1 | 0 |

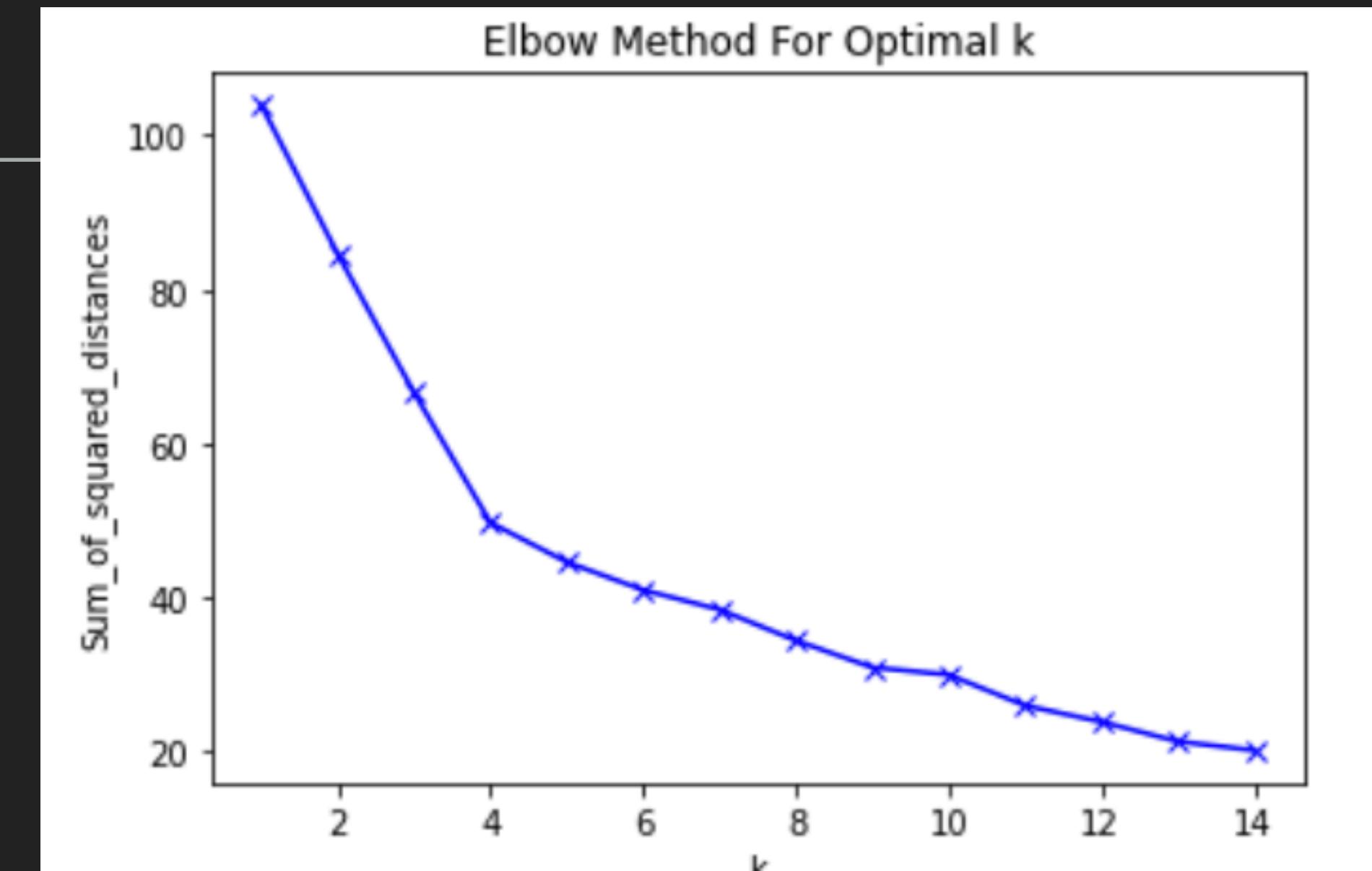
Example: One-hot encoding table for tourism data

K-MEANS CLUSTERING

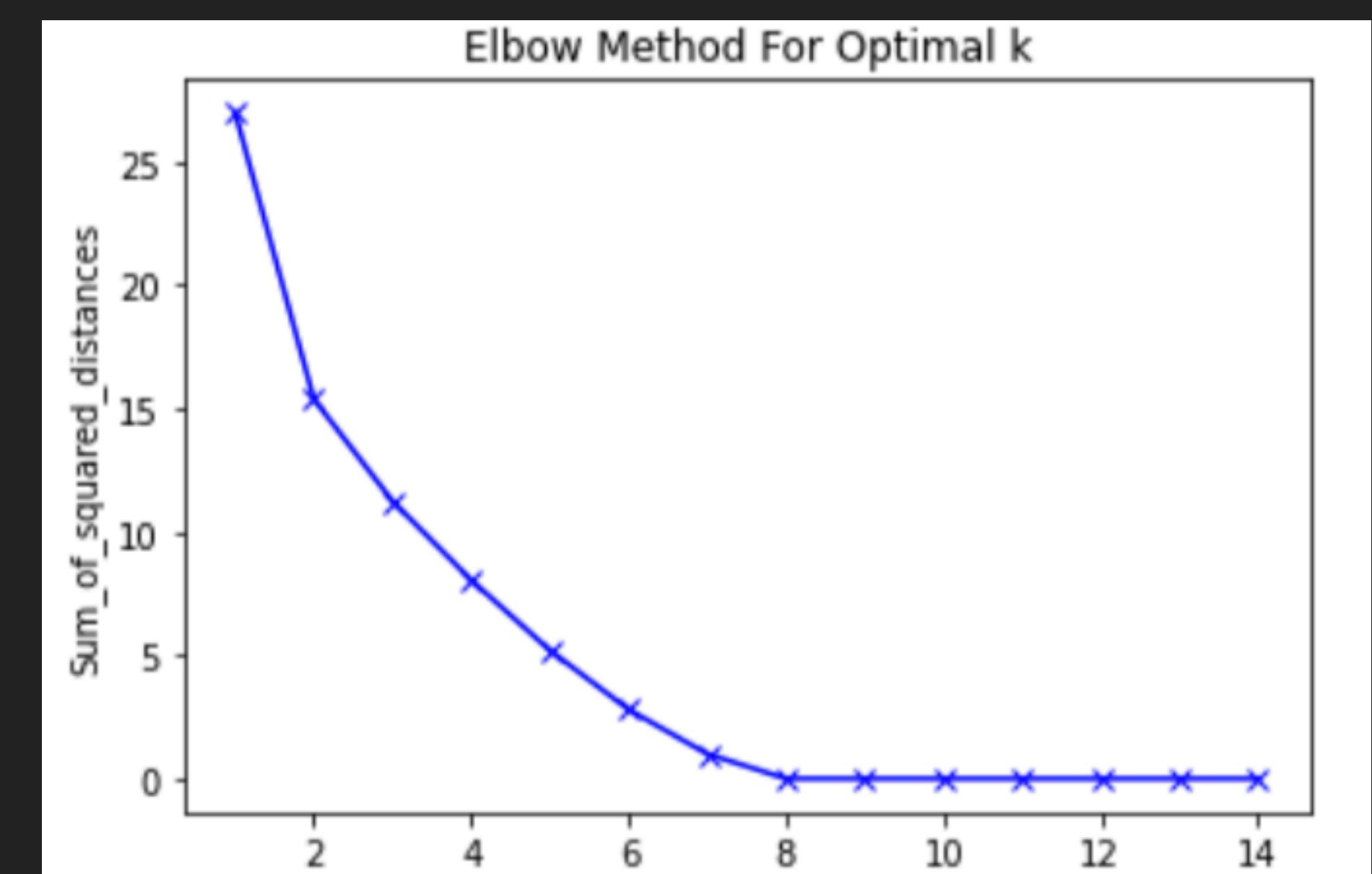
- ▶ Find optimal k using elbow method (k = 4 for restaurants data-frame and k = 8 for tourism data-frame)
- ▶ Apply K-means clustering model and add the generated labels to the data-frames

| | name | id | categories | lat | lng | total likes | total likes_cat | categories_new | label |
|---|--------------------------|--------------------------|----------------|-----------|------------|-------------|-----------------|----------------|-------|
| 0 | Downtown Toronto | 5227bb01498e17bf485e6202 | Neighborhood | 43.653232 | -79.385296 | 316 | above avg | other_tour | 4 |
| 1 | Nathan Phillips Square | 4ad4c05ef964a520a6f620e3 | Plaza | 43.652270 | -79.383516 | 692 | great | other_tour | 5 |
| 6 | LUSH | 4bd0b30d41b9ef3b8fa0fae5 | Cosmetics Shop | 43.653557 | -79.380400 | 22 | poor | shopping | 6 |
| 7 | Textile Museum of Canada | 4ad4c05ef964a520e2f620e3 | Art Museum | 43.654396 | -79.386500 | 25 | poor | other_tour | 0 |
| 8 | CF Toronto Eaton Centre | 4ad77a12f964a520260b21e3 | Shopping Mall | 43.654447 | -79.380952 | 2986 | great | shopping | 2 |

Example tourism data-frame after adding labels



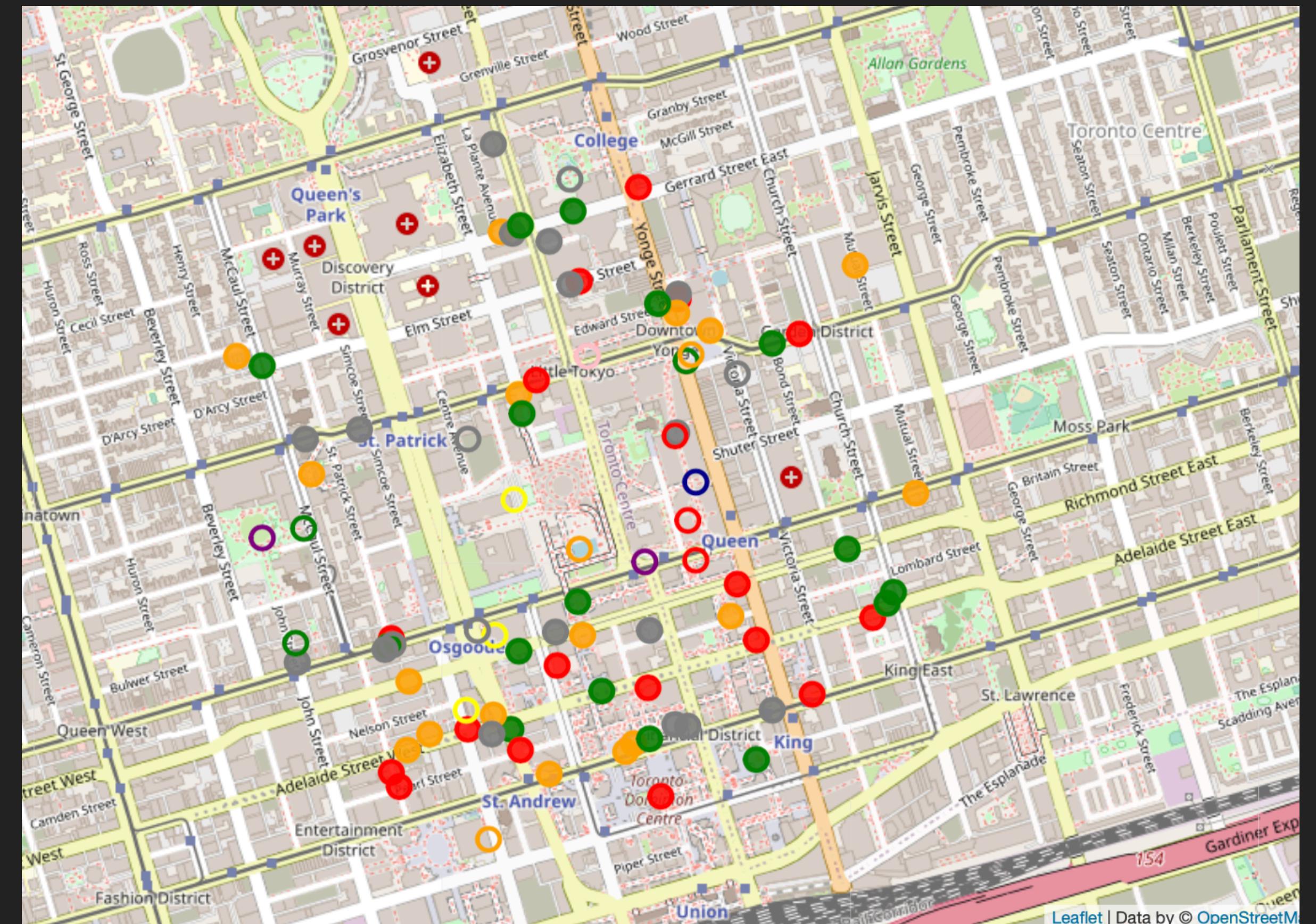
Elbow method for restaurants data



Elbow method for tourism data

TABLES & MAP

- ▶ For the restaurants data, cluster 0 = 'poor' likes, cluster 1 = 'great', cluster 2 = 'below average', and cluster 3 = 'above average' category; gray, red, green, and orange circle marks, filled with the same colors inside the circle, were used for the clusters respectively.
- ▶ For the tourism data, cluster 0 = 'poor' and 'other tour', cluster 1 = 'below average' and 'shopping', cluster 2 = 'great' and 'shopping', cluster 3 = 'below average' and 'other tour', cluster 4 = 'above average' and 'other tour', cluster 5 = 'great' and 'other tour', cluster 6 = 'poor' and 'shopping', and cluster 7 = 'above average' and 'shopping' categories; gray, green, red, purple, yellow, orange, dark blue, pink circle marks, filled without any color inside the circle, were used for the clusters respectively.



DATA GROUPED BY DIFFERENT CLUSTERS - RESTAURANTS

| | name | id | categories | lat | lng | total likes | total likes_cat | categories_new | label |
|----|------------------------------|--------------------------|----------------------------|-----------|------------|-------------|-----------------|--------------------|-------|
| 11 | Crepe Delicious | 4e5d8181a8092f63968617ee | Fast Food Restaurant | 43.654536 | -79.380889 | 8 | poor | fast food | 0 |
| 19 | John & Sons Oyster House | 50ecb1fae4b0beb13294f0aa | Seafood Restaurant | 43.650656 | -79.381613 | 21 | poor | euro italian food | 0 |
| 22 | Bulldog On The Block | 5a3a846af62e0960e9364d11 | Coffee Shop | 43.650652 | -79.384141 | 9 | poor | sweet and beverage | 0 |
| 27 | The Elm Tree Restaurant | 539c6f13498e06f4cc765165 | Modern European Restaurant | 43.657397 | -79.383761 | 28 | poor | euro italian food | 0 |
| 44 | Cafe Plenty | 4f513029e4b07c3382c9fdb9 | Café | 43.654571 | -79.389450 | 19 | poor | sweet and beverage | 0 |
| 46 | The Library Specialty Coffee | 5a6b737b35f98359eed11974 | Coffee Shop | 43.654413 | -79.390902 | 27 | poor | sweet and beverage | 0 |
| 53 | Pilot Coffee Roasters | 59cd51c71b0ea516e9e7b3aa | Coffee Shop | 43.648835 | -79.380936 | 24 | poor | sweet and beverage | 0 |
| 55 | Brick Street Bakery | 4dcbf219d22d7ffe9d39197e | Bakery | 43.648815 | -79.380605 | 16 | poor | sweet and beverage | 0 |
| 61 | Panago | 4ad4cba2f964a520d1fb20e3 | Pizza Place | 43.658258 | -79.384313 | 11 | poor | fast food | 0 |
| 62 | Tokyo Smoke | 5a55078bda5ede6ed9d73a62 | Coffee Shop | 43.657230 | -79.380870 | 5 | poor | sweet and beverage | 0 |
| 63 | Hakata Ikkousha Ramen | 5c9aaa3f66f3cd002c95bd58 | Ramen Restaurant | 43.650299 | -79.388753 | 14 | poor | asia food | 0 |
| 67 | Somethin' 2 Talk About | 4b744336f964a520d8d02de3 | Middle Eastern Restaurant | 43.658395 | -79.385338 | 6 | poor | middle east food | 0 |
| 79 | Omg! Oh My Gyro! | 595d4380c876c841c08f1959 | Souvlaki Shop | 43.650064 | -79.391104 | 17 | poor | middle east food | 0 |
| 88 | Pi Co. | 5a5a59c5a423620ec1dafd41 | Pizza Place | 43.648651 | -79.385874 | 13 | poor | fast food | 0 |
| 89 | Café Plenty | 53c524bd498efaeefb73b291 | Café | 43.649118 | -79.378313 | 20 | poor | sweet and beverage | 0 |
| 99 | NEO COFFEE BAR | 5db70a0c306d6b000861ac9f | Coffee Shop | 43.660130 | -79.385830 | 3 | poor | sweet and beverage | 0 |

Cluster 0

| | name | id | categories | lat | lng | total likes | total likes_cat | categories_new | label |
|----|--|--------------------------|--------------------------|-----------|------------|-------------|-----------------|----------------------|-------|
| 4 | Chatime 日出茶太 | 4e2284b11fc7c0ef9857d143 | Bubble Tea Shop | 43.655542 | -79.384684 | 198 | great | sweet and beverage | 1 |
| 14 | Richmond Station | 506db1a9e4b0a3f3b31412f0 | American Restaurant | 43.651569 | -79.379266 | 248 | great | american food | 1 |
| 21 | The Keg Steakhouse + Bar - York Street | 4ad69511f964a520e40721e3 | Steakhouse | 43.649987 | -79.384103 | 269 | great | american food | 1 |
| 23 | The Queen and Beaver Public House | 4ab2b0b9f964a520e56b20e3 | Gastropub | 43.657472 | -79.383524 | 232 | great | bars | 1 |
| 30 | Dineen Coffee | 514627d1e4b0dba1b85e9ba8 | Café | 43.650497 | -79.378765 | 450 | great | sweet and beverage | 1 |
| 38 | Soho House Toronto | 50322b6ae4b09116a296568c | Speakeasy | 43.648734 | -79.386541 | 219 | great | bars | 1 |
| 39 | Five Guys | 5064c3dde4b07c5a18986283 | Burger Joint | 43.657117 | -79.380853 | 230 | great | fast food | 1 |
| 52 | Cactus Club Cafe | 55fc571f498ec35023360858 | American Restaurant | 43.649552 | -79.381671 | 204 | great | american food | 1 |
| 54 | Banh Mi Boys | 51755dc7498ece19b7261641 | Sandwich Place | 43.659292 | -79.381949 | 311 | great | breakfast | 1 |
| 56 | The Rex Hotel Jazz & Blues Bar | 4b68aed1f964a520de862be3 | Jazz Club | 43.650505 | -79.388577 | 204 | great | bars | 1 |
| 64 | Pai | 529612de11d2ab526191ccc9 | Thai Restaurant | 43.647923 | -79.388579 | 687 | great | thai vietnamese food | 1 |
| 70 | Canoe | 4ad4c05df964a52059f620e3 | American Restaurant | 43.647452 | -79.381320 | 286 | great | american food | 1 |
| 73 | Hokkaido Ramen Santouka らーめん山頭火 | 509e9ef6e4b0ab175389a6c5 | Ramen Restaurant | 43.656435 | -79.377586 | 305 | great | asia food | 1 |
| 74 | Byblos Toronto | 5321f4d9e4b07946702e6e08 | Mediterranean Restaurant | 43.647615 | -79.388381 | 247 | great | euro italian food | 1 |
| 77 | Pizzeria Libretto | 5462ac56498e128ccafe8fea | Pizza Place | 43.648334 | -79.385111 | 209 | great | fast food | 1 |
| 90 | Terroni | 4b49183ff964a520a46526e3 | Italian Restaurant | 43.650927 | -79.375602 | 280 | great | euro italian food | 1 |
| 97 | Beerbistro | 4b5ca7d8f964a5207c3c29e3 | Gastropub | 43.649419 | -79.377237 | 325 | great | bars | 1 |

Cluster 1

| | name | id | categories | lat | lng | total likes | total likes_cat | categories_new | label |
|----|--------------------------------------|--------------------------|-------------------------------|------------|------------|--------------------|------------------------|-----------------------|--------------|
| 5 | Poke Guys | 57bcd3b7498e652a678d0378 | Poke Place | 43.654895 | -79.385052 | 32 | below avg | american food | 2 |
| 13 | M Square Coffee Co | 54132b3b498ee9ca9332e189 | Coffee Shop | 43.651218 | -79.383555 | 32 | below avg | sweet and beverage | 2 |
| 18 | Rosalinda | 5aff06ca6e4650002cc6286b | Vegetarian / Vegan Restaurant | 43.650252 | -79.385156 | 29 | below avg | vegan gluten free | 2 |
| 36 | Hy's Steakhouse | 4bd8cde92e6f0f47dcb20808 | Steakhouse | 43.649505 | -79.382919 | 43 | below avg | american food | 2 |
| 37 | Hailed Coffee | 5a81ae339deb7d369fa7f146 | Coffee Shop | 43.658833 | -79.383684 | 41 | below avg | sweet and beverage | 2 |
| 42 | Burrito Boyz | 55a9bbf9498e00ffd7f4c71f | Burrito Place | 43.656265 | -79.378343 | 29 | below avg | mex southam food | 2 |
| 49 | The Black Canary Espresso Bar | 506a2591e4b0961239b8c825 | Café | 43.657029 | -79.381385 | 44 | below avg | sweet and beverage | 2 |
| 51 | Cafe Landwer | 5b6c842bc36588002c80a934 | Café | 43.648753 | -79.385367 | 42 | below avg | sweet and beverage | 2 |
| 57 | HotBlack Coffee | 59f784dd28122f14f9d5d63d | Coffee Shop | 43.650364 | -79.388669 | 40 | below avg | sweet and beverage | 2 |
| 78 | Mos Mos Coffee | 52138db911d22803b334c641 | Café | 43.648159 | -79.378745 | 54 | below avg | sweet and beverage | 2 |
| 80 | Tim Hortons | 557231e3498e540f05f3083c | Coffee Shop | 43.658570 | -79.385123 | 41 | below avg | sweet and beverage | 2 |
| 82 | Hawthorne Food and Drink | 505c7c30e4b071a17ec2678f | Gastropub | 43.652270 | -79.376318 | 31 | below avg | bars | 2 |
| 87 | Gyu-Kaku Japanese BBQ | 574ad72238fa943556d93b8e | Japanese Restaurant | 43.651422 | -79.375047 | 42 | below avg | asia food | 2 |
| 92 | Versus Coffee | 57963897498e6873394a559d | Coffee Shop | 43.651213 | -79.375236 | 34 | below avg | sweet and beverage | 2 |
| 94 | Kupfert & Kim (First Canadian Place) | 50e44770e4b0e03a48c0d8a8 | Gluten-free Restaurant | 43.648547 | -79.381624 | 47 | below avg | vegan gluten free | 2 |
| 98 | Jimmy's Coffee | 57fe5f64498e08c9fc55cb87 | Café | 43.655827 | -79.392042 | 32 | below avg | sweet and beverage | 2 |

Cluster 2

| | name | id | categories | lat | lng | total likes | total likes_cat | categories_new | label |
|----|------------------------------|--------------------------|-------------------------------|------------|------------|--------------------|------------------------|-----------------------|--------------|
| 2 | Japango | 4ae7b27df964a52068ad21e3 | Sushi Restaurant | 43.655268 | -79.385165 | 161 | above avg | asia food | 3 |
| 25 | Blaze Pizza | 5615b6c4498e3c32c67ad78f | Pizza Place | 43.656518 | -79.380015 | 76 | above avg | fast food | 3 |
| 31 | Assembly Chef's Hall | 5a32970da42362290203a9d6 | Food Court | 43.650579 | -79.383412 | 67 | above avg | fast food | 3 |
| 35 | JaBistro | 509bb871e4b09c7ac93f6642 | Sushi Restaurant | 43.649687 | -79.388090 | 188 | above avg | asia food | 3 |
| 40 | Jimmy's Coffee | 537d4d6d498ec171ba22e7fe | Coffee Shop | 43.658421 | -79.385613 | 162 | above avg | sweet and beverage | 3 |
| 43 | Karine's | 4c90c810ae96a093599f9d46 | Breakfast Spot | 43.653699 | -79.390743 | 60 | above avg | breakfast | 3 |
| 45 | The Chase | 5214e7c111d2a83379eae21f | New American Restaurant | 43.650952 | -79.379422 | 141 | above avg | american food | 3 |
| 60 | Bosk at Shangri-La | 50468014e4b01c18bb731df8 | Asian Restaurant | 43.649023 | -79.385826 | 69 | above avg | asia food | 3 |
| 66 | Maman | 559a8f5a498e31f945041245 | Café | 43.648309 | -79.382253 | 84 | above avg | sweet and beverage | 3 |
| 69 | Chipotle Mexican Grill | 4adc9148f964a520512d21e3 | Mexican Restaurant | 43.656860 | -79.380910 | 171 | above avg | mex southam food | 3 |
| 72 | King Taps | 59603f86112c6c70931c9401 | Gastropub | 43.648476 | -79.382058 | 59 | above avg | bars | 3 |
| 81 | GEORGE Restaurant | 4af618daf964a520220122e3 | Vegetarian / Vegan Restaurant | 43.653346 | -79.374445 | 77 | above avg | vegan gluten free | 3 |
| 83 | Sam James Coffee Bar (SJCB) | 4fccaa8fe4b05a98df3d9417 | Café | 43.647881 | -79.384332 | 71 | above avg | sweet and beverage | 3 |
| 85 | Page One Cafe | 56d4d1b3cd1035fe77e1492c | Café | 43.657772 | -79.376073 | 68 | above avg | sweet and beverage | 3 |
| 86 | Vegetarian Haven | 4aeb711ef964a52017c221e3 | Vegetarian / Vegan Restaurant | 43.656016 | -79.392758 | 58 | above avg | vegan gluten free | 3 |
| 91 | The Burger's Priest | 5578720f498ee3b165b1f32b | Fast Food Restaurant | 43.648643 | -79.387539 | 135 | above avg | fast food | 3 |
| 95 | Copacabana Grilled Brazilian | 52ec621e498ec68fa15ee922 | Brazilian Restaurant | 43.648333 | -79.388151 | 149 | above avg | mex southam food | 3 |

Cluster 3

DATA GROUPED BY DIFFERENT CLUSTERS - TOURISM

| | name | id | categories | lat | lng | total likes | total likes_cat | categories_new | label |
|----|---|--------------------------|-------------------|------------|------------|--------------------|------------------------|-----------------------|--------------|
| 7 | Textile Museum of Canada | 4ad4c05ef964a520e2f620e3 | Art Museum | 43.654396 | -79.386500 | 25 | poor | other_tour | 0 |
| 28 | Friendly Stranger - Cannabis Culture Shop | 4b7ed424f964a5208a0230e3 | Smoke Shop | 43.650387 | -79.388523 | 17 | poor | other_tour | 0 |
| 29 | Canadian Opera Company | 4ad4c062f964a520baf720e3 | Opera House | 43.650660 | -79.386242 | 21 | poor | other_tour | 0 |
| 41 | Jazz Bistro | 514cc159e4b0e4f73af4eced | Music Venue | 43.655678 | -79.379276 | 24 | poor | other_tour | 0 |
| 75 | College Park Area | 4c8facf91664b1f79c90aa2f | Park | 43.659453 | -79.383785 | 9 | poor | other_tour | 0 |

Cluster 0

| | name | id | categories | lat | lng | total likes | total likes_cat | categories_new | label |
|----|--------------------------|--------------------------|------------------------|------------|------------|--------------------|------------------------|-----------------------|--------------|
| 12 | UNIQLO ユニクロ | 57eda381498ebe0e6ef40972 | Clothing Store | 43.655910 | -79.380641 | 114 | below avg | shopping | 1 |
| 26 | Silver Snail Comics | 4ad4c062f964a5200bf820e3 | Comic Shop | 43.657031 | -79.381403 | 115 | below avg | shopping | 1 |
| 47 | Aboveground Art Supplies | 4adf3c01f964a5208f7821e3 | Arts & Crafts Store | 43.652646 | -79.390925 | 35 | below avg | shopping | 1 |
| 71 | Umbra Concept Store | 4ae734bef964a5205ea921e3 | Furniture / Home Store | 43.650417 | -79.391136 | 31 | below avg | shopping | 1 |

Cluster 1

| | name | id | categories | lat | lng | total likes | total likes_cat | categories_new | label |
|----|-------------------------|--------------------------|-------------------|------------|------------|--------------------|------------------------|-----------------------|--------------|
| 8 | CF Toronto Eaton Centre | 4ad77a12f964a520260b21e3 | Shopping Mall | 43.654447 | -79.380952 | 2986 | great | shopping | 2 |
| 24 | Hudson's Bay | 4adf85e1f964a5206e7b21e3 | Department Store | 43.652040 | -79.380391 | 573 | great | shopping | 2 |
| 33 | Apple Eaton Centre | 4ad788c8f964a520e40b21e3 | Electronics Store | 43.652818 | -79.380617 | 493 | great | shopping | 2 |

Cluster 2

| | name | id | categories | lat | lng | total likes | total likes_cat | categories_new | label |
|----|---------------|--------------------------|---------------------|------------|------------|--------------------|------------------------|-----------------------|--------------|
| 10 | Old City Hall | 4ad4c05ef964a5208ef620e3 | Monument / Landmark | 43.652009 | -79.381744 | 73 | below avg | other_tour | 3 |
| 93 | Grange Park | 4b54deadf964a520a3d027e3 | Park | 43.652488 | -79.392053 | 45 | below avg | other_tour | 3 |

Cluster 3

| | name | id | categories | lat | lng | total likes | total likes_cat | categories_new | label |
|----|---|--------------------------|-------------------|------------|------------|--------------------|------------------------|-----------------------|--------------|
| 0 | Downtown Toronto | 5227bb01498e17bf485e6202 | Neighborhood | 43.653232 | -79.385296 | 316 | above avg | other_tour | 4 |
| 9 | Four Seasons Centre for the Performing Arts | 4ad4c062f964a520e5f720e3 | Concert Hall | 43.650592 | -79.385806 | 222 | above avg | other_tour | 4 |
| 32 | Shangri-La Toronto | 4e31b74252b131dcebb08743 | Hotel | 43.649129 | -79.386557 | 181 | above avg | other_tour | 4 |

Cluster 4

| | name | id | categories | lat | lng | total likes | total likes_cat | categories_new | label |
|----|------------------------|--------------------------|-------------------|------------|------------|--------------------|------------------------|-----------------------|--------------|
| 1 | Nathan Phillips Square | 4ad4c05ef964a520a6f620e3 | Plaza | 43.652270 | -79.383516 | 692 | great | other_tour | 5 |
| 20 | Yonge-Dundas Square | 4ad8cd16f964a520c91421e3 | Plaza | 43.656054 | -79.380495 | 823 | great | other_tour | 5 |
| 96 | Roy Thomson Hall | 4ad4c061f964a520b0f720e3 | Concert Hall | 43.646589 | -79.385979 | 318 | great | other_tour | 5 |

Cluster 5

| | name | id | categories | lat | lng | total likes | total likes_cat | categories_new | label |
|---|-------------|--------------------------|-------------------|------------|------------|--------------------|------------------------|-----------------------|--------------|
| 6 | LUSH | 4bd0b30d41b9ef3b8fa0fae5 | Cosmetics Shop | 43.653557 | -79.3804 | 22 | poor | shopping | 6 |

Cluster 6

| | name | id | categories | lat | lng | total likes | total likes_cat | categories_new | label |
|----|-------------|--------------------------|--------------------|------------|------------|--------------------|------------------------|-----------------------|--------------|
| 17 | MUJI | 5479da4f498e8569fb44985c | Miscellaneous Shop | 43.656024 | -79.383284 | 135 | above avg | shopping | 7 |

Cluster 7

YONGE STREET AND THE AREA BETWEEN OSGOODE AND KING IS HIGHLY COMPETITIVE. HOWEVER, THERE IS STILL AN OPPORTUNITY FOR PEOPLE WHO IS GOING TO OPEN CAFE / BUBBLE-TEA STORE.

THERE ARE MANY SHOPPING MALLS / LANDMARKS, AND ONLY FEW RESTAURANTS LOCATED IN THE AREA BETWEEN ST.PATRICK AND QUEEN – GOOD PLACE FOR STARTING A NEW BUSINESS!

FUTURE WORK: ACCESS TO MORE SAMPLES (CURRENTLY ONLY 100 SAMPLES ARE USED) TO IMPROVE THE ACCURACY AND GAIN MORE DETAILED INFORMATION

Conclusion & Future Work