Predicting a Suitable Location for a Tea Restaurant in Alexandria, Virginia

Christopher Dano

August 13, 2019

1 Introduction

Kung Fu Tea is an emerging bubble tea brand that has become a popular venue in Northern Virginia. Living in Alexandria, Virginia; the closest bubble tea location would either be in Arlington or Springfield, Virginia. The Kung Fu Tea Team is looking for business partners to open new locations across the globe. Due to the brands popularity in Northern Virginia, a new opening in Alexandria would help expand Kung Fu Teas presence in this area further. The main issue is deciding on a location. Alexandria is next to the Nations capital and so is already bustling with businesses.

2 Data

Alexandria must be segmented to find ideal bubble tea locations. Data pertaining to the citys zip codes were acquired from the US Zip Code Database. In addition, the zip code for the nearby Springfield Kung Fu Tea location was also included in the dataframe. The database includes zip codes additional features such as latitude and longitude, removing the need to acquire coordinates for each zip code. For each of these zip codes, the categories of 100 venues around the area were obtained using the Foursquare API. The counts for these categories were converted to frequencies and included in the data as features.

3 Methodology

The zip codes were first mapped using Folium to see the locations of the zip codes relative to one another. This was accomplished with the latitude and longitude coordinates provided by the dataset.

Suitable locations for the restaurant were determined through venue similarity. The data was provided without prior labeling, so unsupervised methods were ideal for the analysis. Before analysis began, the venue categories had to be obtained. With the Foursquare API, the categories of the 100 nearest venues were obtained for each zip code. One-hot encoding was done for the venue categories so the feature can be used with any machine learning algorithms. The data was then grouped by zip code and counted. Counts were then converted to frequencies, which were used to determine the top ten venue types for each zip code. The first most common venue type to the tenth most common venue type were added to the data as features to be used as the basis for clustering. For this study, k-means clustering was used to group the areas of Alexandria and the area of Springfield with an existing Kung Fu Tea. Starting with a k of 2, the model was built with increasing k values until there was a value of k that isolated the Springfield location.

4 Results

Figure 1 shows the map of Alexandria with the zip codes marked. The zip code markers in the northeastern section appeared closer together. Elsewhere, the zip codes appeared more dispersed.

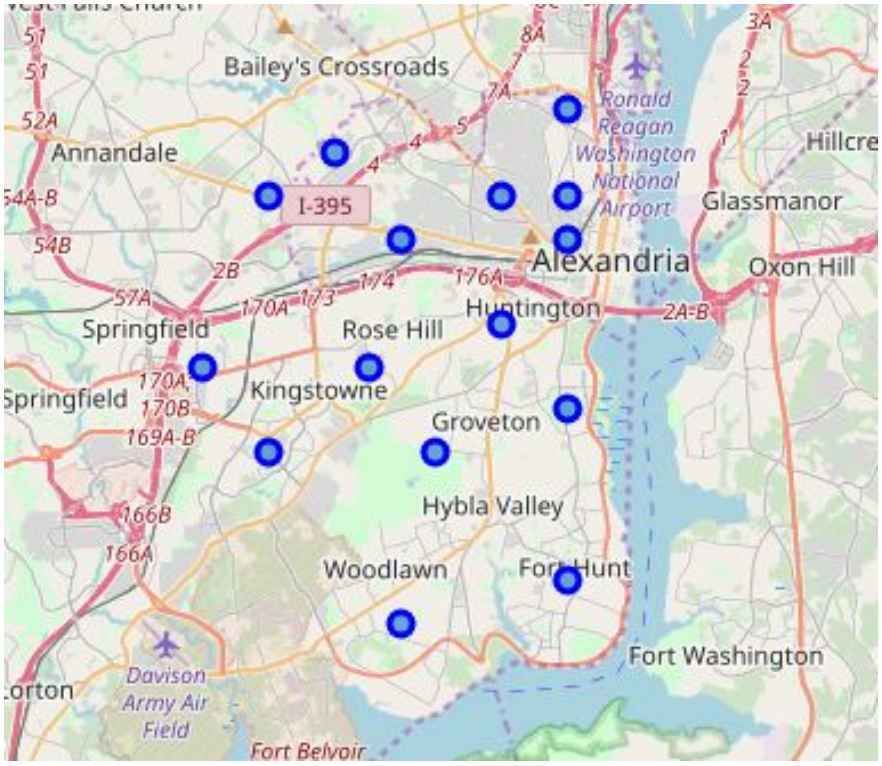

Figure 1: Map of Zip codes for Alexandria and the zip code for the Kung Fu Tea in Springfield

The model with eleven clusters contained the smallest number of similar zip codes to the one with whom a Kung Fu Tea exists. The model with twelve clusters grouped the existing restaurant location by itself. There were three other zip codes clustered with the Springfield Kung Fu Tea; 22301, 22305, and 22314 as shown in figure 2

	zip	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	22301	1	Pizza Place	Spa	Gym / Fitness Center	Coffee Shop	Pharmacy	Mexican Restaurant	Lounge	Playground	Cycle Studio	Convenience Store
4	22305	1	Pizza Place	Grocery Store	Coffee Shop	Furniture / Home Store	Spa	Supermarket	Bank	Ice Cream Shop	Mediterranean Restaurant	Gym / Fitness Center
12	22314	1	Hotel	Sandwich Place	Coffee Shop	Café	Pizza Place	New American Restaurant	Seafood Restaurant	Bookstore	Thrift / Vintage Store	Park
14	22150	1	Clothing Store	Cosmetics Shop	Coffee Shop	Women's Store	Sporting Goods Shop	American Restaurant	Italian Restaurant	Shoe Store	Hotel	Pizza Place

Figure 2: Table showing the cluster containing the Springfield location and similar zip codes.

The map of Alexandria was color-coded to visually show the clustering of the zip codes. The Blue

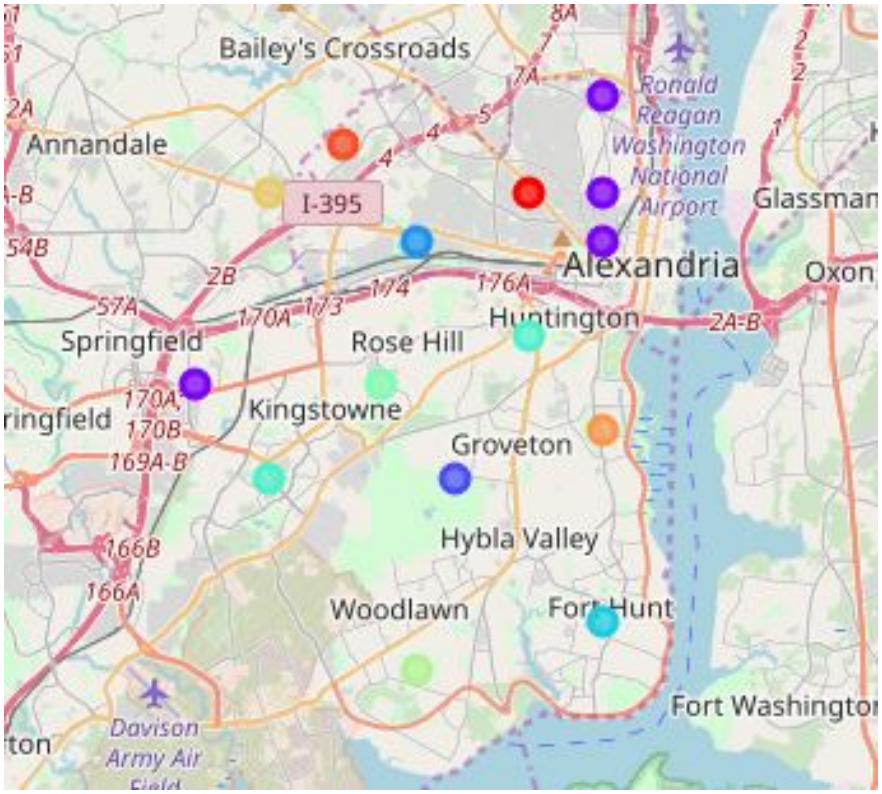

Figure 3: Map of Alexandria with markers colored based on their cluster label.

5 Discussion

The three zip codes clustered with Springfield are interpreted as the most similar areas to the existing locale in terms of most common venue types. Common venue types among all four zip codes include Pizza places and Coffee shops. These types of venues may experience a lot of traffic and can increase exposure of nearby venues. Based on these similarities, it is recommended to open a Kung Fu Tea location within the areas of 22301, 22305, or 22314. One interesting observation is that these locations have Route 1 running through them. In additon; not only are these areas most similar according to the model, but these locations offer a few perks of their own. The zip code area of 22301 houses the Neighborhood of Del Ray, which has numerous eateries and is the location of the Art on the Avenue festival. The area of 22305 is in close proximity to the Crystal City and has route 1 running through it. Lastly, 22314 is the zip code that contains Alexandria's Old Town. Old Town Alexandria is a hotspot containing many venues and experiences high amounts of traffic. Based on personal experience, Areas within 22314 and 22305 can potentially be suitable Kung Fu Tea locations.

6 Conclusion

Zip code areas in Alexandria, Virginia were clustered to find potential locations for Kung Fu Tea. Similarities were explored through the top venue categories around each area. K-means clustering was used to group areas based on top venue types. With the inclusion of an area with a Kung Fu Tea, zip codes can be clustered to see which ones are grouped with the existing location. The areas of 22301, 22305, and 22314 can potentially be suitable locations for another restaurant based on their similarities in venues.