Find a Location for Bill’s Taco Town

Introduction

Bill has just moved to Sun Valley, Los Angeles CA with big dreams. He plans to open up a tacoria; the only problem is that Bill has some stiff competition with the locals. Not only that, Bill’s tacos are not that great but it's the only thing he knows how to make. In order for Bill to succeed he needs to open his restaurant in an area with the least amount of taco competition. Preferable close to his home in Sun Valley so that he is not driving all over town to sell these not so great taco.

The goal of this analysis to find the least taco dense area in Los Angeles CA

Data Source

The first item we will need is Bill’s geo location from geopy.geocoders.

```
The geograpical coordinate of Sun Valley are 34.2204227, -118.3878945.
```

The primary data source for this project is the Foursquare API. This repository includes many of the local venues and restaurants in the Los Angeles area. The foursquare API includes rating, location, and user comments.

The snapshot below is of the pandas data frame of the Foursquare data needed for the analysis.

	name	categories	address	crossStreet	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	f
0	Taco Bell	Fast Food Restaurant	8250 Laurel Canyon Blvd	at Burton St	34.221129	-118.396095	[{'label': 'display', 'lat': 34.22112869754370...	758	91605	US	North Hollywood	CA	United States	
1	Del Taco	Fast Food Restaurant	2515 N Hollywood Way	btwn Avon & Thornton	34.193823	-118.349278	[{'label': 'entrance', 'lat': 34.194329, 'lng'...	4626	91505	US	Burbank	CA	United States	
2	Taco Bell	Fast Food Restaurant	6254 Lexington Ave	at Vine St	34.092357	-118.326281	[{'label': 'display', 'lat': 34.092357, 'lng':...	15344	90038	US	Los Angeles	CA	United States	
3	Taco Bell	Fast Food Restaurant	11694 Magnolia Blvd	Colfax	34.164590	-118.387448	[{'label': 'display', 'lat': 34.16458972609614...	6215	91601	US	North Hollywood	CA	United States	
4	Taco Bell	Fast Food Restaurant	15651 Sherman Way	at Haskell Ave	34.201402	-118.474668	[{'label': 'display', 'lat': 34.20140238194243...	8264	91406	US	Van Nuys	CA	United States	
5	Taco Bell	Fast Food	14115 Ventura	at Hazeltine	[illegible]	[illegible]	[{'label': 'display', 'lat':	[illegible]	[illegible]	US	Sherman	CA	United	

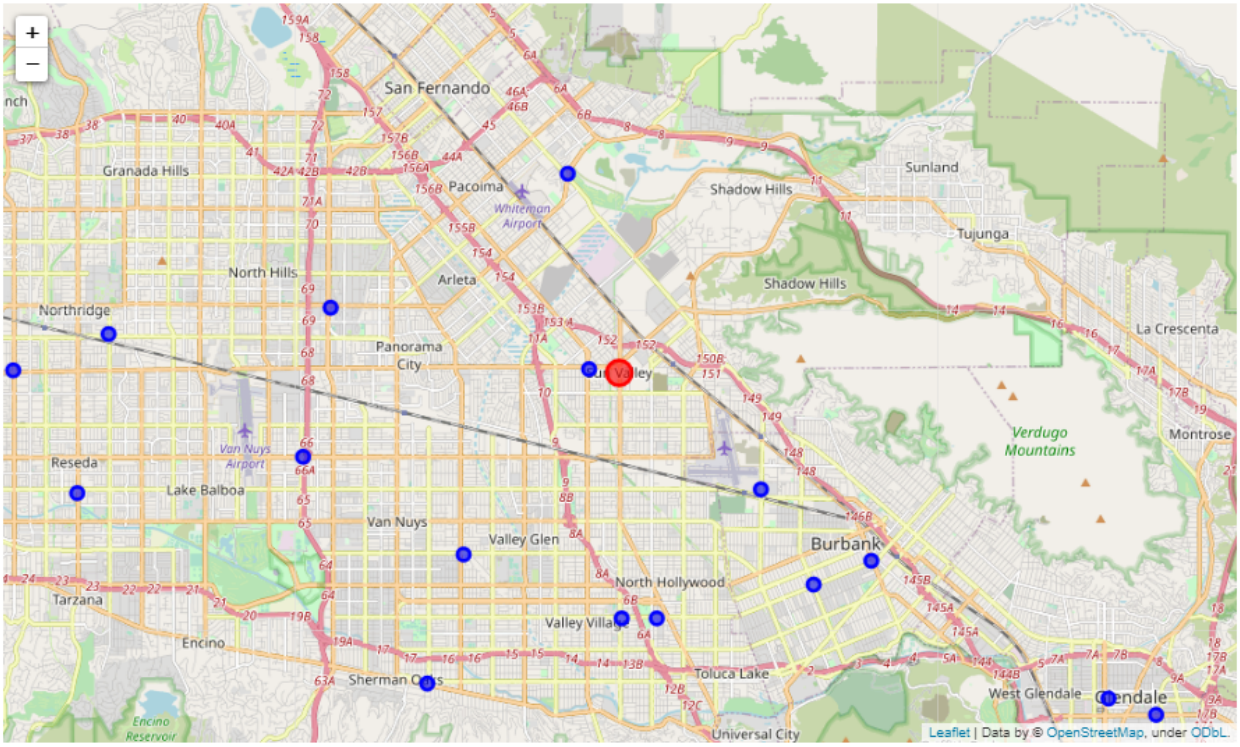
Methodology

We must identify the geo location of Bill. Using that geo location we will tap into the Foursquare API to identify the taco restaurants within a reasonable travel distance from Bill's home in Sun Valley. After all the data is acquired, we perform cleansing and transform the dataset into a panda data frame. The transformed data frame is used primarily to visualize and map the target location using python *folium* library.

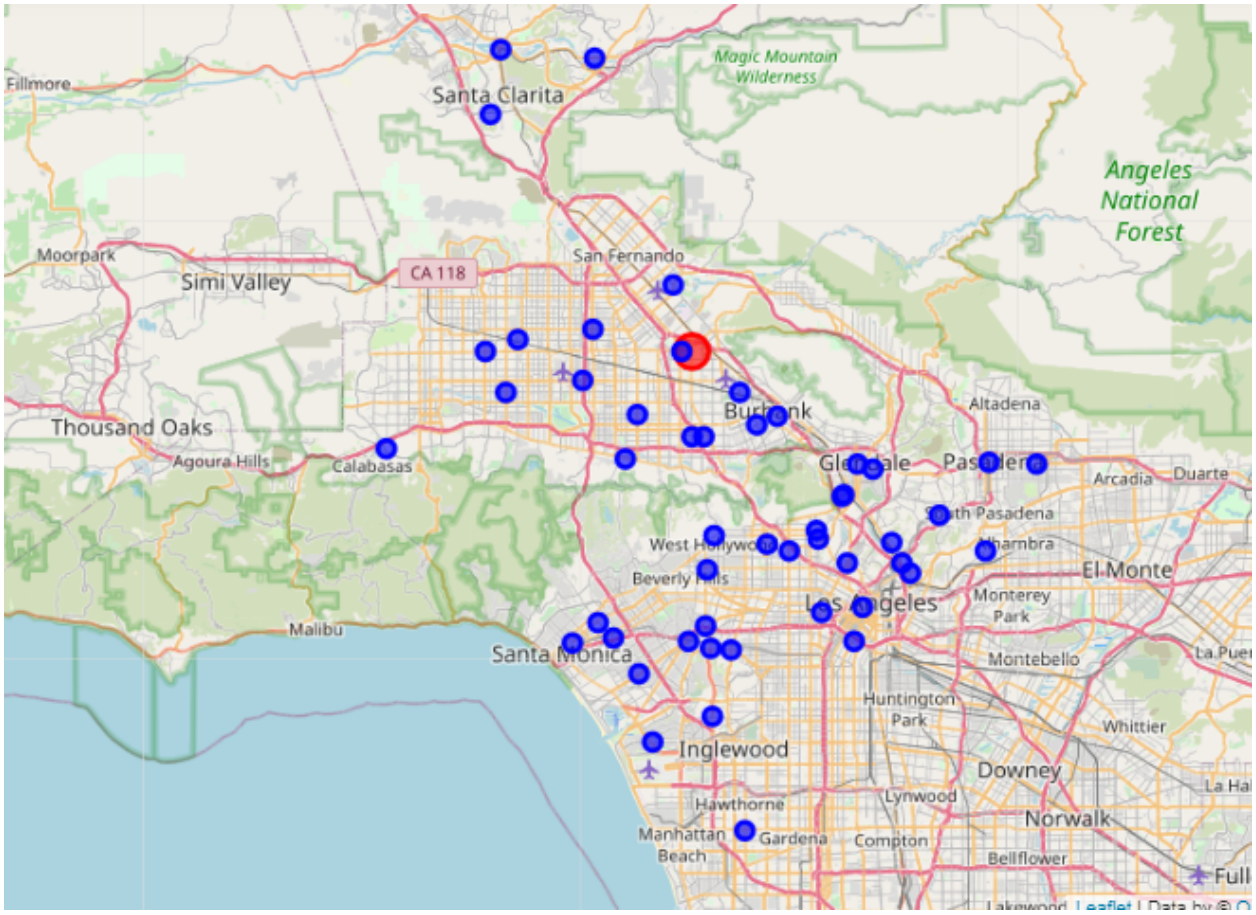
The second essential aspect of this project is to place the taco restaurantes into clusters based on location. We use K-means clustering to partition the venues that we get from Foursquare API into 5 clusters. 5 clusters seemed reasonable based on the size of Los Angeles and the distance a typical person would drive for a taco.

Analysis

Reviewing the local neighborhood it seems to be fairly taco dense. Every couple of miles from Bill's home there are 1 or 2 taco places. Below is a visualization of the local density.



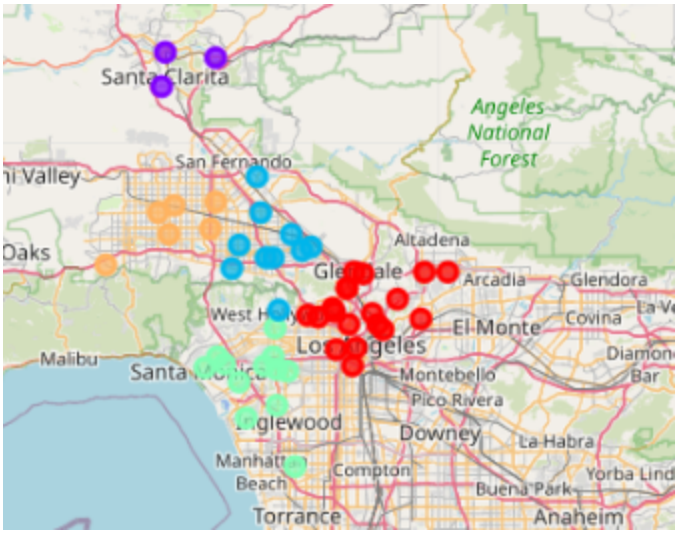
Taking a little wider view it seems that this pattern continues throughout the Los Angeles area. Visually it is difficult to determine where optimal taco location would be.



Placing the taco locations into 5 distinct areas will make it easier to determine the density of a particular area. Below is a visualization of the 5 clusters in Los Angeles. Visually you can see that cluster area number 1 in the purple seems to have the least amount of taco locations. Cluster area number 2 in the blue is where Bill currently lives. Visually this area seems kind of dense so it looks like Bill will have to commute to work. The table on the left gives a count of restaurants per cluster. The counts only confirm our assumptions. Cluster 1 only has 3 taco locations and seems to be our most promising area.

:

Cluster_Labels	name
0	19
1	3
2	10
3	12
4	6



Lastly, let's take a look at what restaurants are in each of the clusters. Cluster1 has 2 Del Taco and 1 Taco Bell in the area. Both of these restaurants are chains so Bill might be able to pull in some customers looking for a “independant” taco. Cluster 4 which only has 6 taco places is mostly fast food but does have Rigo’s Taco 9 which is independent. This might add some additional compensation

```
Out[16]: Cluster_Labels  name
         0               Ave 26 Taco Stand            1
                         Best Fish Taco in Ensenada   1
                         Del Taco                     2
                         El Flamin' Taco              1
                         King Taco                    2
                         King Taco Restaurant         4
                         Taco Bell                    6
                         Taco House #1                1
                         Taco Zone                    1
         1               Del Taco                     2
                         Taco Bell                    1
         2               Casita Taco Al Carbon        1
                         Casita Taco Mexican Grill    1
                         Del Taco                     1
                         Pink Taco                    1
                         Rigo's Taco                  2
                         Taco Bell                    4
         3               Del Taco                     3
                         Taco Bell                    6
                         Taco Bell / KFC              1
                         Taco Bell/KFC                1
                         Taco Plus                    1
         4               Del Taco                     2
                         Rigo's Taco 9                1
                         Taco Bell                    3
         Name: city, dtype: int64
```

Result

Based on our analysis it seems like the best place to open a taco spot would be in cluster number 1, in the Santa Clarita Valley. This location will have the least amount of competition for Bill. This would be a 19 mile commute for Bill and has a total population of 210k. Although one can make a case the taco restaurant per capita for each of these areas is more important, remember that Bill's taco is not that great. So it is more important for Bill to have a location far away from his competitors otherwise all potential customers will just go to his competitors. The sparseness of the Santa Clarita Valley is the perfect location for Bill.