AN ANALYSIS OF FOOD DESERTS AND ECONOMIC OPPORTUNITIES FOR A RETAIL OUTLET (GROCERY STORE) IN HOUSTON,TEXAS.

Introduction

1.1 Background

Houston, inland port city, in Harris, Fort Bend, and Montgomery counties, that is the seat (1836) of Harris county, south-eastern Texas, U.S. It is linked by the Houston Ship Channel to the Gulf of Mexico and the Intracoastal Waterway at Galveston, 50 miles (80 km) southeast. Houston is the state's most populous city and the fourth largest city in the United States. Its skyscrapers rise from the unrelievedly flat Gulf Coastal Plain, which at Houston lies at an elevation of about 55 feet (17 metres) above sea level and is dissected by a series of bayous. The region's climate is warm and humid, and the city is noted for its hot, sticky summers.

In addition to Galveston, other major cities in the Houston metropolitan area include Baytown, League City, Missouri City, Pasadena, Sugar Land, and Texas City. Inc. 1837. Area 601 square miles (1,559 square km). Pop. (2000) 1,953,631; Houston–Sugar Land–Baytown Metro Area, 4,715,407; (2010) 2,099,451; Houston–Sugar Land–Baytown Metro Area, 5,946,800.

Houston's economy since the late 19th century has a broad industrial base in energy, manufacturing, aeronautics, and transportation. Leading in healthcare sectors and building oilfield equipment, Houston has the second most Fortune 500 headquarters of any U.S. municipality within its city limits (after New York City). The Port of Houston ranks first in the United States in international waterborne tonnage handled and second in total cargo tonnage handled. Nicknamed the "Bayou City" "Space City", "H-Town", and "the 713", Houston has become a global city, with strengths in culture, medicine, and research. The city has a population from various ethnic and religious backgrounds and a large and growing international community. Houston is the most diverse metropolitan area in Texas and has been described as the most racially and ethnically diverse major metropolis in the U.S. It is home to many cultural institutions and exhibits, which attract more than 7 million visitors a year to the Museum District. Houston has an active visual and performing arts scene in the Theatre District and offers year-round resident companies in all major performing arts.

Houston has been divided into 88 Super Neighbourhoods where residents of neighbouring communities are encouraged to work together to identify, plan, and set priorities to address the needs and concerns of their community.

The Super Neighbourhood Council serves as a forum where residents and stakeholders can discuss issues, establish priority projects for the area and develop a Super Neighbourhood Action Plan to help them meet their goals. In some cases, more than one of the Super Neighbourhoods have joined together to create a stronger, more active Council.

The Super Neighbourhood initiative is building and improving on past successes and relies upon stakeholder participation and outreach. The people living in the neighbourhoods themselves are best-equipped to know what their needs are, and are the most invested in seeing that those needs are met. They bring these concerns to the City and work together to have them addressed and resolved. Through this initiative, City government has established strong relationships within Houston communities.

A super neighbourhood is a 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.

The boundaries of each super neighbourhood rely on major physical features (bayous, freeways, etc.) to group together contiguous communities that share common physical characteristics, identity or infrastructure.

The super neighbourhood elects a council comprised of area residents and stakeholders that serves as a forum to discuss issues and identify and implement priority projects for the area.

Food deserts according to the CDC are “areas that lack access to affordable fruits, vegetables, whole grains, low-fat milk, and other foods that make up the full range of a healthy diet.” Individuals living in areas classified as food deserts are shown to have an increased risk of developing diabetes, cardiovascular disease, and being obese.

While food deserts can be present in areas typically associated as having low income, they can be present in any area that causes an individual to not have easy access to an available grocery store, and overall, about 23.5 million Americans and a reported 250,000 individuals living in Houston, TX lack easy access to grocery stores according to Rice University.

While living in food deserts are shown to negatively impact one’s health, these areas would present lower competition for new grocery stores and allow for a perfect opportunity for business owners to open new grocery stores that not only

help alleviate public health concerns, but to also serve the demand for fresh food for individuals that live in impacted areas leading to a greater possibility for a successful store.

1.2 Problem

A Retail Store Chain wants to open up an outlet in Houston and wants to know which locations are food deserts in order to select the best location for their new retail outlets that helps to solve the problems of food deserts and also a location that has the best probability of contributing immensely to the success of the business.

1.3 Interest

This project can be used by the following individuals/groups depending on the objectives they want to achieve:

1. Business Organisations
2. Entrepreneurs
3. Government Officials
4. Private Individuals
5. Non-Governmental Organisations

2 Data Acquisition

2.1 Data Sources

Data will be sourced from;

I Data on the 88 Super Neighbourhoods will be sourced from: (http://www.houstontx.gov/planning/Demographics/demograph_docs/income_avgs.htm) which contains the list of neighbourhoods within Houston as well as their location relative to downtown, Texas totalling to 88 super neighbourhoods. BeautifulSoup library for Python will be used to extract the data.

II Geocoding

The File contents will be retrieved and fit into a Pandas data frame. The Longitude and Latitude of the neighbourhoods will be retrieved using a geocoder (geolocator) and the geometric locations received will be inserted into the initial data frame.

III Foursquare API

From the location data obtained after Web Scraping and Geocoding, the venue data is found out by passing in the required parameters to the Foursquare API, and creating another Data Frame to contain all the venue details along with the respective neighbourhoods.

3 Methodology

The Wikipedia webpage along with the information contained in the table will be scraped using BeautifulSoup and geographical coordinates such as latitude and longitude data will be found for each neighbourhood using a geocoder. Further data for each neighbourhood will be collected using the Foursquare API.

The Foursquare API will allow for the gathering of venue data for each neighbourhood such as the frequency and occurrence of groceries stores each neighbourhood. Once the frequency of grocery stores are found in each Houston super neighbourhood using foursquare, the data will be combined with the location data of the neighbourhoods. After combining the two data sets it will be clustered using K-means and mapped using the merged data set.

Lastly each cluster will be examined to determine the best location to place a new grocery store.

Further data analysis on the clusters will be done using bar graphs to determine the best clusters to open a new grocery store.

4 Results

Houston averages 0.024497 grocery stores overall within its 88 super neighbourhoods.

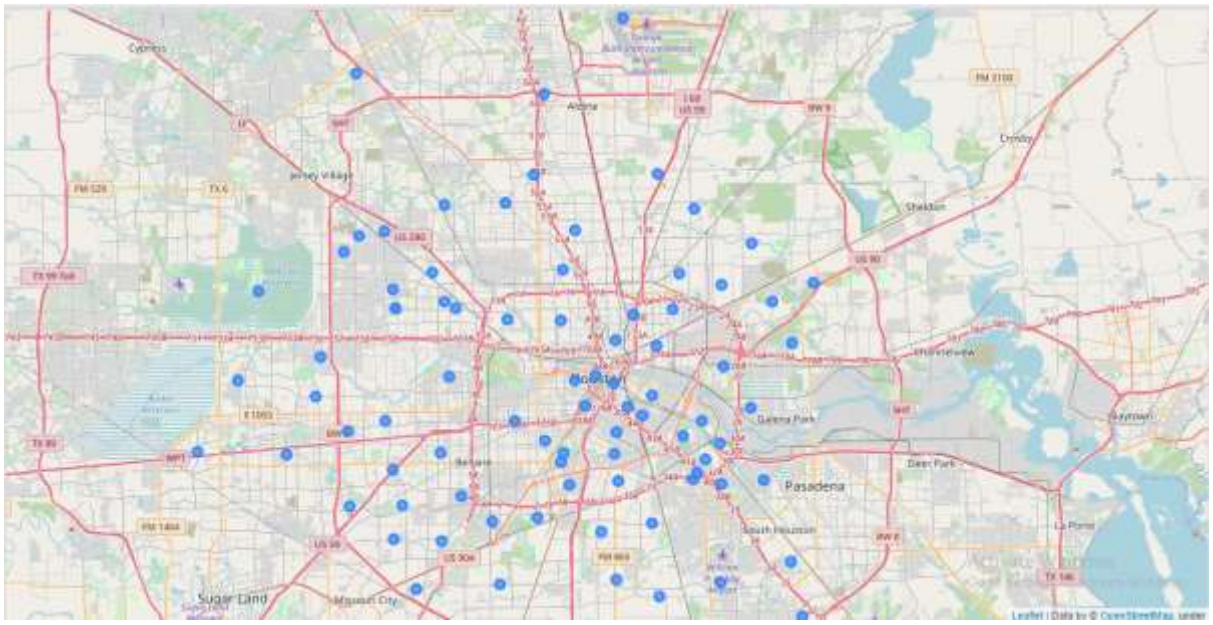
The super neighbourhoods were split into a total of 4 clusters

• Cluster 0 (red): Neighbourhoods with moderate frequency of grocery stores

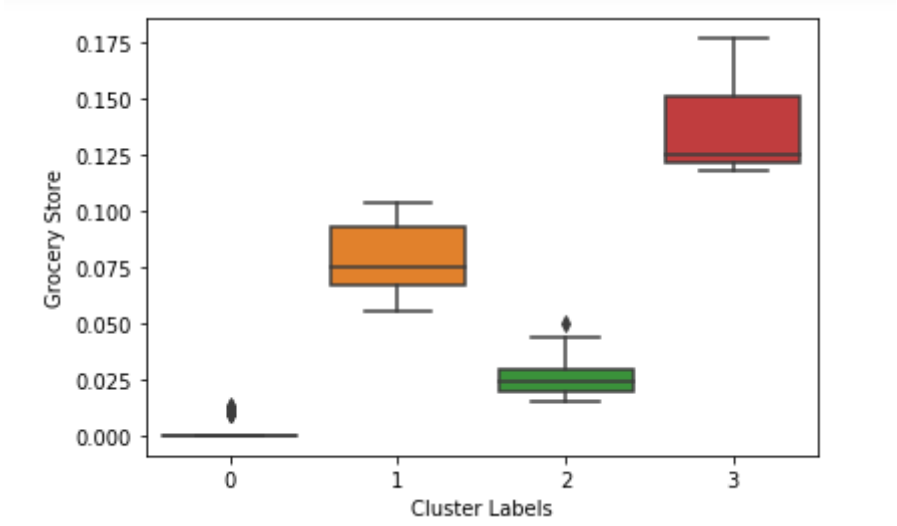
• Cluster 1 (blue): Neighbourhoods with high frequency of grocery stores

• Cluster 2 (light blue): Neighbourhoods with low frequency of grocery stores

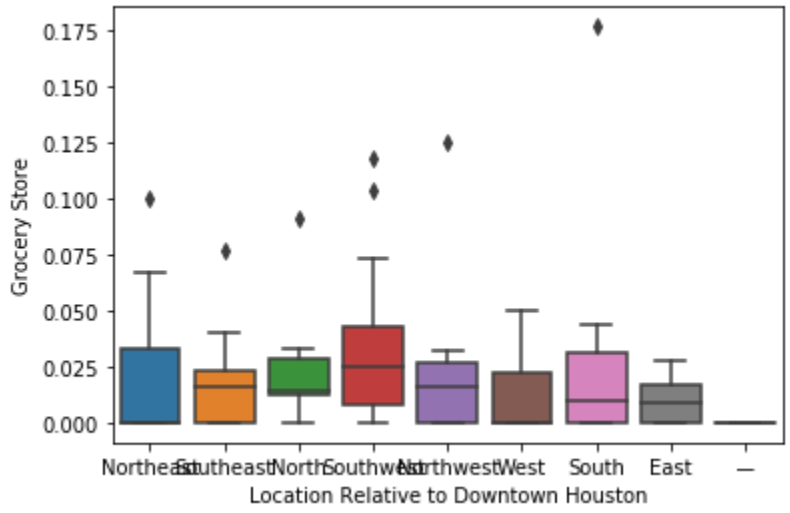
• Cluster 3 (green): Neighbourhoods with very high frequency of grocery stores



A Clustered Map of Grocery Store Frequency



Box Plot for Grocery Stores in Each Clustered Neighbourhood



Box Plot of Location Frequency for Grocery Stores

5 Discussion

Discussion

The analysis of the results shows that Houston overall has a low frequency of grocery stores within its super neighbourhoods, having a mean frequency of grocery stores being 0.024497.

In terms of the clustered data, cluster 2 has the lowest overall frequency of grocery stores by comparison to other groups.

Cluster 3 is shown to have the highest frequency of grocery stores by a large margin when compared to other clusters, making it a possible outlier cluster group for grocery stores in super neighbourhood within Houston, Texas.

In further looking at the frequency of grocery stores based on location relative to downtown Houston, areas located to the Northeast, and Southwest have the largest frequency of grocery stores. The lowest frequency of grocery stores relative to downtown Houston is central Houston showing 0 grocery stores in the area.

Overall the best location to set up a new grocery store would be within cluster 2, as it has the lowest frequency of grocery stores and therefore the lowest amount of competitors present. **The overall best location in cluster 2 would be areas in either central Houston or areas within the loop**. *The worst location would be the Northeast or Southwest as they would be closer to competitors or any areas within cluster 1 or 3.*

6 Conclusion

The goal of this project was to analyse the 88 super neighbourhoods that are located within Houston Texas in order to see possible food desert locations as well as determine the best location to open a new grocery store that would have the least competition leading to a possible more successful store. Neighbourhood data was collected using a web scrape of Houston neighbourhoods located on Wikipedia web page.

This data was mapped to check for accuracy. The frequency of grocery stores was then found for each of the 88 super neighbourhoods using the Foursquare API. The data sets were then merged and clustered in order to better analyse each super neighbourhood. Overall the conclusion of the project is that Houston has a low frequency of grocery stores and a large amount of food deserts, and **the best place to open a new grocery store would be in central Houston cluster 2.**

This project was made using the Foursquare API Sandbox Tier Developer Account, this limits the number of API calls and therefor limits the depth of our search and the results returned. Before a final decision should be concerning the projects results, further consideration should be done that looks at areas other than frequency or competition data before setting up a new store. Other areas such as income level and population density within the defined neighbourhoods should also be considered to best determine the optimal location for a grocery store.