

# The Battle of Neighborhoods: Exploration of Austin Food Venues

May 28, 2019

## 1 Introduction

Located in Texas, Austin is one of the most rapidly growing cities in the United States. According to Census Bureau estimates released in mid-April 2019 [1], the Austin-Round Rock metro's population reached 2,168,000 as of July 1, 2018, which resulted in 2.5% jump from the previous year. The city of Austin also showed a 26.3% increase in population since 2010 and is now ranked as 11-th largest in the country. According to the newest forecasts, [2], this trend will only continue and Austin, with the population of 981,035 on January 1, 2019, will reach 1 million by 2020.

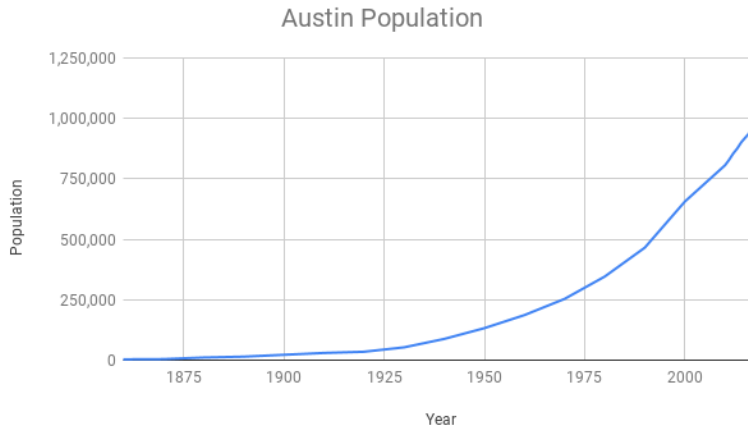


Figure 1: Dynamics of Austin city population

Despite fast-paced growth, the city does not suffer from increasing crime and poverty rates. Indeed, U.S. News & World Report, [3], recently ranked Austin as the best place to live in the United States with an overall score of 7.8 out of 10. Steady job growth, stable housing market, and the desirable lifestyle of the city, not to mention *the booming tech industry*, have been the reasons to move for many people in places like California, metro Chicago area and New York.

Millennials and retiring baby boomers are now making up the majority of new Austinites, all being drawn in for the city's shopping, museums, entertainment, and, last but not the least, restaurants. Austin is of course synonymous with barbecue, as well as killer bars, and all manner of food trucks. And while several years ago the idea of traveling here solely for great restaurants would have been strange, Austin is now becoming one of the top food destination in the US, [4]. This, together with city's demographic trends suggests a great opportunity for a new business: *open a restaurant, which would feature a more elegant selection of food and drinks, comparing to BBQ, tacos and beer, that at the same time would be interesting and attractive for both locals and tourists.*

**The purpose of this project is to conduct the analysis of Austin food venues and suggest the best unusual category for a new restaurant, together with the optimal location, based**

on the additional analysis of the city's neighborhoods.

## 2 Data

## 2.1 Data sources

In order to answer the proposed question, two types of information were required:

1. Data, describing the current food venues available in the city. Such information is used in the analysis to discover the most and the least represented, unusual, categories. Furthermore, more detailed information about the unusual restaurants, such as rating, is used to determine the optimal category for a new restaurant. All the required data is obtained through the Foursquare API, [5].
2. Data, summarizing the geographical representation of the city itself, including the list of its neighborhoods, together with the their coordinates, as well as some extra information, such as each neighborhood's area and population. City-Data website, [6], provides general information and is used to collect data about Austin's neighborhoods. Further, areaConnect website, [7], is used for translating neighborhoods' zipcodes into the corresponding geographical coordinates.

## 2.2 Neighborhoods data wrangling

The neighborhoods data was scraped using requests and BeautifulSoup libraries. The starting point was the City-Data website main page, that looks as following:

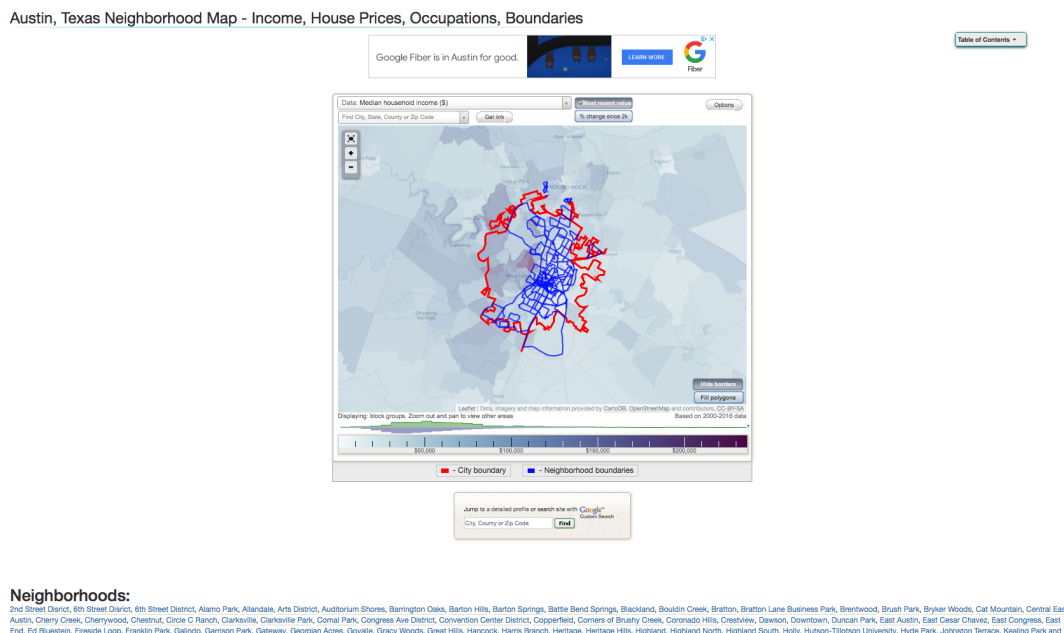


Figure 2: City-Data webpage

While parsing this webpage, a part of the code detected and followed the links to a separate webpage for each neighborhood (see the list on the bottom of the webpage, as shown of Figure 2.) to obtain the information, such as neighborhood names, population, area and zipcodes. One can notice that some of these links contain more detailed information, including, for instance, residents' age and income. While such information can be useful for clustering analysis of the neighborhoods of Austin, for quite some of the neighborhoods this information was missing and it was decided to not use these features at all.

The obtained data was further transformed: a) features' values were casted into appropriate data types; b) neighborhoods sharing same zipcodes, were combined into one big neighborhood with population and area equal to the sum of all smaller ones.

The next step was to obtain geographical coordinates for each neighborhood. This was done straightforwardly using areaConnect website. For those neighborhoods, that correspond to several zipcodes, the resulting coordinates were computed as the average values. The resulting partitioning of the city into neighborhoods is illustrated on Figure 3.

## 2.3 Venues data wrangling

The neighborhoods data was scraped using requests and BeautifulSoup libraries. The starting point was the City-Data website main page, that looks as following:

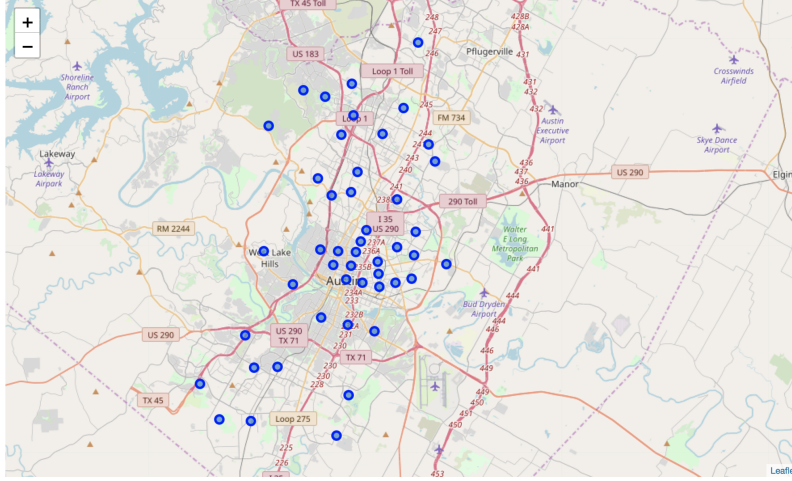


Figure 3: Centers of Austin's neighborhoods

The venues' information was scraped using calls to Foursquare API. For each of the neighborhoods, data about top 100 venues was downloaded, using the coordinates of the centers of the neighborhoods and the radius, computed under the assumption that every neighborhood can be approximated as a circle. Such an assumption may not be very realistic, but, on the other hand, using a fixed value for the radius may lead to even more inaccuracy, since some of the neighborhoods are quite different in size. In order to account for the errors due the assumption, the radius was further adjusted by 20%.

For each venue name, id, category and the corresponding geographical coordinates were stored. This data contained the information about all popular places, not limited to restaurants/food places, so the later were further picked out of all available data. Later in the analysis, after the candidate unusual categories of the restaurants were detected, customers' likes and venues' ratings were further downloaded using venue id for a small group of restaurants.

## 3 Methodology

### 3.1 Determine optimal venue category

The first step of the analysis was to determine an attractive unusual restaurant category. For this purpose, for every neighborhood frequency of each type of cuisines/food places was computed, and, based on this, the top 20 and bottom 20 most frequent categories were identified. Figure 4 summarizes the resulting top and bottom venue types.

Next, the venues from the least frequent categories were analyzed in more details. For each venue, the rating and customers' likes were added. The decision was made based on the top rated places with

sufficient amount of likes, i.e., rating at least 8.0 and more than 50 likes. One particular category stood out in the resulting group of venues, **Tapas Restaurant**: one out of two such venues had the highest ratings, while the other one had the second to largest amount of likes, as shown of Figure 5. Moreover, this category looked attractive as the goal was to determine a type of restaurants with elegant food and drinks in contrast to the everyone’s favorite food trucks and taco places.

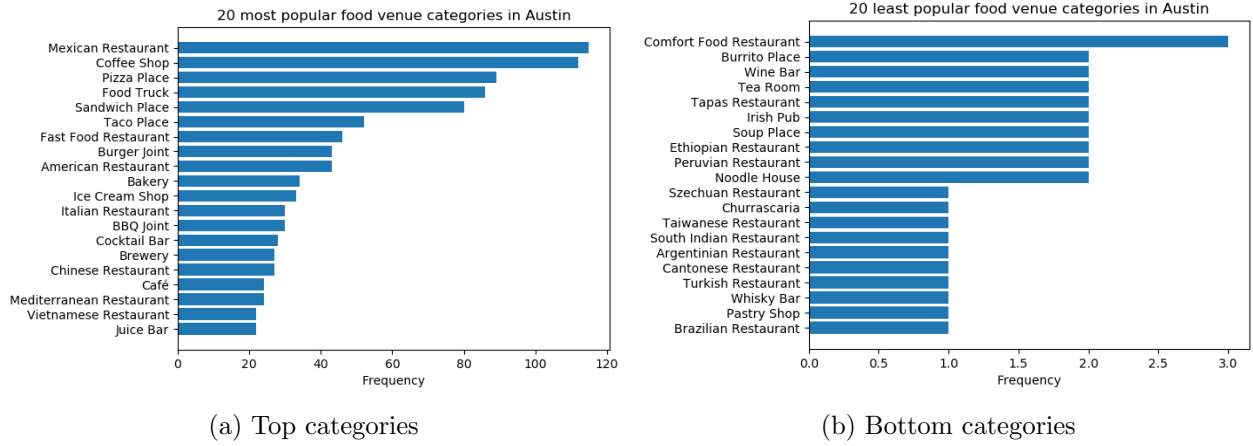


Figure 4: The most and the least frequent venue types

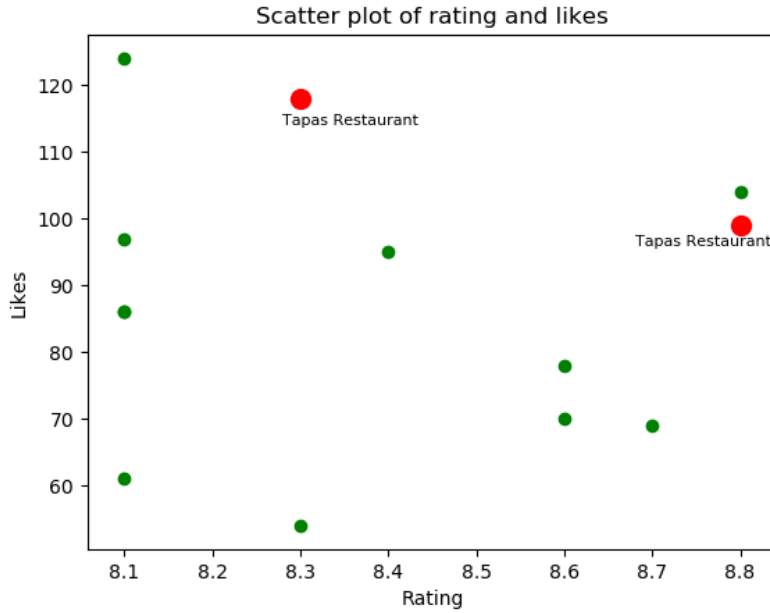


Figure 5: Ratings and likes for selected venues

### 3.2 Choosing the best location

After the best venue category was determined, it was left to suggest the optimal place to open a new restaurant. As it followed from the data, both Tapas Restaurants were located in the same neighborhood (see Figure 6). Clearly, opening another restaurant of the same, quite rare, category in the same neighborhoods would not be the smartest strategy. Therefore, the next step of the analysis was dedicated to studying the city neighborhoods.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue ID	Venue Latitude	Venue Longitude	Venue Category	Rating	Likes
678	Auditorium Shores, Barton Springs, Bouldin Cre...	30.2440	-97.7630	Winebelly	52169f0711d26dd8be94f30a	30.241435	-97.758936	Tapas Restaurant	8.8	99
575	Bryker Woods, Old Enfield, Windsor Road	30.2930	-97.7490	Cabo Bob's	5261db9511d2b9ed254233d6	30.295861	-97.744015	Burrito Place	8.8	104
687	Auditorium Shores, Barton Springs, Bouldin Cre...	30.2440	-97.7630	Soup Peddler Real Food & Juice Bar	4ff0f46a011c9ad2fa340562	30.246298	-97.756045	Soup Place	8.7	69
2830	North Burnet	30.3925	-97.7355	Cru Wine Bar	49e71bedf964a52079641fe3	30.401997	-97.726415	Wine Bar	8.6	70
2445	Georgian Acres, Heritage Hills, North Lamar, W...	30.3710	-97.6720	New Fortune Chinese Seafood Restaurant	5468ecfe498e87b191079be1	30.377848	-97.688096	Cantonese Restaurant	8.6	78
2238	Hyde Park, North Loop, Triangle State	30.3080	-97.7250	Vino Vino	49bed457f964a520e5541fe3	30.306973	-97.735281	Wine Bar	8.4	95
1916	Southeast Austin (SE Austin)	30.1575	-97.7500	Teapioca Lounge	4f348c3f7bebf16206886b93	30.160661	-97.791289	Tea Room	8.3	54
719	Auditorium Shores, Barton Springs, Bouldin Cre...	30.2440	-97.7630	Barlata	519523f1498ec02c62bb194a	30.252468	-97.764229	Tapas Restaurant	8.3	118
1269	Upper Boggy Creek	30.2960	-97.6990	Xian Sushi & Noodle	525d8fa5498e73e322f6ebe0	30.302841	-97.699082	Noodle House	8.1	86
1310	Oriens Park, Pecan Springs-springdale, Rmma, U...	30.3070	-97.6830	Xian Sushi & Noodle	525d8fa5498e73e322f6ebe0	30.302841	-97.699082	Noodle House	8.1	86

Figure 6: Top 10 venues from rare categories, sorted by rating

To measure the similarity between neighborhoods, the clustering analysis, using k-means algorithm, was performed. The features included the frequencies of all possible popular venues (not limited to restaurants, as in the previous subsection), as well as population and area of each neighborhood. Below are the results of clustering of 48 Austin's neighborhoods using k-means with  $k = 7$ . The neighborhood that already has two Tapas Restaurants is marked by a red circle.

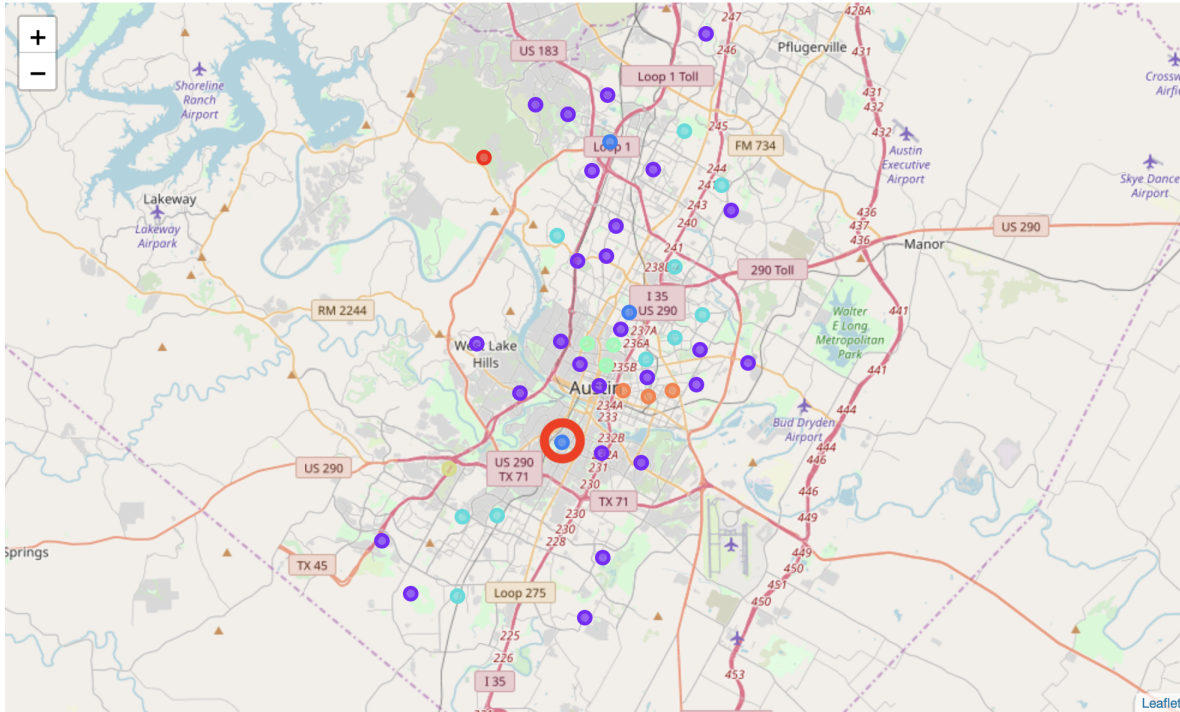


Figure 7: Clusters of Austin's neighborhoods

The cluster of interest contains 3 neighborhoods, including the one with existing Tapas Restaurants. Hence, based on this analysis, the other two neighborhoods are optimal for opening a new restaurant of

this category. These neighborhoods (consisting of smaller ones) include:

- – Hyde Park
- North Loop
- Triangle State
- North Burnet

## 4 Results and Discussion

Based on the performed analysis, the best category turned out to be a **Tapas Restaurant** and the optimal location - one of the two neighborhoods:

- – Hyde Park
- North Loop
- Triangle State
- North Burnet

The choice of restaurant type seems reasonable. First of all, it is suitable both as an upscale food destination for the tourists, as well as the locals, who would like to try something unusual. Moreover, since the Spanish cuisine is up to some extent close to the Mexican cuisine, such a restaurant would still possess some of the flavour of the local famous and favorite food. Therefore, a Tapas Restaurant fits nicely, provided one chooses the location optimally.

According to the available data, there are currently two Tapas Restaurants in Austin, both located in the same neighborhood. In order not to get unnoticed when opening a new restaurant and to avoid competition with the existing places, the new restaurant should be opened in another neighborhood, which nevertheless possesses similarities with the one that has the restaurants. As the results of the clustering analysis, it was determined that the neighborhoods with existing restaurants form a cluster with two other ones, which are, therefore, a good choice for the new restaurant location. Interestingly, there are *Wine Bars* in both of these neighborhoods, which can be viewed as a good sign, since this category is quite close to the Tapas Restaurants and, hence, the new restaurant should be of interest for the residents.

Although, the conclusion based on the performed analysis seems persuasive enough, there are several inaccuracies, that may introduce errors and lead to misleading results:

- The neighborhoods were represented geometrically as circles, based on the area, and the radius of each circle (with some adjustments) was used to the Foursquare calls. Clearly, such assumption doesn't sound very realistic, as the neighborhoods can be irregularly shaped. However, a more accurate procedure seems to be unavailable, since Foursquare calls use geographical coordinates of the center of a neighborhood and a radius.
- The neighborhoods sharing same zipcode were combined into bigger ones and the corresponding coordinates were computed as the average values. Although, this might be inaccurate and introduce some errors into the analysis, a more accurate geographical segmentation of the city is out of the scope of this project.

Finally, one possible improvement can be based on gathering more data the neighborhoods and their population statistics, including average resident's age, income and so on. Such detailed information might improve the segmentation accuracy.

## 5 Conclusions

In this project the analysis of Austin's neighborhoods and available restaurants was performed. According to the data, the most popular venue types include Mexican Restaurants, Coffee Shops, Pizza Restaurants and other, while the places such as Tapas Restaurants, Wine Bars, and exotic cuisines, such as Ethiopian and Peruvian, are not very common. Based on the current demographic trends and the information about existing food venues, it was concluded that a Tapas Restaurant would be a perfect food venue category for a new restaurant. It was further shown through the segmentation of Austin's neighborhoods that there are certain locations that are the most suitable for such new restaurant. The project was summed up by discussing the results, pointing out existing weaknesses, as well as possible improvements.

## References

- [1] <https://www.census.gov/newsroom/press-releases/2019/estimates-county-metro.html>
- [2] <http://austin.culturemap.com/news/city-life/01-09-19-population-of-austin-prepared-to-pop-to-1-million-in-2020/>
- [3] <https://realestate.usnews.com/places/texas/austin>
- [4] <https://www.zagat.com/b/30-most-exciting-food-cities-in-america-2017>
- [5] <https://developer.foursquare.com/>
- [6] <http://www.city-data.com/nbmaps/neigh-Austin-Texas.html#N22>
- [7] <http://austintx.areaconnect.com/zip2.htm?city=Austin&q=TX&searchtype=bycity>