

INVESTIGATION OF HOUSTON CITY NEIGHBORHOODS

HAO LI

2019-10-14

INTRODUCTION

- **Background**

- The city of Houston is one of the largest cities in Texas. It is the fourth most populous solo city in the United States. Houston is famous for its economy in broad industries such as energy, manufacturing, aeronautics, and medical.
- Houston is the largest city in the United States by total area. It has many neighborhoods. Houston city has been described as the most racially and ethnically diverse major metropolis in the U.S.A. It is home of many cultural institutions, restaurants and millions of people from all over the world.

- **Business Problem**

- Houston is a big city and heavily diversified. Most people don't know the features of the neighborhoods outside their living/working area. It will be interesting to investigate the neighborhoods, hence people can quickly find places to eat, shop, entertainment, etc. This project is going to investigate the venues within neighborhoods of Houston city and make a thorough report.

- **Interest**

- Then of thousands of people visit Houston over the world each year. Knowing more about the city before arrival is always helpful. This investigation will be very handy for visitors to get familiar with the city and plan activities. Besides that, Houston residents also benefit from the results and can enjoy more about this great city.

DATA ACQUISITION

- The data acquisition steps are:
 1. Get the neighborhood names from the Wikipedia page of Houston.
 2. Use the Geocoder Python package to get the latitude and the longitude coordinates of each neighborhood.
 3. Use the Foursquare to collect the venues of the neighborhoods.
- The data process and analysis steps are:
 1. Clean the data and prepare for clustering
 2. Check the venues and plot on the map
 3. Data Example

RAW DATA EXAMPLE

#		Name	Location relative to Downtown 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...

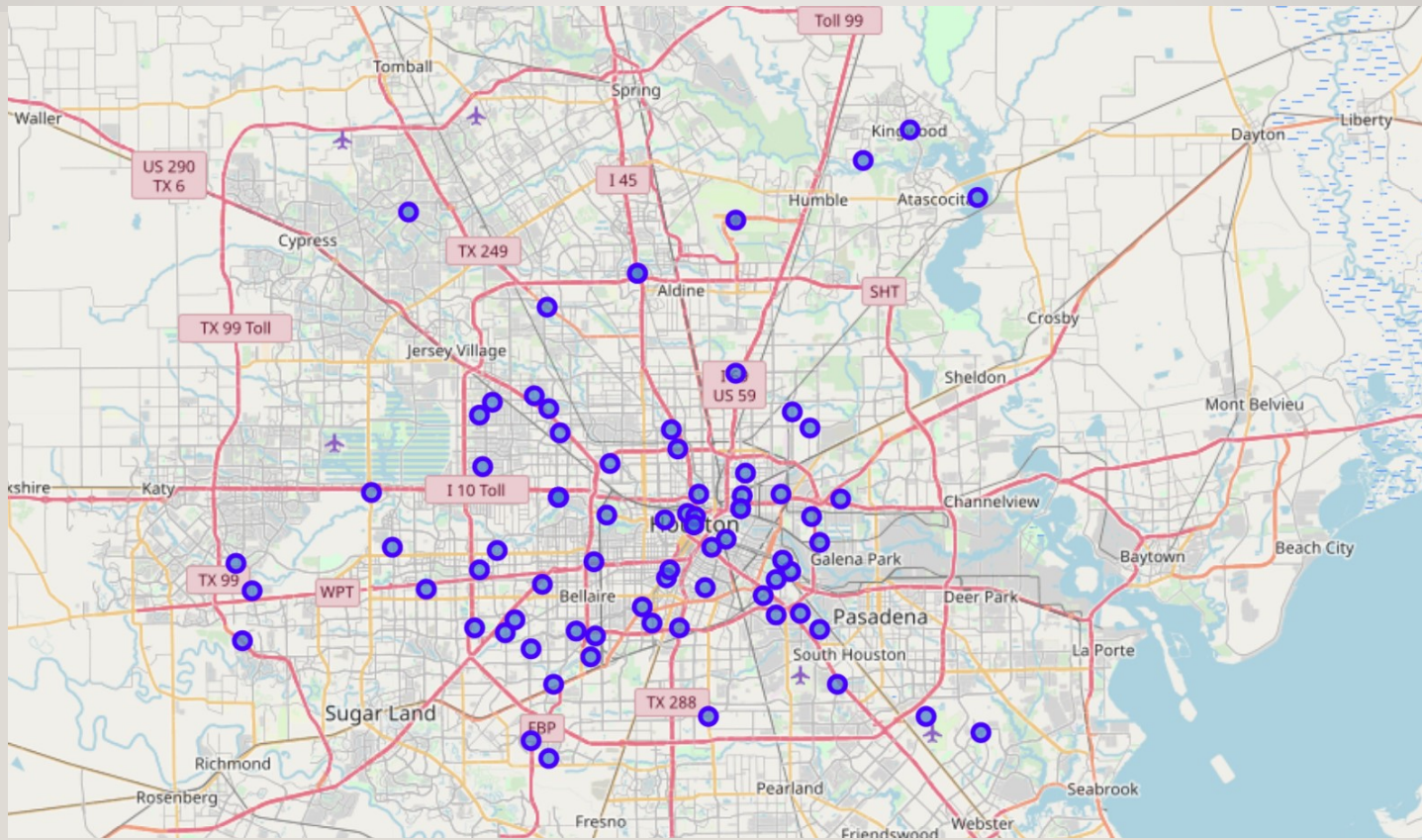
METHODOLOGY

- To perform the analysis, we will work on key steps as below:
 - Exploratory Data Analysis (EDA)
 - Data Collection of Latitude/Longitude from Geocoder
 - Houston Map Using Latitude/Longitude Collected
 - Define Foursquare Credentials and Version
 - Using Foursquare to Investigate All the neighborhoods
 - Houston Neighborhoods Analysis Using K-Means Method

PREVIEW OF PROCESSED DATA

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Acres Home	Health Food Store	Intersection	Pharmacy	Park	Pub
5	Braeburn	Bar	Food	Indian Restaurant	Bank	School
10	Central Northwest (formerly Near Northwest)	Pizza Place	Chinese Restaurant	Mexican Restaurant	Spa	Big Box Store
15	East Houston	BBQ Joint	Museum	Convenience Store	Moving Target	Seafood Restaurant
20	Eldridge / West Oaks	Home Service	Convenience Store	Paper / Office Supplies Store	Clothing Store	Pet Store
25	Fourth Ward	Park	Hotel Bar	Restaurant	Rock Club	Salon / Barbershop
30	Greater Inwood	Pet Store	Bakery	Bed & Breakfast	Sporting Goods Shop	Fish & Chips Shop
35	Gulfgate Riverview / Pine Valley	Accessories Store	Baseball Field	Park	Skate Park	Zoo Exhibit

RESULT ANALYSIS: MAP BEFORE CLUSTERING



DISCUSSION

- For Houston, the neighborhoods are clustered into 6 groups.

The number size of each group:

- Group1 - 11
- Group2 - 37
- Group3 - 7
- Group4 - 24
- Group5 - 2
- Group6 – 6

DISCUSSION

- The biggest group has 37, and it shows the wide spread of the restaurants. This makes good sense as Houston is a city with name for its diversity and good food. Due to weather conditions, Houston is not a place with a lot of travel sites. The second largest one has 24, and it shows the shopping is a popular activity. The third one has 11, which indicates the activities around the Banking. Considering it as the oil Capital city, one can imagine how busy the banks are. The fourth group shows the coffee and fast food. The fifth one which has a number as 6, shows the real life of Houstonians, they go to supermarkets, home furniture stores, etc. The smallest one, which only has a number 2, reflects the only popular visiting area of Houston, the Memorial park area. It has a big zoo, a good outdoor park, and museums nearby. This is the place people looking for relaxation with family

CONCLUSION

- The project gives good results with useful information. As we know, everyday many people come to Houston over the world. Houston is so big and diversified. The visitors would want to know more about the city before they arrive. This result is very useful for them to get familiar with the city and plan for their trip. The local people can get benefits as well.
- Looking forward, this project has space to be further improved. One example is that it may combine with zip code along with the population information. As the project is based on the neighborhoods, and the neighborhoods varies in the size, population, etc. One example is that if we weighs the population factor, that probably will improve the analysis accuracy.