SEPTEMBER 2, 2018

# THE BATTLE OF NEIGHBORHOODS REPORT
## COURSERA – CAPSTONE PROJECT

TIMBAKE DANG

# TABLE OF CONTENTS

There are many existing venues in neighborhoods in New York as well as Toronto city. How can we know which venue by category and city have the most interesting from people as well as have the most number of users checked-in? Then how can we know the competition of the venues of the same category in a city or between neighborhoods in a city bases on the statistic number of users, visits, checked-ins, likes, rating?

Through by answering this questions, we can have insight about the competition between venues of the same category in a city or in a neighborhood, then we can find out the reasons could make these venues to be attractive to the people.

Finally, after we know reasons and have the accurate statistic data about the existing competitors, the behavior of users, the user culture, the demands are not served or are not served well, the needs in future, we will develop strategies for changes in the market, sustainable growth as well as expansion business to new market.

# DATA - STATISTICS & REPORTS

The data used in these statistics and reports was resources retrieved from:

- Toronto data: Wikipedia (Toronto postal code) combined with neighborhoods, venues data by using Foursquare APIs.

- New York data: included neighborhoods, venues data gathered from https://ibm.box.com/shared/static/fbpwbovar7lf8p5sgddm06cgipa2rxpe.json

Before using Foursquare APIs to get neighborhoods, venues data, I used Geocoder library to get coordinates (latitude, longitude) for each neighborhood, instead of Google Geoencoding APIs.

Then I combined 2 data frames Toronto and New York together in order to perform comparison, statistic and create reports.

## TOP 20 VENUES RECOMMENDED BY CATEGORY IN CITY, NEIGHBORHOODS

TABLE 1 TOP 20 CATEGORIES HAS THE MOST VENUES IN NEW YORK, TORONTO

|  | New York | Toronto | Total |
| --- | --- | --- | --- |
| Pizza Place | 434.0 | 49.0 | 483.0 |
| Coffee Shop | 264.0 | 198.0 | 462.0 |
| Italian Restaurant | 322.0 | 48.0 | 370.0 |
| Deli / Bodega | 274.0 | 18.0 | 292.0 |
| Bakery | 231.0 | 48.0 | 279.0 |
| Café | 164.0 | 103.0 | 267.0 |
| Bar | 223.0 | 41.0 | 264.0 |
| Chinese Restaurant | 204.0 | 25.0 | 229.0 |
| Sandwich Place | 186.0 | 41.0 | 227.0 |
| Park | 171.0 | 45.0 | 216.0 |
| American Restaurant | 176.0 | 31.0 | 207.0 |
| Grocery Store | 179.0 | 24.0 | 203.0 |
| Mexican Restaurant | 170.0 | 19.0 | 189.0 |
| Pharmacy | 169.0 | 18.0 | 187.0 |
| Donut Shop | 168.0 | 2.0 | 170.0 |
| Gym / Fitness Center | 144.0 | 21.0 | 165.0 |
| Ice Cream Shop | 141.0 | 20.0 | 161.0 |
| Restaurant | 87.0 | 61.0 | 148.0 |
| Sushi Restaurant | 113.0 | 30.0 | 143.0 |
| Hotel | 96.0 | 45.0 | 141.0 |

Top 20 Categories Has Most Recommended Venues in New York and Toronto



- New York
- Toronto

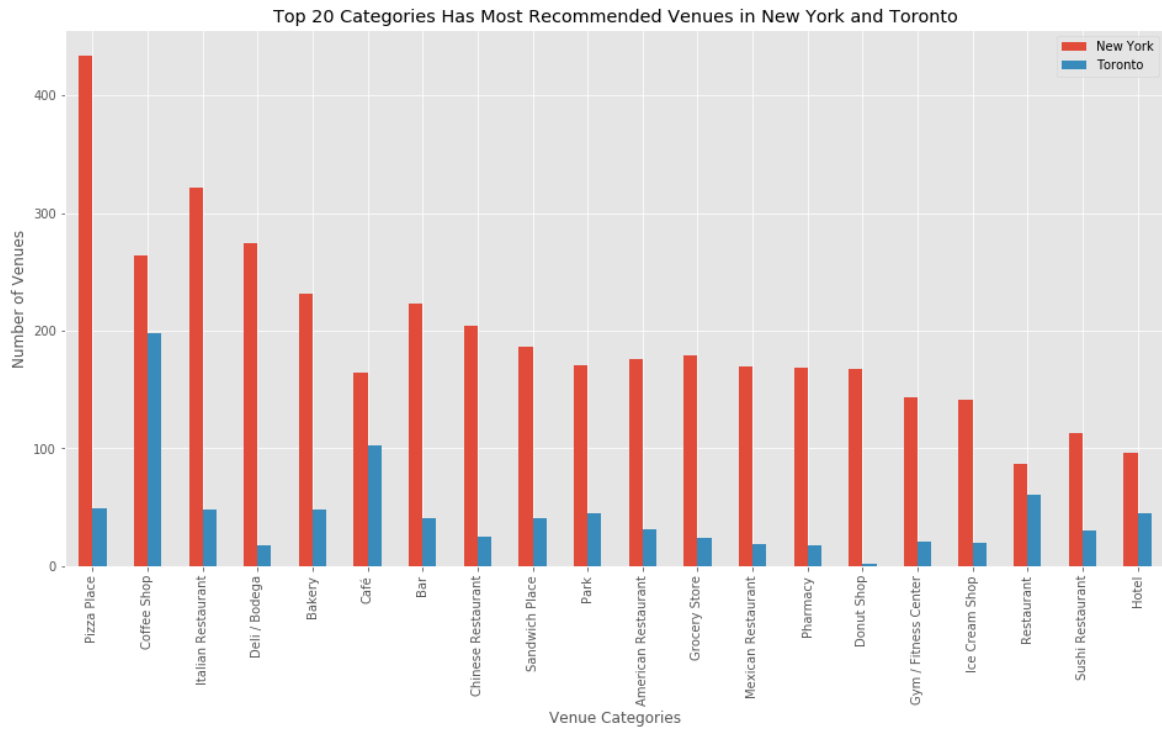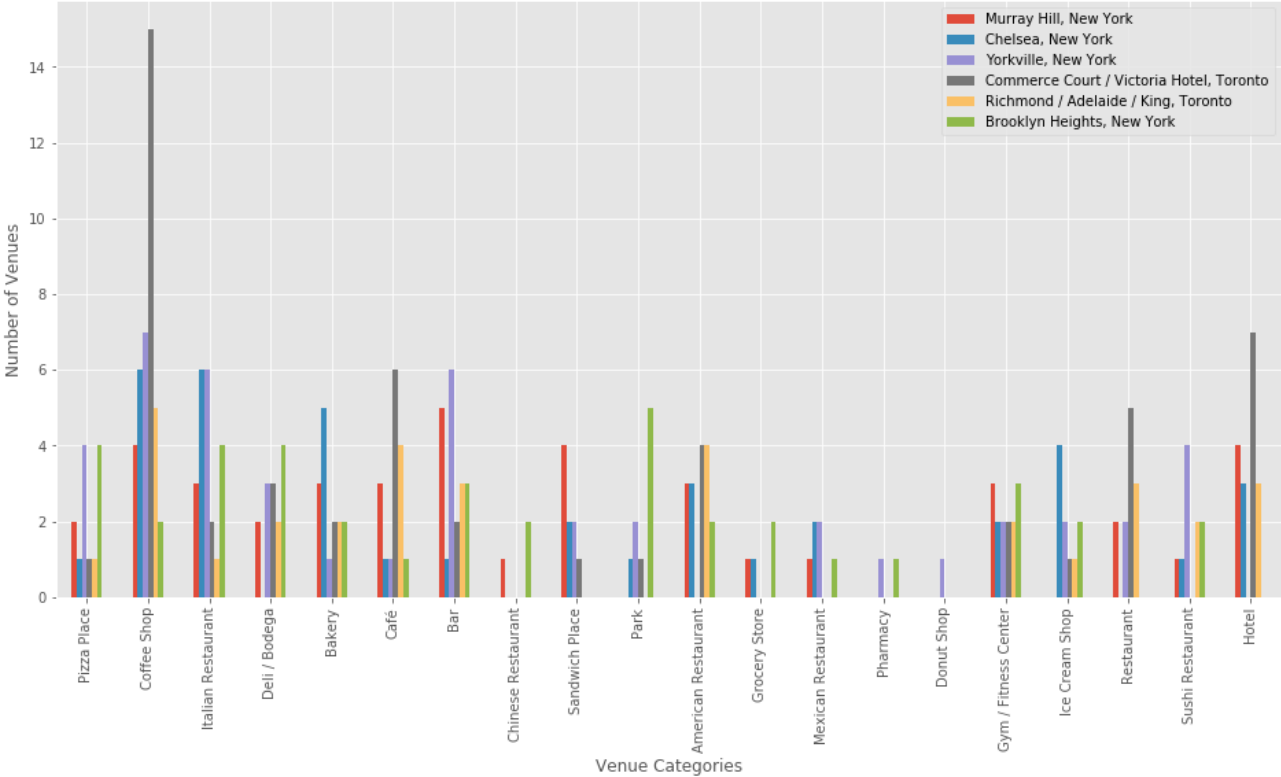(Y axis: Number of Venues; X axis: Venue Categories)

TABLE 2 TOP 20 CATEGORIES HAVE THE MOST VENUES IN 6 NEIGHBORHOODS

| | Murray Hill, New York | Chelsea, New York | Yorkville, New York | Commerce Court / Victoria Hotel, Toronto | Richmond / Adelaide / King, Toronto | Brooklyn Heights, New York | Total |
|---|---|---|---|---|---|---|---|
| Pizza Place | 2 | 1 | 4 | 1 | 1 | 4 | 483 |
| Coffee Shop | 4 | 6 | 7 | 15 | 5 | 2 | 462 |
| Italian Restaurant | 3 | 6 | 6 | 2 | 1 | 4 | 370 |
| Deli / Bodega | 2 | 0 | 3 | 3 | 2 | 4 | 292 |
| Bakery | 3 | 5 | 1 | 2 | 2 | 2 | 279 |
| Café | 3 | 1 | 1 | 6 | 4 | 1 | 267 |
| Bar | 5 | 1 | 6 | 2 | 3 | 3 | 264 |
| Chinese Restaurant | 1 | 0 | 0 | 0 | 0 | 2 | 229 |
| Sandwich Place | 4 | 2 | 2 | 1 | 0 | 0 | 227 |
| Park | 0 | 1 | 2 | 1 | 0 | 5 | 216 |
| American Restaurant | 3 | 3 | 0 | 4 | 4 | 2 | 207 |
| Grocery Store | 1 | 1 | 0 | 0 | 0 | 2 | 203 |
| Mexican Restaurant | 1 | 2 | 2 | 0 | 0 | 1 | 189 |
| Pharmacy | 0 | 0 | 1 | 0 | 0 | 1 | 187 |
| Donut Shop | 0 | 0 | 1 | 0 | 0 | 0 | 170 |
| Gym / Fitness Center | 3 | 2 | 2 | 2 | 2 | 3 | 165 |
| Ice Cream Shop | 0 | 4 | 2 | 1 | 1 | 2 | 161 |
| Restaurant | 2 | 0 | 2 | 5 | 3 | 0 | 148 |
| Sushi Restaurant | 1 | 1 | 4 | 0 | 2 | 2 | 143 |
| Hotel | 4 | 3 | 0 | 7 | 3 | 0 | 141 |

Top 20 Categories Has Most Recommended Venues Of Top 6 Neighborhoods

From above data, I'm going to use FOURSQUARE API get detail information of venues such as likes, rating, the number of users, etc... However, because API endpoint Details of FOURSQUARE is PREMIUM this endpoint limits only 50 calls per day. It certainly doesn't adapt for my above data set with 12,374 different venues in New York and Toronto. I decided to use API endpoint Likes as an alternative method to get the number of likes.

Besides, because the limitation of Foursquare Regular Accounts as well as the number of venues of categories in New York is more than in Toronto, I also decided to get data frame of venues by top 10 categories and in New York only. This filtering help me to decrease the size of data set to 3916 records (venues).

In this project, I only retrieve the number of User Likes from Foursquare for first 500 venues in my data set.

TABLE 3 TOP 10 CATEGORIES HAVE THE MOST TOTAL USER LIKES IN NEW YORK CITY

| | Category Name | Likes |
|---|---|---|
| 0 | Italian Restaurant | 1134 |
| 1 | Pizza Place | 1040 |
| 2 | Bar | 816 |
| 3 | Donut Shop | 551 |
| 4 | Mexican Restaurant | 449 |
| 5 | Deli / Bodega | 376 |
| 6 | Pharmacy | 368 |
| 7 | American Restaurant | 346 |
| 8 | Bakery | 297 |
| 9 | Park | 293 |

TOP 10 CATEGORIES HAVE THE MOST TOTAL USER LIKES IN NEW YORK CITY
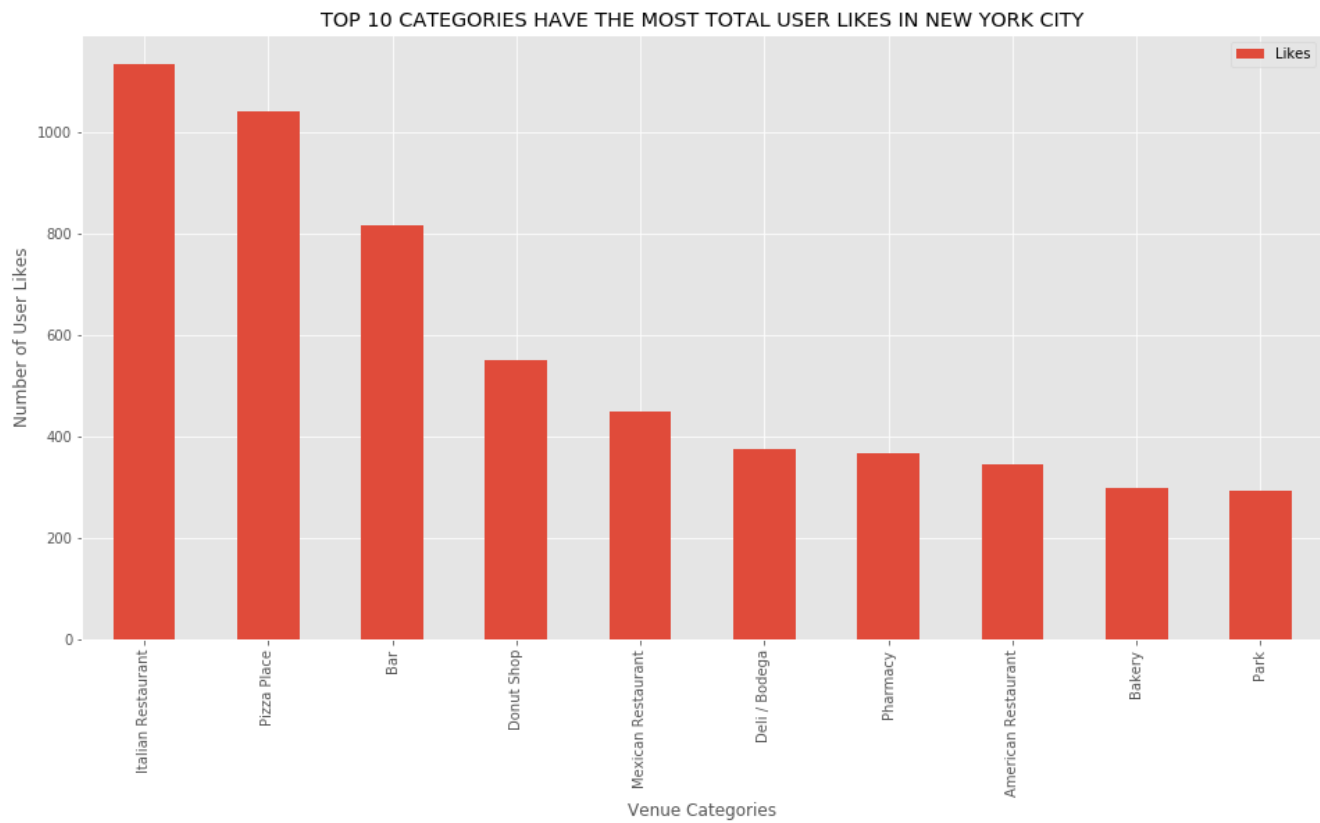


TABLE 4 TOP 10 VENUES HAVE THE MOST TOTAL USER LIKES IN NEW YORK CITY

| | Venue Name | Likes |
|---|---|---|
| 0 | Tosca Café | 334 |
| 1 | Dunkin' Donuts | 284 |
| 2 | Dunkin Donuts | 267 |
| 3 | Zero Otto Nove | 217 |
| 4 | Rite Aid | 164 |
| 5 | Planet Fitness | 160 |
| 6 | Tino's Delicatessen | 137 |
| 7 | The Bronx Beer Hall | 129 |
| 8 | Taqueria Tlaxcali | 122 |
| 9 | The Kettle Black | 115 |

TOP 10 VENUES HAVE THE MOST TOTAL USER LIKES IN NEW YORK CITY
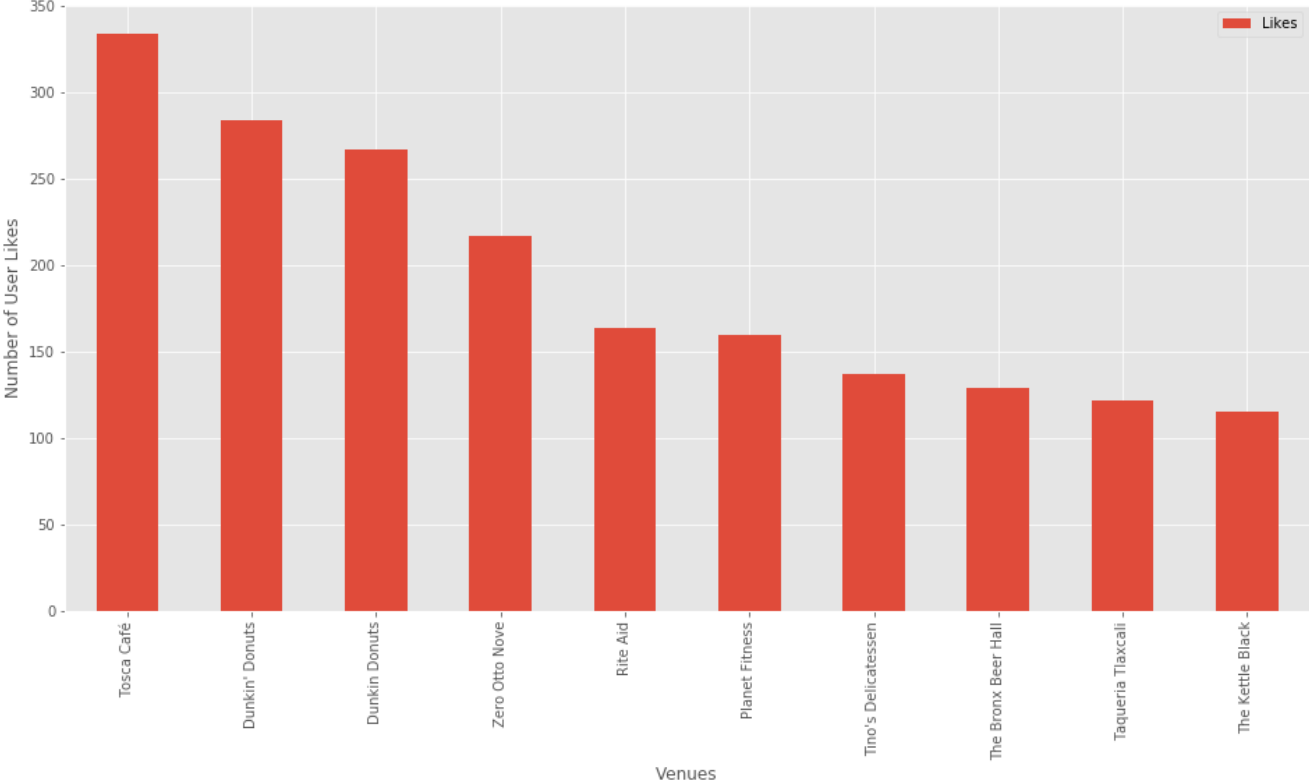
## TOP 5 VENUES HAVE THE HIGHEST RATING BY CATEGORY IN CITY, NEIGHBORHOODS

## TOP 5 VENUES HAVE THE MOST VISITS, USERS BY CATEGORY IN CITY, NEIGHBORHOODS

## TOP 10 CATEGORIES HAVE THE MOST VISITS, USERS BY CITY, NEIGHBORHOODS

https://github.com/timbake/TheBattleOfNeighborhoods

- NewYorkRVenues.csv: was from ibm
  https://ibm.box.com/shared/static/fbpwbovar7lf8p5sgddm06cgipa2rxpe.json combined with venues info
  retrieved from Foursquare using API endpoint Explore.
- TorontoRVenues.csv: was scrapped coordinates form Wikipedia
  https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M , then also combined with venues info
  retrieved Foursquare.
- RecommendedVenues.csv: was merged from 2 data set dfNewYorkRVenues (NewYorkRVenues.csv) and
  dfTorontoRVenues (TorontoRVenues.csv).
- 500NewYorkFilteredRVenues.csv