BATTLE OF THE NEIGHBORHOODS: A COMPARISON OF TORONTO, NEW YORK AND PARIS

Ms. Thin Zar Myo Lwin

Applied Data Science Capstone Project Presentation

27th July 2021

1 BUSINESS PROBLEM

Business Problem

- This is an fact-finding analysis report intending to establish whether the city of Toronto is more similar to New York or Paris. The report is to recommend the entrepreneur, who is willing to open a restaurant, where they should start their own business among these three cities.

- Overall, these report will help the contractors or capitalists to make the correct decision regarding the business analysis to start a new one, including the different kinds of business and diversion destinations that a city might have to deal.

What is the Solution?

- Getting geometric data of New York, Toronto and Paris
- Exploring the locations of nearby venues in these cities
- Finding out which neighborhoods food venues are concentrated in Frequency of food venue types in neighborhoods
- Finding the boroughs and locations to open a restaurant
- Finding the most popular food venues in these cities

2 DATA ACQUISITION

NEW YORK

- **New York Neighborhood Data**:
  - *New York Neighborhood coordinates will be collected using the prepared geospatial coordinates New York City (JSON) file.*
- **Foursquare API:**
  - *The number, type, location, postal code and distance of the food venues in each neighborhood will be obtained using the Foursquare API.*

TORONTO

- **Toronto Neighborhood Data**:
  - *Information such as the postal code, borough, and neighborhood of the city of Toronto will be collected by scraping the Wikipedia page.*
- **Geospatial Coordinates:**
  - *Toronto Neighborhood coordinates will be collected using the prepared geospatial coordinates csv file.*
- **Foursquare API:**
  - *The number, type, location, postal code and distance of the food venues in each neighborhood will be obtained using the Foursquare API.*

PARIS

- **Paris Neighborhood Data**:
  - *Information such as the postal code, borough, and neighborhood of the city of Paris will be collected by scraping the Wikipedia page.*
- **Foursquare API:**
  - *The number, type, location, postal code and distance of the food venues in each neighborhood will be obtained using the Foursquare API.*

3 METHODOLOGY

ANALYSIS TECHNIQUE

- **City Analysis**:
  - *To discover how many place of venues are there, based on this result we can know the venue measure of each city*
  - *examine the diversity of venues in each city, which may tell us how varied the urban venues can be and how interrelated between cities.*
- **Neighborhood Analysis:**
  - *to investigate the top 10 venues in each district using Foursquare figures*
  - *To discover the neighborhood market may have adequate capacity and keenness to opportune a new restaurant if the previously current restaurants here are located.*
- **Clustering Analysis:**
  - *clustering (using k-mean clustering) will be established to offer stakeholders different alternatives in terms of neighborhood.*

4 DATA ANALYSIS

CITY ANALYSIS

- **The number of records based on the city group**:
  - *According to the analysis, we found out that the density of the venues per unit of the neighborhood and the numbers are 33, 20 and 66 in New York, Toronto and Paris respectively.*
  - *The following figure is shown the number of records grouping by the City and there are 10501, 2122 and 1921 boroughs in the cities respectively.*

```
[343]: analyze_venue_cateogry_df.groupby('City').count().sort_values(by=['Venue Category'], ascending=False).head()
```

[343]:

City	Borough	City_code	Latitude	Longitude	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Postal Code	Venue	Venue Category	Venue Latitude	Venue Longitude
New_york	10501	10501	10501	10501	10501	10499	10499	0	10499	10499	10499	10499
Toronto	2122	2122	2122	2122	2122	2119	2119	2122	2119	2119	2119	2119
Paris	1921	1921	1921	1921	1921	1921	1921	0	1921	1921	1921	1921

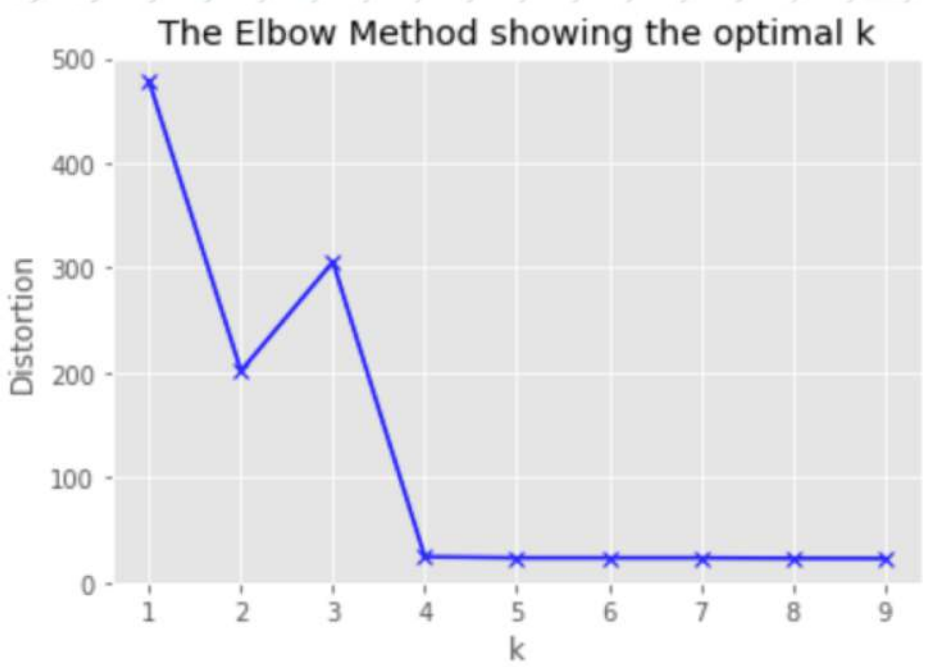
NEIGHBORHOOD ANALYSIS

- **The 10 most common venue in New York, Toronto and Paris:**
  - *Based on these data, the report shows Pizza Place is the most popular place in New York and French restaurant and Coffee shop for Paris and Toronto city correspondingly.*

	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	New_york	Pizza Place	Coffee Shop	Italian Restaurant	Deli / Bodega	Bakery	Chinese Restaurant	Bar	Grocery Store	Park	Sandwich Place
1	Paris	French Restaurant	Hotel	Italian Restaurant	Bakery	Bar	Café	Bistro	Japanese Restaurant	Plaza	Coffee Shop
2	Toronto	Coffee Shop	Café	Restaurant	Park	Bakery	Pizza Place	Italian Restaurant	Japanese Restaurant	Sandwich Place	Clothing Store

CLUSTERING ANALYSIS

- **The Elbow Method showing the optimal K:**
  - *K-means clustering analysis method is used to classify the area/venue to know the nature of the neighborhoods. First, I will run K-Means to cluster the boroughs into 4 clusters because when I analyze the K-Means with elbow method it ensured me the 4 degree for optimum k of the K-Means.*

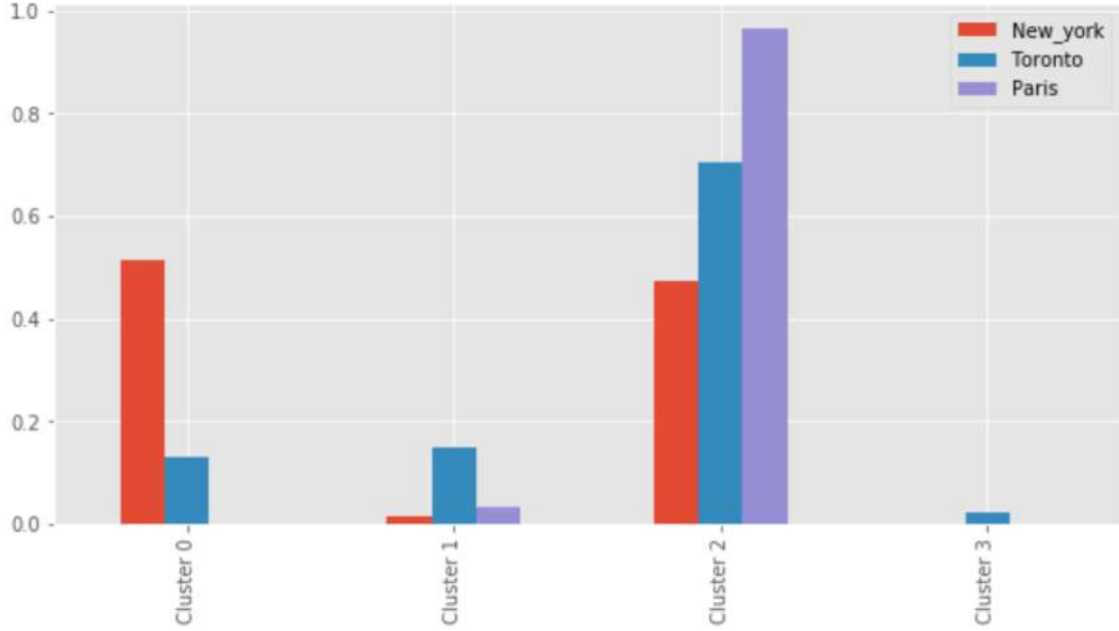


CLUSTERING ANALYSIS

- **Clustering the Neighborhood**:
  - *We used K-means clustering analysis method to classify the area/venue to know the nature of the neighborhoods. According to the clustering analysis, we got four clusters and identify the area names based on the venues data.*
- **Cluster 1** => **Urban Pioneer (up-and-coming) Area** where is near downtown area and inner-ring suburbs.
- **Cluster 2** => **Residential Area** where large single family homes are situated and this is active neighborhood.
- **Cluster 3** => **Urban Core Downtown Area** where is downtown area which is the heart of the major metros. This area is close to nightlife and has the city attraction. There is little to no public parking, typically has higher rates of crimes and transients.
- **Cluster 4** => **Suburbs Area** where is far from the downtown area of the city and it has space and privacy.

```
[569]: new_cluster_df.plot(kind='bar',figsize=(10,5))

[569]: <matplotlib.axes._subplots.AxesSubplot at 0x7fd1729e94e0>
```



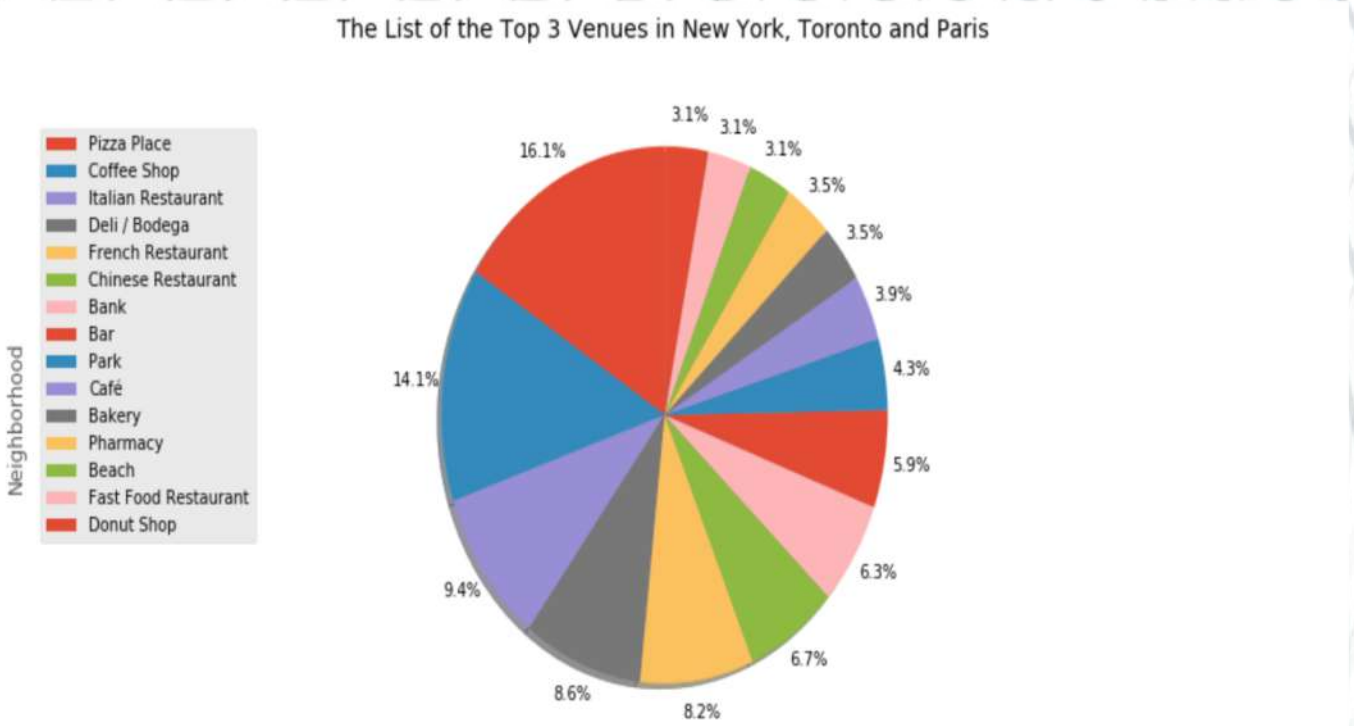
The bar chat for four clusters in New York, Toronto and Paris

5 RESULTS & DISCUSSION

TOP 3 VENUES ANALYSIS

- **Top 3 popular venues in the cities** :

- Only from this evaluation it is assumed that these zones are appreciated regions for food locations. Although these are appropriate place, it may be a suitable choice for some food venues.
- Based on the analysis, we found that Pizza shop, French restaurant and Coffee shop are most popular places in New York, Paris and Toronto respectively.



The list of the top 3 venues in New York, Toronto and Paris

NEW YORK CLUSTERS ANALYSIS

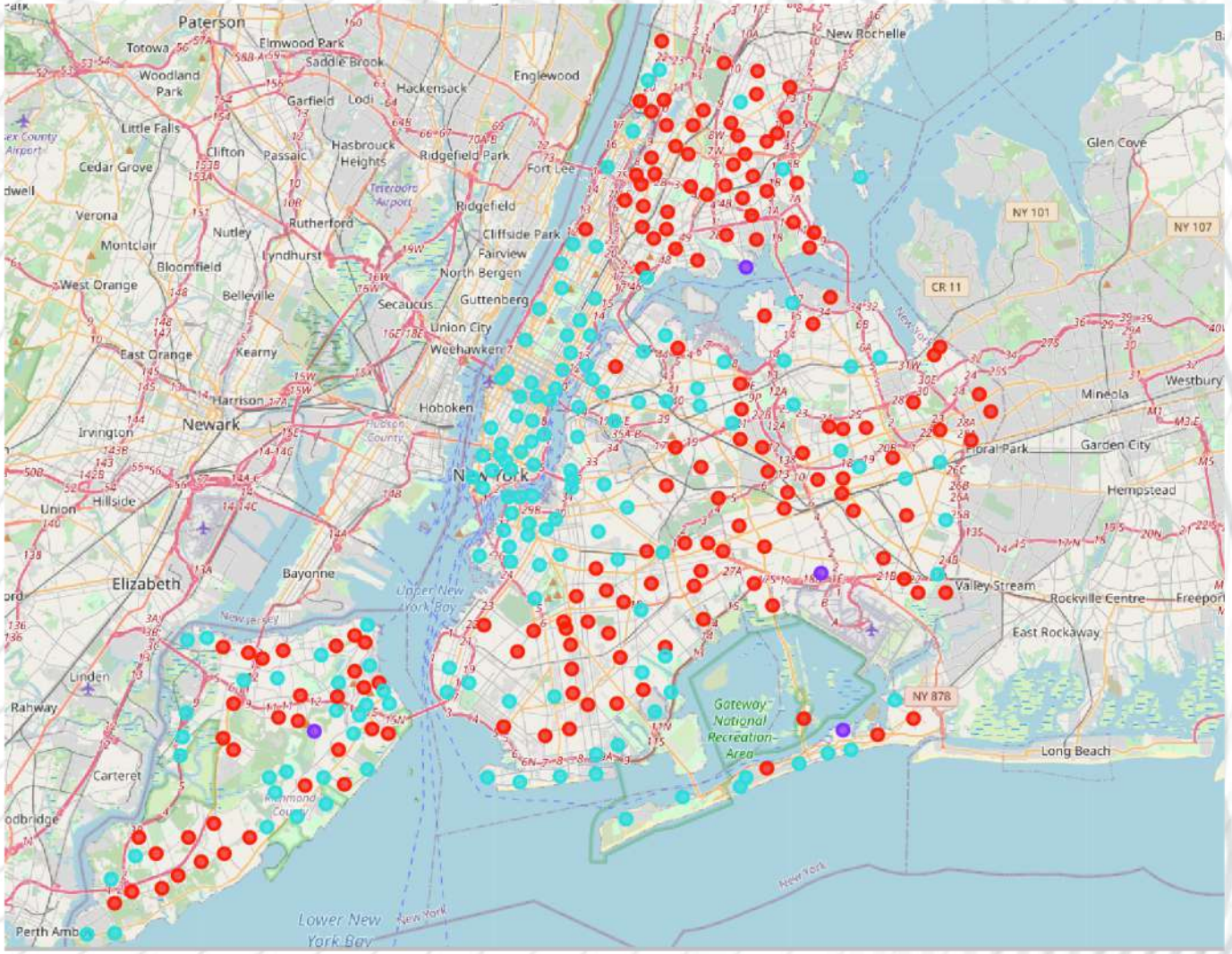
- **Data Analysis result of New York City :**
  - *Based on the analysis, we defined the following boroughs in the respective clusters / defined areas. According to the analysis, New York is defined in the three clusters.*

- **Cluster 1** (Red marker in the map) **:** We can define the **some areas of Bronx, the Brooklyn, Queens, Staten Island** are in the **clusters 1** and these areas are urban pioneer area so that the entrepreneurs may open their restaurants in this area.

- **Cluster 2** (Purple marker in the map) **:** We can define the **some areas of Bronx, Queens, Staten Island** are in the **clusters 2** and these areas are residential area so that the convenience store and food truck is more suitable to open there.

- **Cluster 3** (Cyan color in the map) **:** We can define the **some areas of Brooklyn and Manhattan** are in the **clusters 3** and these areas are heart of city downtown core area so that the entrepreneurs should open their restaurants in this area to make profit. We recommend this area is the best suitable and profitable place to start up the business. However, we must consider the other factors such as rent, competitors, etc.

- **Cluster 4 :** There is no neighborhoods in this cluster area.



The map is shown by cluster analysis result in New York City

TORONTO CLUSTERS ANALYSIS

Data Analysis result of Toronto City :

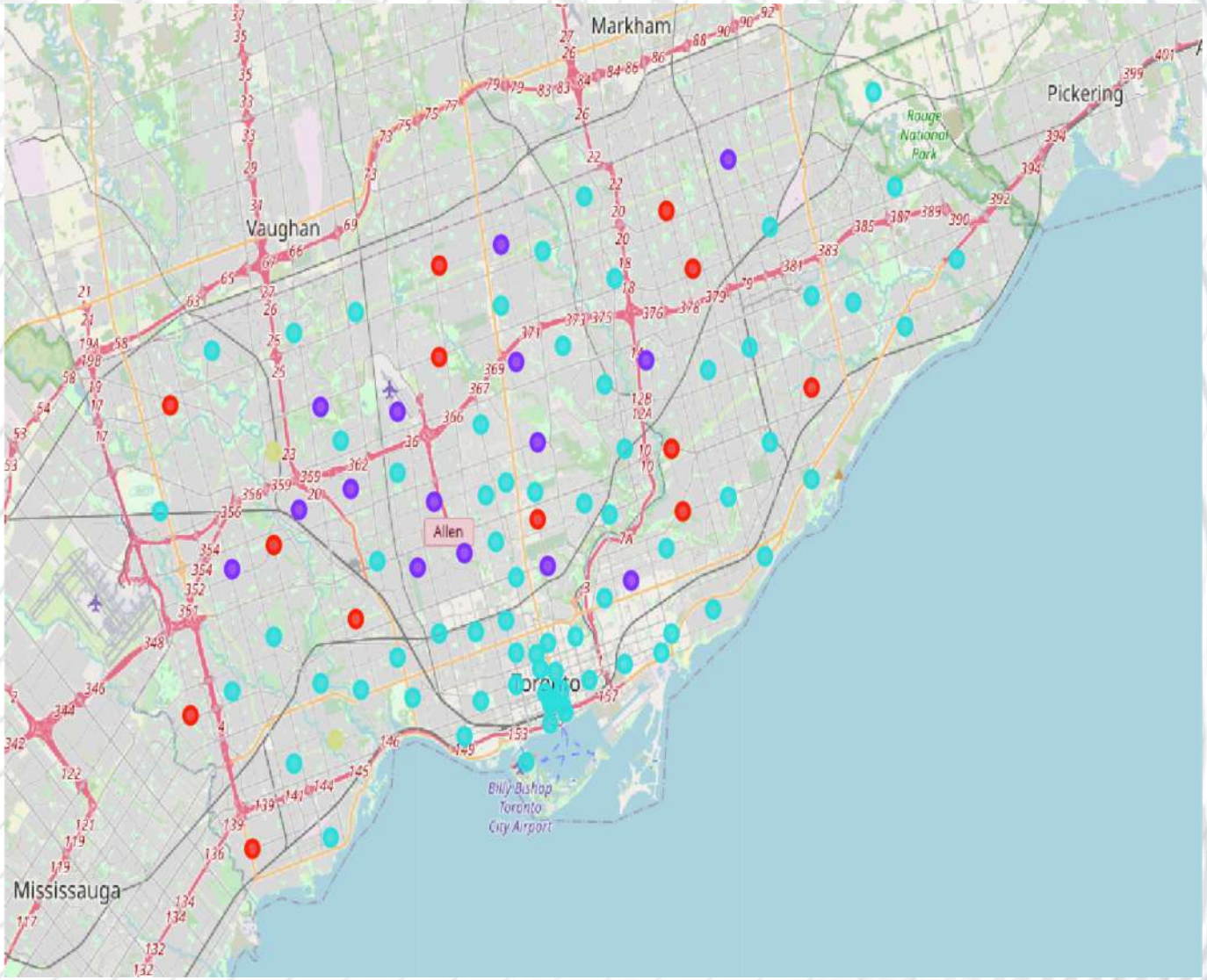
- *Based on the analysis, we defined the following boroughs in the respective clusters / defined areas. According to the analysis, Toronto is defined in the four clusters.*

- **Cluster 1** (Red marker in the map)**:** We can define the **some areas of Etobicoke, North York, Scarborough** are in the **clusters 1** and these areas are urban pioneer area.

- **Cluster 2** (Purple marker in the map)**:** We can define the **some areas of North York and York** are in the **clusters 2** and these areas are residential area so that the grocery store and food truck is more suitable to open there.

- **Cluster 3** (Cyan marker in the map)**:** We can define the **some areas of Downtown Toronto, North York, East Toronto, Scarborough, West Toronto, Central Toronto** are in the **clusters 3** and these areas are core area so that the entrepreneurs should open their restaurants in this area to make profit. This is the recommended place for an entrepreneurs and business owners.

- **Cluster 4** (Gold color marker in the map)**:** We can define the **some areas of Etobicoke and North York** are in the **clusters 4** and these areas are suburb area so that we cannot make much profit to start the business there.



The map is shown by cluster analysis result in Toronto City

PARIS CLUSTERS ANALYSIS

Data Analysis result of Paris City :

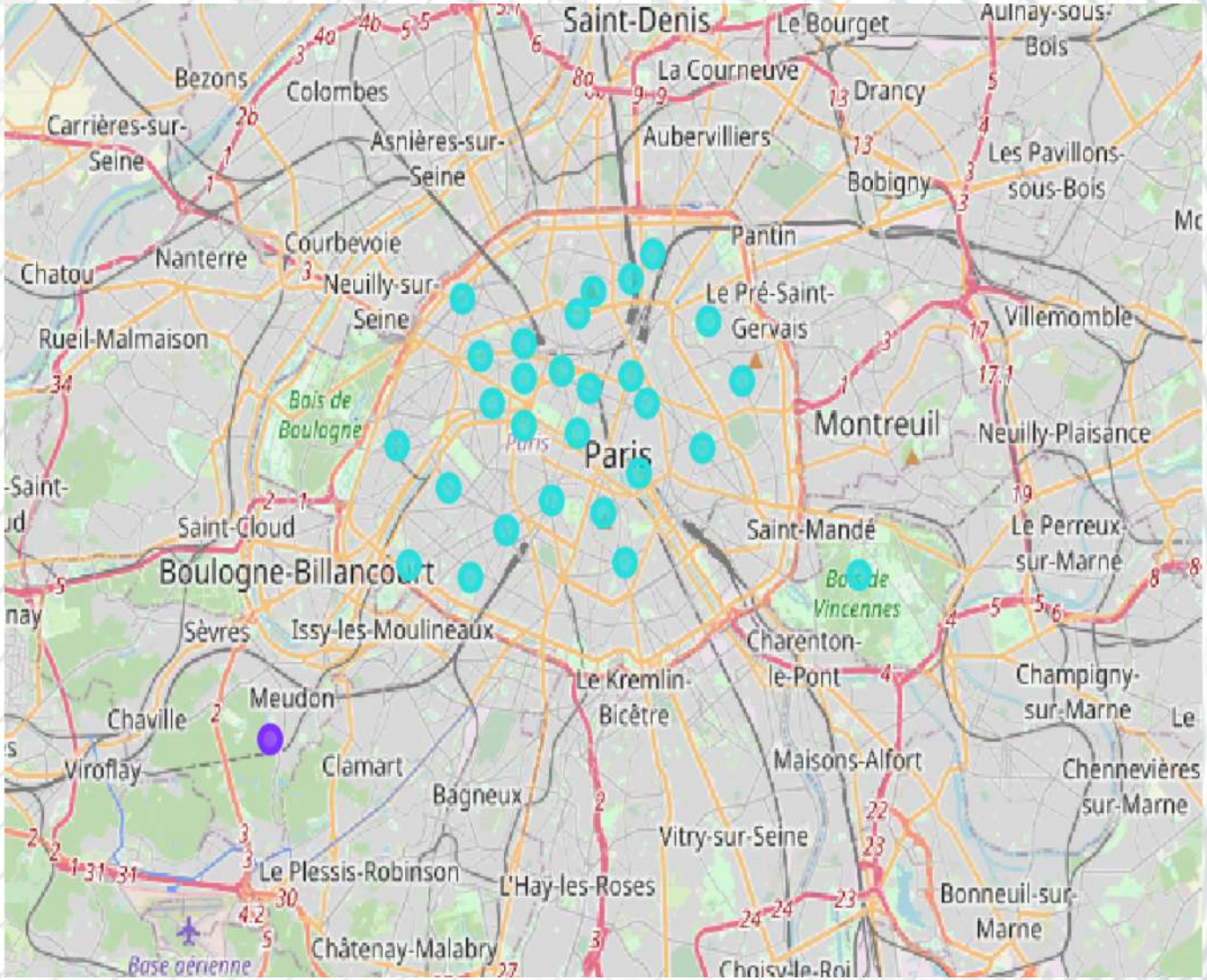
- *Based on the analysis, we defined the following boroughs in the respective clusters / defined areas. According to the analysis, Paris is defined in the two clusters.*

- **Cluster 1 :** There is no neighborhoods in this cluster area.

- **Cluster 2** (Purple maker in the map) **:** We can define the **Obervatoire** is in the **clusters 2** and these areas are residential area so that the convenience store and food truck is more proper to open there.

- **Cluster 3** (Cyan maker in the map) **:** We can define the **areas of Butte-Montmartre, Elysee, Vaugiard** are in the **clusters 3** and these areas are city downtown core area where is the recommended place which is the best appropriate and commercial place to start up the business.

- **Cluster 4 :** There is no neighborhoods in this cluster area.



The map is shown by cluster analysis result in Paris City

BUSIEST AREA

(The borough/ location where a restaurant/business should be started)

- Based on the data analysis, we determined these following busiest area in terms of food venues in neighborhoods connected to the downtown area of Toronto, New York City and Paris, where are suitable to open a restaurant and this place is the recommended one.
  - *specific areas of New York (**Brooklyn, Manhattan**)*
  - *Toronto ('Downtown **Toronto','East Toronto','West Toronto','Central Toronto**') and*
  - *Paris (**'Butte-Montmartre','Elysee','Vaugirard**')

6 CONCLUSION

CONCLUSION

- The goal of this report was to help investors trying to find venues to open a restaurant or start up a business in the city of New York, Toronto and Paris. There are many causes that affect the situation of a food venue. In this analysis, studies were prepared considering only the location of the competition factor.
- It should be noticed that the investors should pay attention to all aspects such as human resources, rental fees, demography, labor costs and minimum wage, health regulations, etc. when selecting the place for the food venue

THANK YOU