

Data Science Capstone

Where to put an Asian supermarket in San Antonio?

Introduction

San Antonio, despite being the 7th most populous city in the country and the 2nd most populous city in the state of Texas, has no "China town", or Asian shopping district. According to the 2019 U.S. Census, San Antonio also has a similar Asian population to Dallas (45k vs 50k) which has its own Asian food district. Even Katy, TX has a "China town" with its small Asian population of 867. San Antonio does have some smaller Asian grocery stores scattered about the city but has no mainstream Asian superstores such as H-Mart or 99-Ranch. At the end of 2019, a medium-sized Asia Supermarket opened on the East Side of town, but its debut was quickly met with the COVID pandemic. Most seekers of East Asian cuisine and ingredients have to drive hours to Austin or Houston, TX. The goal of this study is to find an ideal part of San Antonio to open/build a large Asian market in hopes to nucleate a "China town" and increase the cultural diversity of the city.

Dataset, feature selection

The demographic data used in this work comes from the 2019 U.S. Census (freely available from data.census.gov), more specifically the ZTA5 data. The Asian population, the Hawaiian and other Pacific Islander population, and the total population were selected from this dataset. The geospatial data came from <http://download.geonames.org/export/zip/>. Data for Asian restaurants was collected from Foursquare.

Table 1. Features used and not used in this study

Kept features	Dropped Features	Reason for dropping features
Asian pop, Hawaiian Pacific Islander, Total pop	All other demographics	non-AAPI are not the primary target demographic
ZIP Code, Latitude, Longitude	County, City, State	For Foursquare data acquisition, latitude and longitude are sufficient

Venue Category	Other tags, tips, etc	ZIP is not available for every venue
----------------	-----------------------	--------------------------------------

Methodology

The number of Asian American Pacific Islander (AAPI) residents in each ZIP code was calculated by adding the number of Asian residents and the number of Hawaiiin/Pacific Islander residents. The decimal fraction of AAPI residents was obtained by dividing the AAPI population by the total population of the ZIP code.

Venue data was obtained from the Foursquare API. A cursory venue search was accomplished through the ‘explore’ feature. A more specific query of the Asian restaurants in each ZIP code was performed with the ‘search’ feature and the categoryId for ‘Asian restaurant’. The venue categories were tabulated for each ZIP code. Each ZIP code was clustered according to the frequency of the types of Asian restaurants via K-means in order to understand any similarities between ZIP codes.

Folium was used to create a map of the clustered ZIP codes. Another bubble map was also made using the number of Asian venues within 1200 meters of each ZIP code’s geo-coordinates and the AAPI population of each ZIP code.

Results

One limitation of census data is that the census is voluntary and not checked for accuracy. This means the AAPI population of a ZIP code is more accurately the “number of residents who claimed to be AAPI.” Another limitation of census data is the way of recording residents with more than one race. Labels such as “two or more races”, “two races including some other race and three or more races”, and “two races excluding some other race and three or more races” do not specify any of the races that constitute them. This study assumes any half-Asian (or Pacific Islander) population is insignificant, and that AAPI people will be the main customers of such a supermarket. According to the 2019 census, the ZIP code with the most self-identifying mono racial Asians (AAPI) is 78240. A table of the top 5 ZIP codes is shown below.

Table 2. Top 10 San Antonio ZIP codes based on AAPI residents

ZIP Code	AAPI residents
78240	1627.0
78258	856.0

78250	836.0
78254	796.0
78251	766.0
78249	668.0
78245	650.0
78230	631.0
78247	585.0
78221	522.0

Table 3. Top 5 San Antonio ZIP codes based on AAPI fraction of population

ZIP Code	Fraction AAPI
78243	0.200980
78256	0.033814
78266	0.032403
78257	0.029451
78240	0.026704

Table 4. Number of Asian restaurants within 1200 m of ZIP code latitude and longitude

ZIP Code	# Asian restaurants within 1200 m
78229	18
78205	18
78212	13
78218	11
78251	9
78240	9
78227	9
78215	8
78217	7

For the Foursquare data, one limitation is the radius of the call-request. Some ZIP codes - such as the areas downtown - are significantly smaller than those on the perimeter of the city. A radius too large can result in venues in the center of San Antonio being counted more than once and venues in larger ZIP codes to be left out. This work used a radius of 1200 meters, which had some (but minimal) double counting downtown and some undercounting outside of downtown. The ZIP code with the most Asian restaurants was 78229 (the Medical Center to the Northwest) followed by two downtown ZIP codes (78212 and 78205). Interestingly, the ZIP code clustering had Northwest San Antonio and downtown in the same cluster. 78240 is lower in terms of the number of Asian restaurants (sixth) but is relatively close to the Medical Center. According to the maps generated in this study, many Asian restaurants are in downtown San Antonio or in the Northwest side of town near the Medical Center. Though the central ZIP codes have a higher concentration of Asian venues, the Northwest ZIP codes have more AAPI residents.

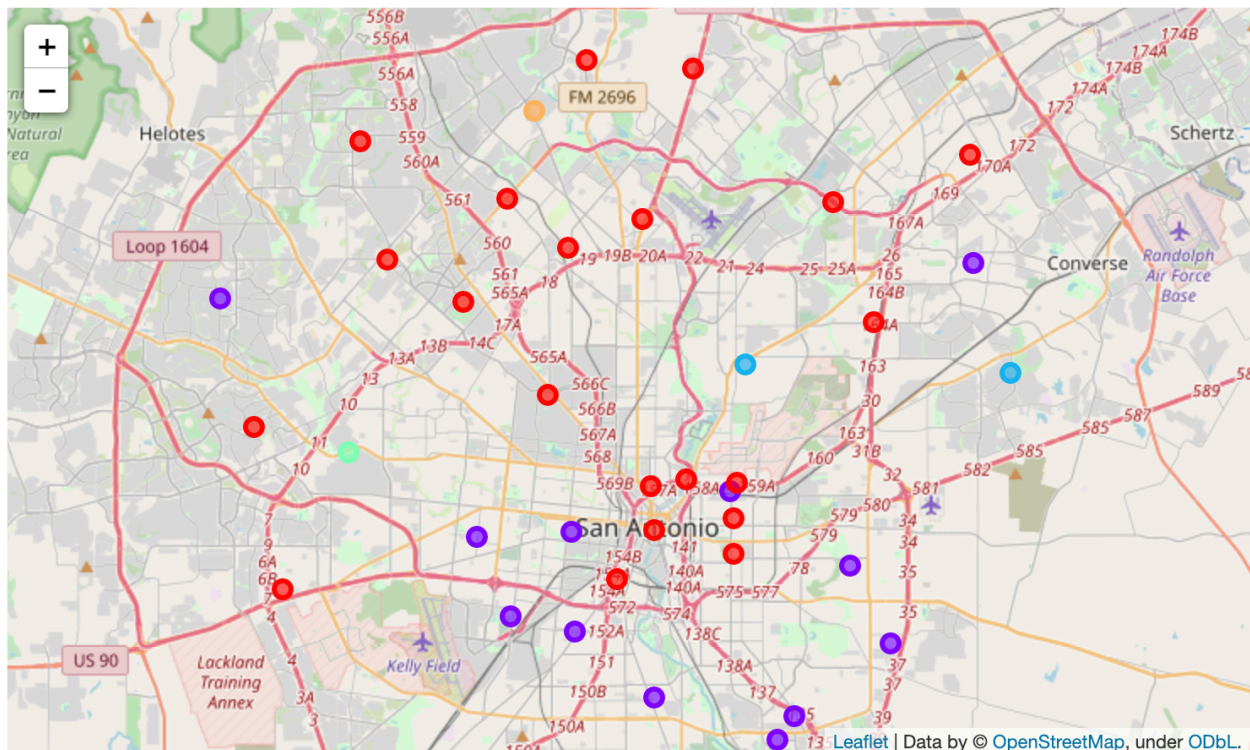


Figure 1. Map of K-means clusters of San Antonio ZIP codes based on types of Asian restaurants

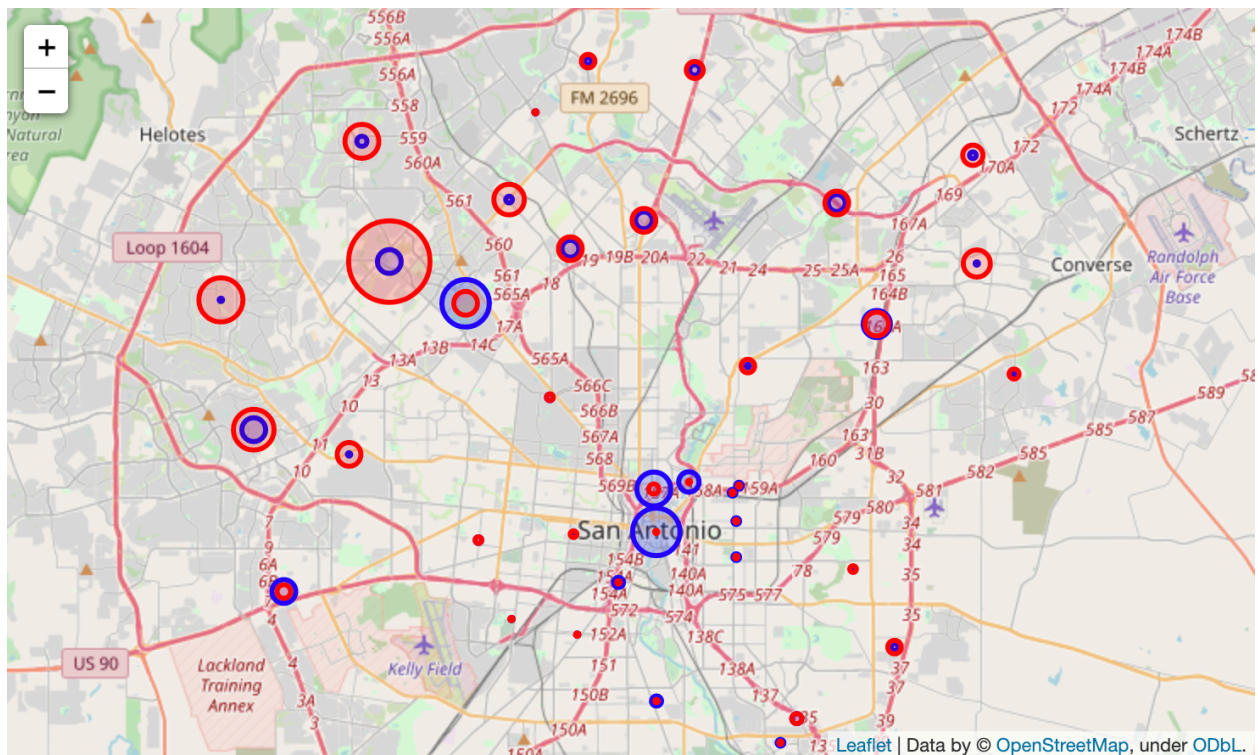


Figure 2. Map of San Antonio ZIP codes. The size of the blue bubbles represents the number of Asian restaurants, and the size of the red bubbles represents the number of AAPI residents.

Discussion and Conclusions

Downtown and Northwest San Antonio appear to be excellent areas for a H-Mart or 99-Ranch. Clustering revealed the two areas' ZIP codes to be similar in terms of the types of Asian restaurants contained within them. Downtown has a higher concentration of Asian restaurants than Northwest San Antonio, but the Northwest has more AAPI residents, the likely customers of such a store. With the Medical Center (78229) having a high number of Asian restaurants and the adjacent 78240 having the highest number of AAPI residents (and the sixth highest density of AAPI residents), Northwest San Antonio is the best part of the city to build a large Asian supermarket. Further research on available real estate and distribution channels should be looked into, but this initial study gives a good starting point.