Analysis of Locations for a Business

Anil Sasidharan

1. Introduction

1.1 Background

Success of any business, whatever the domain might be, depends on a number of common factors. It is imperative that the owner of the business gets it right on most of them, if not all, to ensure success. Among these factors, the opening of the said business, as an event, carries a huge significance, both in terms of contributing to the success and also in financial management. As a result, this is almost always treated as very important by entrepreneurs. More the information they can acquire that can help towards the successful opening of their business, the better planned they can be. This proves how useful information is to any business, and in turn, data. For this particular project report, I am concentrating on specifically location data. One of the important parts in opening a business is selecting its location. So I will be demonstrating how to achieve a data driven result.

Toulouse, the city where I am currently residing since 2016, has many facets. It is a small and calm typical French city, with 18th century architecture and beautiful cafes and restaurants wherever you go. The different neighborhoods, or 'quartiers' as they are referred to here, each have their own features, along with the different behaviors and lifestyles of the people there. I am using location data from Foursquare to get an insight into these properties so as to obtain good candidate locations for a business, which is – a Games and Recreation center.

1.2 Problem

The aim of this analysis is to show that data analysis can be used in important factors of opening a business. It can uncover certain aspects that tend to be hidden – for example, for a particular public business opening in a particular location, how will the residents of the locality perceive it, what percentage of the population are welcome to its idea, and is it in line with the lifestyles of the residents. These are factors which normally come to light a while after the business has been opened. And there is also a chance of a negative impact. The aim of data analysis is to provide an advantage to business owners in this regard.

1.3 Stakeholders

As noted above, entrepreneurs would be the most interested in this analysis. Existing business owners would like the idea of being a step ahead when opening branches or franchises in unfamiliar regions. Also, this kind of analysis has a broader scope. Gauging public opinions, generating public feedback, comparing growth of businesses, and even measuring economic growth are some of the uses. And so, the stakeholders list will enlarge to include survey companies, reviewing companies, economic departments, government administration and last but not least, social media.

2. Data for the Analysis

2.1 Principle of Use

The data required for this project is classified into two parts – Location data of the different neighborhoods of Toulouse, and Foursquare data for these locations.

Location data will include compiling a comprehensive list of the neighborhoods and the area they cover, and the geographical coordinates. This will be the primary database and a foundation for the analysis. The Foursquare data will generate information on trends in lifestyle and the popularity of different businesses. This data for each location can be added progressively to the primary database to obtain a final version for analysis.

2.2 Source

Data on neighborhoods or quartiers in Toulouse has been taken from the site data.toulouse-metropole.fr. The link for the CSV file given below:

https://data.toulouse-metropole.fr/explore/dataset/recensement-population-2016-grands-quartiers-familles/download/?format=csv&timezone=Europe/Berlin&lang=fr&use_labels_for_header=true&csv_separator=%3B

The file has to be cleaned and prepared to get a final list of neighborhoods and coordinates. The Foursquare data has to be obtained by repeated calls to the Foursquare API. The different API calls will be explained in the Methodology section as we go forward in the analysis. To obtain results with Foursquare API calls, I have my free account with my exclusive credentials.

2.3 Data Preparation and Feature Selection

The initial database that I have downloaded is a census database with population numbers classified under several categories. Thus, it has a lot of features not useful in this analysis.

Firstly, I have cleaned the data by removing all the population numbers and also the features “Index”, “GRD_Quart” and “Dep” numbers. Secondly, I have renamed the features to English as “Coordinates” and “Neighborhood”. Thirdly, I have split the data in label “Coordinates” into “Latitude” and “Longitude” data. After this step, the database is primed for use, to generate interactive maps, as well as generating Foursquare API calls.

3. Methodology
4. Results
5. Discussion
6. Conclusion