Coursera Capstone

IBM Applied Data Science Capstone

Opening a New Healthy Fast-Food Restaurant in Houston, United States

By: Kimberley Wagner

March 2021

Introduction

Many people have the goal of improving their health through diet. However, with the challenges of balancing home and work life, it can be difficult to find time to make healthy food at home and many people turn to fast-food restaurants. Providing nutritious fast-food options allows people to meet their goals of improving their diets without the time demands of home cooking. Health focused fast-food restaurants can be a lucrative business, especially in neighbourhoods where healthy convenient options are limited. Opening a restaurant is, however, always a risky decision and a thorough analysis should be undertaken to increase the chance of sustaining a successful business. The location of the restaurant is one of the most important decisions that will determine whether the business will be a success or failure.

Business Problem

The objective of this capstone project is to analyse and select the best locations in the city of Houston, Texas to open a new healthy fast-food restaurant. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the business question: In the city of Houston, Texas, if a potential owner is looking to open a new healthy fast-food restaurant, where would you recommend that they open it?

Target Audience

This project is particularly useful to potential owners and investors looking to open or invest in a restaurant in Houston, Texas. As the rate of obesity continues to increase across the country, and especially in the Southern states like Texas, the demand for convenient and healthy options increases. This provides opportunity for investors to fill the gap of nutritious fast food options in a neighbourhood that might be dominated by unhealthy, meat-dominated, or slow food options only.

Data

Neighborhood data was scraped from Wikipedia (‘https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Houston’), latitude and longitude pulled from Google, and venue data acquired from the FourSquare API. The area of investigation was limited to a ~5 mile radius around downtown Houston, in order to target the higher population density areas.

Methodology

To begin, the data was scraped from Wikipedia. The data was well behaved and didn’t require any cleaning prior to working with it. The geographic coordinates were acquired from Google so that the

neighborhoods could be located on a map, then a 5 mile radius around downtown was defined and used to filter out the neighborhoods outside the area of interest. The remaining neighborhoods were plotted on a map (figure 1).

After this, venue data was pulled from the FourSquare API. As this pulls all venue data, it had to then be filtered to just contain restaurant information.

A required step before we can run the clustering algorithm is to use the one-hot encoding technique which converts the categorical values into dummies so they can be used for machine learning. K-Clustering was used in order to form clusters of neighborhoods, grouped by the most common types of restaurants (figure 2). K-Clustering was chosen as it proved useful in previous labs which required similar analysis.

Results

Details of the results can be found in the notebook provided on Github, but I have summarized the key points within this report.

88 neighborhoods were pulled from the Wikipedia page. Once filtered down to a ~5 mile radius around downtown, this was reduced to 21.

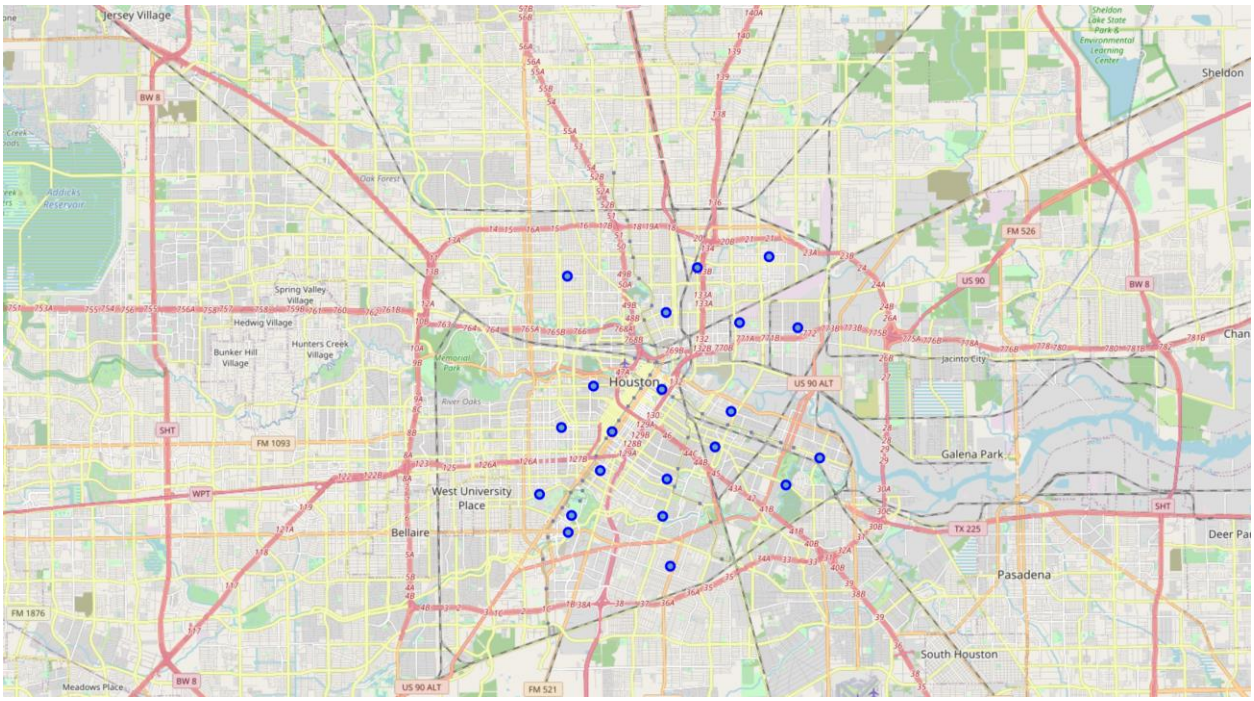

Figure 1. Neighborhoods within a 5 mile radius of downtown.

Table 1 shows the most common restaurant types per neighborhood, this was the data that was used to cluster the neighborhoods. The K-means clustering algorithm clusters the data by similar restaurant demographics, giving us 5 clusters.

Cluster 0: American Restaurant dominant

Cluster 1: Fast Food dominant

Cluster 2: A wide variety of food options (little fast food)

Cluster 3: American Restaurant dominant

Cluster 4: Fast Food dominant

	#	Neighborhood	Location relative to Downtown Houston	Approximate boundaries	Latitude	Longitude	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
0	15	Greater Heights	Northwest	East of White Oak Bayou, south of Interstate 6...	29.796250	-95.39487	2.0	Mexican Restaurant	New American Restaurant	Vietnamese Restaurant	Seafood Restaurant	Cajun / Creole Restaurant	Greek Restaurant	Vegetarian / Vegan Restaurant	Restaurant	American Restaurant	Southern / Soul Food Restaurant
1	24	Neartown / Montrose	Southwest	Straddles Montrose Boulevard east of Shepherd ...	29.742810	-95.39728	2.0	Italian Restaurant	Mediterranean Restaurant	Mexican Restaurant	New American Restaurant	Japanese Restaurant	Ramen Restaurant	Greek Restaurant	Vietnamese Restaurant	Szechuan Restaurant	Seafood Restaurant
2	28	University Place	Southwest	East of Kirby Drive, south of Interstate 69, w...	29.719490	-95.40643	2.0	American Restaurant	Italian Restaurant	Seafood Restaurant	French Restaurant	Greek Restaurant	Mexican Restaurant	New American Restaurant	Restaurant	Mediterranean Restaurant	Southern / Soul Food Restaurant
3	32	Braeswood	Southwest	Along north bank of Brays Bayou immediately ea...	29.705829	-95.39473	0.0	American Restaurant	Restaurant	Mexican Restaurant	Chinese Restaurant	Fast Food Restaurant	Japanese Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Dumpling Restaurant
4	33	Medical Center	South	South of Hermann Park between Main Street and ...	29.711790	-95.39315	0.0	American Restaurant	Restaurant	Fast Food Restaurant	Mexican Restaurant	Japanese Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Dumpling Restaurant
5	46	Jensen	Northeast	North of Interstate 610 between Hardy Toll Roa...	29.799362	-95.34218	1.0	Fast Food Restaurant	Mexican Restaurant	Vietnamese Restaurant	Latin American Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Dumpling Restaurant	French Restaurant
6	51	Northside Village	North	Enclosed by Interstate 10, Interstate 45, Inte...	29.783340	-95.35496	1.0	Fast Food Restaurant	Mexican Restaurant	Vietnamese Restaurant	Latin American Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Dumpling Restaurant	French Restaurant
7	52	Kashmere Gardens	Northeast	Inside Interstate 610 west of Elysian Street a...	29.803110	-95.31298	4.0	Fast Food Restaurant	Chinese Restaurant	Seafood Restaurant	Vietnamese Restaurant	Latin American Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Dumpling Restaurant	French Restaurant
8	55	Greater Fifth Ward	Northeast	North of Buffalo Bayou, east of Elysian Street...	29.780000	-95.32484	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9	56	Denver Harbor / Port Houston	East	East of Lockwood Drive, south of Liberty Road,...	29.778120	-95.30118	1.0	Fast Food Restaurant	Mexican Restaurant	American Restaurant	Latin American Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Dumpling Restaurant	French Restaurant
10	60	Fourth Ward	West	East of Taft Street, south of Buffalo Bayou, a...	29.757620	-95.38449	2.0	Vietnamese Restaurant	Seafood Restaurant	Italian Restaurant	Japanese Restaurant	Vegetarian / Vegan Restaurant	Mexican Restaurant	Portuguese Restaurant	Restaurant	Mediterranean Restaurant	American Restaurant
11	61	Downtown	—	Enclosed by Interstate 45 to the south and wes...	29.756260	-95.35656	2.0	Italian Restaurant	Mexican Restaurant	New American Restaurant	American Restaurant	Southern / Soul Food Restaurant	Cajun / Creole Restaurant	Dumpling Restaurant	French Restaurant	Hawaiian Restaurant	Indian Restaurant
12	62	Midtown	South	South of Interstate 45, north and west of Inte...	29.741550	-95.37687	2.0	Vietnamese Restaurant	Mexican Restaurant	French Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Dumpling Restaurant	Fast Food Restaurant	Italian Restaurant	Mediterranean Restaurant
13	63	Second Ward	East	North of Harrisburg Boulevard, east of BNSF li...	29.748480	-95.32843	1.0	Mexican Restaurant	Fast Food Restaurant	Vietnamese Restaurant	Latin American Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Dumpling Restaurant	French Restaurant
14	64	Greater Eastwood	Southeast	Between Interstate 45 and Harrisburg Boulevard...	29.735990	-95.33498	2.0	Mexican Restaurant	Thai Restaurant	Fast Food Restaurant	Italian Restaurant	Latin American Restaurant	Vietnamese Restaurant	Japanese Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant
15	66	Museum Park (formerly Binz)	South	Between Interstate 69 and State Highway 288 no...	29.727690	-95.38150	2.0	Southern / Soul Food Restaurant	American Restaurant	Mexican Restaurant	Vegetarian / Vegan Restaurant	French Restaurant	Fast Food Restaurant	Seafood Restaurant	Sushi Restaurant	Caribbean Restaurant	Chinese Restaurant
16	67	Greater Third Ward	South	South of Interstate 45 and east of Interstate 69	29.724790	-95.35440	4.0	Fast Food Restaurant	Vietnamese Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Seafood Restaurant	Latin American Restaurant	Caribbean Restaurant	Chinese Restaurant	Dumpling Restaurant	French Restaurant
17	68	Greater OST / South Union	Southeast	North of Interstate 610, east of State Highway...	29.693880	-95.35301	3.0	American Restaurant	Latin American Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Dumpling Restaurant	Fast Food Restaurant	French Restaurant	Greek Restaurant
18	82	Magnolia Park	East	East of Union Pacific line, south and west of ...	29.732280	-95.29223	1.0	Mexican Restaurant	Chinese Restaurant	Fast Food Restaurant	Restaurant	Vietnamese Restaurant	Japanese Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Dumpling Restaurant
19	83	MacGregor	Southeast	Along Brays Bayou south of Blodgett Street, ea...	29.711500	-95.35603	4.0	Chinese Restaurant	Fast Food Restaurant	Vietnamese Restaurant	Latin American Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Dumpling Restaurant	French Restaurant	Greek Restaurant
20	88	Lawndale / Wayside	Southeast	Northeast of Interstate 45 along Brays Bayou b...	29.722700	-95.30594	4.0	Fast Food Restaurant	Seafood Restaurant	Vietnamese Restaurant	Latin American Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Dumpling Restaurant	French Restaurant

Table 1. Neighborhoods and their most common restaurant type

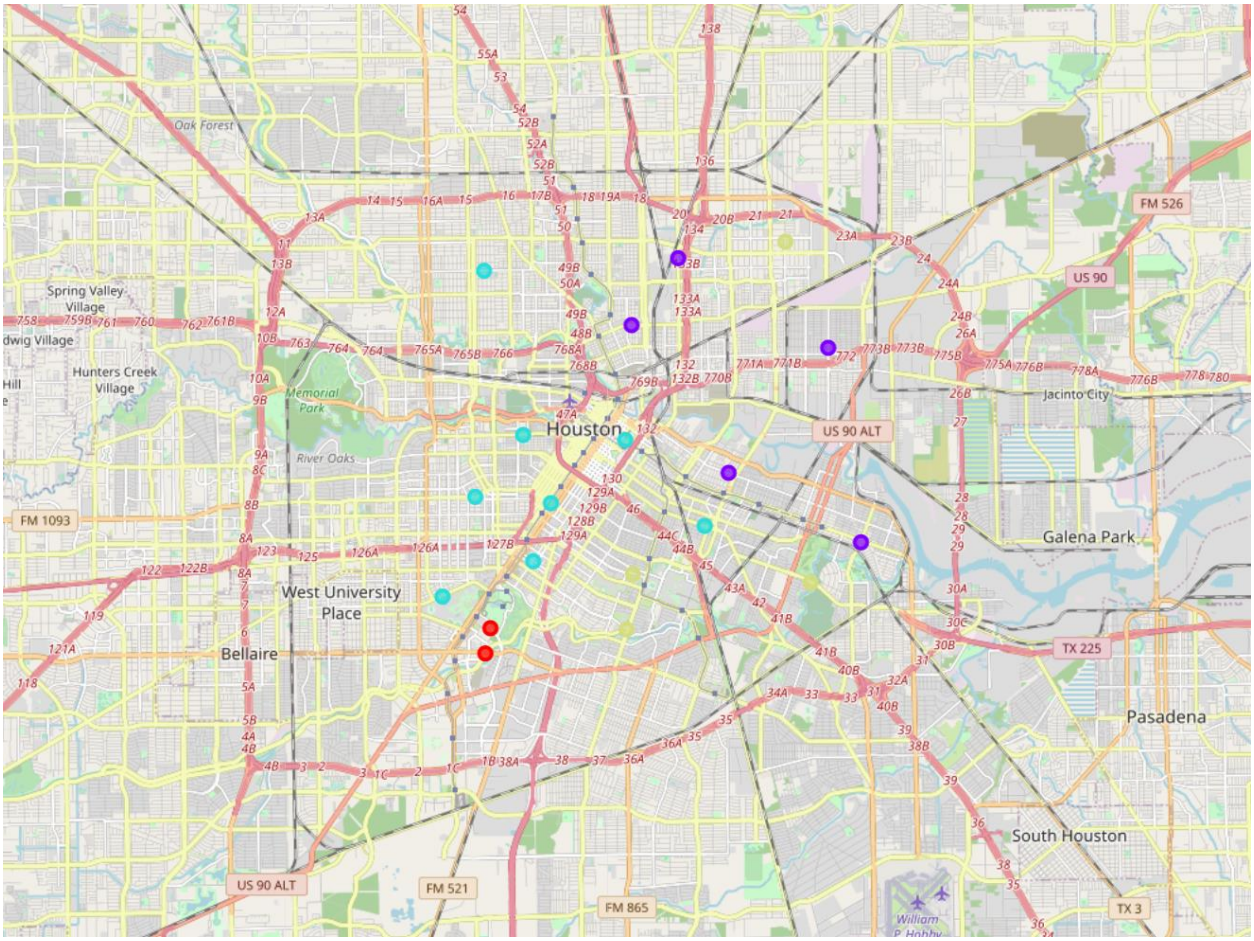

Figure 2. Neighborhoods clustered by their most common restaurant types.

Discussion

After investigating the clusters, it was discovered that clusters 1 and 3 (purple and yellow) already had fast food restaurants as their most common type. This makes them less than ideal for opening a new fast food restaurant, as there's already enough supply to meet the demand. Cluster 2 (blue) stands out as the ideal cluster for considering a new fast food restaurant, as it contains several neighborhoods, few fast food restaurants, and a wide variety of food options. This last point means that there appears to be plenty of public interest in new and different food options. This contrasts with clusters 0 (red) which is dominantly American Restaurants; this can indicate a population that is less likely to try foods that are different than what they are used to.

Conclusion

In conclusion, Cluster 2 is the recommended area to further investigate for the ideal location of a new healthy fast food restaurant (figure 3). This analysis provided a high-level overview of the restaurant distribution around downtown Houston, highlighting areas already saturated with fast food options, areas where a new healthy restaurant may under perform, and an area where this sort of business may prosper. Further analysis would be need to find the ideal location within the Cluster 2

area, but we were able to successfully remove unwanted areas which will help to reduce the time and cost of future analysis.

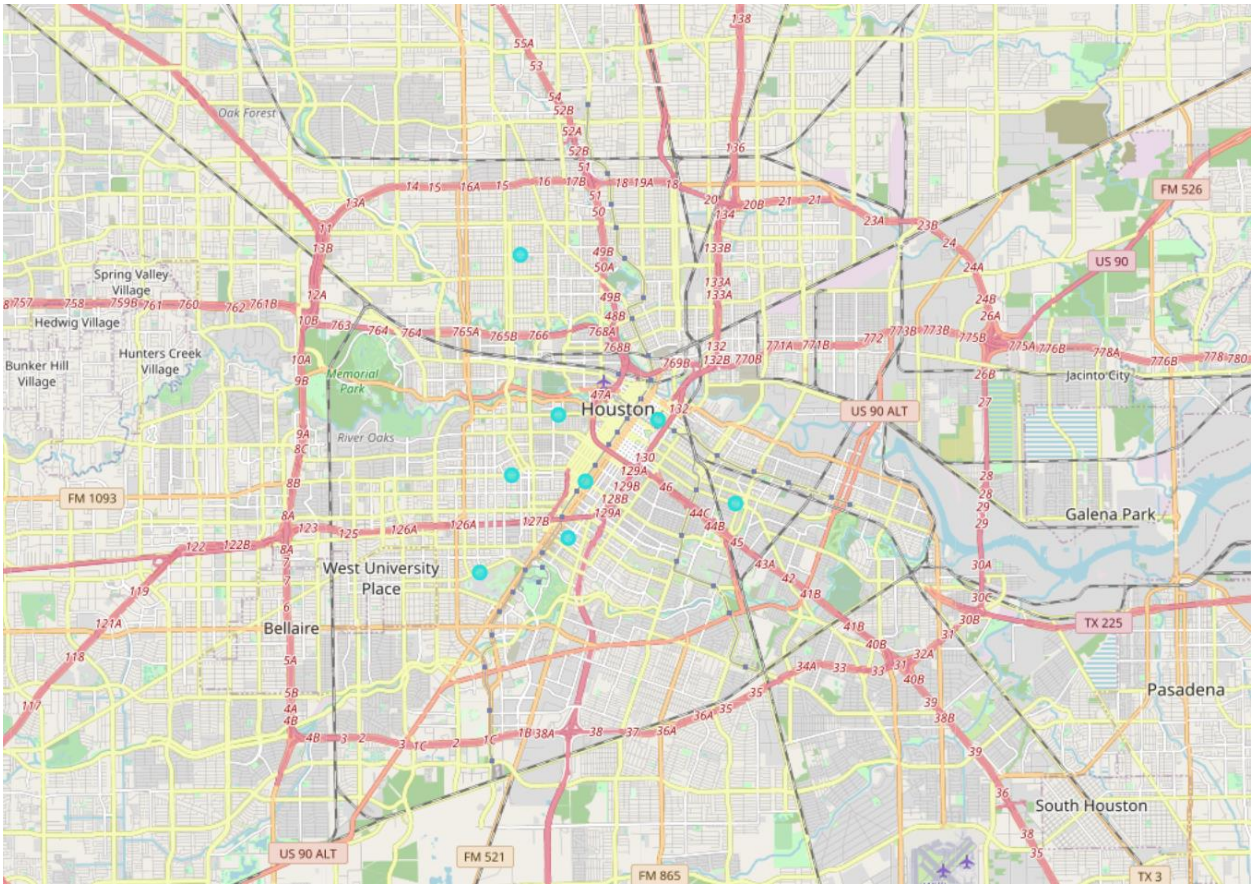

Figure 3. Recommended cluster of neighborhoods for a new healthy fast food restaurant.