Applied Data Scinece Capstone

Gyan Prakash Singh

February 2020

1 Introduction

1.1 Background

Using data to solve is a business problem is one of the things which makes data science useful. Deciding the location of a new business venture is one of the crucial decisions which can have critical consequences for the future of a business. However, the factors which affect the suitablity of a location for a new business venture vary and as businesses have unique requirements to ensure success. I will be In this report, I will be exploring and implementing

In this report, we will explore Restaurant business in Delhi, India. Delhi is the capital of India and a very populous city. People of Delhi are known for their love of food which makes Restaurants flourish in general. However, with growth in income, tastes of people are getting more refined and it is making a new restaurant venture a challenging task.The task is particularly more difficult for fine dining restaurants. With high income inequality combined with high average income compared to rest of the country, finding a suitable location for a fine dining restaurant in delhi is a challenging but most crucial task in planning a new restaurant business.

1.2 Problem

Core problem is use of data science tools to find 3 candidate locations for setting up a fine dining restaurant in delhi. Factors which may lead to higher footfall such as number of high quality fine dining restaurants in the locality, other landmarks bazaars, museums, art galleries, shopping centers, parking lots etc. need to be identified and explored to make a suitable model which may solve this problem.

2 Data

Foursquare location data will be utilized as core of our analysis. It will be used to extract information such as nearby venues of interest. Based on the definition of the problem, data points which I would like to explore are:

- number of existing restaurants in the neighborhood (any type of restaurant)
- number of and distance to fine dining restaurants in the neighborhood, if any

- distance of neighborhood from major landmarks such as museums, shopping centers, art galleries etc.
- distance from coparatively affluent residential localities of the city.
- tranport facilities nearby.

We will also use data published by Office of the Registrar General Census Commissioner, India Ministry of Home Affairs, Government of India under national census. Census data contains extensive information ranging from population to households amenities and assets. This data is avaliable at Tehsil (roughy translated to locality) level. For example we can use census data to find out the average car ownership in a locality. It may be noted that car ownership is considered a sort of luxary and not everyone owns a car as is normal in developed countries. It may be assumed that potential patrons of a fine dining restaurant will be car owners. We can first select a suitable number of districts/localities in delhi based on relative affluence of the neighbourhood and limit further analysis to these districts/localities only.