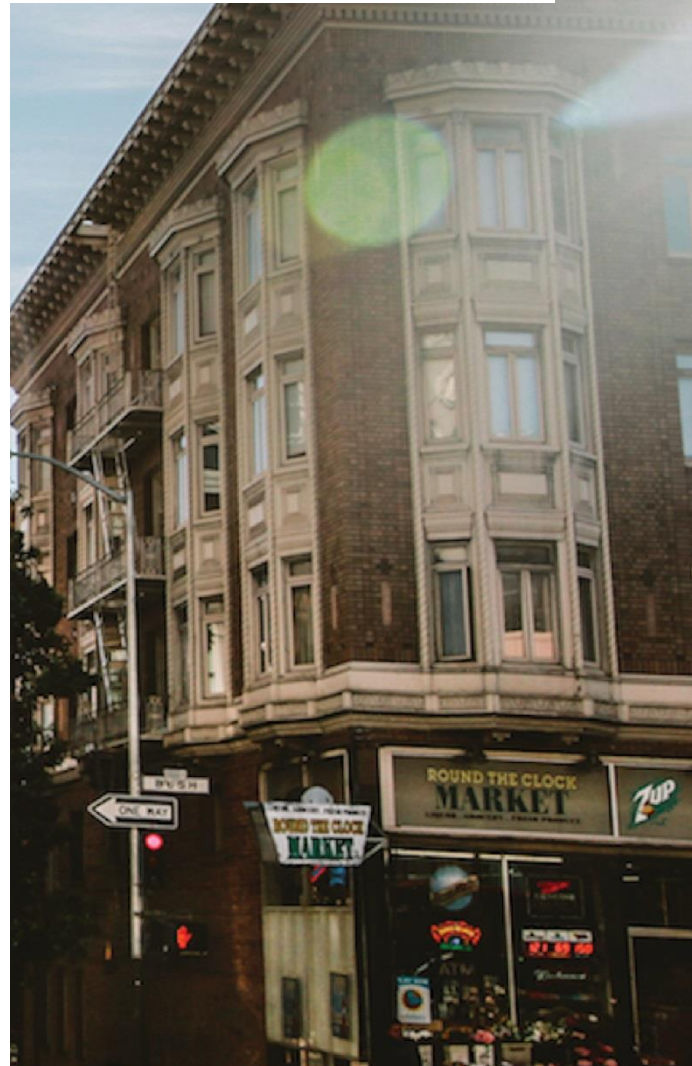


# CAPSTONE PROJECT REPORT ON THE BATTLE OF NEIGHBORHOODS

By Hla Tun



---

## Introduction/Business Problem

The main purpose of this project is to find a proper location for a restaurant in Texas. Assuming my client is an American from an other state (not from Texas), I am trying to work for him/her. Since, we don't know what type of restaurants (such as american, indian, asian etc.) are popular in Texas so first we would like to study on it. Then we will study the population density and per capita income in different cities in Texas. In addition to this, crowded and popular locations of the city is judged by the availability of the other venues such as restaurants, shopping malls, and other recreational venues. In essence, the summary of the project will be divided into two parts as follow:

- Part 1: Data collection
- Part 2: Solving the problem based on population, population density, per capita income, and popular venues. Foursquare API is used in this work to access the different venues in the Texas state.

## Description of Data

### *Wikipedia Sources*

- [https://en.wikipedia.org/wiki/List\\_of\\_United\\_States\\_cities\\_by\\_population](https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population)
- [https://en.wikipedia.org/wiki/List\\_of\\_Texas\\_locations\\_by\\_per\\_capita\\_income](https://en.wikipedia.org/wiki/List_of_Texas_locations_by_per_capita_income)

### *Foursquare API*

- For different venues in the city

## Methodology

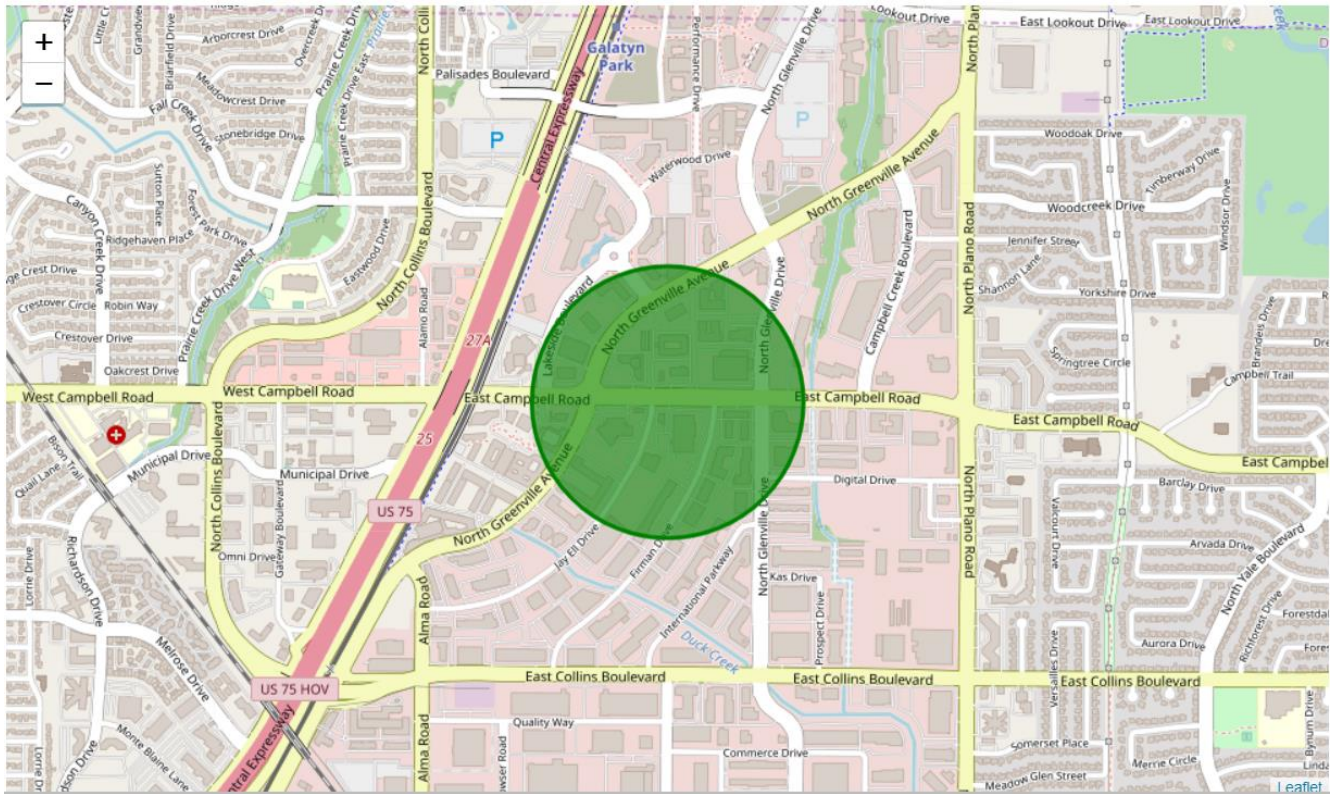
- Wikipedia data is used to get the population density, per capita income, population, cities, and the locations of the cities. Using beautiful soup we can easily get these information from the Wikipedia.
- The raw tables from the Wikipedia is post-processed. Then a simple dataframe is prepared.
- Graphical visualization technique was used to see it on the Texas map.
- Foursquare API is accessed by using my client ID and Client secret.

- The locations (latitude and longitude) of the cities are used to extract the venues from Foursquare API.
- A word cloud of the restaurants was prepared to visualize the most popular restaurant types in Texas.
- Benchmarks are set for the venues for further calculations.
- Based on benchmarks the most suitable city for restaurant is selected using K means algorithm.
- Then I estimate the most suitable location of the city for the restaurant.

- Mexican restaurants are the most popular restaurants in Texas.



- Irving is the most suitable city in Texas for opening a new restaurant. The exact coordinate for the restaurant location: latitude = 32.975230 and longitude = -96.709309.



## Discussions and Improvements

In this work, we choose the benchmark very roughly to estimate the most suitable city in Texas. However, for better estimation, we need to choose the benchmark very carefully. Furthermore, It is also necessary to study the ethnicity/race in the green region to choose weather the opening of the Mexican restaurant is profitable or not. The restaurant business might also depends on the grocery stores of the nearby areas. For extensive study, we might need to consider it.

## Conclusions

Based on the brief study on this project, we can choose the green circles in the map to open a Mexican restaurant in Irving, TX.