Measuring the Level of Convenience of a Neighborhood

Introduction

This assignment will use the Foursquare location data to solve a problem in Toronto, Canada.

Business Problem

Toronto is divided into various neighborhoods. In this assignment, the neigbourhoods are defined by the postal-codes. In this assignment, we are trying out to find out which neighbourhoods are convenient to live, and present the results on the map of Toronto.

How to measure convenience

We can define the meaning of convenience by looking at the convenience stores. What makes a store a convenience stores? It is about the variety of goods that sold in the stores. Borrowing this concept, we can say that a neighbor is convenient if there are different variety of shops in it.

How to measure variety

We can borrow the concept of entropy from Statistical Mechanics.

The entropy is defined as S = $-\sum_{i=1}^{n} p_i log p_i$

Here, p_i is the frequency of each location type.

Approach

For each neighborhood, its entropy is calculated. The value represent the variety of locations. In the map, for each neighborhood, a circle is drawn and the size of the circle is related to its entropy.

Source Data for the Postal Code of Toronto

Here is the URL:

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

Latitude and Longitude of Neighbors.

A data-frame is prepared with the latitude and longitude of each neighbor.

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

Location Data for Neighbors

For each neighbor, the location data from Foursquare is retrieved to count the number of stores in each category.

	Neighbourhood	Accessories Store	Adult Boutique	Afghan Restaurant	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	...	Train Station	Vegetarian / Vegan Restaurant	Video Game Store	Video Store	Vietnamese Restaurant	Warehous Stor
0	Adelaide, King, Richmond	0	0	0	0	0	0	0	0	0	...	0	1	0	0	0	
1	Agincourt	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	
2	Agincourt North, L'Amoreaux East, Milliken, St...	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	
3	Albion Gardens, Beaumond	0	0	0	0	0	0	0	0	0	...	0	0	0	1	0	

Measure the Entropy for Each Neighbor

For each neighbor, the entropy value is calculated.

The entropy is defined as $S = -\sum_{i=1}^{n} p_i \log p_i$

Here, p_i is the frequency of each location type.

	PostalCode	Borough	Neighbourhood	Latitude	Longitude	entropy
0	M1B	Scarborough	Rouge, Malvern	43.806686	-79.194353	0.773489
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497	0.098514
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711	0.082909
3	M1G	Scarborough	Woburn	43.770992	-79.216917	0.189245
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476	0.176899
5	M1J	Scarborough	Scarborough Village	43.744734	-79.239476	0.265976
6	M1K	Scarborough	East Birchmount Park, Ionview, Kennedy Park	43.727929	-79.262029	0.098514
7	M1L	Scarborough	Clairlea, Golden Mile, Oakridge	43.711112	-79.284577	0.329958
8	M1M	Scarborough	Cliffcrest, Cliffside, Scarborough Village West	43.716316	-79.239476	0.543755

Draw the Result on a Map

Use Folium to draw on a map using bubble marker where the size is proportional to its entropy value (convenience).