



CAPSTONE PROJECT

Final Report

VEGETARIAN/VEGAN RESTAURANT EXPLORER

Find a suitable place to live where there are vegetarian restaurants nearby in Downtown Toronto

Report by Vasanth T

Contents

Introduction/Business Problem	2
Data Section	2
Downtown Toronto – Data	2
Downtown Toronto – Map	3
Methodology.....	3
Clustering Approach through K-Means	3
Code Snippet	3
Most Common Venues near the Neighbourhood	4
Results Section	4
Visualization of the Clusters	4
Cluster Suitable for Raj	5
Discussion Section	5
Conclusion Section	5

Introduction/Business Problem

Raj, who is based out of Chennai, India, is working in an MNC IT firm. He is unmarried and stays at a shared accommodation with his friends.

Based on his performance, he is provided with an opportunity to migrate to Canada on work permit for a prestigious multi-year project, which is expected to last 10 years. The company has the office setup in Toronto (Downtown Toronto). Raj is very much excited about this project and migration to Canada, but at the same time, he is concerned about the food choices that Toronto has to offer, as he is a pure vegetarian by birth.

His only concern is that he wants to find a place where his place of stay is very close to Vegetarian/Vegan restaurants in particular.

The objective of this project is to solve Raj's problem, by exploring and comparing the nearby places, where there are Vegetarian/Vegan restaurants nearby, where he can put himself up, during his stay at Canada

Data Section

Our data source would be [Canada Postal Codes](#)

We will scrap the data pertaining to the location "Downtown Toronto" and use the following libraries to resolve the problem:

- ✓ JSON/XML
- ✓ Foursquare
- ✓ Pandas
- ✓ Scikit Learn
- ✓ Folium
- ✓ Geocoder
- ✓ Beautiful Soup, Requests
- ✓ Matplotlib: Python Plotting Module

Downtown Toronto – Data

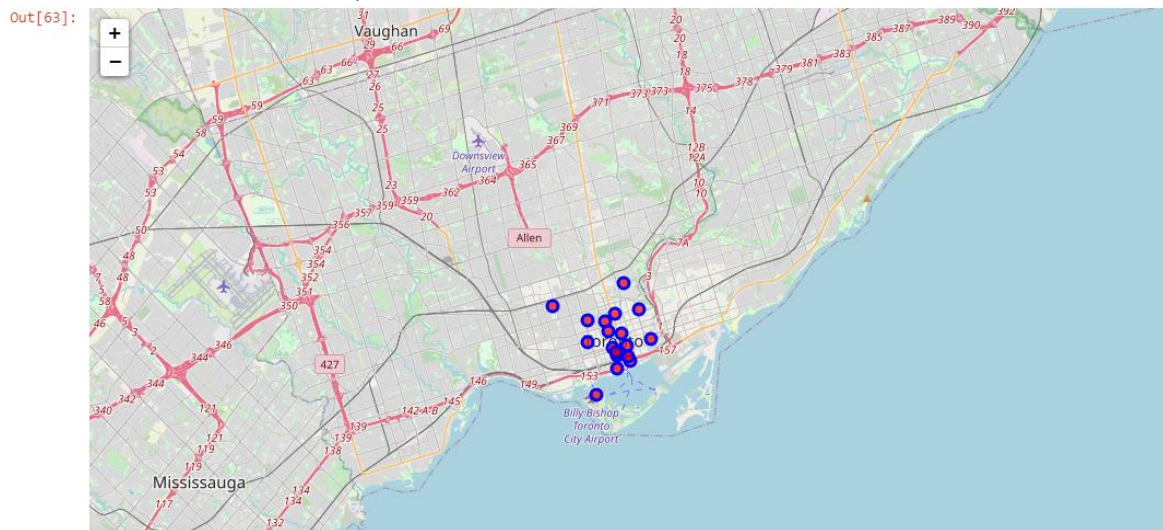
Isolating the data related to Downtown Toronto into a separate data frame

```
In [6]: dt_toronto = df[df['Borough'] == 'Downtown Toronto'].reset_index(drop=True)
dt_toronto.head(5)
```

Out[6]:

	Postal Code	Borough	Neighbourhood	Latitude	Longitude
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
1	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494
2	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937
3	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418
4	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306

Downtown Toronto – Map



Methodology

Clustering Approach through K-Means

In order to zero-in on the neighbourhood where Vegetarian/Vegan restaurants are nearby, we need to group the venues into clusters through a clustering approach. The approach that is commonly used is K-Means algorithm which is a kind of unsupervised learning.

We also used Foursquare API for mining the near-by restaurants of the neighbourhoods.

Code Snippet

Here is a code snippet that creates 7 different clusters with associated labelling

Clustering the neighbourhoods by setting the number of clusters as 7

```
In [80]: # setting the number of clusters as 7
kmean_clusters = 7

dt_toronto_clustering = dt_toronto_onehot_grouped.drop('Neighbourhood', 1)

# Let's run k-means clustering
k_means = KMeans(n_clusters=kmean_clusters, random_state=0).fit(dt_toronto_clustering)

# Let's check cluster labels generated for each row in the dataframe
k_means.labels_[0:10]
#dt_toronto_clustering
```

Out[80]: array([0, 5, 3, 2, 0, 0, 5, 0, 1, 6], dtype=int32)

```
In [40]: # assigning the value to a new variable so that the original is unaltered when we add cluster Labels
neighbourhoods_venues_sorted_with_labels = neighbourhoods_venues_sorted
neighbourhoods_venues_sorted_with_labels.insert(0, 'Cluster Labels', k_means.labels_)
neighbourhoods_venues_sorted_with_labels.head(5)
```

Out[40]:

	Cluster Labels	Neighbourhood	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	0	Berczy Park	Restaurant	Seafood Restaurant	Comfort Food Restaurant	Greek Restaurant	Vegetarian / Vegan Restaurant	Japanese Restaurant	Eastern European Restaurant	French Restaurant	Sushi Restaurant	Thai Restaurant
1	5	Central Bay Street	Italian Restaurant	Thai Restaurant	Japanese Restaurant	Portuguese Restaurant	Falafel Restaurant	Indian Restaurant	Vegetarian / Vegan Restaurant	Korean Restaurant	Middle Eastern Restaurant	Modern European Restaurant
2	3	Christie	Italian Restaurant	Restaurant	Dumpling Restaurant	Gluten-free Restaurant	German Restaurant	French Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant
3	2	Church and Wellesley	Sushi Restaurant	Japanese Restaurant	Restaurant	Fast Food Restaurant	Mediterranean Restaurant	Italian Restaurant	Mexican Restaurant	Caribbean Restaurant	Chinese Restaurant	Indian Restaurant
4	0	Commerce Court, Victoria Hotel	Restaurant	Italian Restaurant	Seafood Restaurant	Japanese Restaurant	American Restaurant	Thai Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant	Gluten-free Restaurant	Molecular Gastronomy Restaurant

Most Common Venues near the Neighbourhood

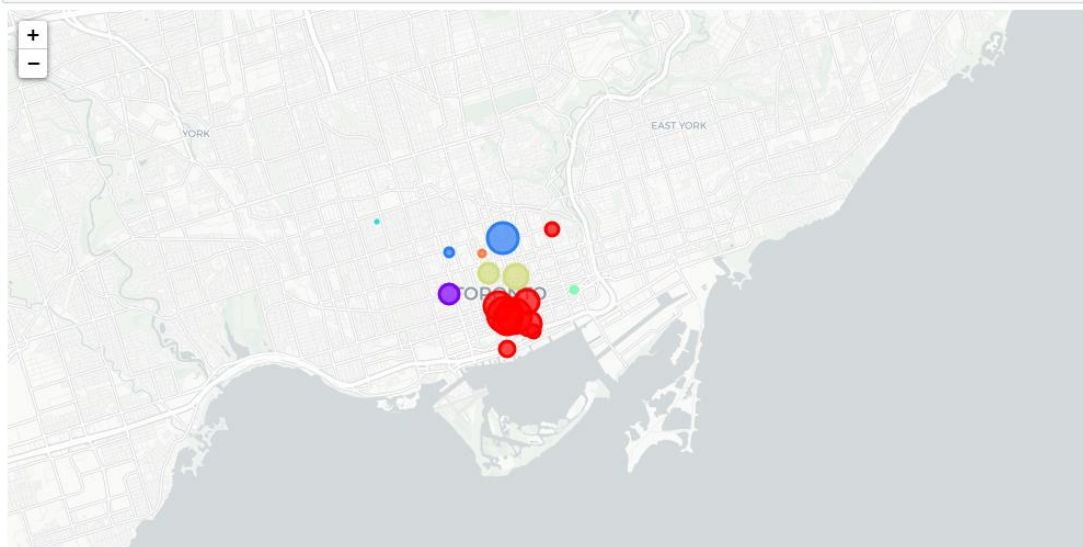
Out[83]:

	Postal Code	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	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
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	4	Restaurant	Asian Restaurant	French Restaurant	Mexican Restaurant	Greek Restaurant	Vietnamese Restaurant	Eastern European Restaurant
1	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	6	Sushi Restaurant	Italian Restaurant	Japanese Restaurant	Mexican Restaurant	Eastern European Restaurant	German Restaurant	French Restaurant
2	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	5	Italian Restaurant	Japanese Restaurant	Middle Eastern Restaurant	Fast Food Restaurant	Ramen Restaurant	Modern European Restaurant	Chinese Restaurant
3	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	0	Italian Restaurant	Seafood Restaurant	Moroccan Restaurant	Restaurant	American Restaurant	Comfort Food Restaurant	Japanese Restaurant
4	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306	0	Restaurant	Seafood Restaurant	Comfort Food Restaurant	Greek Restaurant	Vegetarian / Vegan Restaurant	Japanese Restaurant	Eastern European Restaurant

Results Section

Visualization of the Clusters

Out[84]:



Cluster Suitable for Raj

Of all the 7 clusters, we see that Cluster 2 has the second most common venue as Vegetarian/Vegan restaurant. So, we can conclude that this cluster could be an ideal place for Raj's stay.

Examining Cluster 2

```
In [85]: #examining cluster 2
dt_toronto_merged_updated.loc[dt_toronto_merged_updated['Cluster Labels'] == 1, dt_toronto_merged_updated.columns[[1] + list(range(2, dt_toronto_merged_updated.shape[1]))]]
```

Out[85]:

	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	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
12	Downtown Toronto	Kensington Market, Chinatown, Grange Park	43.653206	-79.400049	1	Vietnamese Restaurant	Vegetarian / Vegan Restaurant	Mexican Restaurant	Caribbean Restaurant	Belgian Restaurant	Japanese Restaurant	Comfort Food Restaurant

Discussion Section

The objective of this project is to solve Raj's problem, by exploring and comparing the nearby places, where there are Vegetarian/Vegan restaurants nearby, where he can put himself up, during his stay at Canada.

The availability of data did create a bottleneck as the Wikipedia source for Toronto postal codes suddenly changed midway, and I have to look out for alternatives that consumed quite a lot of time. I then explored the history of that Wiki page and was able to find the previous version, which I ultimately used in completing this project.

That minor glitch apart, throughout this project, problem solving was a breeze mainly due to the plethora of libraries that Python has offered for various tasks like, mining, extracting, transforming and visualizing the data. I gained tremendous amount of experience throughout the course and I thank Coursera for that.

Conclusion Section

The business problem identified during the beginning of the project has been successfully achieved. Raj can now have a sigh of relief as he no longer has to worry about his place of dwelling. Raj must definitely be pleased with the outcome of this project.

Thank You !