

IBM Applied Data Science Capstone

Recommendations on location and food type of a restaurant to open in Toronto

Boyan Gu

1. Introduction

1.1 Background

Toronto is the most populous city in Canada, and is an international center of business, finance, arts, and culture. It is recognized as one of the most multicultural and cosmopolitan cities in the world. Its economy is highly diversified with strengths in technology, design, education, arts, tourism, and etc. Therefore, it will be of great chance to open a restaurant successfully and attractive customers, especially tourists.

1.2 Business Problem

The target audiences of this problem are investors who want to open a new restaurant targeting tourists. The problems for the investors now are where to open a restaurant is least competitive and can make a great profit, and which food types are popular and attractive for tourists in Toronto. We could simply assume that restaurants visited most are those located near hotels, which are convenient for tourists to get there. Therefore, we could locate hotels in Toronto and do clustering and KNN to know the competition levels among restaurants near hotels. And then analyze the trending type of food that are favorable by visitors based on foot-traffic data.

2 Data

This report would require location data of hotels, restaurants, and food-traffic data in restaurants, and their types. We would use Foursquare API to get these data for our analysis. We would use **Search** calls to query **Hotels** in Toronto, and get names and location data of hotels. Next, we would use **Explore** calls to query trending venues and extract **Restaurants** based on venue's categories to find the trending food type. K-means Clustering would be applied to cluster hotels, and locate the centers of clusters as location where there are most tourists. Moreover, we would use K-Nearest Neighbors Classification to analyze the competition levels of restaurants around those cluster centers.