

# Capstone Project

## The Battle of Neighborhoods, Sushi Restaurants in Houston

Noriko Yao

December 20, 2020

### 1. Introduction

#### 1.1. Problem Statement

A Japanese Sushi Chain Restaurant Owner wants to increase the business outside of Japan. Sushi has been one of the popular cuisines in US and known as bringing a full of healthy benefit such as protein, vitamins, antioxidant, and omega 3 fatty acids. Based on her market research data, Houston has been selected to develop business opportunity in the Southern city in the United State. To understand the locations and what types of venues are available in Houston and where they are located will be essential information for business development team to select the neighborhood in Houston.



Figure 1 (Left) Traditional Sushi Restaurant, Figure 2 (Right) Evolving Sushi Restaurant in Houston

Houston is the fourth most populous city in the United States with the estimated population of 2,320,268 as of July 1st, 2019. The city is in Southwest Texas and the most populous city in Texas as of July 1st, 2020 Data. This Southern City has been increasing the population growth rate 21.6% between 2010 through 2019 and expected to grow with the rate of 17.2% in 2020-2029. As you can see, there have been huge population surge in Houston over a decade and expected to continue. In 2019, Houston was ranked as the most diverse city in the United State according to the new report (<https://abc13.com/houston-diversity-most-diverse-city-in-america-the-united-states-big/5265339/>). As restaurant chain owners, it's potential opportunity to increase the numbers of their restaurants in this populous and the diverse city of Houston.



Figure 3 Houston Downtown

## 2. Data Requirements

The data used in this project are:

- Foursquare API: <https://api.foursquare.com>

Extracting the venue information for restaurants in Houston, Foursquare API was used for this project. This data includes latitude and longitude information for each venue.

- Houston ZIP Code and latitude/longitude Information

This data was not easily obtained from websites. The data was extracted from several websites as shown below. "Houston\_zip\_lat\_long.csv" file (Uploaded in Github) gathered and combined and evaluated for this project.

<https://public.opendatasoft.com/explore/dataset/us-zip-code-latitude-and-longitude/table/>

<https://cohweb.houstontx.gov/ZipCodePublic/>

- OPEN DATA HOUSTON

<http://data.houstontx.gov/>

<https://datausa.io/profile/geo/houston-tx#housing>

- List of Houston neighborhoods

This information contains many neighborhoods, ranging from planned community to historic wards. City of Houston recognize a lost of 88 super neighborhoods which encompass broadly recognize regions. According to the city, a super neighborhood is *"geographically designated area where residents, civic organizations, institutions and businesses work together to identify, plan, and set priorities to address the needs and concerns of their community"*.

[https://en.wikipedia.org/wiki/List\\_of\\_Houston\\_neighborhoods](https://en.wikipedia.org/wiki/List_of_Houston_neighborhoods)

### 3. Data Acquisition and Explorations

Many data are available in many different formats in order to analyze and explore useful information. But also, the age of data and the accuracy of the data source is important to review prior to utilize and also lead a decision for business.

#### 3.1.1. Data Source

Many useful data are available by the [Open Data Portal](#). The Open Data Portal is City of Houston's platform to host City-Wide open data. This open portal includes the more than 200 datasets such as Zip Codes in the regions, Houston Super Neighborhood boundaries. This data was used to construct the database for identifying potential locations of the restaurants. Household Income information in neighborhood is critical information to identify the best spots to open restaurants to get customers. Since restaurant business is all about location, location, and location, gathering data set from DATA USA, and constructing data frame from the datasets was necessary steps.

#### 3.1.2. Data Cleaning and Formatting

After constructing datasets into dataframe, cleaning up and formatting data and dealing with missing data is another step to take. Firstly, identify the locations of the sushi restaurants in Houston, limited to 5,000, within 100,000 meters away from the center of Houston official address, 901 Bagby, Houston, Texas for obtaining GeoJSON. At the same time, obtaining the list of restaurants in general with the same criteria to check if anything outstands in comparison.

Secondly, the venue data from Foursquare include 25 columns of information which includes useless data for this project being removed. However, the location data such as latitude and longitudes, address of the restaurants, cross streets and categories were filtered. Also, datasets which recognized as redundant, too obvious, unnecessary have been removed from this analysis. All NaN row of data was intentionally removed, since most of information was not numerical for this project.

### 4. Exploratory Data Analysis

#### 4.1. Restaurant Information in Houston SN

First, according to City of Houston, there are 88 neighborhoods and historical wards shown below. The population is 2,320,268 as of 2019 in a total area of 637.4 square miles (1,651 km<sup>2</sup>).

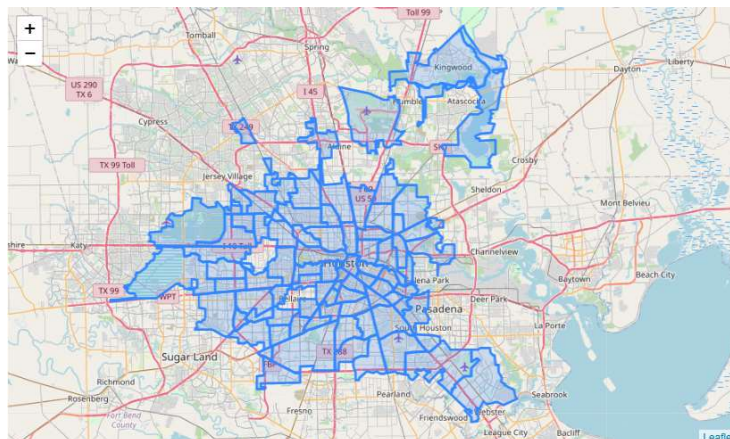


Figure 4, 88 Super Neighborhoods in Houston

The 88 neighborhoods are called Super Neighborhood or SN in this report. The shapefile was obtained from City of Houston, <https://cohweb.houstontx.gov/ZipCodePublic/> via API access. The approximate boundaries of each SN relative to Downtown Houston is extracted from Wikipedia as shown below.

#	Neighborhoods	Location in Houston	Approximate boundaries
0 1	Willowbrook	Northwest	Along Texas State Highway 249 northwest of Bel...
1 2	Greater Greenspoint	North	Around the junction of Beltway 8 and Interstat...
2 3	Carverdale	Northwest	South of the junction of Beltway 8 and U.S. Ro...
3 4	Fairbanks / Northwest Crossing	Northwest	Along U.S. Route 290 between Interstate 610 an...
4 5	Greater Inwood	Northwest	North of Fairbanks / Northwest Crossing and ea...
...	...	...	...
83 84	Spring Branch North	Northwest	East of Beltway 8, south of Clay Road, west of...
84 85	Spring Branch Central	Northwest	Between Blalock Road and Bingle Road north of ...
85 86	Spring Branch East	Northwest	Enclosed by Blalock Road, Highway 290, Interst...
86 87	Greenway / Upper Kirby	Southwest	Along Interstate 69 between Bissonnet Road and...
87 88	Lawndale / Wayside	Southeast	Northeast of Interstate 45 along Brays Bayou b...

88 rows × 4 columns

Figure 5 The dataframe scraped from Wikipedia Houston Super Neighborhood Table

First, utilizing Foursquare.com the venue data of any restaurants in Houston SN within 100,000 ft range was extracted. The center of Houston address was obtained from City of Houston Official address, which is 901 Bagby, Houston TX (latitude: 29.760179, longitude: -95.3693746509527).

```
[2]: CLIENT_ID = '5TSQV2RFYETFWILHHRYGS4EP3NZRLUHWDXEFPOC1BWZXNMW' # your Foursquare ID
CLIENT_SECRET = 'RI2EKFYIORRBPYMLRX4YXLD21TZGKGPLRBCVV5C3K5ESVPKY' # your Foursquare Secret
ACCESS_TOKEN = 'HSDZDTVDTJ0A1UPNRPDCTWXAGIVNBOBEYZSLV4AVXZN4EZBF' # your FourSquare Access Token
VERSION = '20180604'
LIMIT = 100000
print('Your credentials:')
print('CLIENT_ID: ' + CLIENT_ID)
print('CLIENT_SECRET: ' + CLIENT_SECRET)

Your credentials:
CLIENT_ID: 5TSQV2RFYETFWILHHRYGS4EP3NZRLUHWDXEFPOC1BWZXNMW
CLIENT_SECRET: RI2EKFYIORRBPYMLRX4YXLD21TZGKGPLRBCVV5C3K5ESVPKY
```

Figure 6 Access to Foursquare.com for the venue information

Based on the same criteria above, Figure 6 bar chart was generated based on the venue “Restaurant” response from Foursquare.com is shown below.

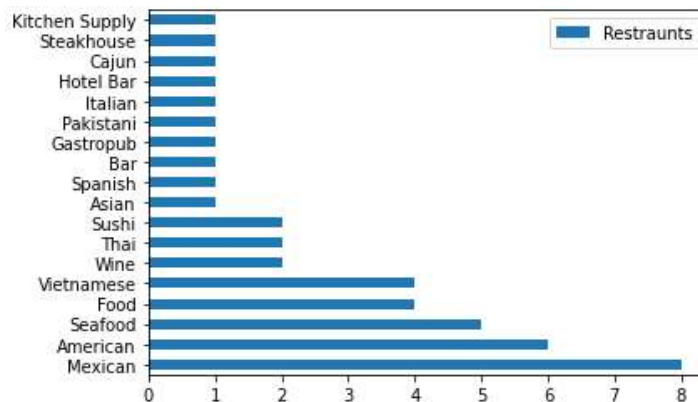




Figure 6 The types of the venues in Houston

The total response was 49 responses as the max limit, and the number of Sushi restaurant is 1 (2%). Next, “Sushi” venue search on Foursquare.com was extracted with the same criteria within 100,000 ft from 901 Bagby, Houston TX (latitude: 29.760179, longitude: -95.3693746509527), the 49 names of the restaurants were found as shown in Figure 7.

```
[14]: 0 Spindletop Restaurant at Hyatt Regency Houston
      1 Andalucia Tapas Restaurant and Bar
      2 Kim Son Restaurant - Downtown
      3 Harry's Restaurant
      4 Cuchara Restaurant
      5 Mai's Restaurant
      6 Trofi Restaurant
      7 R.K.G. New Orleans Cajun Po-Boy Restaurant
      8 Njoy Thai Restaurant
      9 Dover's Restaurant
     10 Bouray's Burrito Bar and Vietnamese Restaurant
     11 Aquarium Restaurant
     12 Huynh Restaurant
     13 Restaurant Depot
     14 Padthai Thai Restaurant
     15 Ninfa's Restaurant
     16 Merida Mexican Restaurant
     17 Chin Tao Chinese Restaurant
     18 Lorenzo's Restaurant
     19 The Fish Restaurant & Sushi Bar
     20 Charlie's Restaurant and Lounge
     21 Spanish Flowers Mexican Restaurant
     22 Alma Latina Mexican Restaurant
     23 Metro Restaurant
     24 The Pearl Restaurant & Bar
     25 Oishii Japanese Restaurant & Sushi Bar
     26 Zula Restaurant
     27 Brazos Restaurant & Bar
     28 Picos Restaurant
     29 House of Blues Restaurant & Bar
     30 Brazos Restaurant (Crowne Plaza Hotel)
     31 Sage Event Center & Restaurant
     32 Quatro Bar & Restaurant
     33 New York Houston Restaurant
     34 Morningside Thai Restaurant
     35 Fifteen Restaurant and Lounge
     36 Monarch Restaurant
     37 American grill restaurant
     38 Connie's Seafood Market & Restaurant
     39 Tony's Mexican Restaurant
     40 Wasabi Japanese Restaurant
     41 Shiva Indian Restaurant
     42 Sonoma Wine Bar & Restaurant
     43 El Jardin Mexican Restaurant
     44 Juanita's Mexican Restaurant and Cantina
     45 Segari's Seafood Restaurant and Bar
     46 Himalaya Restaurant
     47 Aga's Restaurant & Catering
     48 Don Jose Mexican Restaurant
     49 Korea Garden Restaurant
      Name: name, dtype: object
```

Figure 7 the names of 49 Restaurants in Houston Super Neighborhood

In this study, the correctness of venue categorization was confirmed. Sometimes Sushi restaurants were identified Seafood restaurants or Asian restaurants instead. By this checking, those types of errors are insignificant and ignore another data cleaning steps for this analysis. The locations of those restaurants are mapped via folium in Python as shown as below.

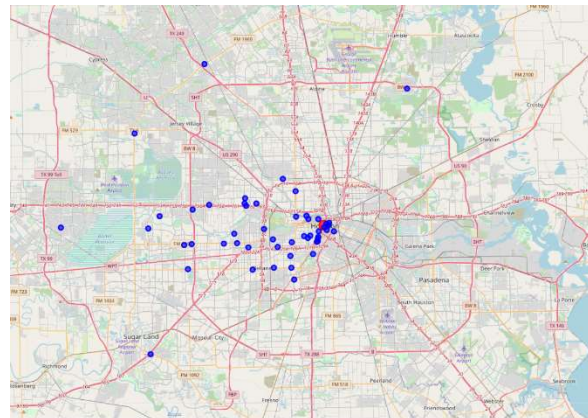
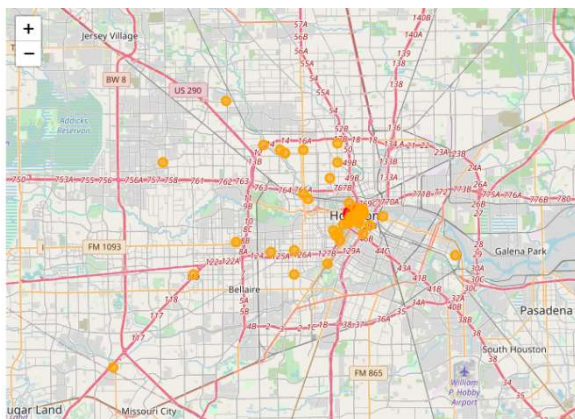


Figure 8 (Left) the location of restaurants in general, Figure 9 (right) the location of Sushi restaurants found from Foursquare.com

As you can see, most of restaurants are located in Northwest to Southwest of Houston at large. At the east side of Houston, related to the port of Houston activities and also Chemical plants and oil refineries are located. Those area are not ideal locations to open Sushi restaurants where business people, hung with friends and famili members in the evnining and on weekends.

## 4.2. Restaurant Information in Houston SN

Understanding the best locations to open new restaurants, understanding the median household income in the neighborhood is also critical. From City of Houston Open Data Portal, each SN median household income information was extracted. The data was published in 2015, and slightly old data. However, this information was only publicly available for free of charge, it's decided to use for this analysis.

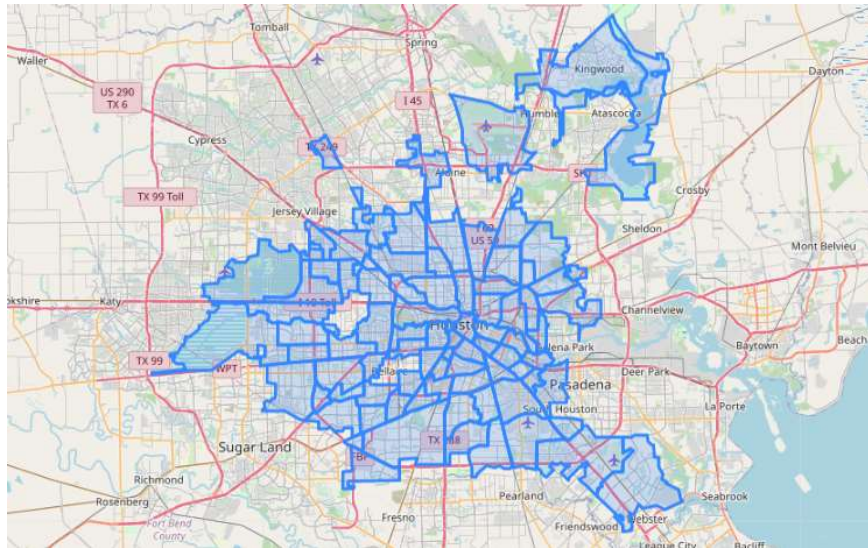


Figure 10, 88 SNs (Super Neighborhood) in Houston

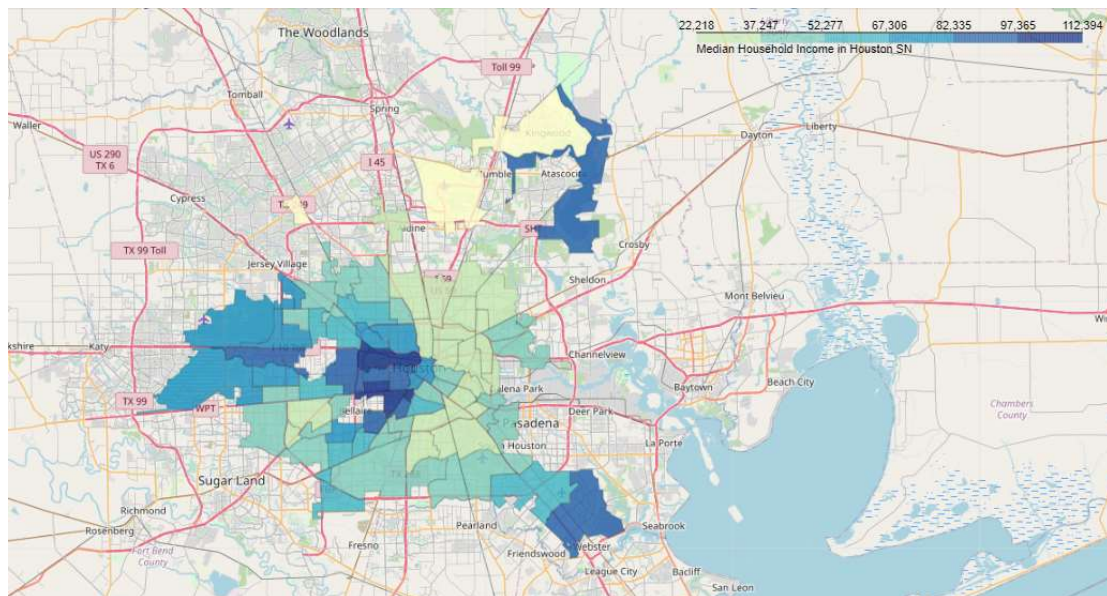


Figure 11 Median Household Income in Houston SNs

Comparing Figure 8 and 9 with Figure 11, as you can see the locations of the restaurants follows Median Household Income information. Most of Sushi restaurants seem to be in the areas beyond \$67,306 USD (Data as of 2015). The areas recognized for potential SN where future Sushi restaurants to be located are shown below Table 1.

SNNAME	population	income
UNIVERSITY PLACE	16342	111510
MUSEUM PARK	5509	101703
GREENWAY / UPPER KIRBY AREA	21120	100274
WASHINGTON AVENUE COALITION / MEMORIAL PARK	29033	99302
AFTON OAKS / RIVER OAKS AREA	14518	95682
LAKE HOUSTON	22280	95074
MEMORIAL	47604	94927
BRAESWOOD	21835	93748
FOURTH WARD	4085	88231
NEARTOWN - MONTROSE	31037	85296
CLEAR LAKE	620268	84710
GREATER UPTOWN	50731	83399
MEDICAL CENTER AREA	2717	82830
MEYERLAND AREA	21445	81408
WESTBRANCH	3633	81240
ADDICKS PARK TEN	19683	80584
GREATER HEIGHTS	41362	80048
LAZYBROOK / TIMBERGROVE	13099	80032
ELDRIDGE / WEST OAKS	72347	79787
MIDTOWN	8597	77261
BRIAR FOREST	43018	76462
DOWNTOWN	12088	71666
SPRING BRANCH WEST	31878	71534
CARVERDALE	3903	68107

Top 10 Median Household Income in Houston SN in 2015

Top 20 Median Household Income in Houston SN in 2015

Table 1 The SNs where the median household income greater than \$67,306 USD

Out of this top 20 highest household income list, narrowing the areas to study in deeper level to understand the quantity of sushi restaurants in each SN is conducted as below.

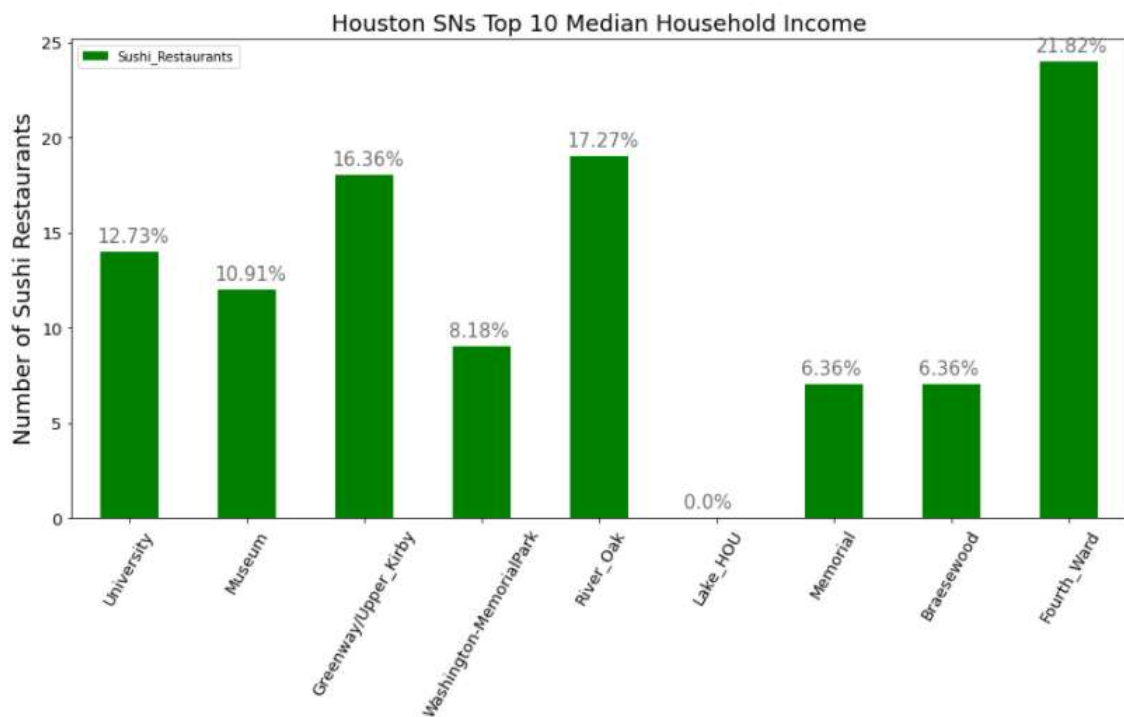


Chart 1 The number of Sushi Restaurants in the top 10 median household income SNs



## 5. Results

In this analysis, the top priority is to identify the locations of opening new sushi restaurants in Houston neighborhoods. Therefore, at first, I did research on the popularity of Sushi Restaurants in order to see if there is demand in this neighbor. Sushi restaurants were 8<sup>th</sup> popular restaurants in Houston, which was recognized as “popular” enough to say there is a demand. Secondly, it was one of the important to understand “where” to open them but had to start from understand how the city is arrange in borough and neighborhood, and then map out where they are typically located at in Houston neighborhoods. Interesting enough, this analysis has shown the most of Sushi restaurants were located in westside of Houston (Northwest, West, and Southwest) due to the east side is mainly occupied by industrial areas such as the port activities, Chemical plants, and oil refineries along Gulf of Mexico. Interestingly, where the restaurants in general located were matching where Sushi restaurants are located. I did not find anything significant areas where sushi restaurants are heavily located in this analysis. Next, in order to identify the locations to open a new sushi restaurant, the median household income in Houston Super Neighborhood is another critical piece of the information to narrow down the candidate locations. As you can see in Chart 1, there was the top median household income, but there were zero sushi restaurants recognized by this analysis in Lake Houston area. It could a potential location by marketing strategy but there could be a reason why no sushi restaurants in the neighborhood due to expensive tenant rents or no tenants’ buildings, neighborhood surrounding age and culture background, population, and commute distance from crowds. However, Fourth Quarter, River Oak, and Greenway/ Upper Kirby neighborhoods are the Top 3 Sushi restaurants populated in the top 10 median household income neighborhoods in Houston, which means these neighborhoods have certain demands of sushi. However, on the other hand, there could be some tough competition and potentially monthly tenant rent could be significantly higher than other locations.

## 6. Conclusion

As the world moving faster, as an entrepreneur we are forced to make critical decisions quickly and executing actions on the spot to survive in high competitive business fields nowadays. Currently, we need and depend on data from government, data management & service companies, connecting to social media and digital technology service companies via API has been one of the key process to obtain the latest data and digest to present the findings visually to help us make decision day to day basis in many industries now. During this project, I was surprised to see how difficult to acquire the latest, accurate, and reliable data for free. Most of the case, they ask for some fee to data utilizations. Other thing was the inputs from subject matter experts on our analysis will be a key not to derail the course of analysis paths. Knowing the subject at the same time knowing data analysis will be very powerful to detect what’s going on in matters, and also pick the best solutions to a problem we are supporting as a data scientist. In this project, I did feel the need to obtain the tenant rent information in those super neighborhoods to tune in and have more depth of the analysis. Unfortunately, I found some information on several websites, but did not seem to be reliable or data was too old to use for this project, since population increasing in Houston rapidly and restaurant business landscape may be different from 5 year ago. Especially COVID, I am sure it will be different from the last year due to city lockdown. However, during this project, I could experience how the data can be obtain and digest from the available resources, and how much it will be useful tool to make decisions. But at the same time, I could peak their challenges to obtain data, quality of the data, and so on. Once those challenges are solved, it will speak by itself and will be more eloquent in business or technology scenes.