Yelp business data analysis and exploration of restaurant business in the neighborhood.

Sneha Videkar

May 12, 2020

1 Introduction

Yelp [1][2] is a company which publishes crowd-sourced reviews about businesses using the company website[2]. This site captures reviews related to business of various categories like Dentist,Restaurant, etc. Along with reviews, location details of every business is listed in Yelp database. Business data is captured from around the world.

2 Business Problem

Success of new business depends on many factors. Location is the key factor which should be considered while opening new business. For this purpose, location attributes present in the Yelp dataset[4] will help find a place to start a new business. This project aim to help investors/entrepreneur to find an appropriate location for their prospective restaurant business. With the help of Yelp Business dataset, this project will locate most common restaurant business in a neighborhood which ultimately guide investors while starting a new restaurant business.

3 Data Description

3.1 Data acquisition

For this problem below dataset will be used:

1. Business.csv will be used from Yelp Dataset. This dataset is taken from Kaggel website[3] and Yelp dataset website[4]. For modeling purpose, business with "Restaurant" as category are considered.

2. Forsquare API will be used to get the most common venues for above mentioned business and in Toronto area. Only near by venues of type restaurants will be considered for further analysis to achieve the project goal.

3.2 Data Preprocessing

Yelp Business data was loaded in the IBM Watson Studio and accordingly csv file was read in dataframe called businessDF.

Various categories of the business data is captured in Yelp business dataframe. Figure 1 gives glimpse of the most important categories available in the businessDF.

Figure 1: Various business categories from Yelp Dataset

References

[1] https://en.wikipedia.org/wiki/Yelp

[2] https://www.yelp.com/about

[3] https://www.kaggle.com

[4] https://www.yelp.com/dataset/documentation/main