Coursera Capstone – The Battle of Neighborhood

Recommendation to open an International Grocery Store in Toronto

Introduction:

According to Wikipedia "**Toronto** is an international center of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world. The diverse population of **Toronto** reflects its current and historical role as an **important** destination for immigrants to Canada." This diversity helps to run an international grocery store in the Toronto. This store will bring together international famous food items together.

Problem Description:

According to blogto.com website **Toronto's** population is made up of 51 per cent of residents born outside of Canada, but is also home to 230 different **nationalities**, which makes it far more diverse than Dubai. As more national people live in **Toronto's,** so it's a great opportunity for us to open big grocery store. More international people will buy more grocery items at the same time we will make give more emphasis to educate about international food. Purpose of this report is to proposed a profitable proposal to open an internal grocery store.

Data Explanation:

We need to know more about the City of Toronto's. we have to make the segmentation, clustering about City. To make this research more convenient, we scrap data from the Wikipedia web page. So, we scrap the data than organized the data in an structed way.

1. We implement geocoding to retrieve latitude and longitude of Toronto's neighborhood. Geo data represent, organized and visualize the Toronto's neighborhood.

Explanation of secondary data:

```
#merge two table on the column "PostalCode"
Toronto_df_new = Toronto_df_grouped.merge(coordinates, on="PostalCode", how="left")
Toronto_df_new.head(20)
```

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M1B	Scarborough	Rouge, Malvern, Rouge, Malvern	43.806686	-79.194353
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union, Highla...	43.784535	-79.160497
2	M1E	Scarborough	Guildwood, Morningside, West Hill, Guildwood, ...	43.763573	-79.188711
3	M1G	Scarborough	Woburn, Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae, Cedarbrae	43.773136	-79.239476
5	M1J	Scarborough	Scarborough Village, Scarborough Village	43.744734	-79.239476

2. Population data helps to make segmentation of population country wise. The higher the number of populations from a specific country or region, we have to concentrate food items from that region. Population data also give good understanding of demographic features of population. So, we can look at region wise venue in Toronto's to understand the number of regional venues in that region.

```
Toronto_merged.loc[Toronto_merged['Cluster Labels'] == 2,
 Toronto_merged.columns[[1] + list(range(5, Toronto_merged.shape[1]))]]
```

	Borough	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	East Toronto	2	Pub	Trail	Asian Restaurant	Neighborhood	Health Food Store	Dumpling Restaurant	Donut Shop	Eastern European Restaurant	Department Store	Doner Restaurant
23	Central Toronto	2	Jewelry Store	Trail	Mexican Restaurant	Sushi Restaurant	Yoga Studio	Dim Sum Restaurant	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant

3. Another key indicator is the venue latitude and venue longitude. We are going to get an overall picture nearby venue and postal code.

```
# convert the venues list into a new DataFrame
venues_df = pd.DataFrame(venues)

# define the column names
venues_df.columns = ['PostalCode', 'Borough', 'Neighborhood', 'BoroughLatitude', 'BoroughLongitude', 'VenueName', 'VenueLatitude', 'VenueLongitude', 'VenueCategory']

print(venues_df.shape)
venues_df.head(10)
```

```
(1714, 9)
```

	PostalCode	Borough	Neighborhood	BoroughLatitude	BoroughLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	M4E	East Toronto	The Beaches, The Beaches	43.676357	-79.293031	Glen Manor Ravine	43.676821	-79.293942	Trail
1	M4E	East Toronto	The Beaches, The Beaches	43.676357	-79.293031	The Big Carrot Natural Food Market	43.678879	-79.297734	Health Food Store
2	M4E	East Toronto	The Beaches, The Beaches	43.676357	-79.293031	Grover Pub and Grub	43.679181	-79.297215	Pub
3	M4E	East Toronto	The Beaches, The Beaches	43.676357	-79.293031	Upper Beaches	43.680563	-79.292869	Neighborhood
4	M4E	East Toronto	The Beaches, The Beaches	43.676357	-79.293031	Seaspray Restaurant	43.678888	-79.298167	Asian Restaurant
5	M4K	East Toronto	The Danforth West, Riverdale, The Danforth Wes...	43.679557	-79.352188	Pantheon	43.677621	-79.351434	Greek Restaurant
6	M4K	East Toronto	The Danforth West, Riverdale, The Danforth Wes...	43.679557	-79.352188	MenEssentials	43.677820	-79.351265	Cosmetics Shop
7	M4K	East Toronto	The Danforth West, Riverdale, The Danforth Wes...	43.679557	-79.352188	Cafe Fiorentina	43.677743	-79.350115	Italian Restaurant
8	M4K	East Toronto	The Danforth West, Riverdale, The Danforth Wes...	43.679557	-79.352188	Mezes	43.677962	-79.350196	Greek Restaurant
9	M4K	East Toronto	The Danforth West, Riverdale, The Danforth Wes...	43.679557	-79.352188	Messini Authentic Gyros	43.677827	-79.350569	Greek Restaurant