

Applied Data Science Capstone

Exploring Potential Locations of New Franchises for Torchy's Tacos Food Chain

21-June-2020

Background

Austin is the capital city of the U.S. state of Texas. It is the 11th-most populous city in the United States, the fourth-most-populous city in Texas.

Torchy's Tacos is a famous restaurant in Texas with more than 60 physical locations.

It was founded in Austin in 2006 by former corporate chef Michael Rypka, who was inspired by a love for tacos and a passion for experimenting with food to bring an untraditional taco experience to the people of Texas.



Problem

This project focus on a *hypothetical* business plan.

Assuming someone wants to open another franchise of Torchy's Tacos in Austin. Given the huge amount of investment involved, it become vital to determine the best location to get the highest rate of return in the investment.

Interested Audience

- Any potential investors who plan to open a new franchise in the Austin
- Existing owners of the chains to investigate why certain restaurants are performing better than other.



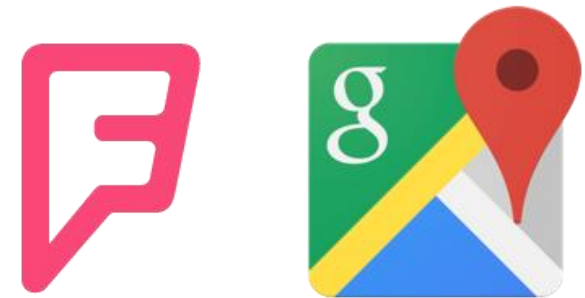
Data

FourSquare API: It was used to obtain information about nearby venues based on the geo (latitude and longitude) data.

Store Locations: The store locations and addresses were accessed using the webpage of the food chain.

Google Maps: The latitude and longitude data were obtained via google maps.

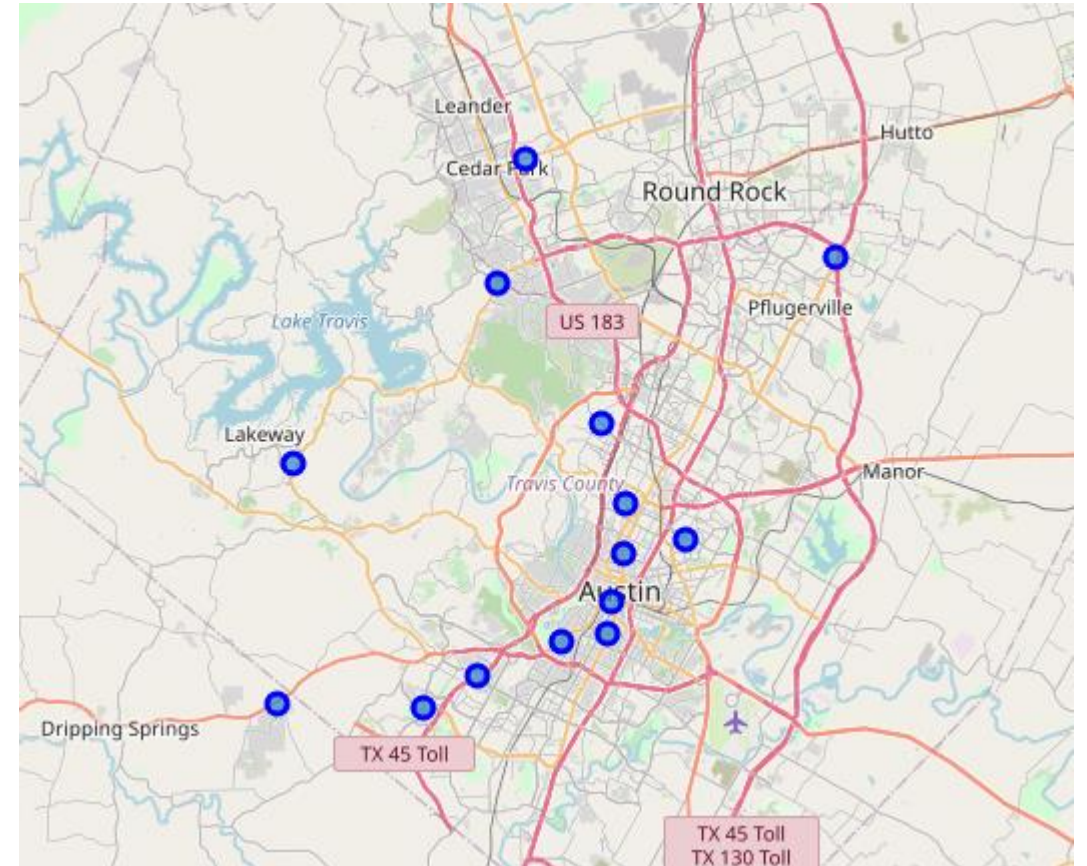
Yelp Rating: The yelp rating of the food chain at different locations along with the numbers of reviews posted online was used to rate the performance of the different locations.



Methodology

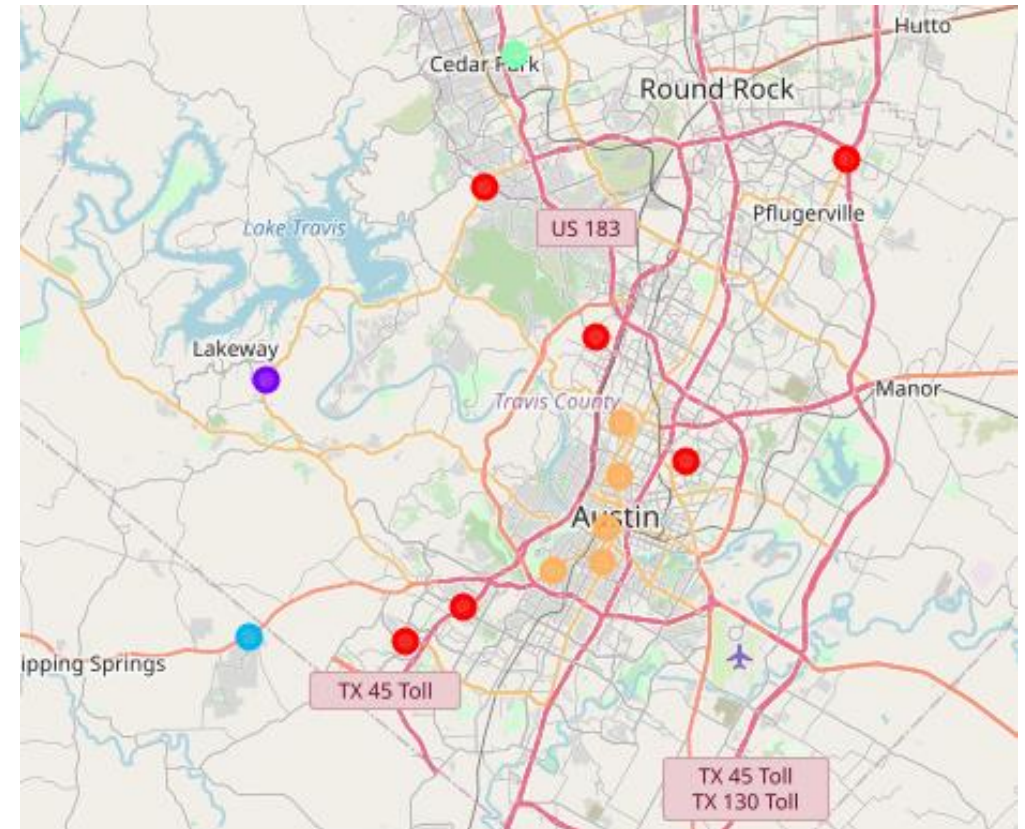
Exploratory Data Analysis (EDA)

EDA was performed to visualize the existing locations of the food chain- Torchy's Tacos in the Austin, TX region.



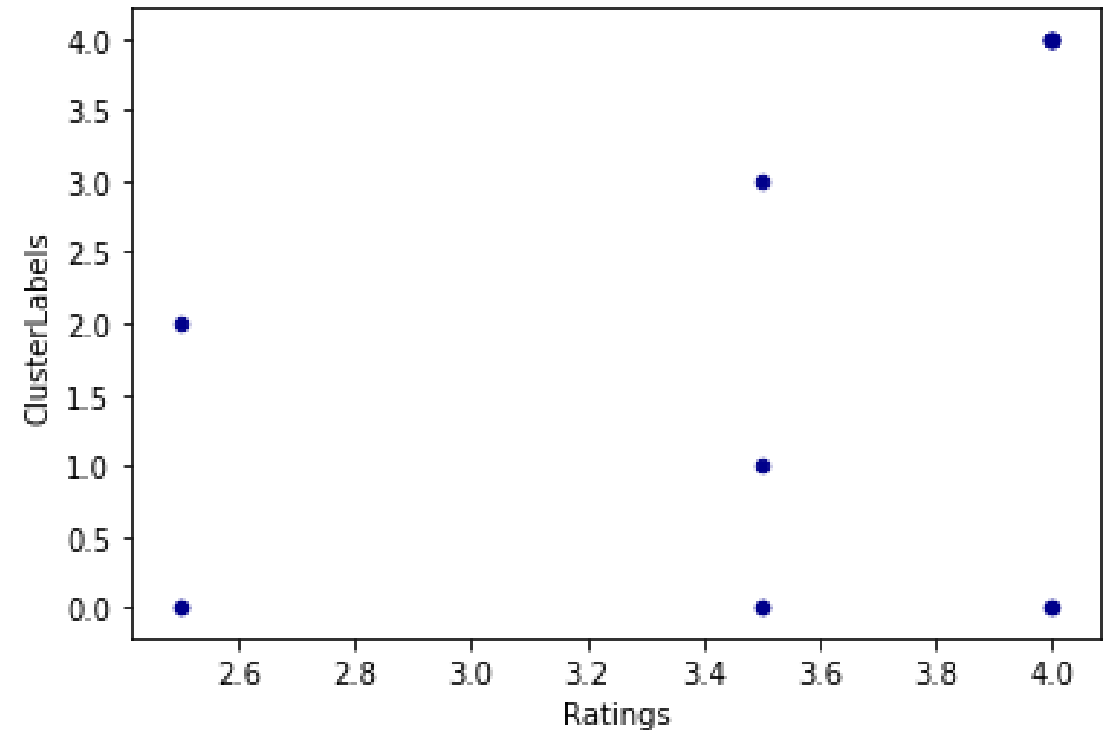
Methodology

K-Mean Clustering: This is a method of vector quantization that aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean (cluster centers or cluster centroid), serving as a prototype of the cluster. Python in-built function was used to create color coded clusters encompassing the different location based on the similarity of the different venues around them.



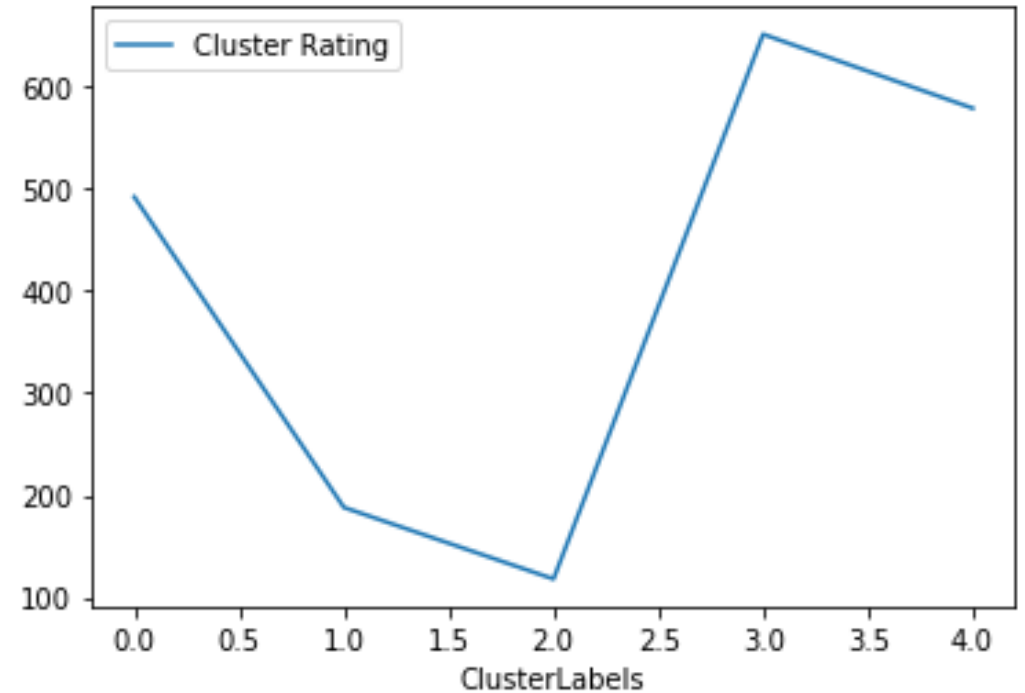
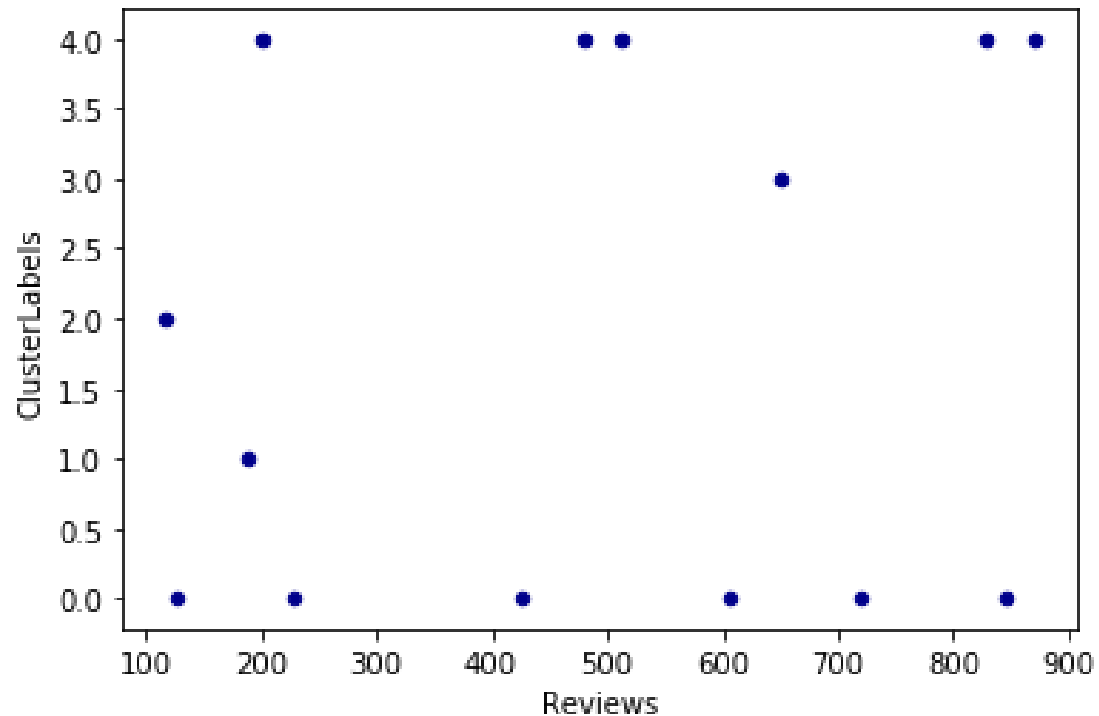
Results & Discussion

- After clustering, the next step was to determine which cluster is performing best. Since, the financial data of these stores are not available, we rely on the Yelp data (rating and reviews) of these locations which is publicly available.
- The ratings of the restaurant are much strong function of the service provide in that specific location. Thus, it may not be a good indicator while choosing a new location as service quality can be improved in the new location
- Thus, the number of reviews was chosen as the success criteria under the assumption that a more popular location will gain more social engagement and hence, higher reviews



Results & Discussion

- Let's plot the number of reviews and average review for cluster and the determined clusters label.



Results & Discussion

- Best locations of the franchise should be close to:
- Cluster 3: This cluster has the higher cluster rating so a nearby location in that region is a good candidate for the new franchise. However, the data to support this is limited as only one restaurant falls into this category.
- Cluster 4: This cluster already has 5 restaurants which are doing well based on the social ratings. A new franchise in this cluster region can also be a good candidate. It's also located in the heart of Austin downtown area which is very popular.

