

Data Science Capstone Project

Exploring the Cities in Texas

Business Problem

- Texas is one of the most diverse states in the US for both ethnic and economic diversity, making it a prospective location for restaurant business especially Thai cuisine
- Explore the potential of opening a Thai restaurant in Texas
- Focus on the five most populous cities in Texas: Houston, San Antonio, Dallas, Austin, and Fort Worth.
- Consider demographic and economic information, current competition, people's interests, and other common venues in the cities

Data

1. Geographical Coordinates of Each City - Geopy

2. Demographic and Economic Information - Web Scraping

- Statistical information from the United States Census Bureau website (<https://www.census.gov/quickfacts/fact/table/US/PST045219>)
- Parameters of interest are %Asian, Median household income, and Median gross rent.

3. Restaurants Search and Venues Exploring - Foursquare location data

- Search for restaurants to study the competition
- Explore common venues in the city to study the interests of people

Methodology

1. Import, Transform, and Explore Data

1.1 Obtain Geographical Coordinates from Geopy

1.2 Import Demographic and Economic Information from the United States Census Bureau website

1.3 Search Thai Restaurants using Foursquare API

1.4 Search All Restaurants using Foursquare API

1.5 Explore the city using Foursquare API

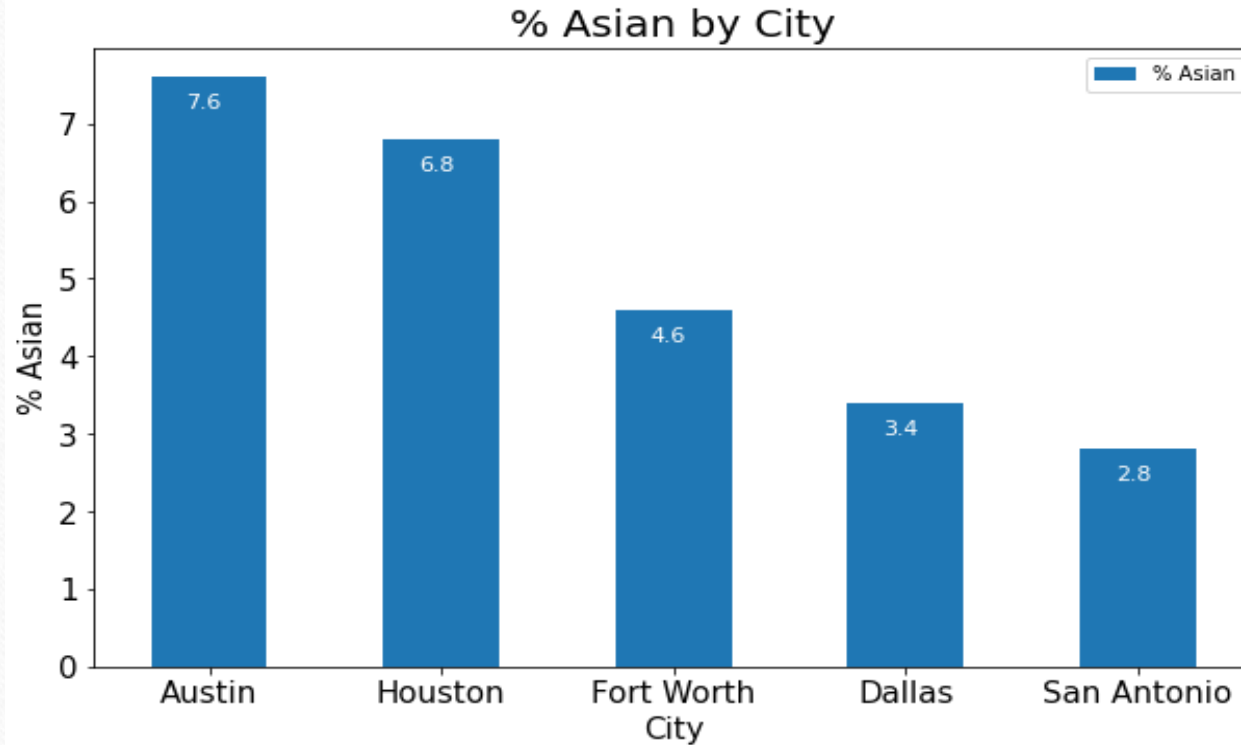
2. Analyze and Visualize Data

- Create charts and maps using Pyplot and Folium

3. Cluster the Data

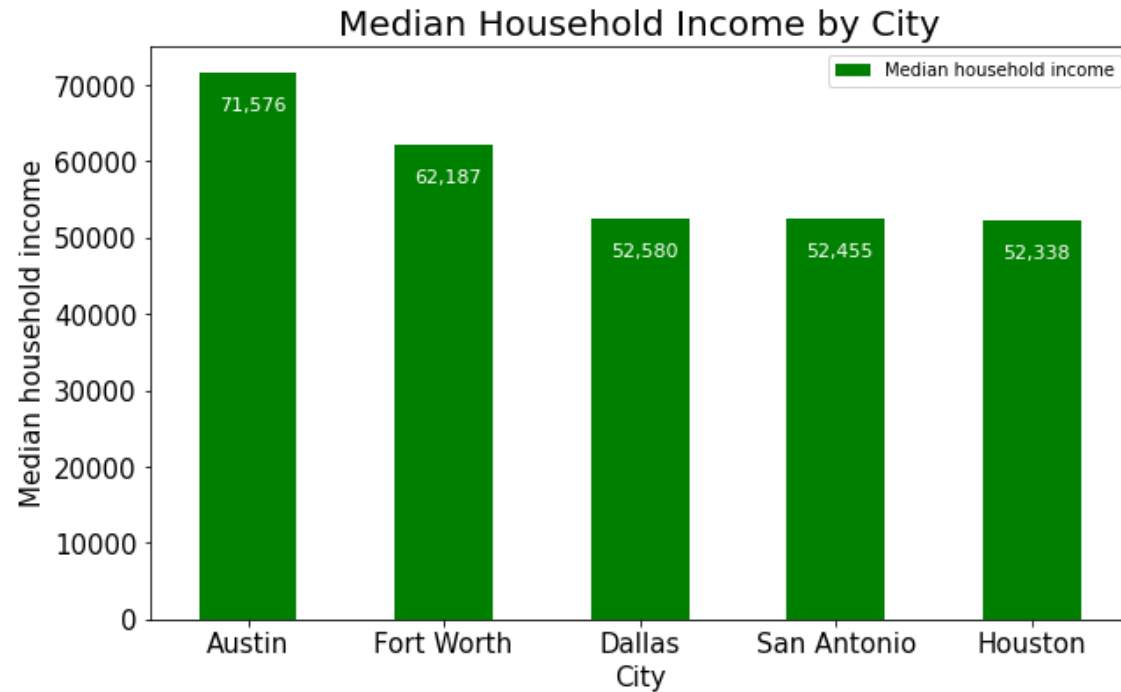
- Cluster cities based on venues using K-Means clustering

Results



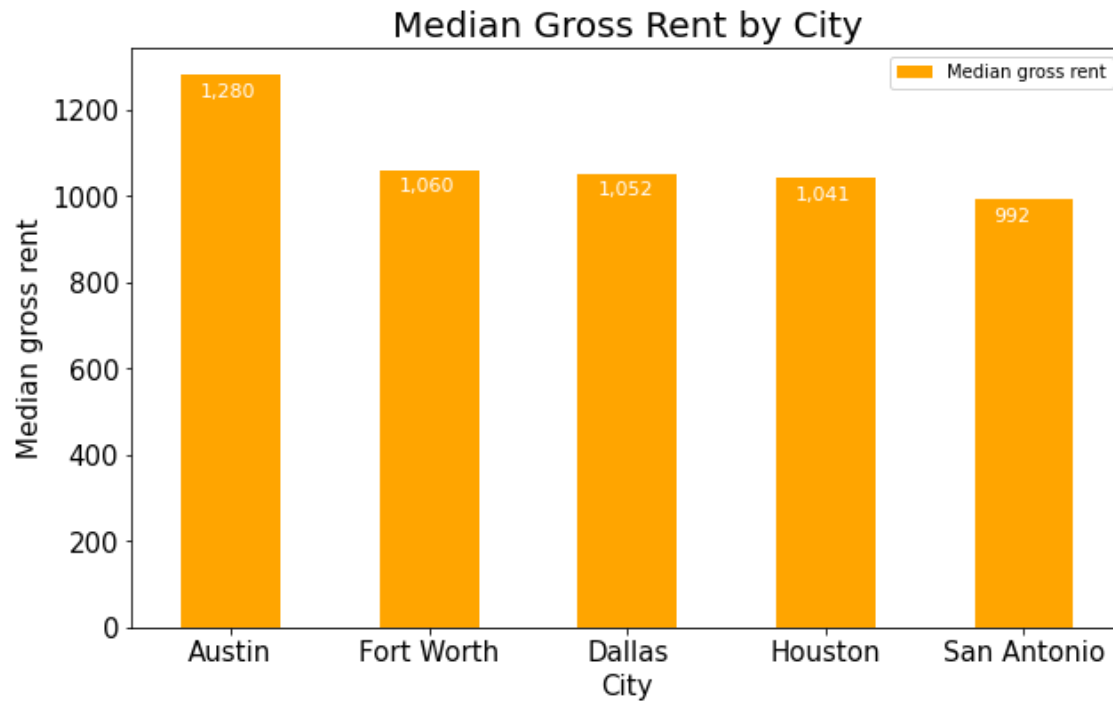
- % Asian in each city from largest to smallest
- Target group of customers are primarily Asians and Americans with Asian descent.
- Austin has the highest % Asian at 7.6% and San Antonio has the lowest % Asian at 2.8%
- Data obtained from the United States Census Bureau website

Results



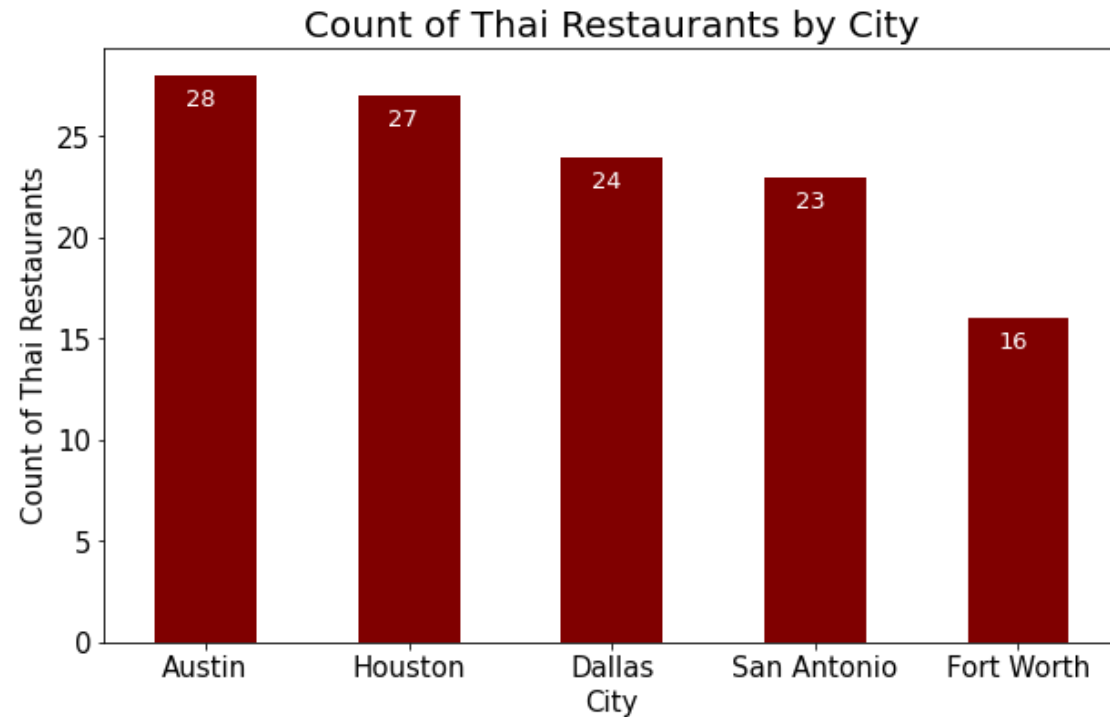
- Median household income from highest to lowest
- Measures economic condition of the city and affordability of the restaurant
- Austin has the highest median household income while Houston has the lowest median household income
- Data obtained from the United States Census Bureau website

Results



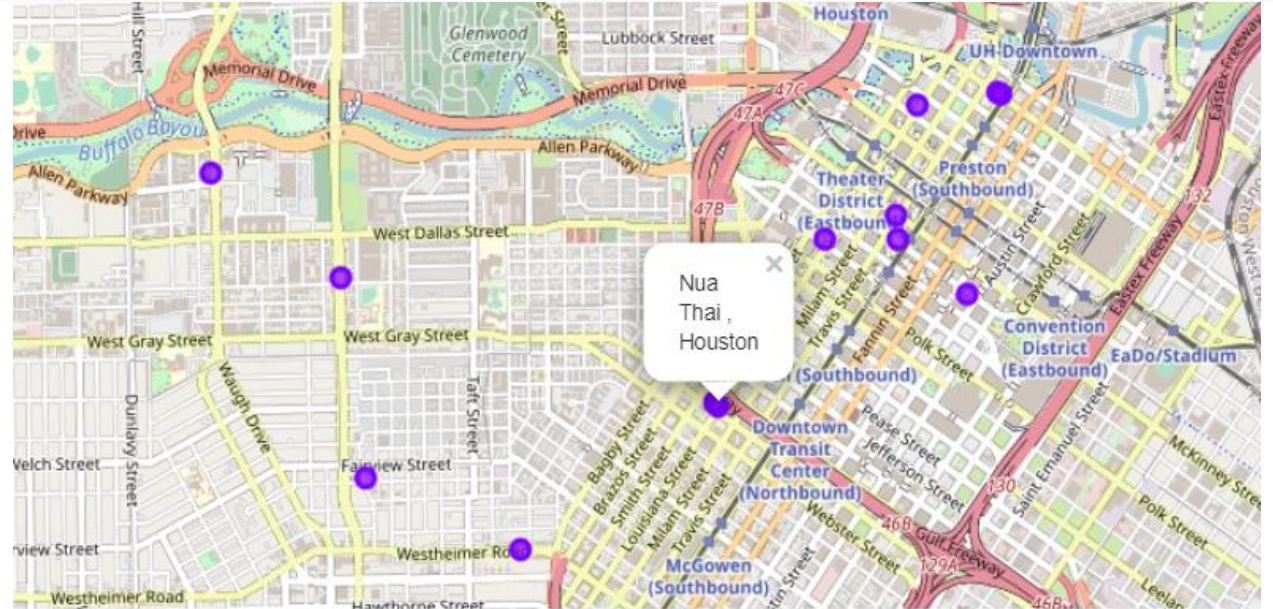
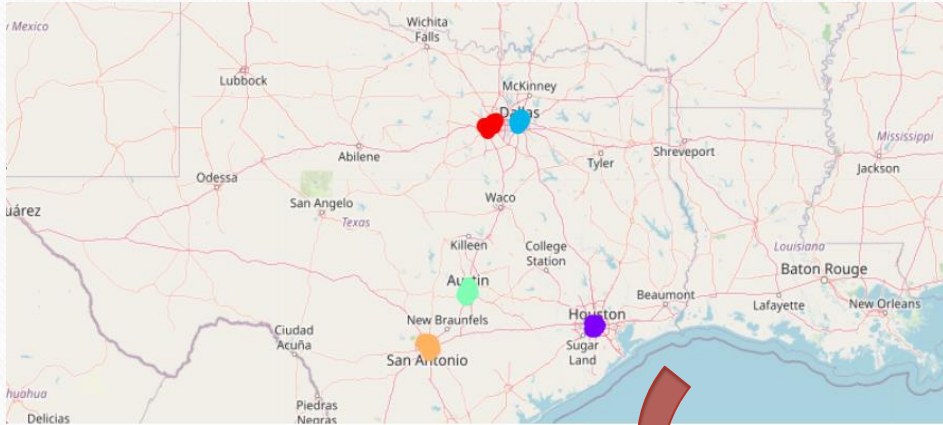
- Median gross rent from highest to lowest
- Actual amount for restaurant lease will vary but the trend should still hold for cost comparison between cities
- The higher the amount will result in the higher fixed costs for operating the restaurant in the future
- Austin has the highest median gross rent while San Antonio has the lowest one

Results



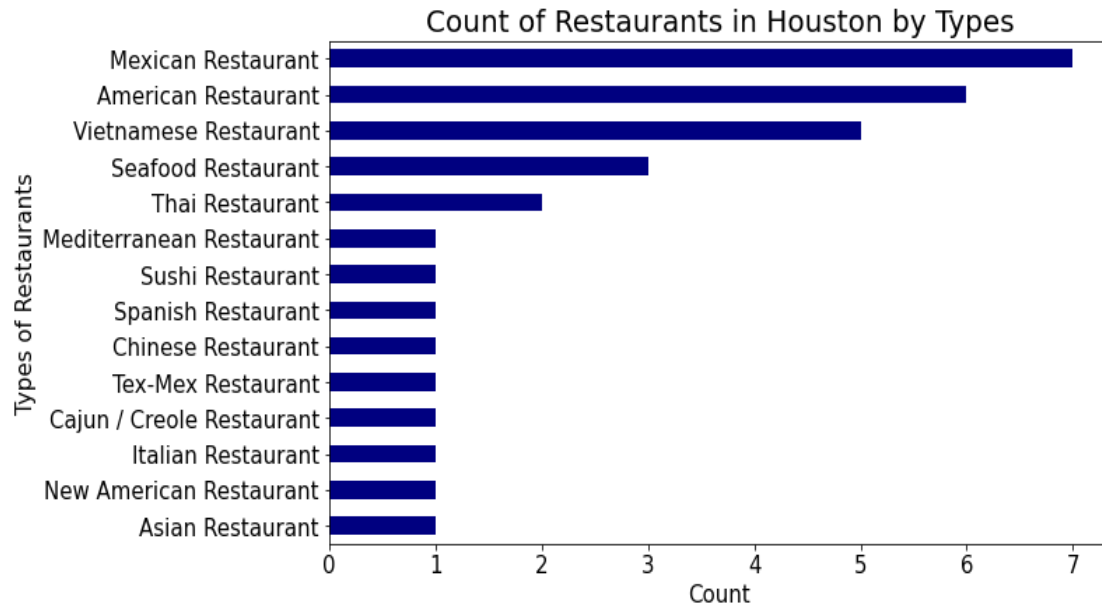
- Study our direct competitors which are other Thai restaurants
- Chart shows the number of Thai restaurants in each city
- Data obtained from the search for Thai restaurants using Foursquare API
- Highest competition in Austin with 28 Thai restaurants within 10 km radius while lowest competition in Fort Worth with 16 Thai restaurants

Results



- Leaflet map showing the location of Thai restaurants in each city
- Can zoom in and click on the markers to display the name and city of restaurants
- Data obtained from the search for Thai restaurants using Foursquare API

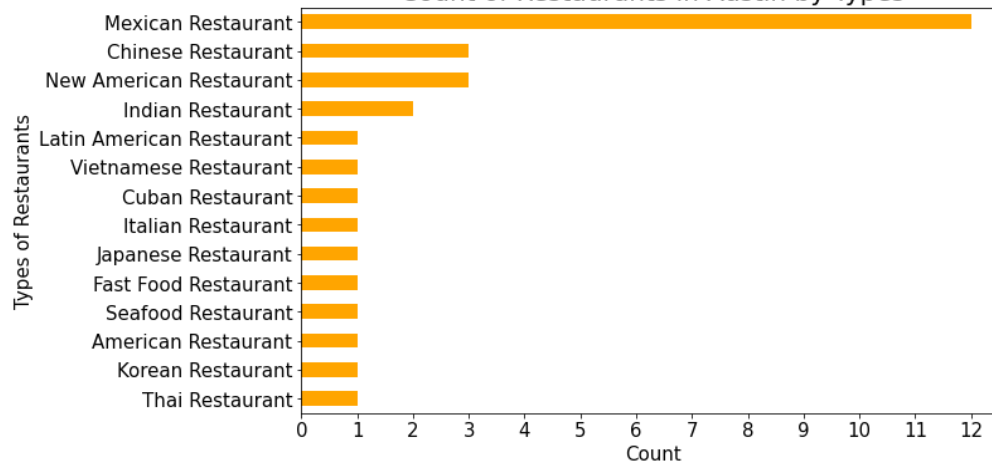
Results



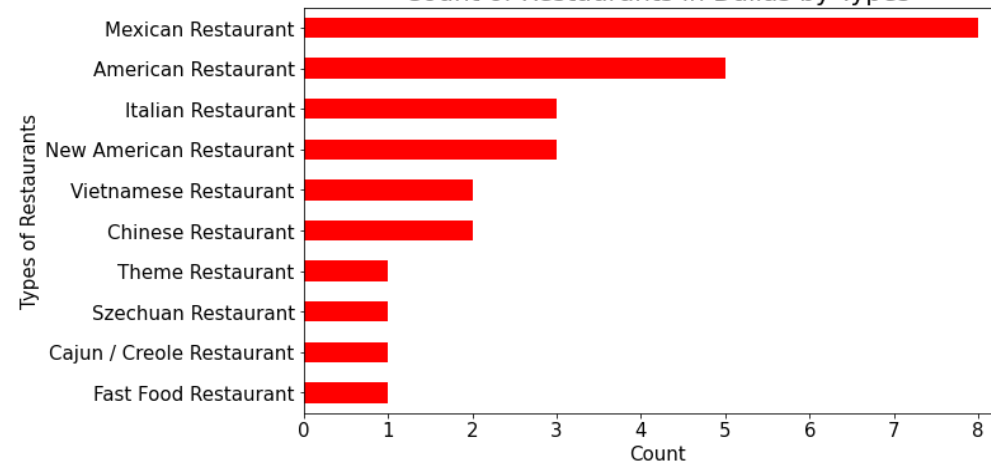
- Evaluate restaurant business beyond Thai restaurants
- Create charts that show the number of restaurants in each city by types
- Data obtained from the search for all restaurants using Foursquare API
- Since limit for search is only 50 venues, not all the restaurants are captured. But it should give a good idea of the availability and proportion of different types of restaurants in each city
- Radius for search is within 10 km

Results

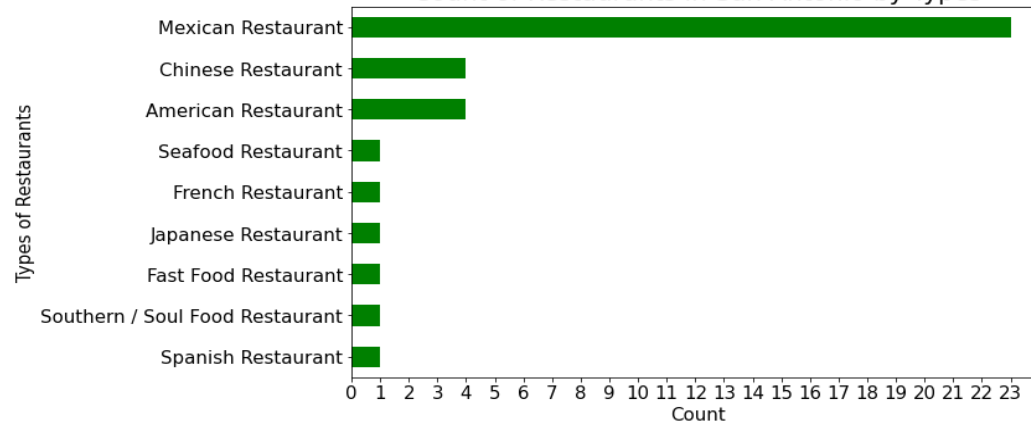
Count of Restaurants in Austin by Types



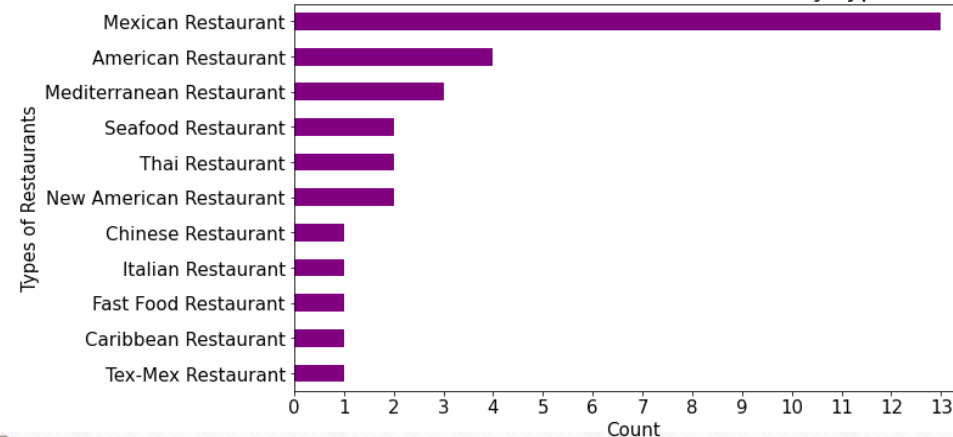
Count of Restaurants in Dallas by Types



Count of Restaurants in San Antonio by Types



Count of Restaurants in Fort Worth by Types



Results

- Study popular spots and learn the locals' interests for future marketing plans
- Analyze common venues and display the top 10 venues in each city
- Data obtained from exploring the city using Foursquare API

	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Austin	Hotel	Coffee Shop	Pizza Place	Park	Food Truck	Seafood Restaurant	Cocktail Bar	Italian Restaurant	Performing Arts Venue	Movie Theater
1	Dallas	Coffee Shop	Hotel	Brewery	Burger Joint	Park	Steakhouse	Bar	New American Restaurant	Cocktail Bar	Mexican Restaurant
2	Fort Worth	Mexican Restaurant	Brewery	American Restaurant	Pizza Place	Seafood Restaurant	Coffee Shop	Burger Joint	Taco Place	Café	Art Museum
3	Houston	Coffee Shop	Bar	Beer Garden	Vietnamese Restaurant	Park	Café	Mexican Restaurant	Trail	Grocery Store	Brewery
4	San Antonio	Hotel	Ice Cream Shop	Plaza	Beer Garden	Brewery	Theater	Coffee Shop	Cocktail Bar	Restaurant	Lounge

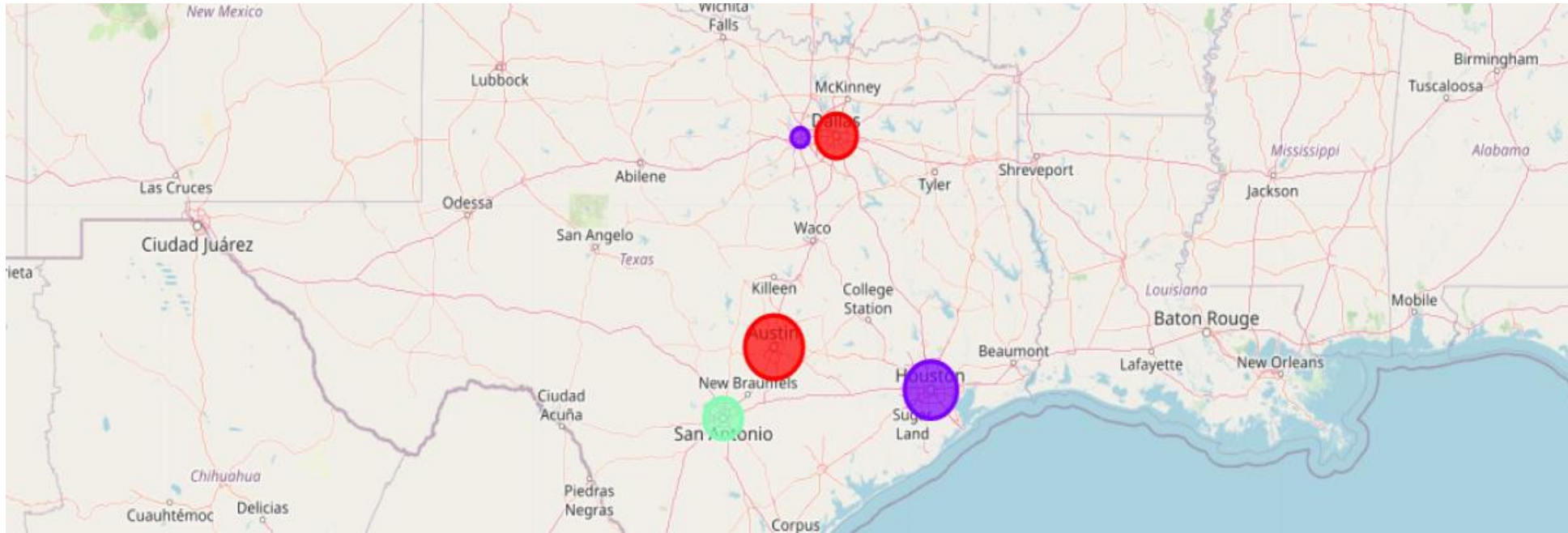
Results

- Cluster the cities based on their venues using K-Means clustering
- Austin and Dallas are clustered together with their top venues being hotels and coffee shops
- Fort Worth and Houston are in another cluster with restaurants and bars being their common venues
- San Antonio is separated in a different cluster with a lot of hotels, ice cream shops and plazas

	City	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	% Asian	Median household income	Median gross rent	Latitude	Longitude
0	Austin	2	Hotel	Coffee Shop	Pizza Place	Park	Food Truck	Seafood Restaurant	Cocktail Bar	Italian Restaurant	Performing Arts Venue	Movie Theater	7.6	71576	1280	30.271129	-97.743700
1	Dallas	2	Coffee Shop	Hotel	Brewery	Burger Joint	Park	Steakhouse	Bar	New American Restaurant	Cocktail Bar	Mexican Restaurant	3.4	52580	1052	32.776272	-96.796856
2	Fort Worth	0	Mexican Restaurant	Brewery	American Restaurant	Pizza Place	Seafood Restaurant	Coffee Shop	Burger Joint	Taco Place	Café	Art Museum	4.6	62187	1060	32.753177	-97.332746
3	Houston	0	Coffee Shop	Bar	Beer Garden	Vietnamese Restaurant	Park	Café	Mexican Restaurant	Trail	Grocery Store	Brewery	6.8	52338	1041	29.758938	-95.367697
4	San Antonio	1	Hotel	Ice Cream Shop	Plaza	Beer Garden	Brewery	Theater	Coffee Shop	Cocktail Bar	Restaurant	Lounge	2.8	52455	992	29.424600	-98.495141

Results

- Visualization of clusters in leaflet map using Folium
- Colors distinguish clusters and size of marker represents the number of Thai restaurants in that city



Discussion and Conclusion

- Fort Worth appears to have lowest competition
 - Lowest number of Thai restaurants within specified radius, while having Thai restaurants as the most popular Asian restaurants in the city
 - Second highest median household income and decent percentage of Asian population
- Should drill down further to study the neighborhoods in each city and also consider other factors such as distance from offices, proximity to public transportation, and distance from attractions
- Should be able to use the insights we gained from studying top venues to help in making future marketing plans.
 - For instance, since coffee shops and ice cream shops are popular in these cities, we might plan to include them in the menus or open a mini dessert bar to attract more customers in the future
- Should also be able to use Foursquare API to explore trending venues with highest foot traffic at a particular time to help in determining specific locations for the restaurant or where to promote the restaurant