CAPSTONE PROJECT - THE BATTLE OF NEIGHBORHOODS

Find the best location to open Italian restaurant in Austin neighborhood.

OPPORTUNITIES FOR NEW RESTAURANT

There's an influx of tech companies moving to Austin. Lower costs, laid-back lifestyle continue to draw tech companies to Austin, Texas. According to the Austin Chamber of Commerce, 58 major companies relocated to the Austin area in 2019 alone – not including tech giants such as Apple, Amazon, and Google, who opened new offices in the region.

All those jobs are going to require smart, motivated, skilled workers to fill them. And those workers need places to live and restaurants or food joints to eat.

By using data science and machine learning methods such as clustering, this project will recommend a best suitable location to open a new Italian restaurant.

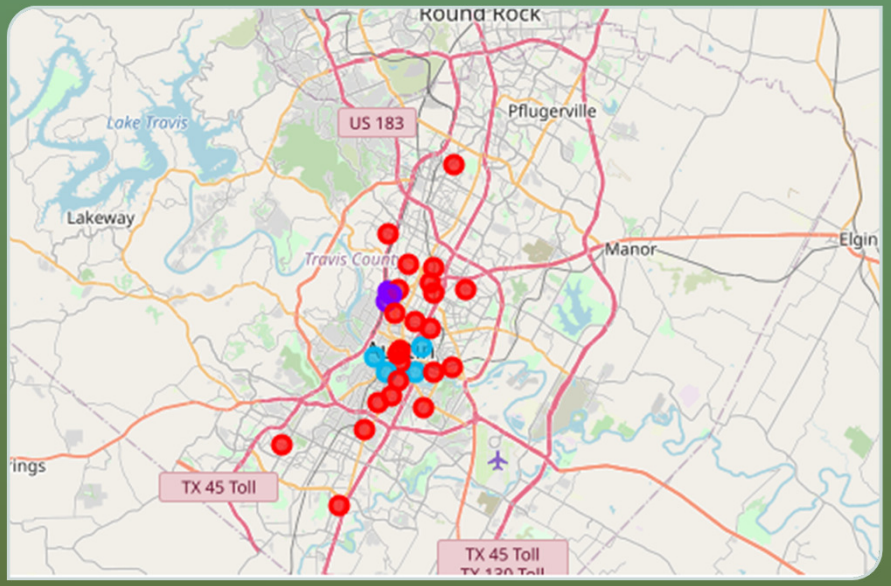
DATA ACQUISITION AND CLEANING

Following data is required for this project:

- List of Austin neighborhoods scraped from Wikipedia page that contains list of Austin neighborhoods
- Latitude and Longitude of these neighborhoods, which can be obtained from Geocoder package
- Venue data related to these neighborhoods that can be obtained using Foursquare API

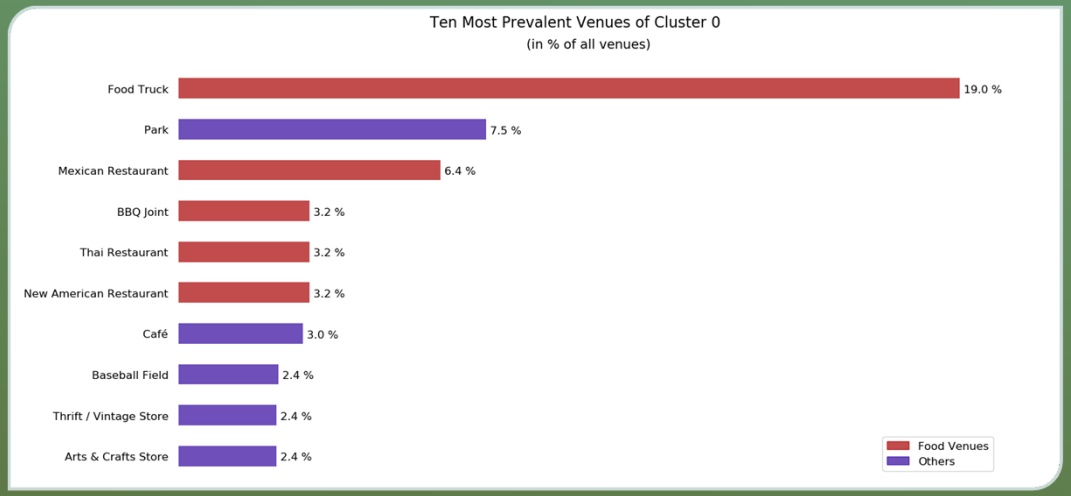
Data scraped from Wikipedia is filtered to retrieve names of Austin neighborhoods, duplicate entries removed, and special characters cleaned to prepare the final dataset for Austin neighborhood.

APPROACH



- Location coordinates of each neighborhood is obtained using GeoPy Nominatim geolocator and appended to the neighborhood data. Using this data, a folium map of Austin neighborhoods is created.
- Explore each of neighborhoods and their venues using Foursquare location data.
- Using K-means clustering locations are grouped into clusters based on their similarities.

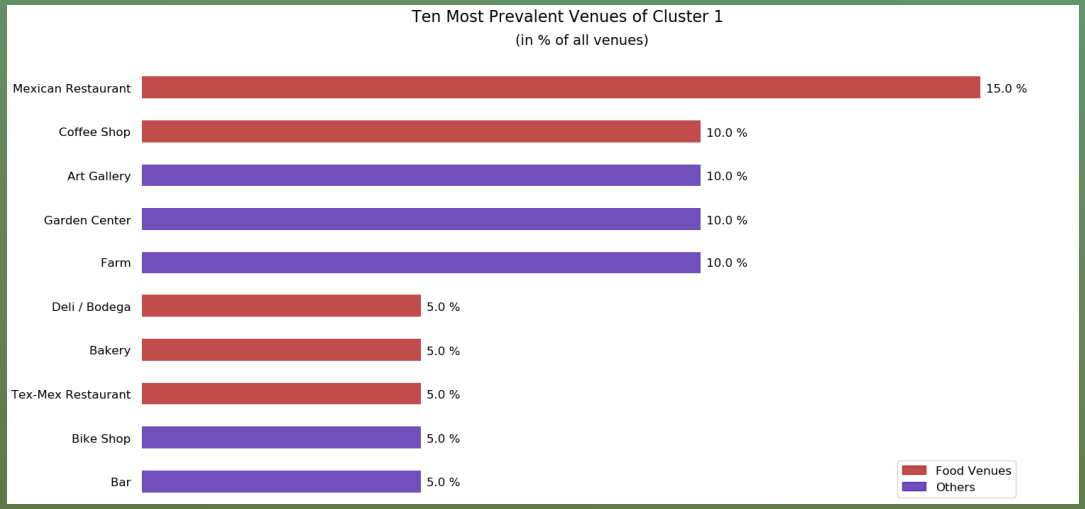
CLUSTER 0 - ANALYSIS



Cluster 0 has 5 food venues among the top 10, with Food Truck and Mexican Restaurants making up nearly 25% of all venues.

This poses considerable competition for a new food establishment hence it is not a best option to explore further in terms of setting up a new restaurant.

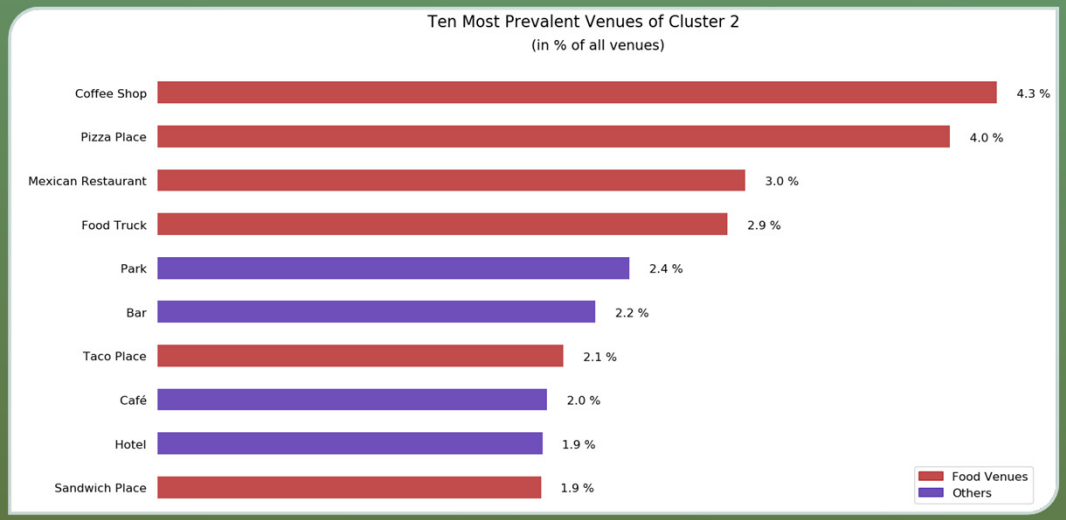
CLUSTER 1 - ANALYSIS



Cluster 1 has 5 food venues among the top 10, with Mexican Restaurants making up a huge majority (nearly 20%) of all venues.

Adding to this other than 1 art gallery there are not enough venues to create required foot traffic for a new restaurant, so Cluster 1 is not the best option

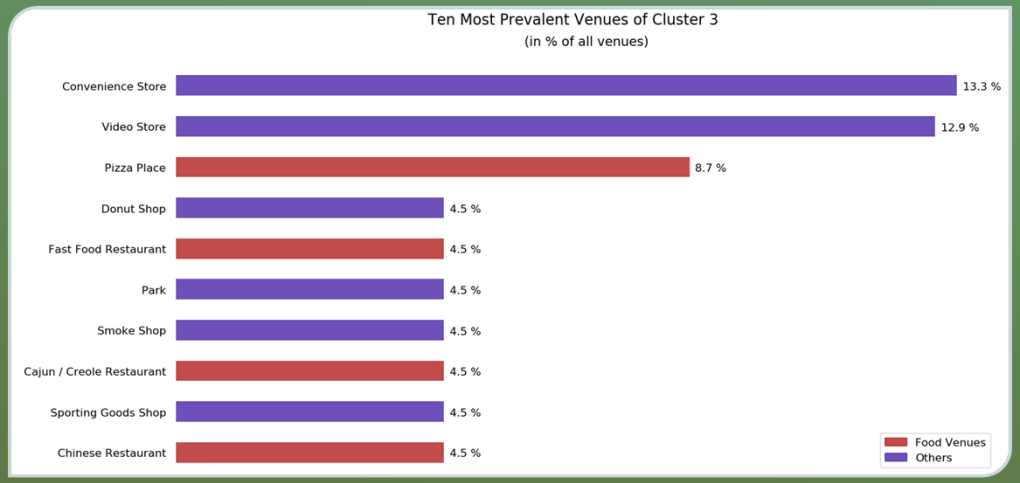
CLUSTER 2 - ANALYSIS



Cluster 2 has 6 food venues in the top 10 venues with only 5% full-service Restaurants which makes it viable. But this cluster has only 1 Park and no other office / business establishments to bring in foot traffic needed for a new restaurant.

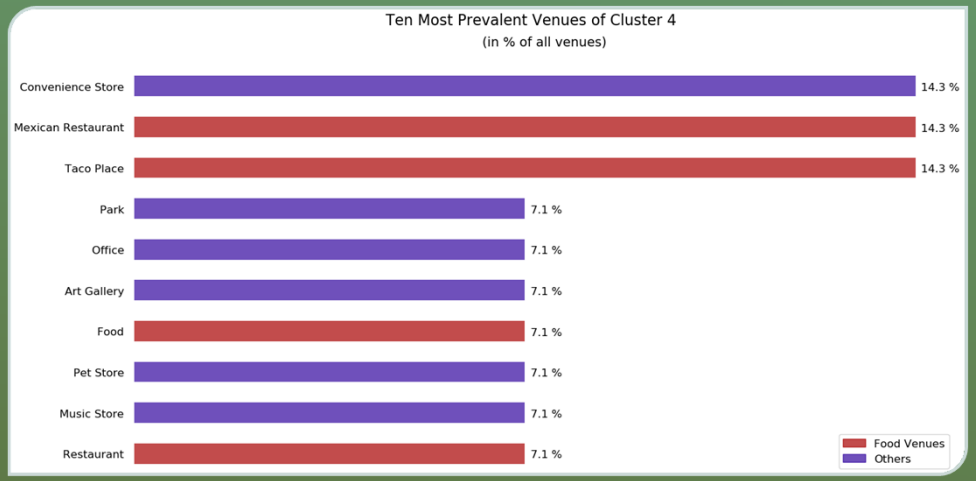
Analyzing the top venues, it appears the fast food and other food establishments in this cluster are around the park, so a new full-service Italian restaurant will not be viable in this cluster.

CLUSTER 3 - ANALYSIS



Cluster 3 has 4 food venues in the top 10 mostly fast food, but not many restaurants. There is a Park and shopping places which are favorable indicators of foot traffic, which can be optimal for new restaurant.

CLUSTER 4 - ANALYSIS



Cluster 4 has 4 food venues in the top 10 with only 2 full serving restaurants. Additionally, compared to cluster 3 there are a lot of public venues in this cluster - venues that see a lot of foot traffic such as parks, art gallery, office and department stores.

The presence of these high traffic places indicates cluster 4 as optimal place for the new restaurant.

CONCLUSION

- Identified data required, extracted and prepared the data
- Performing machine learning by utilizing k-means clustering to group similar neighborhoods into clusters
- Further analyzed the clusters and recommended a viable cluster of locations suitable to address the business problem – best location for a new Italian restaurant in Austin neighborhood.