

Recommending Similar Location in Another City

Battle of Neighbourhoods



Introduction

- People migrate to other places in search of better opportunities and standards of living.
- The project aims towards developing a model to recommend localities in the other city based on venues in the current location.
- It will help people migrating from one city to another to find a locality with similar venues, points of interest, and facilities.

Data acquisition

- **Toronto Data:** The postal codes, borough, and neighborhood data are scrapped from the Wikipedia using BeautifulSoup. After cleaning and formatting the dataset, it merged with Longitude and Latitude Dataset provided by Coursera. With the help of Foursquare API, the venues nearby a given neighborhood is obtained.
- **New York Data:** The postal codes, borough, and neighborhood data are obtained from the given source (Cognitive Class). The data is in JSON which will be converted to a pandas dataframe. The data of the borough, neighborhood, latitude, and longitude will be used in this project.

Categories of Venues

- The number of unique venue types in Toronto is 321
- The number of unique venue types in New York is 489

Number of common venue categories in both data are : 293

Number of different venue categories in New York city are: 196

Number of different venue categories in Toronto city are: 28

Categories of Venues

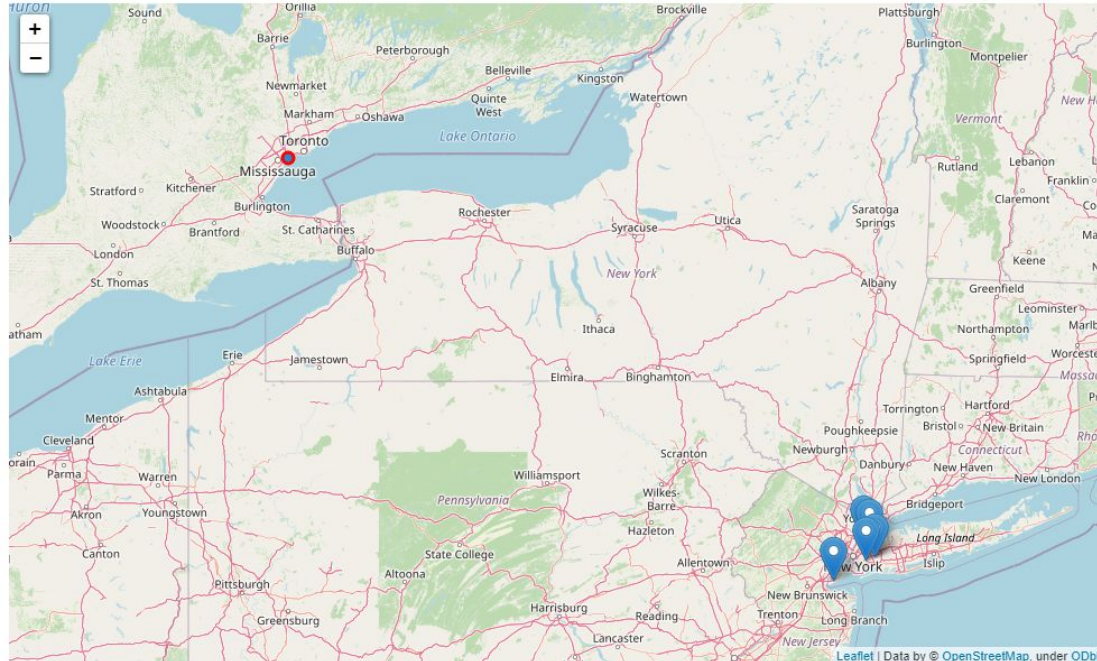


Finding Similar Location

Cosine Similarity is used for finding the most similar location in the other city. The top n number of most similar locations will be selected based on user details including latitude, longitude, and the venues near the location.

Use Case

Enter the current City
City: Toronto
Enter the current Borough
Borough: Etobicoke
Enter the current Neighbourhood
Neighbourhood: Alderwood, Long Branch



Conclusion

In this project, Cosine similarity is used to find similar locations in two different cities to help people migrating from one city to another. This model can be implemented for any location or any scale provided the required data is provided. It only considers venues of the location as the factor for finding similar localities but environmental and social factors are also important and play a major role in decision making.



Thank You