



CAPSTONE PROJECT

Battle of Cities for Business

Abstract

If you had to open a business, in what major Texas city would be best and what business?

Elias Adame

Introduction

When an entrepreneur wants to open a new business, or diversify their business holdings, there are two main questions: Where and What? In this project we will attempt to analyze select US cities to discover what business market is popular among them, and if one city lags behind the others in that business. This will provide a market opportunity to capitalize on the potential growth that the business has in that city.

Data

The data used in this project are:

-) Name of US cities
-) Latitudes and Longitudes of each city from:
<https://www.latlong.net/category/cities-236-15.html>
-) Data pulled from Foursquare API which includes:
 - Venues in a radius from a latitude and longitude
 - Venue Names
 - Venue Categories
 - Venue Categories can be thought of as business markets

Methodology

We utilized Python to perform the analysis and utilized the Pandas library to organize data into data frames, and the requests library to pull data from Foursquare's API. We did some exploratory data analysis such as:

-) Get counts of Venues pulled for each city
-) Unique categories pulled

We then utilized a method called Hot Encoding to get a weighted average for each venue category for each city to determine which venues were popular or unpopular in each city. We grouped all the scores for each venue for each city and sorted to get the most/least popular venue

categories in each city. We then got some simple statistics. We pulled what Venue category was the most popular amongst all the cities sampled, as well as pulling the cities where that same venue was # 10 most popular.

We then used the K-Means cluster Machine Learning technique to cluster cities together by their similarities of popular venues. We used this technique because when attempting to decide where to set up a business and what that business should be, we need to see what cities are most like each other.

Results

As we processed and analyzed the data, the most popular venue in the selected cities was Fast Food Restaurants with 9 cities having it as the #1 Most Popular Venue and 33 cities having it as a Top 10. There were 6 cities where Fast Food Restaurant was at the #10 spot, signifying that there is potential growth in these cities since they're similar (same cluster) as the cities where Fast Food Restaurant is #1. Those cities are:

-) College Station, TX
-) Thornton, CO
-) Kansas City, KS
-) Newport News, VA
-) Vancouver, WA
-) Montgomery, AL

Discussion

As mentioned in the Results section, there are 6 cities that show growth potential for the Fast Food industry. As we dig deeper into these 6 cities, we notice that College Station, TX, Kansas City, KS, and Newport News, VA have “American Restaurant” in their top 3 venues, thus showing possible competition for the Fast Food industry in these cities. Out of the other 3, Vancouver, WA has a Pizza Place as it's #2, and Thornton, CO has “Food Truck” as it's #4. The only city

that does not show competition to a Fast Food Restaurant in its Top 5 spots is Montgomery, AL. Thus, we recommend this city and a Fast Food restaurant as a possible business opportunity with growth potential.

Conclusion

As we analyze select cities in the United States, we look at their most popular venues and realize that based on similarities in cities, and dissimilarities in what business they have in the #1 venue, there are growth potentials in several cities and several businesses. In this project we concluded that Montgomery, AL has growth potential for its Fast Food Industry.