Will Lee
5/18/2020

Opening a new restaurant in Nashville

1. Introduction

1.1 Background

Nashville is one of the fastest growing cities in America. If you are a resident in the city, the growth is very apparent. Constructions are prevalent all around the metro area, and the downtown area seems to attract increasingly larger crowd each year. Even if you are not located at, the numbers can show you a glimpse of what it's like. In 2018, Nashville saw a record breaking 15.2 million tourists. The city also enjoyed its low unemployment rate of 3.2% in 2019 thanks to the influx of new businesses and assets.

Given the facts mentioned above, this report will aid the potential business owner in deciding the location and the type. This report will focus on the restaurant business.

1.2 Business Problem

The city's ability to attract significant number of new residents is enticing for new businesses. If one were to open up a new restaurant in Nashville, relevant questions that could be posed are:

- *Where should the restaurant be located?*
- *What type of restaurant should it be?*

1.3 Parties of Interest

This report will be useful for prospective restauranteurs that are interested in evaluating the playing field in Nashville.

2. Data

To answer the questions posed above, the following list of data will be used:

- *Neighborhoods of Nashville*
- *Geospatial data of Nashville and its neighborhoods*
- *Food venues near the neighborhoods*
- *Details of all food venues*

2.1 Sources

The list of neighborhoods will be collected from the Wikipedia page (https://en.wikipedia.org/wiki/Category:Neighborhoods_in_Nashville,_Tennessee) using a web scraper, namely BeautifulSoup. The following photo shows the Wikipedia page showing all the neighborhoods.

The following 20 pages are in this category, out of 20 total. This list may not reflect recent changes (learn more).

A
- Antioch, Tennessee

B
- Bakers, Tennessee
- Bellevue, Tennessee

D
- Donelson, Tennessee

E
- East Nashville, Tennessee

G
- Green Hills, Nashville, Tennessee
- The Gulch, Nashville, Tennessee

H
- Hermitage, Tennessee
- Hillsboro Village
- Hopewell, Davidson County, Tennessee

I
- Inglewood, Nashville, Tennessee

J
- Joelton, Tennessee

L
- Lakewood, Tennessee
- Lockeland Springs

M
- Madison, Tennessee

O
- Old Hickory, Tennessee

P
- Pasquo, Tennessee

R
- Richland–West End Historic District

T
- Tusculum, Nashville, Tennessee

W
- Whites Creek, Tennessee

Geospatial data will be collected using ArcGIS of the Geocoder package on Python. (ArcGIS: https://www.arcgis.com/index.html). This will allow us to determine the latitude and longitude values for the city of Nashville and its neighborhoods.

Food venues and their respective details will be collected using Foursquare API (https://foursquare.com/). The results will be more accurate as the geospatial data will be provided in tandem to the addresses.

2.2 Collection

The resulting dataset collected from the Wikipedia page using BeatifulSoup and converted into a pandas dataframe is as follows:

	Neighborhood
0	Antioch, Tennessee
1	Bakers, Tennessee
2	Bellevue, Tennessee
3	Donelson, Tennessee
4	East Nashville, Tennessee
5	Green Hills, Nashville, Tennessee
6	The Gulch, Nashville, Tennessee
7	Hermitage, Tennessee
8	Hillsboro Village
9	Hopewell, Davidson County, Tennessee
10	Inglewood, Nashville, Tennessee
11	Joelton, Tennessee
12	Lakewood, Tennessee
13	Lockeland Springs
14	Madison, Tennessee
15	Old Hickory, Tennessee
16	Pasquo, Tennessee
17	Richland–West End Historic District
18	Tusculum, Nashville, Tennessee
19	Whites Creek, Tennessee

The data was then cleaned (Refer to next section) to provide the Geocoder package with a more accurate address for better results.

Geospatial data collected is as follows:

	Neighborhood	Latitude	Longitude
0	Antioch, Nashville, Tennessee	36.04256	-86.64971
1	Bakers, Tennessee	36.37219	-86.75999
2	Bellevue, Nashville, Tennessee	36.07383	-86.93535
3	Donelson, Tennessee	36.16743	-86.66331
4	East Nashville, Tennessee	36.17256	-86.75972
5	Green Hills, Nashville, Tennessee	36.10465	-86.81473
6	The Gulch, Nashville, Tennessee	36.15273	-86.78371
7	Hermitage, Tennessee	36.19985	-86.62067
8	Hillsboro Village	36.13467	-86.80647
9	Hopewell, Davidson County, Tennessee	36.22463	-86.62561
10	Inglewood, Nashville, Tennessee	36.21432	-86.73062
11	Joelton, Tennessee	36.31515	-86.86564
12	Lakewood, Nashville, Tennessee	36.24291	-86.63536
13	Lockeland Springs	36.17764	-86.74181
14	Madison, Tennessee	36.26398	-86.70834
15	Old Hickory, Tennessee	36.26029	-86.64595
16	Pasquo, Tennessee	36.03537	-86.96860
17	Richland West End Historic District, Nashville...	36.15405	-86.85818
18	Tusculum, Nashville, Tennessee	36.05894	-86.71650
19	Whites Creek, Tennessee	36.26645	-86.83090

Food venue data will be collected from Foursquare API by using the *venues/explore* request (Documentation : https://developer.foursquare.com/docs/api-reference/venues/explore/). Categoryid of food venues will also be passed in as a parameter to get a more relevant result.

Part of venue information collected from Foursquare API response:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Antioch, Nashville, Tennessee	36.04256	-86.64971	Wendy's	36.044098	-86.649371	Fast Food Restaurant
1	Antioch, Nashville, Tennessee	36.04256	-86.64971	New Century Buffet	36.040905	-86.650686	Chinese Restaurant
2	Bakers, Tennessee	36.37219	-86.75999	Daddy O's Pizza	36.368933	-86.757378	Pizza Place
3	Bellevue, Nashville, Tennessee	36.07383	-86.93535	McDonald's	36.070457	-86.936172	Fast Food Restaurant
4	Bellevue, Nashville, Tennessee	36.07383	-86.93535	Laissez Les Bon Temps Rouler Eu Cajun's	36.077054	-86.935122	Cajun / Creole Restaurant

2.3 Cleaning

Certain names of the neighborhoods had to be edited as they were too vague for the geospatial locator package to effectively return exact Lat/Long values. In order to address this, some of the data were accessed directly and edited accordingly. The following pictures are - code snippet for data cleaning and the new resulting dataset.

```
nash_df.iloc[0, nash_df.columns.get_loc('Neighborhood')] = 'Antioch, Nashville, Tennessee'
nash_df.iloc[2, nash_df.columns.get_loc('Neighborhood')] = 'Bellevue, Nashville, Tennessee'
nash_df.iloc[12, nash_df.columns.get_loc('Neighborhood')] = 'Lakewood, Nashville, Tennessee'
nash_df.iloc[17, nash_df.columns.get_loc('Neighborhood')] = 'Richland West End Historic District, Nashville, Tennessee'

nash_df
```

	Neighborhood
0	Antioch, Nashville, Tennessee
1	Bakers, Tennessee
2	Bellevue, Nashville, Tennessee
3	Donelson, Tennessee
4	East Nashville, Tennessee
5	Green Hills, Nashville, Tennessee
6	The Gulch, Nashville, Tennessee
7	Hermitage, Tennessee
8	Hillsboro Village
9	Hopewell, Davidson County, Tennessee
10	Inglewood, Nashville, Tennessee
11	Joelton, Tennessee
12	Lakewood, Nashville, Tennessee
13	Lockeland Springs
14	Madison, Tennessee
15	Old Hickory, Tennessee
16	Pasquo, Tennessee
17	Richland West End Historic District, Nashville...
18	Tusculum, Nashville, Tennessee
19	Whites Creek, Tennessee

2.4 Feature Selection

The Foursquare API's response has gratuitous information that will not be needed for analysis. The following table showcases the features from the response fields that were selected/unselected.

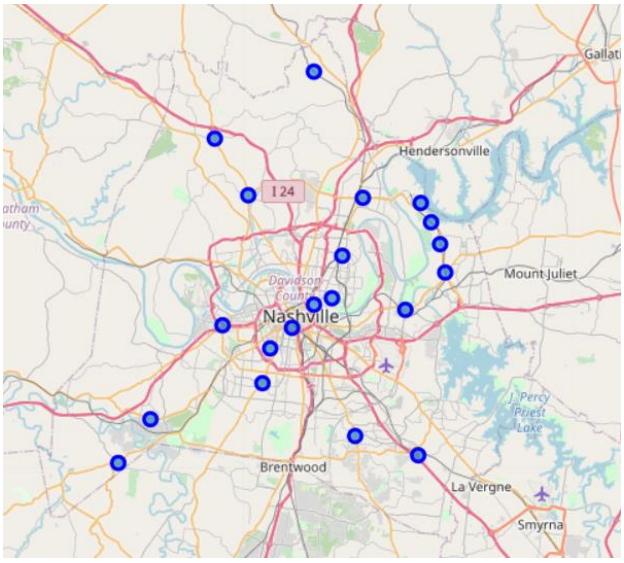
Warning	Dropped
Groups	Dropped
SuggestedRadius	Dropped
HeaderLocation	Dropped
HeaderFullLocation	Dropped
HeaderMessage	Dropped
ID	Dropped
Name	Selected
Location	Selected
Categories	Selected

3. Methodology

3.1 Exploratory Data Analysis

3.1.1 Visualizing neighborhoods

Using the geospatial data with Folium, a visualization of the neighborhoods was created as follows:



As seen above in the map, the locations and the number of the neighborhoods are correctly shown. One observation to be made from the map is that more developed area of the city has a higher density of neighborhoods.

3.1.2 Venue Datasets

The resulting food venue dataset from the Foursquare API had 172 rows and 7 columns. The dataset was grouped by neighborhoods to count the different venue categories.

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Antioch, Nashville, Tennessee	2	2	2	2	2	2
Bakers, Tennessee	1	1	1	1	1	1
Bellevue, Nashville, Tennessee	2	2	2	2	2	2
Donelson, Tennessee	7	7	7	7	7	7
East Nashville, Tennessee	15	15	15	15	15	15
Green Hills, Nashville, Tennessee	39	39	39	39	39	39
Hermitage, Tennessee	5	5	5	5	5	5
Hillsboro Village	8	8	8	8	8	8
Hopewell, Davidson County, Tennessee	1	1	1	1	1	1
Inglewood, Nashville, Tennessee	6	6	6	6	6	6
Joelton, Tennessee	1	1	1	1	1	1
Lakewood, Nashville, Tennessee	3	3	3	3	3	3
Lockeland Springs	5	5	5	5	5	5
Madison, Tennessee	10	10	10	10	10	10
Pasquo, Tennessee	4	4	4	4	4	4
Richland West End Historic District, Nashville, Tennessee	7	7	7	7	7	7
The Gulch, Nashville, Tennessee	47	47	47	47	47	47
Tusculum, Nashville, Tennessee	5	5	5	5	5	5
Whites Creek, Tennessee	4	4	4	4	4	4

As seen in the resulting dataframe above, the Gulch and Green Hills, the most popular neighborhoods in Nashville, have a wide variety of food venues as expected.

3.2 Preprocessing

The dataset above contains categorical variables that need to be analyzed. In order to work with categorical variables, one hot encoding will be implemented to change them to numerical variables. A part of the resulting data set after applying one hot encoding is as follows:

	Neighborhood	American Restaurant	Asian Restaurant	BBQ Joint	Bagel Shop	Bakery	Bistro	Breakfast Spot	Buffet	Burger Joint	...
0	Antioch, Nashville, Tennessee	0	0	0	0	0	0	0	0	0	...
1	Antioch, Nashville, Tennessee	0	0	0	0	0	0	0	0	0	...
2	Bakers, Tennessee	0	0	0	0	0	0	0	0	0	...
3	Bellevue, Nashville, Tennessee	0	0	0	0	0	0	0	0	0	...
4	Bellevue, Nashville, Tennessee	0	0	0	0	0	0	0	0	0	...

Then the dataset was grouped by neighborhoods and each value in the columns are consolidated as averages. This resulting dataset, a part of it shown below, allows for examination of the most common food venues closest to each of the neighborhoods.

	Neighborhood	American Restaurant	Asian Restaurant	BBQ Joint	Bagel Shop	Bakery	Bistro	Breakfast Spot	Buffet	Burger Joint	...
0	Antioch, Nashville, Tennessee	0.000000	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.000000	...
1	Bakers, Tennessee	0.000000	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.000000	...
2	Bellevue, Nashville, Tennessee	0.000000	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.000000	...
3	Donelson, Tennessee	0.000000	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.000000	...
4	East Nashville, Tennessee	0.066667	0.000000	0.066667	0.000	0.066667	0.000000	0.000000	0.000000	0.066667	...
5	Green Hills, Nashville, Tennessee	0.102564	0.025641	0.000000	0.000	0.051282	0.000000	0.025641	0.000000	0.000000	...
6	Hermitage, Tennessee	0.200000	0.200000	0.000000	0.000	0.000000	0.000000	0.200000	0.000000	0.000000	...
7	Hillsboro Village	0.125000	0.125000	0.000000	0.125	0.000000	0.000000	0.000000	0.000000	0.000000	...

For easier analysis of each neighborhood, the dataset was then sorted to show the ten most common food venues in each of the neighborhood as follows:

	Neighborhood	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	Antioch, Nashville, Tennessee	Fast Food Restaurant	Chinese Restaurant	Wings Joint	Gastropub	Food Truck	Food	Donut Shop	Deli / Bodega	Cajun / Creole Restaurant	Café
1	Bakers, Tennessee	Pizza Place	Wings Joint	Food Truck	Food	Fast Food Restaurant	Donut Shop	Deli / Bodega	Chinese Restaurant	Cajun / Creole Restaurant	Café
2	Bellevue, Nashville, Tennessee	Fast Food Restaurant	Cajun / Creole Restaurant	Wings Joint	Gastropub	Food Truck	Food	Donut Shop	Deli / Bodega	Chinese Restaurant	Café
3	Donelson, Tennessee	Pizza Place	Seafood Restaurant	Sandwich Place	Greek Restaurant	Gastropub	Bakery	Food	Fast Food Restaurant	Donut Shop	Deli / Bodega
4	East Nashville, Tennessee	Restaurant	Greek Restaurant	Pizza Place	BBQ Joint	Bakery	Burger Joint	Café	Donut Shop	Food	Fried Chicken Joint

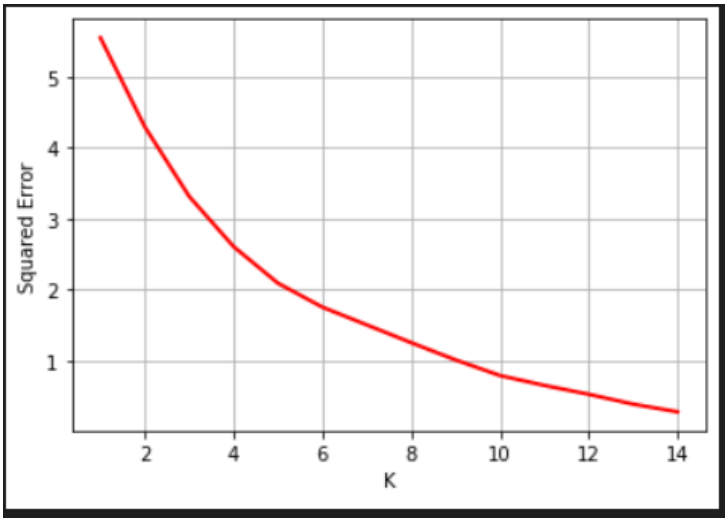
This dataframe will be pivotal for applying K-Means clustering algorithm to classify different categories of neighborhoods.

Will Lee
5/18/2020

3.3 K-Means Clustering

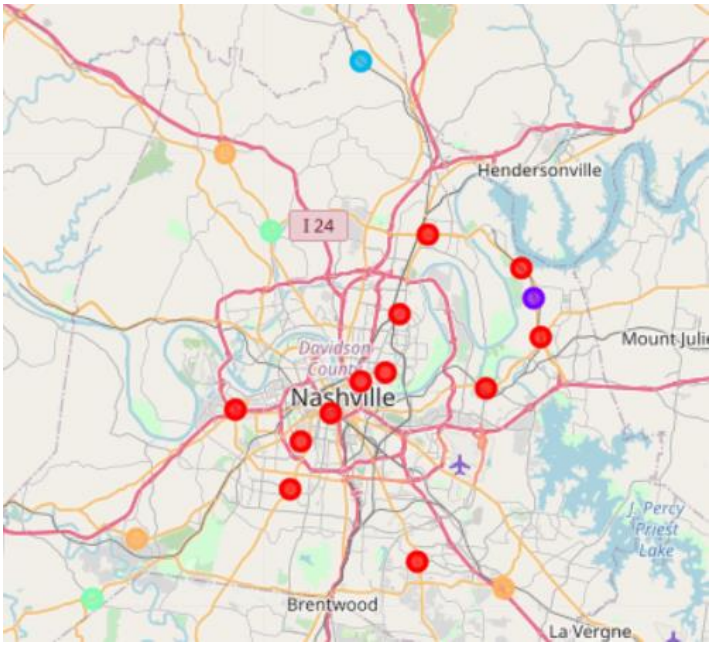
K-Means clustering was applied to the dataset in order to classify the neighborhoods with similar characteristics. The resulting clusters will be examined and allow the question of *where* and *what type* to be answered.

To identify the optimal number of clusters K for the dataset, the elbow method was used. By iterating through K = 1 from K = 10, graphing the squared error and identifying the "elbow", the optimal K could be found.



Although the elbow is not extremely apparent, upon closer inspection, K = 5 has the largest delta. K-means clustering was then applied to the dataset with K = 5.

After the application of K-means clustering, the neighborhoods were labeled on the map with different color schemes for each cluster.



Will Lee
5/18/2020

4. Results

4.1 Cluster 1

	Neighborhood	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
3	Donelson, Tennessee	Pizza Place	Seafood Restaurant	Sandwich Place	Greek Restaurant	Gastropub	Bakery	Food	Fast Food Restaurant	Donut Shop	Deli / Bodega
4	East Nashville, Tennessee	Restaurant	Greek Restaurant	Pizza Place	BBQ Joint	Bakery	Burger Joint	Café	Donut Shop	Food	Fried Chicken Joint
5	Green Hills, Nashville, Tennessee	Pizza Place	American Restaurant	Sandwich Place	Mediterranean Restaurant	Mexican Restaurant	Bakery	Café	Restaurant	Greek Restaurant	Fast Food Restaurant
6	The Gulch, Nashville, Tennessee	Restaurant	American Restaurant	Steakhouse	Café	Sandwich Place	Deli / Bodega	Burger Joint	Gastropub	Pizza Place	Mexican Restaurant
7	Hermitage, Tennessee	American Restaurant	Breakfast Spot	Food Truck	Fried Chicken Joint	Asian Restaurant	Bagel Shop	Bakery	Bistro	BBQ Joint	Buffet
8	Hillsboro Village	American Restaurant	Bagel Shop	Hot Dog Joint	Italian Restaurant	Noodle House	Fried Chicken Joint	Asian Restaurant	Sushi Restaurant	Bakery	Chinese Restaurant
10	Inglewood, Nashville, Tennessee	Pizza Place	Fast Food Restaurant	Seafood Restaurant	Mexican Restaurant	Fried Chicken Joint	Burrito Place	Food	Donut Shop	Deli / Bodega	Chinese Restaurant
12	Lakewood, Nashville, Tennessee	Thai Restaurant	Bistro	Breakfast Spot	Cajun / Creole Restaurant	Fried Chicken Joint	Food Truck	Food	Fast Food Restaurant	Donut Shop	Deli / Bodega
13	Lockeland Springs	American Restaurant	Pizza Place	Food	Bakery	Mexican Restaurant	Fast Food Restaurant	Donut Shop	Deli / Bodega	Chinese Restaurant	Cajun / Creole Restaurant
14	Madison, Tennessee	Wings Joint	Pizza Place	BBQ Joint	Bakery	Chinese Restaurant	Fast Food Restaurant	Fried Chicken Joint	Italian Restaurant	American Restaurant	Seafood Restaurant
17	Richland West End Historic District, Nashville...	Fast Food Restaurant	Chinese Restaurant	Breakfast Spot	Buffet	Mexican Restaurant	Pizza Place	Wings Joint	Café	Food	Donut Shop
18	Tusculum, Nashville, Tennessee	Mexican Restaurant	Food Truck	Pizza Place	Chinese Restaurant	Food	Fast Food Restaurant	Donut Shop	Deli / Bodega	Cajun / Creole Restaurant	Café

The neighborhoods included in the cluster are within the more developed part of the city. As they generally have more customers and different needs, the spectrum of cuisines varies. However, a general pattern can be seen: American Restaurants are the most common, and in certain places where specific ethnicities are located, the fact is reflected by the type of the common food venues.

4.2 Cluster 2

	Neighborhood	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
9	Hopewell, Davidson County, Tennessee	Hot Dog Joint	Wings Joint	Café	Food Truck	Food	Fast Food Restaurant	Donut Shop	Deli / Bodega	Chinese Restaurant	Cajun / Creole Restaurant

The neighborhood included in cluster 2, Hopewell, located near the Hermitage Farm and Cumberland River has fast food around, as they serve people on the road.

4.3 Cluster 3

	Neighborhood	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
1	Bakers, Tennessee	Pizza Place	Wings Joint	Food Truck	Food	Fast Food Restaurant	Donut Shop	Deli / Bodega	Chinese Restaurant	Cajun / Creole Restaurant	Café

The neighborhood included in cluster 3, Bakers, is located North of Nashville. There are few attractions located, which makes it suboptimal for dine-in restaurants to be profitable. This fact is located in the dataframe above, as the neighborhood has mostly venues focused on delivery or takeout.

4.4 Cluster 4

	Neighborhood	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
16	Pasquo, Tennessee	Southern / Soul Food Restaurant	BBQ Joint	Food	Wings Joint	Café	Food Truck	Fast Food Restaurant	Donut Shop	Deli / Bodega	Chinese Restaurant
19	Whites Creek, Tennessee	American Restaurant	Food	Southern / Soul Food Restaurant	Cajun / Creole Restaurant	Burrito Place	Food Truck	Fast Food Restaurant	Donut Shop	Deli / Bodega	Chinese Restaurant

Neighborhoods in cluster 4, Pasquo, a neighborhood that is expanding thanks to the westward expansion of the Nashville suburbs and Whites Creek, serve Southern Food.

4.5 Cluster 5

	Neighborhood	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	Antioch, Nashville, Tennessee	Fast Food Restaurant	Chinese Restaurant	Wings Joint	Gastropub	Food Truck	Food	Donut Shop	Deli / Bodega	Cajun / Creole Restaurant	Café
2	Bellevue, Nashville, Tennessee	Fast Food Restaurant	Cajun / Creole Restaurant	Wings Joint	Gastropub	Food Truck	Food	Donut Shop	Deli / Bodega	Chinese Restaurant	Café
11	Joelton, Tennessee	Fast Food Restaurant	Wings Joint	Gastropub	Food Truck	Food	Donut Shop	Deli / Bodega	Chinese Restaurant	Cajun / Creole Restaurant	Café

The neighborhoods in cluster 5 has numerous fast food restaurants as they are located at the outskirts of Nashville.

5. Discussion

If one were to open up a restaurant in Nashville, neighborhoods in cluster 1 would be optimal if one has enough capital. The neighborhoods in cluster 1 is developed and has a high number of potential customers. A fusion of American and non-American cuisine seems be a good option, as there are numerous American restaurants which reflect the palate of the general population. However, as seen in the dataset Chinese food, Italian food and other non-American food venues are quite common in a few places. To have less competition and to appeal to the general public, it would be smart to open a fusion restaurant.

Another choice would be neighborhoods in cluster 5 if one has limited capital. The neighborhoods in cluster 1 will most likely have expensive rent rates – that's why neighborhoods in cluster 5 is also a strong choice. In addition, neighborhoods in cluster 5 is not as saturated in terms of food venues as cluster 1, leading to a more flexible choice when it comes to the type of the restaurant. For cluster 5, it seems to be a good choice to open an American restaurant as dine-in restaurants are not as prevalent according to the data.

6. Conclusion

The result could be even more accurate if the dataset analyzed *most popular food venues, population density* and *distribution of ethnicities.* In addition, it seems the Foursquare database on smaller cities such as Nashville seems to have less information compared to big cities such as Los Angeles or New York. If there was a database that contains all the food venues located in Nashville, the result would be more accurate.

Overall, however, it seems that the result drawn from the project is non-trivial nonetheless.

Will Lee
5/18/2020

Bibliography

Reicher, M., 2020. [online] Tennessean.com. Available at: <https://www.tennessean.com/story/news/2019/04/18/nashville-population-growth-slows-2018/3498194002/> [Accessed 18 May 2020].

“Tennessee's Unemployment Rate Hits Record Low in February.” Tennessee State Government - TN.gov, www.tn.gov/workforce/general-resources/news/2019/3/21/tennessees-unemployment-rate-hits-record-low-in-february.html.