

New Profitable Start ups of Toronto

Author: Jeyakumar B

Issued: March 2020

Table of Contents

1. Introduction
2. Data Mining
 - 2.1 Goals
 - 2.2 Selecting data
 - 2.3 Data Processing and Transformation
3. Data Analysis Methodology
 - 3.1 Comparison of Toronto and New York Neighbourhood
 - 3.2 Comparison of Toronto and New York Businesses
 - 3.3 Comparison of Toronto and New York Demographics
 - 3.4 Relation of distance between City Centre and Common Venues
 - 3.5 Identifying more optimal location to start business
 - 3.6 More Insights around Yonge-Dundas Square
4. Results
5. Discussion
6. Conclusion

1. Introduction

NxtFut Group is New York based Multi-industry company, which is recognised for its quality, sustained growth and values to the New York community. NxtFut is doing various businesses including Restaurants, Donut Shops, Food delivery, travel, Health and hospitality. They want to start their new branch in Toronto. They desperately want to make their first business in Toronto a huge success. This success will help them to create a strong brand image and they can build their other businesses easily in Toronto. Hence NxtFut wants to find what is the best business to start in Toronto and how the Toronto receive it.

We are carrying out our data science project to find the solution of below problems.

1. What is the best business to start in Toronto
2. How the community will receive it
3. what is the best location to start the branch office.

This data science project will not only help the NxtFut group, this project will help all the new entrepreneurs or start-ups who want to start a new business in Toronto and to understand the community of Toronto.

2. Data Mining:

2.1 Goals:

As part of this data science project we are going to collect and analyse data by establishing below data mining goals to solve the problem that we have.

1. Identify the most common places people gather
2. Identify the most popular business on New York missing in Toronto
3. Identify the best place to start the business around centre of city of Toronto
4. Compare New York and Toronto demographics data & identify will the Toronto accept the selected business

2.2 Selecting Data:

This project is going to use the data collected from foursquare, demographics of Toronto and New York.

Foursquare will be used to collect Borough, Neighbourhood, latitude-longitude of each neighbourhood, top 100 venues of each neighbourhood and their respective latitude-longitude, venue categories.

We will be using demographics data from 2016-2017 census of Toronto and New York to collect the details of Age and racial/origin distribution of the respective cities.

We will be using other necessary python libraries to calculate the top most common venues of each neighbourhood, distance of each neighbourhood from the city centre of Toronto.

2.3 Data Processing and Transformation:

In this phase I had to overcome processing of few messy data to determine the appropriate format that can be easy for further analysis.

1. We had a lot of categories in demographics of cities. Those includes age, race, profession, income, education, etc., I had taken only age and racial distribution for our analysis as the other data is not necessary for our identifying the acceptability of start-up and it may deviate us from our goal.
2. I had grouped age and racial data to transform the large messy categories into abstract and effective data sets. Peoples from same age group or origin will possess almost same kind of expectations and living style that we want to analyse.

For example all the religious people in Asian region will like to have spices in their every meal. All the middle age people will like to save money for future rather than over spending it. All young people will like to travel and explore various aspects of life.

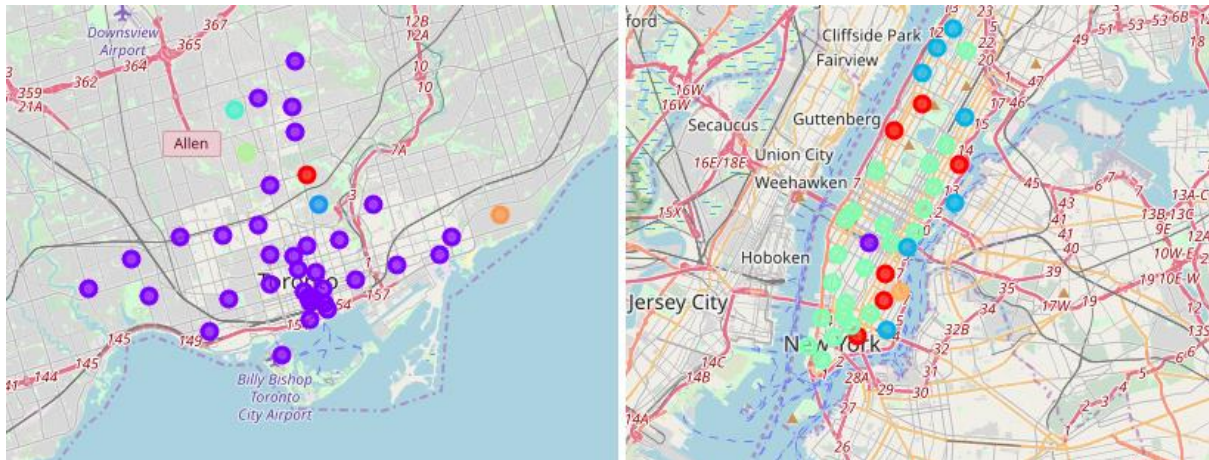
3. We have to create one single table to contain all the neighbourhood, latitude-longitude details, various venue categories, most common places, distance between the venues and city centre. Because this will be useful when we plot the geographical diagrams for all the venues and clustering those venues.

3. Data Analysis Methodology:

3.1 Comparison of Toronto and New York neighbourhoods

From the foursquare data of venues and neighbourhood let we chart all the venues of both cities in geographical diagram and create the clusters out of manhattan and old

Toronto Boroughs. We have chosen the two financial hub of New York and Toronto to identify the best business and location to start the business.

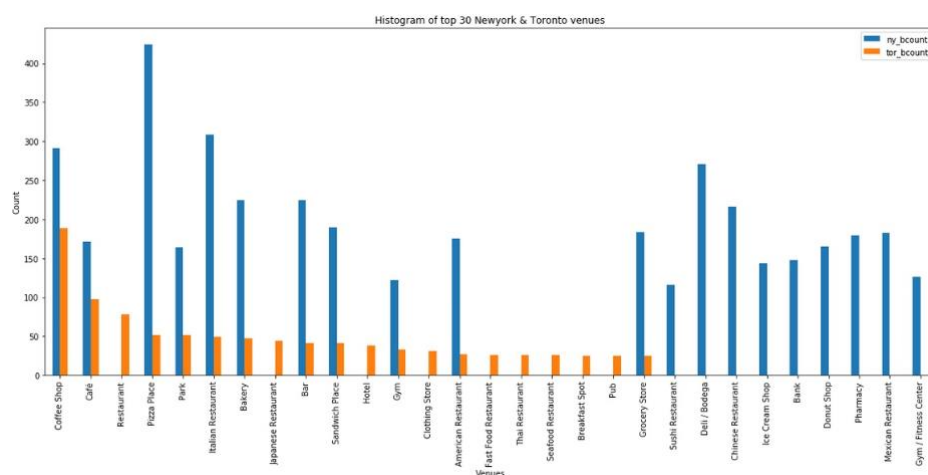


When we compare both clusters of Manhattan and Toronto they both looks very different in terms of placement of neighbourhood clusters and distributions. And geographically they both are very different in structure and in placement. Hence the as both the cities are financial Hub we can't compare based on the shape or region.

3.2 Comparison of Toronto and New York Businesses

It is very important to understand the businesses, various industries, interested areas, gatherings in both New York and Toronto to identify what could be the right entrepreneurship or start ups that will be profitable in the region.

So let we collect the top 30 most common venues accessed by the people of Toronto and New York and compare both of them.



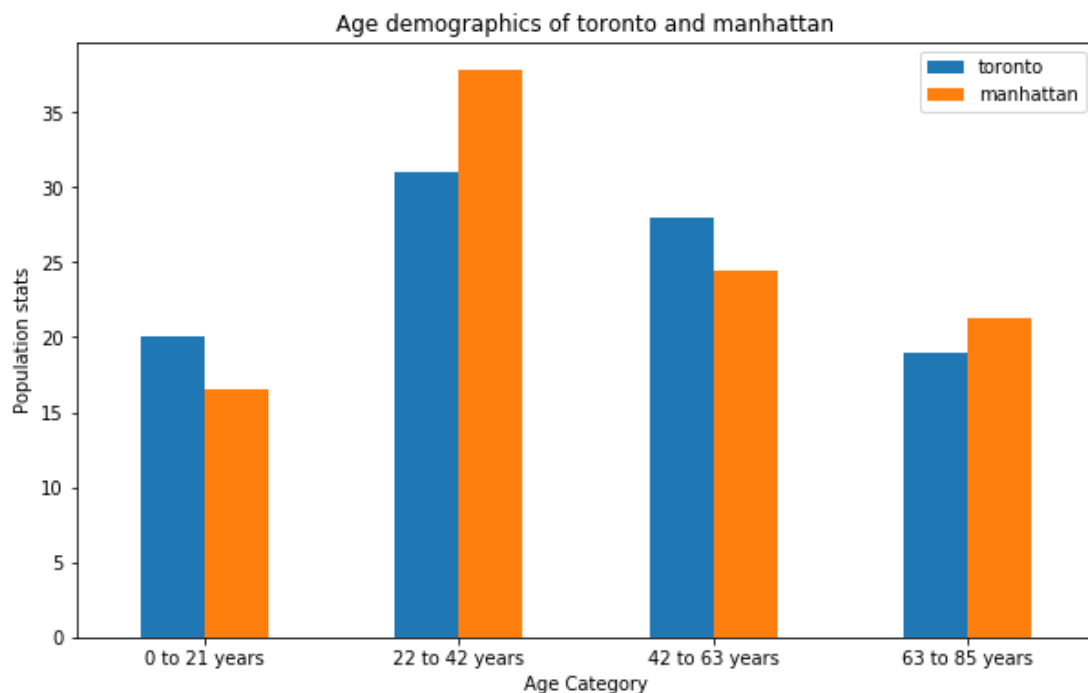
Venue	ny_bcount
Deli / Bodega	271.0
Chinese Restaurant	216.0
Mexican Restaurant	183.0
Pharmacy	179.0
Donut Shop	165.0
Bank	148.0
Ice Cream Shop	144.0
Gym / Fitness Center	126.0
Sushi Restaurant	116.0

By analysing the above chart we can easily identify the list of businesses those are very popular in New York do not have lot of venues in Toronto. And these are the business that may have the great scope in Toronto to start with.

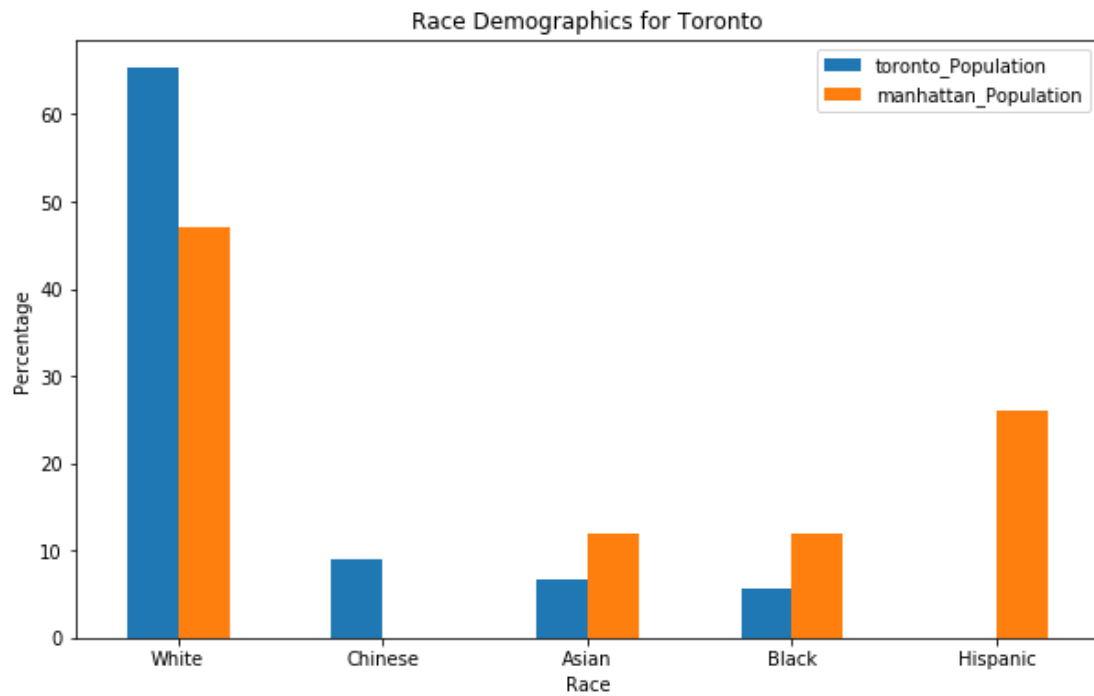
3.3 Comparison of Toronto and New York Demographics

Before we proceed further through the analysis of identified businesses let we compare the demographics data of both cities to understand how the people are similar or dissimilar.

Understanding the origins of the local community will gives us the better knowledge to know how different the markets of both city and possibility of receiving the new start ups by the people.



By comparing the demographics of age it is evident that there is no major changes in distribution of the students, young, middle age, retired peoples in both Toronto & New York.



By comparing the demographics of Race we can see few notable differences in both the cities.

- Toronto has more population of White than Manhattan
- Toronto has Chinese population where as Manhattan has Hispanic population
- Toronto has comparatively less population of Asian and Black than Manhattan

Collectively Toronto has more white and Chinese population compared to Manhattan. Hence any business those are relative to these two categories has great scope of getting success.

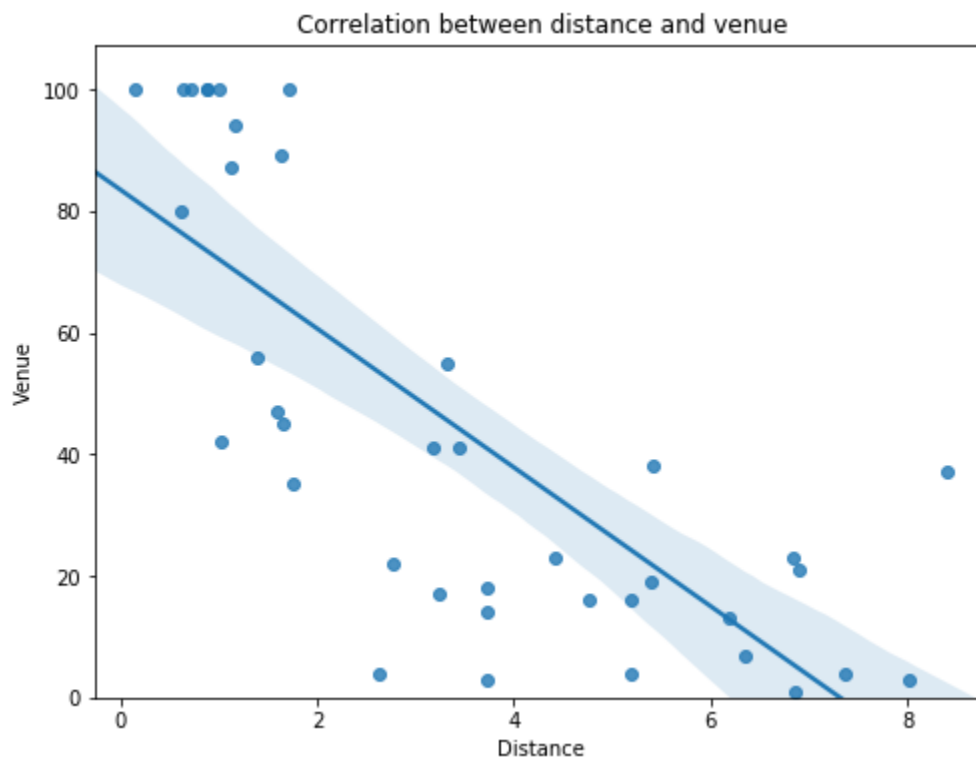
Based on these analysis, from this list of identified businesses let we pick the businesses that can do well in Toronto at the same time those are related to our customers portfolio and majority population of Toronto.

- Chinese Restaurant
- Mexican Restaurant
- Donut Shop
- Ice Cream Shop
- Pharmacy

3.4 Relation of distance between City Centre and Common Venues

Centre of city of Toronto is Yonge-Dundas Square, Canada and latitude and longitude of the location is 43.6561162, -79.38016329. Let we calculate the distance between the

most popular venues of old Toronto and to the centre of the city and plot the correlation diagram between these variables.



We have used **Linear Regression ML model** to find the correlation, as we have single dependent value and independent value.

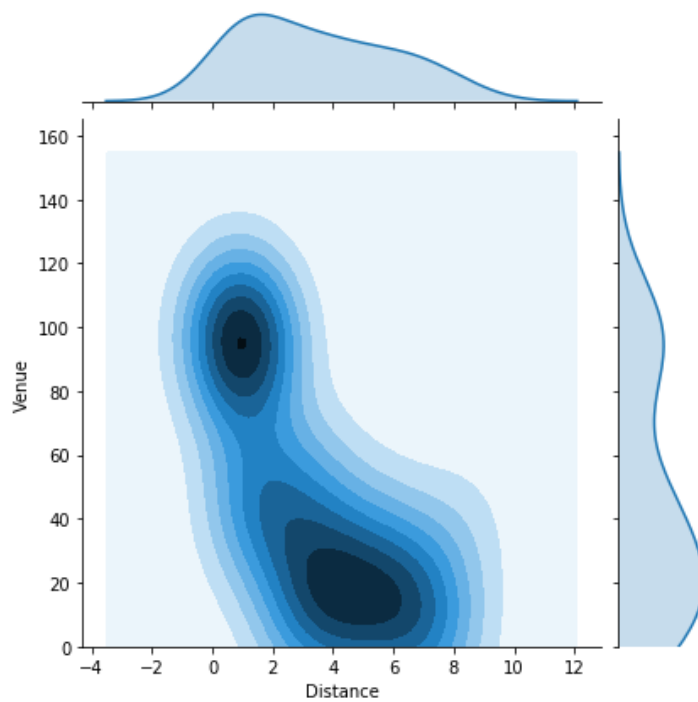
Distance and most common places of Toronto are negatively correlated. Based on this information most common places are getting reduced when we go far from Yonge-Dundas Square. It is suggestable to identify some neighbourhoods near to the Yonge-Dundas Square to start the business and attracting more people to the store.

3.5 Identifying more optimal location to start business

We found that the store has to be closer to Yonge-Dundas Square. Yet we need to find the optimal places where the people are gathering more to get more benefit of the location that we chose. Let we plot the kernel Density Diagram to get the more insights of the location data out of these 39 neighbourhoods and 1715 number of venues in old Toronto.

From this chart we can decide that the best place to start a business is

- Within 2 kms of Yonge-Dundas Square or
- Between 4 km to 6 km of Yonge-Dundas Square



3.6 More Insights around Yonge-Dundas Square

The above findings on density estimation data looks logically convencing, but strategically we are not able to provide any justification on the most common places are grouping between 2km to 4km.

Let we plot the geographical diagram of the most common venues available between these region of 0 - 2 km and 4 - 6 km. This may give us more insights on how the venues are placed.

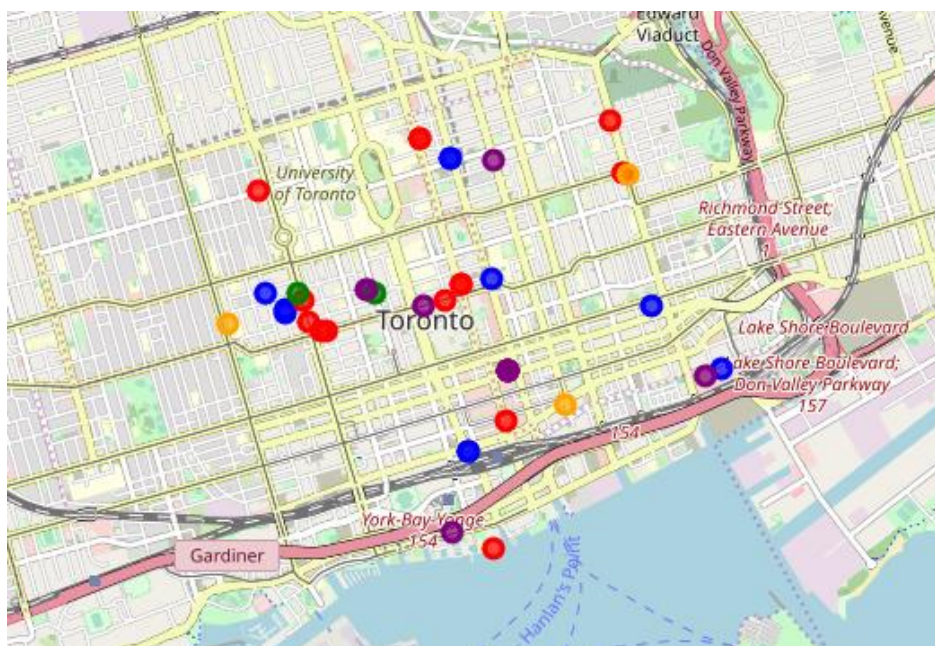


This explains us that all the venues located between 4 – 6 Kms are scattered around various places of Toronto. It is not the good decision to place the store between 4 – 6 kms from Yonge-Dundas Square. Hence the best place to start the identified business is within 2kms of Yonge-Dundas Square.

4. Results:

Based on the analysis carried out the comparison of New York and Toronto businesses gives us the knowledge on identifying the right business to start in Toronto. The demographics data states that both the cities are similar other than Toronto has more white and Chinese people. Hence all these identified businesses will be doing good in Toronto.

Linear Regression Model and Density Estimation diagram helps us to identify the right place where the most common venues are available and to decide the right location to place our store that can attract more customers. And now let we generate the geographical diagram of selected business to see the current competitors on the area.



● Chinese Restaurant ● Mexican Restaurant ● Donut Shot ● Ice Cream Shop ● Pharmacy

This diagram shows us the various competitors based on the business type on the region where our Customer wants to start the business.

5. Discussion:

With this analysis we have identified that below set of business are having great scope of profitability in Old Toronto.

- Chinese Restaurant
- Mexican Restaurant
- Donut Shop
- Ice Cream Shop
- Pharmacy

Manhattan and Old Toronto both are demographically very similar and the added population of white and Chinese community will play a major role in stable and sustained growth in terms of restaurant and donut shops. Even though Pharmacy is not in the portfolio of our customer it is also one of the good businesses to start with in Toronto. We have a lot of medical centres and Hospitals in the region but we have a very less number of pharmacies.

As our customer base portfolio matches with Restaurants and food industry, and can invest considerable amount of principal I suggest to start a Food Complex within the range of 1.5kms from Yonge-Dundas Square.

The Food Complex can hold both Donut and Ice cream Shop along with large space for chat and games at the ground floor. Mexican and Chinese restaurants at first floor. Toronto's half of the population is kids and young peoples, those will be attracted to donut and Ice cream shop, that will bring more families to the restaurant. And we can include the home deliveries. In a long run this will add more customers and more profit to the business.

I have observed that further analysis on below areas are necessary in a regular fashion to predict the future of the business.

- Identify future growth and shape of the business
- Identify other best provinces to expand the business

6. Conclusion

As part of this Data Science project we have identified the best strategic location and profitable business to start with in Toronto. We have seen the similarities and dissimilarities of Toronto and Manhattan. We still have identified future scope of analysis for our assured success in the industry on this dynamic and changing world.