Optimal Location for a Vietnamese Restaurant in Austin, Texas

Cathy Le

August 14, 2019

1. Introduction

1.1 Background

Beloved for its quirky Texan charm, booming employment, scenic nature, and live music, Austin is met with an increasing number of visitors as well as permanent residents every year. On top of being a rapidly growing and upcoming tech city, Austin has an avid food culture, many of which start small as food trucks or modest venues.

New businesses are met with great potential and opportunity in Austin due to the city's support of local startups. The success of the business itself will depend largely on its location as there needs to be a balance of competition and demand from the surrounding area.

Austin is seen as a young and trendy city, fueled by the creative energy of the young generation. Representative of this is the prestigious University of Texas which prides itself on being the best college campus in the state for student life and a place where impactful change is made. Businesses would be wise to cater to the thousands of students that attend this thriving campus.

1.2 Problem

The business problem that this report will be catering to is a group of stakeholders interested in opening a Vietnamese restaurant in the heart of Austin, Texas. Specifically it is aimed towards the downtown region where the students of The University of Texas reside. We will find the optimal location for a business and to do so we must consider areas that are proven to be successful for similar venues.

There are other factors to consider concerning this problem such as the desire for low competition while retaining high demand. Competition comes from restaurants which also serve Vietnamese or similar cuisines. Demand can be gauged by the density of restaurants in the area, specifically ones that serve opposing cuisines, as we will want a location with high foot traffic to draw people into our business. Competition with other Vietnamese restaurants can be helpful to some extent in bringing in customers to the area, so for this project we will use the presence of Vietnamese restaurants as an indicator for a promising location.

1.3 Interest

Those who would be interested in this problem are stakeholders who would like to bring Vietnamese culture in the form of food to Austin, Texas. Austin puts a large emphasis on diversity and culture, so a Vietnamese restaurant that is both new and authentic can be successful in this city. Austin has a small but passionate Asian as well as Vietnamese community that would be quick to either support an upcoming Vietnamese business or start one themselves.

2. Data

2.1 Data Source

We will solely be using Foursquare API and geographical location data to solve this problem.

2.2 Data Usage

Our problem requires us to find high ranking Vietnamese restaurants in Austin and visualize them on a map. Afterwards, making clusters of these restaurants will help to identify areas where success is more likely.

With Foursquare API we will gather information on venues and local businesses. We will be collecting the number of restaurants, types, and locations in the corresponding Austin neighborhood. Foursquare will also be used to categorize the businesses based on what food they serve, differentiating Vietnamese restaurants from other cuisines.

Using geographic location to cluster the restaurants, we can explore and observe which areas are the most dense in demand for food venues. We will be using the "explore" request in Foursquare and passing specific parameters for Vietnamese or similar cuisines.

3. Methodology

3.1 Exploratory Data Analysis

Our solution will start by us finding the Vietnamese restaurants in Austin as this will be the basis of what areas are successful.

We will first get the address of the University of Texas after connecting to Foursquare API.

```
#University of Texas address
address = 'Austin, TX 78712'

geolocator = Nominatim()
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of The University of Texas are {}, {}.'.format(latituc
```

```
/opt/conda/envs/Python36/lib/python3.6/site-packages/ipykernel/__main__.py:4: Deprec
ationWarning: Using Nominatim with the default "geopy/1.20.0" `user_agent` is strong
ly discouraged, as it violates Nominatim's ToS https://operations.osmfoundation.org/
policies/nominatim/ and may possibly cause 403 and 429 HTTP errors. Please specify a
custom `user_agent` with `Nominatim(user_agent="my-application")` or by overriding t
he default `user_agent`: `geopy.geocoders.options.default_user_agent = "my-applicati
on"`. In geopy 2.0 this will become an exception.

The geograpical coordinate of The University of Texas are 30.2711286, -97.7436995.
```

Then we will get the Vietnamese restaurants nearby the college campus by querying the search results for "Vietnamese Restaurant" and setting the radius to 5000 meters which is a reasonable distance from campus that will be easy to get transportation to.

```
search_query = 'Vietnamese Restaurant'
radius = 5000
print(search_query + ' .... OK!')

url2 = 'https://api.foursquare.com/v2/venues/explore?client_id={}&client_secret={}&ll={
url2
```

```
Vietnamese Restaurant .... OK!
```

```
]: 'https://api.foursquare.com/v2/venues/explore?client_id=5BURNXFD4JD3OUXRRQXNRC3XZCK5
HBFEH03N4KSIP0F1CP1H&client_secret=51A4KROSWJOX04TKEEXZSKQAIVNVQTEXC5T4CDLCDEQB1UPR&
ll=30.2711286,-97.7436995&v=201902012&query=Vietnamese Restaurant&radius=5000&limit=
100000'
```

The resulting json file will be transformed into a pandas dataframe.

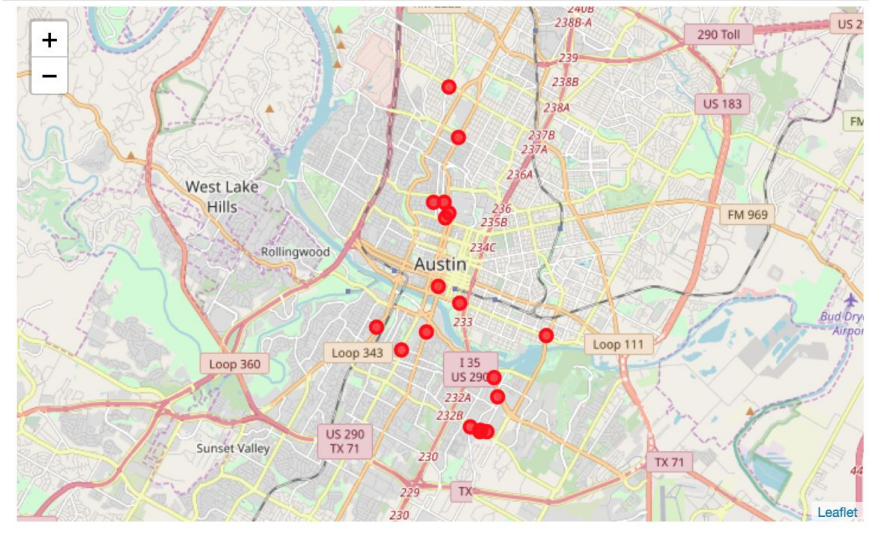
```
results2 = requests.get(url2).json()
```

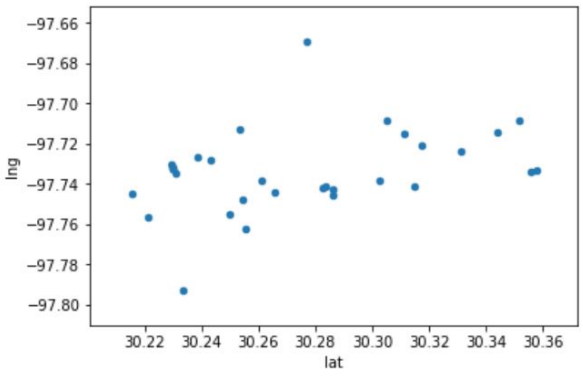
```
venues = results['response']['groups'][0]['items']

dataframe = json_normalize(venues)
```

```
df=pd.DataFrame()
df['venue_ID']=dataframe['venue.id']
df['name']=dataframe['venue.name']
df['lat']=dataframe['venue.location.lat']
df['lng']=dataframe['venue.location.lng']
df.head()
```

	venue_ID	name	lat	lng
0	4ef4c212b8f77e0f983c04e4	Elizabeth St. Café	30.249828	-97.754829
1	56a2e11e498e4ab3feb17449	Pho Please	30.243062	-97.728082
2	4a9ebac6f964a520f63a20e3	888 Vietnamese Restaurant	30.229627	-97.730117
3	4a357b73f964a520099d1fe3	Hai Ky	30.230977	-97.735033
4	5816307538fa252712aba520	Heo Eatery	30.331042	-97.723462

From this we can use a Folium map to observe where the Vietnamese restaurants are spatially. From a scatter plot we can also see the relative distances as well.





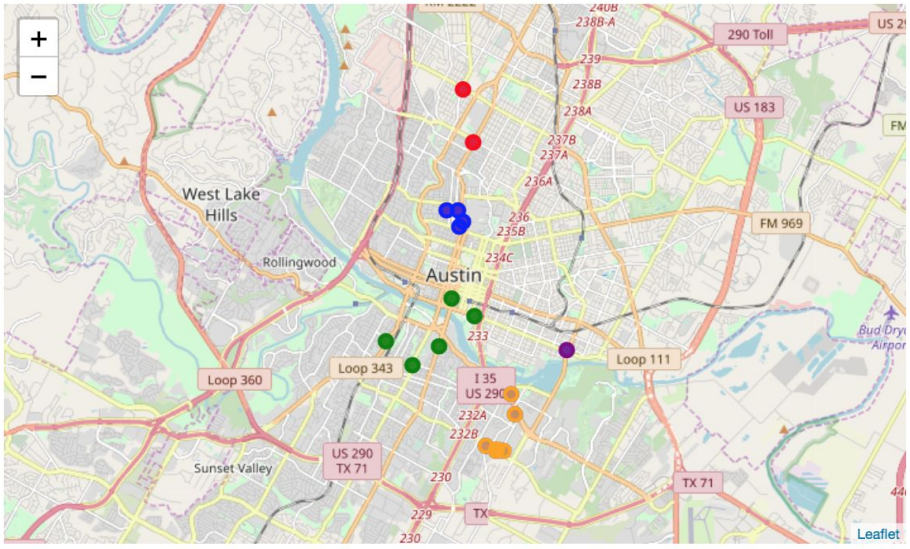
3.2 Machine Learning

K-means clustering is used to organize the locations into 5 clusters.

```
k_means = KMeans(init = "k-means++", n_clusters = 5, n_init = 12)
df_array=np.array(df[['lat','lng']])
k_means.fit(df_array)
k_labels=pd.DataFrame(k_means.labels_)
k_labels.info()
df['cluster']=k_labels
df.groupby('cluster').count()
```

cluster	venue_ID	name	lat	lng
0	9	9	9	9
1	10	10	10	10
2	5	5	5	5
3	6	6	6	6
4	1	1	1	1

We can visualize the clusters on the Folium map.



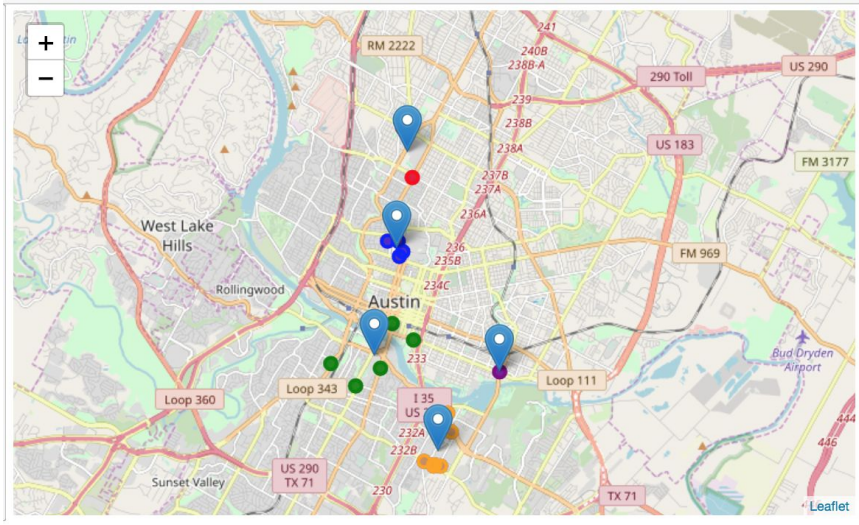
The recommended locations are derived by getting the geographic average of the coordinates of each cluster.

cluster	lat	lng
0	30.308689	-97.739885
1	30.233123	-97.730973
2	30.257286	-97.749436
3	30.284638	-97.742914
4	30.253432	-97.713051

4. Results

4.1 Discussion

Our analysis shows us that there are a small number of Vietnamese areas that are a convenient distance from the University of Texas. Of these 18 restaurants, many are scattered with no common area; however, there are some areas that are more dense with Vietnamese restaurants than others.



4.2 Observations and Recommendations

We see that the purple cluster is the closest to the University of Texas and one of the more prominent clusters. This would be a good area to start a business as it will be close to the college campus and is within an area that has seen success from other similar businesses. By finding the mean coordinates of this purple cluster, this gives us a good starting point of where we should look when choosing a location.

The second most dense cluster is the yellow cluster, of which three spots are on the same street. The algorithm does not pick up on this and thus gave us the mean coordinates of those three and two nearby restaurants. This area would also be a good place for a vietnamese restaurant. The mean location given is sandwiched between multiple spots, so it avoids close competition while still being in an area with Vietnamese presence.

The green cluster is the most spread out, but this may be due to the fact it is located around the river downtown which is a restaurant dense area. Though less dense this can actually be an advantage as we get the benefits of a downtown location while being not too close to our competitors.

The purple and red clusters seem to be outliers in that they only have one or two restaurants. It would be riskier to start a business here as there is little experience to back up the success of these locations.

5. Conclusion

The purpose of this project was to identify an area close to the University of Texas where our stakeholders could open a Vietnamese Restaurant. The optimal location was found by calculating areas of dense Vietnamese restaurants from Foursquare data. From this we identified five clusters whose mean location are potentially fruitful based on the already existing Vietnamese restaurants that preside there.

This analysis was solely based on geographic data and the use of other Vietnamese restaurants. There were many things that were assumed from the visual data that can be supported by extra analysis such as the density of the cluster. Because there were only 18 restaurants from which to draw data from, the clusters are apt to be not as representative and more difficult to cluster. We used 5 clusters in this analysis, but using more or less will change the mean location coordinates.

These are not definite locations but rather sones of interests and starting point for exploration that can be further elaborated upon by the stakeholders.

This analysis can be expanded by taking into consideration other factors such as proximity to roads, attractions, noise level, land availability, presence of other venues, prices, and ratings of the Vietnamese restaurants.