VenueFinder API for Toronto City

Karan Pillay

04/26/2021

1. **Introduction**

   VenueFinder allows you to find all the venues located in the city of Toronto. Explored neighborhoods of Toronto city with the help of Foursquare API and later used this feature to group the neighborhoods into clusters. Used k means to complete this task. Finally used Folium library to visualize the neighbors in the Toronto city and their emerging clusters.

2. **Business Problem**

   One of the most difficult parts of planning your wedding, birthday parties or any celebration is choosing the right venue. There are many factors which can impact the search like budget, number of guests, aesthetics, and accommodations. There are so many things to consider, and people to please that what used to be a fun experience has turned into more of a daunting task. Therefore, the primary aim of VenueFinder is to effortlessly find and recommend venues in Toronto city.

3. **Data Acquisition and Data Cleaning**

   The dataset used in this project is the Toronto neighborhood data, a Wikipedia page exists that has all the information needed to explore and cluster the neighborhoods in Toronto based on other demographics. Firstly, the dataset was scarped from Wikipedia page using beautiful soup object later it was wrangled, cleaned, and structured into a pandas dataframe. Also used a geospatial_coordinates.csv which includes latitudes and longitudes data as geocoder package was not responding. Finally merged the two datasets to obtain a final dataset which contains the following columns.

a) PostalCode
b) Borough
c) Neighborhood
d) Latitude
e) Longitude

In data cleaning process and manipulation process observed that the final_df contains 103 rows and 5 columns. Dropped ‘Postal Code’ column as it was similar to ‘PostalCode’ later checked for any NULL values in the dataset using isnull().sum() function in pandas fortunately there were no NULL values. On further analysis found that there are 15 Boroughs and 103 Neighborhoods in Toronto city.