Recommending Restaurant Location

Magnus Billups

April 2, 2019

1. Introduction

1.1 Background

A prospective business owner wants to open a new Asian Restaurant in Austin, Texas. The owner needs to know what part of Austin would be ideal for opening their new location.

1.2 Problem

Opening a business in the proper location can be pivotal to success, if you open in the wrong location you won't get the right amount of customer traffic to stay in business no matter how good your restaurant is. In this report we will be analyzing which part of Austin would be best based on similarity of popular restaurants in that area*. We will be looking at metrics such as venue locations, popularity, and similarity of food served.

*Normally I would also look at the cost of property in that area. This would give a more realistic idea of the risk factor for opening a business. However, due to the lack of available pricing data I am ignoring this aspect.

1.3 Interest

Any business owner would be interested in knowing the best spot to open their new venue. It helps to reduce risk of dealing with competitors and/or not having enough customer traffic.

2. Data Sources, Cleaning, and Features

2.1 Data Sources

To perform our analysis I used venue data pulled from Foursquare's Places API. To get a reasonable amount of data points I pulled data from the search and explore endpoints.

2.2 Data Cleaning

Most of the data that was provided by the search results was delivery information or redundant location data. The data provided by the explore results was more concise but still included redundant or unneeded venue information. After reviewing both data result sets I was able to narrow down the necessary information to create a combined result set shown in the figure below.

Fig. 1

	name	categories	latitude	longitude	postalCode
0	Broth & Basil	[Vietnamese Restaurant]	30.468596	-97.595634	78660
1	Sushi Nini	[Sushi Restaurant]	30.479040	-97.671817	78664
2	Saigon El Vendeur	[Vietnamese Restaurant]	30.265481	-97.744288	78701
3	Koriente Restaurant	[Korean Restaurant]	30.267108	-97.736029	78701
4	P.F. Chang's	[Chinese Restaurant]	30.263782	-97.741413	78701

2.3 Feature Selection
In order to give an accurate estimate of the best location I needed to be able to visualize the data, so I decided to focus on features that would lend to that. The data provided by Foursquares APIs included both the latitude and the longitude of every venue it returned. This was used for creating a map to visualize the density of competition in the region.

The categories feature was also returned by Foursquare and was used to get a more concise representation of what types of Asian restaurants are in the region and the density of each type. The postalCode was similarly used to visualize where the highest density of competition is, as these will likely be areas to avoid due to too much competition.