**Battle of Neighborhood – Relocating to Toronto**

1. **Introduction**

**1.1 Background**

A big company in New York is relocating to Toronto Canada as part of its expansion. The company offered the employees who would like to move to Toronto same salary package and relocation assistance.

* 1. **Problem**

As part of the relocation assistance, the company would like to do the research and provide the employees analysis on the Toronto neighborhood details, to take that stress of the employees and the employees will have the readymade analysis for them on the Toronto neighborhoods so they can make the decision on which neighborhood would be their new home.

* 1. **Interest**

All the employees who are considering moving to Toronto along with the company would benefit from this analysis as it will provide them with necessary groundwork to decide which neighborhood suits their needs the best

**2. Data acquisition and cleaning**

**2.1 Data sources**

To get the neighborhood data we will use the below wikipedia page to get all the boroghs and neighborhoods and their latitude and longitudes.

<https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>

<http://cocl.us/Geospatial_data/Geospatial_coordinates.csv>

**2.2 Data Cleaning**

Extracted the Boroughs information table from the XML using BeautifulSoup package

Cleaned the data by deleting 'Not Assigned' Borough. If Neighborhood is blank, it was set same as borough. Merged the neighborhood with the same postalcode. Cleaned up the duplicates.

Using the Postalcode from the table and http://cocl.us/Geospatial\_data/Geospatial\_coordinates.csv file, get latitude and longitude for each Borough and Neighborhood

Using FourSquare API got a list of 100 nearby venues within 500m radius for each neighborhood

1. **Methodology**

The approach we took to explore and cluster the neighborhoods is as follows

#### For all the neighborhoods in Toronto, we checked how many venues were returned for each neighborhood

#### 

#### Then found out how many unique categories can be curated from all the returned venues

#### 

#### Then we grouped the data by neighborhood and by taking the mean of the frequency of occurrence of each category.

#### 

#### 

#### Then we Run k-means to cluster the neighborhood into 5 clusters

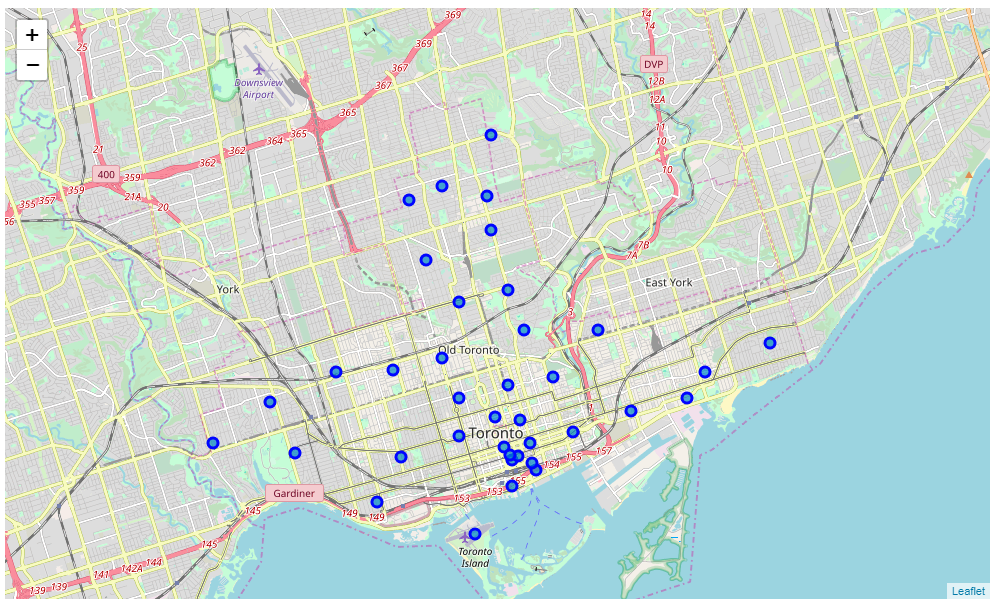
#### And created a new dataframe that includes the cluster as well as the top 10 venues for each neighborhood.

#### 

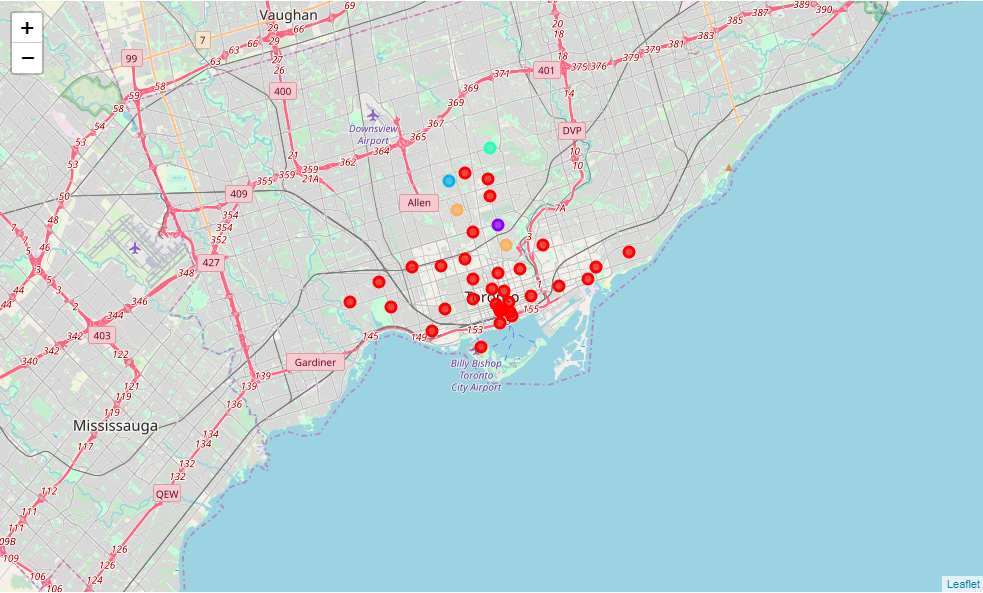
1. **Results**

Below are different neighborhoods in Toronto

|  |  |  |
| --- | --- | --- |
| **PostalCode** | **Borough** | **Neighborhood** |
| M4S | Central Toronto | Davisville |
| M4P | Central Toronto | Davisville North |
| M5P | Central Toronto | Forest Hill North, Forest Hill West |
| M4N | Central Toronto | Lawrence Park |
| M4T | Central Toronto | Moore Park, Summerhill East |
| M4R | Central Toronto | North Toronto West |
| M5N | Central Toronto | Roselawn |
| M4V | Central Toronto | Summerhill West, South Hill, Rathnelly, Deer P... |
| M5R | Central Toronto | Yorkville, The Annex, North Midtown |
| M5H | Downtown Toronto | Adelaide, Richmond, King |
| M5E | Downtown Toronto | Berczy Park |
| M4X | Downtown Toronto | Cabbagetown, St. James Town |
| M5G | Downtown Toronto | Central Bay Street |
| M5T | Downtown Toronto | Chinatown, Grange Park, Kensington Market |
| M6G | Downtown Toronto | Christie |
| M4Y | Downtown Toronto | Church and Wellesley |
| M5S | Downtown Toronto | Harbord, University of Toronto |
| M5A | Downtown Toronto | Regent Park, Harbourfront |
| M4W | Downtown Toronto | Rosedale |
| M5B | Downtown Toronto | Ryerson, Garden District |
| M5V | Downtown Toronto | South Niagara, Island airport, Railway Lands, ... |
| M5C | Downtown Toronto | St. James Town |
| M5W | Downtown Toronto | Stn A PO Boxes 25 The Esplanade |
| M5K | Downtown Toronto | Toronto Dominion Centre, Design Exchange |
| M5J | Downtown Toronto | Toronto Islands, Harbourfront East, Union Station |
| M5X | Downtown Toronto | Underground city, First Canadian Place |
| M5L | Downtown Toronto | Victoria Hotel, Commerce Court |
| M7Y | East Toronto | Business Reply Mail Processing Centre 969 Eastern |
| M4K | East Toronto | Riverdale, The Danforth West |
| M4M | East Toronto | Studio District |
| M4E | East Toronto | The Beaches |
| M4L | East Toronto | The Beaches West, India Bazaar |
| M6K | West Toronto | Brockton, Exhibition Place, Parkdale Village |
| M6H | West Toronto | Dovercourt Village, Dufferin |
| M6P | West Toronto | High Park, The Junction South |
| M6J | West Toronto | Little Portugal, Trinity |
| M6R | West Toronto | Parkdale, Roncesvalles |
| M6S | West Toronto | Swansea, Runnymede |



Clustered Neighborhoods



Cluster 1 :



**Cluster 2 :**

**Cluster 3:**



**Cluster 4:**

**Cluster 5:**



1. **Discussion**

The data collection and analysis gives a good understanding of the neighborhood and its vicinity. With this detail analysis, we can see each neighborhood and the top venues in those neighborhoods.

* If the client needs proximity to pharmacy, super market, bank, park then West Toronto neighborhood is the best
* If the client has small kids and needs proximity to kids activities then Downtown toronto is good neighborhood
* If the client needs proximity to Airport, may be because he still has family back in New York and needs to visit them frequently, then Downtown Toronto is good neighborhood
* If the client has pets, East Toronto is a good neighborhood.

**6. Conclusion**

In this study and analysis,all the neighborhoods in Toronto were explored in detailed to give a complete picture of the neighborhood and the venues nearby. It also gave a detailed analysis of the top 10 venue categories. This will help the client decide which neighborhood is best for his/her specific needs. This study will surely take some stress out of employees relocation anxiety and help them deal with the stress better.