Davao City, Philippines: A Booming City in the South

1. Introduction/Business Problem

Recent changes in the Philippine government administration paved the way for decentralization of development plans. What was formerly the mindset of focusing on beautifying and advancing the capital of the country has now been changed into strengthening the other potentially progressive cities and municipalities. One such rising city in the south is Davao City.

Those who have resided in Davao City for a long time can attest to the rapid developments taking place in the city. High-rise buildings and business establishments which have previously been rare are now appearing like mushrooms, not to mention the increase in traffic flow and congestion as more and more citizens are being financially capable of owning private vehicles. People from surrounding cities and municipalities swarm because of the education and job opportunities the city has to offer. For a business man who wants to invest and take advantage of the flourishing economy in this city, a good understanding of the existing venues (or lack thereof) spread all over the city would be a good starting point. Thus, this project aims to provide a geospatial analysis of the venue categories found in the barangays in Davao City. Specifically, the barangays will be clustered according to the venue categories in order to identify potential business ventures as well as the location of such.

This project will provide valuable information to businessmen who want to determine which category of venue will be most promising for a particular area, considering the existing venues nearby. Furthermore, the results of the project may also help city planners identify barangays which can potentially cater to certain categories of venues, thereby contributing to strategic planning of new venues.

2. Data Description

All barangays in Davao City, Philippines will be included in the analysis. The list of barangays can be found in a Wikipedia page (https://en.wikipedia.org/wiki/Districts_of_Davao_City) which will be scraped through the BeautifulSoup package in Python. The barangays are grouped according to the district they are assigned to. There are 3 districts in Davao City, and each is subdivided according to district centers. For example, the 1st District has two centers: Poblacion with 40 barangays and Talomo with 14 barangays.

The geographical coordinates of the barangays will be obtained through geocoder. Particularly, the latitude and longitude data per barangay will be retrieved. These coordinates

will be used to determine 100 venues within a 500-meter radius per barangay. These venues can be retrieved from Foursquare. The venues and venue categories, as well as the geographical coordinates per venue will be included in the analysis. For instance, for Poblacion in the 1st District, one of the venues within its 500-meter radius is the famous "Roxas Street Foods" under the category "BBQ Joint" with latitude and longitude equal to 7.063038 and 125.611376, respectively. The barangays and the venues will then be mapped through Folium.

3. Methodology

3.1 Putting data into a dataframe

The data on the list of barangays in Davao City scraped from the Wikipedia webpage were put into a dataframe in Python for it to be more amenable to analysis. The dataframe columns were first constructed as follows:

```
In [6]:
#Defines the dataframe column names
column_names = ['District','District Center','Barangay']

import pandas as pd
dvobrgys_df = pd.DataFrame(columns = column_names)
dvobrgys_df
```

Out[6]:

District	District Center	Barangay

Next, the dataframe was populated with the list of barangays from the Wikipedia page and their respective geographical coordinates (latitude and longitude) using geocoder:

```
In [13]:
for row in range(dvobrgys_df.shape[0]):
    address = '{}, Davao City'.format(dvobrgys_df.loc[row,'Barangay'])
    geolocator = Nominatim(user_agent="ny_explorer")
    location = geolocator.geocode(address)
    latitude = location.latitude
    longitude = location.longitude
    dvobrgys_df.loc[row,'Latitude'] = latitude
    dvobrgys_df.loc[row,'Longitude'] = longitude

dvobrgys_df.head()
```

Out[13]:

	District	District Center	Barangay	Latitude	Longitude
0	1st District	Poblacion (40)	Poblacion	7.064524	125.608414
1	1st District	Talomo (14)	Bago Aplaya	7.040826	125.535749
2	1st District	Talomo (14)	Bago Gallera	7.054778	125.512349
3	1st District	Talomo (14)	Baliok	7.046495	125.499887
4	1st District	Talomo (14)	Bucana	7.049276	125.603363

Some barangays had to be dropped because geocoder returns 'NaN' whenever their geographical coordinates are fetched. In particular, the following barangays were dropped:

Drops the following barangays because coordinates cannot be found:

1. Centro (San Juan)
2. Kap. Tomas Monteverde, Sr.
3. Wilfredo Aquino
4. San Isidro (Licanan)
5. Baguio Proper
6. Tambubong
7. Marilog Proper
8. Toril Proper

```
In [12]: dvobrgys_df = dvobrgys_df.drop(dvobrgys_df[dvobrgys_df['Barangay']==dvobrgys_df.loc[16,'Barangay']].index).reset_index(drop=True)
         dvobrgys_df = dvobrgys_df.drop(dvobrgys_df[dvobrgys_df['Barangay']==dvobrgys_df.loc[18,'Barangay']].index).reset_index(drop=True)
         dvobrgys_df = dvobrgys_df.drop(dvobrgys_df[dvobrgys_df['Barangay']==dvobrgys_df.loc[23,'Barangay']].index).reset_index(drop=True)
         dvobrgys_df = dvobrgys_df.drop(dvobrgys_df[dvobrgys_df['Barangay']==dvobrgys_df.loc[43,'Barangay']].index).reset_index(drop=True)
         dvobrgys_df = dvobrgys_df.drop(dvobrgys_df[dvobrgys_df['Barangay']==dvobrgys_df.loc[57,'Barangay']].index).reset_index(drop=True)
         dvobrgys_df = dvobrgys_df.drop(dvobrgys_df[dvobrgys_df['Barangay']==dvobrgys_df.loc[61,'Barangay']].index).reset_index(drop=True)
         dvobrgys_df = dvobrgys_df.drop(dvobrgys_df[dvobrgys_df['Barangay']==dvobrgys_df.loc[90,'Barangay']].index).reset_index(drop=True)
         dvobrgys_df = dvobrgys_df.drop(dvobrgys_df[dvobrgys_df['Barangay']==dvobrgys_df.loc[116,'Barangay']].index).reset_index(drop=True)
```

But then for ‘Baguio Proper’, ‘Marilog Proper’, and ‘Toril Proper’, removing the word ‘Proper’ enables geocoder to return the geographical coordinates of the barangays. Thus, these barangays were returned to the dataframe but the address passed on to geocoder did not include the word ‘Proper’:

Return the following barangays (dropped the word 'Proper' in order to find the coordinates):

1. Baguio Proper
2. Marilog Proper
3. Toril Proper

```
In [14]: geolocator = Nominatim(user_agent="ny_explorer")
         baguio = geolocator.geocode('Baguio, Davao City')
         marilog = geolocator.geocode('Marilog, Davao City')
         toril = geolocator.geocode('Toril, Davao City')

         dvobrgys_df.loc[135] = ['3rd District','Baguio (8)','Baguio Proper',baguio.latitude,baguio.longitude]
         dvobrgys_df.loc[136] = ['3rd District','Marilog (12)','Marilog Proper',marilog.latitude,marilog.longitude]
         dvobrgys_df.loc[137] = ['3rd District','Toril (25)','Toril Proper',toril.latitude,toril.longitude]

         dvobrgys_df.tail()
```

Out[14]:

	District	District Center	Barangay	Latitude	Longitude
133	3rd District	Tugbok (18)	Tugbok Proper	7.109754	125.484458
134	3rd District	Tugbok (18)	Ula	7.131920	125.492772
135	3rd District	Baguio (8)	Baguio Proper	7.173743	125.401937
136	3rd District	Marilog (12)	Marilog Proper	7.297476	125.303010
137	3rd District	Toril (25)	Toril Proper	7.019342	125.497599

The dataframe used for the entire analysis hence comprised of 138 barangays in Davao City.

```
In [15]: dvobrgys_df.shape
Out[15]: (138, 5)
```

3.2 Creating a map of Davao city with its barangays

Using geocoder, the latitude and longitude of Davao city were obtained. These were then passed as arguments to Folium in order to create a map of Davao city and its barangays.

3.3 Obtaining the list of nearby venues and their categories

Using the geospatial data from foursquare.com, the nearest 100 venues within a 500-meter radius per barangay were obtained, along with their latitudes, longitudes, and venue categories. The following function was defined and used on the data from the dataframe constructed previously:

```
In [20]: def getNearbyVenues(names, latitudes, longitudes, radius=500, LIMIT=100):

    venues_list=[]
    for name, lat, lng in zip(names, latitudes, longitudes):
        #print(name)

        # create the API request URL
        url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            lat,
            lng,
            radius,
            LIMIT)

        # make the GET request
        results = requests.get(url).json()["response"]['groups'][0]['items']

        # return only relevant information for each nearby venue
        venues_list.append([(
            name,
            lat,
            lng,
            v['venue']['name'],
            v['venue']['location']['lat'],
            v['venue']['location']['lng'],
            v['venue']['categories'][0]['name']) for v in results])

    nearby_venues = pd.DataFrame([item for venue_list in venues_list for item in venue_list])
    nearby_venues.columns = ['Barangay',
                  'Barangay Latitude',
                  'Barangay Longitude',
                  'Venue',
                  'Venue Latitude',
                  'Venue Longitude',
                  'Venue Category']

    return(nearby_venues)
```

```
In [21]: dvobrgys_venues = getNearbyVenues(names=dvobrgys_df['Barangay'],
                                          latitudes=dvobrgys_df['Latitude'],
                                          longitudes=dvobrgys_df['Longitude']
                                         )
```

However, after using the function, it turns out that not all barangays have venue records:

```
In [23]: print('There are {} barangays covered in the entire Davao City.'.format(len(dvobrgys_venues['Barangay'].unique())))
```

```
There are 62 barangays covered in the entire Davao City.
```

Since the aim of this project is to perform analysis on all the barangays, the other barangays were inserted into the ***dvobrgys_venues*** dataframe but the values were set to 'None', as shown below:

```
In [27]: dvobrgys_venues.shape
```

```
Out[27]: (499, 7)
```

```
In [28]: for i in range(dvobrgys_df.shape[0]):
             if (sum(dvobrgys_venues['Barangay'] == dvobrgys_df.loc[i,'Barangay'])==0):
                 dvobrgys_venues.loc[dvobrgys_venues.shape[0],'Barangay'] = dvobrgys_df.loc[i,'Barangay']
         dvobrgys_venues.shape
```

```
Out[28]: (575, 7)
```

```
In [29]: dvobrgys_venues.tail()
```

Out[29]:

	Barangay	Barangay Latitude	Barangay Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
570	New Valencia	NaN	NaN	NaN	NaN	NaN	NaN
571	Tagakpan	NaN	NaN	NaN	NaN	NaN	NaN
572	Talandang	NaN	NaN	NaN	NaN	NaN	NaN
573	Ula	NaN	NaN	NaN	NaN	NaN	NaN
574	Marilog Proper	NaN	NaN	NaN	NaN	NaN	NaN

```
In [30]: dvobrgys_venues.fillna('None', inplace=True)
         dvobrgys_venues.tail()
```

Out[30]:

	Barangay	Barangay Latitude	Barangay Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
570	New Valencia	None	None	None	None	None	None
571	Tagakpan	None	None	None	None	None	None
572	Talandang	None	None	None	None	None	None
573	Ula	None	None	None	None	None	None
574	Marilog Proper	None	None	None	None	None	None

3.4 Preparing the data for clustering

The clustering algorithms require numerical inputs, so for the dataset at hand, dummy variables were created for each venue category, including the 'None' value:

```
In [30]: # one hot encoding
         dvobrgys_onehot = pd.get_dummies(dvobrgys_venues[['Venue Category']], prefix="", prefix_sep="")

         # add barangay column back to dataframe
         dvobrgys_onehot['Barangay'] = dvobrgys_venues['Barangay']

         # move barangay column to the first column
         fixed_columns = [dvobrgys_onehot.columns[-1]] + list(dvobrgys_onehot.columns[:-1])
         dvobrgys_onehot = dvobrgys_onehot[fixed_columns]

         dvobrgys_onehot.head()
```

Out[30]:

	Barangay	Airport	Asian Restaurant	Athletics & Sports	Australian Restaurant	BBQ Joint	Bakery	Bar	Basketball Court	Beach	...	Steakhouse	Supermarket	Theme Park	Theme Park Ride / Attraction	Thrift / Vintage Store	University	Vacation Rental
0	Poblacion	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
1	Poblacion	0	0	0	0	1	0	0	0	0	...	0	0	0	0	0	0	0
2	Poblacion	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
3	Poblacion	0	0	0	1	0	0	0	0	0	...	0	0	0	0	0	0	0
4	Poblacion	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0

Since we are interested in clustering similar barangays and analysing each cluster, we have to group together those barangays whose proportions of venue categories are more similar. To do this, the mean was obtained per category:

```
In [32]: dvobrgys_vcfreq = dvobrgys_onehot.groupby('Barangay').mean().reset_index().round(2)
         dvobrgys_vcfreq.head()
```

Out[32]:

	Barangay	Afghan Restaurant	Airport	Arcade	Asian Restaurant	Athletics & Sports	Australian Restaurant	Auto Workshop	BBQ Joint	Bakery	...	Sports Club	Steakhouse	Supermarket	Theme Park	Theme Park Ride / Attraction	Thrift / Vintage Store	University
0	Acacia	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	Agdao Proper	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.25	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	Alambre	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	Alejandra Navarro (Lasang)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	Alfonso Angliongto Sr	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0

5 rows × 118 columns

```
In [33]: dvobrgys_vcfreq.shape
```

Out[33]: (138, 118)

Finally, the latitude and longitude data were also inserted into the dataframe:

```
In [36]: to_cluster = dvobrgys_vcfreq
         for i in range(to_cluster.shape[0]):
             for j in range(dvobrgys_df.shape[0]):
                 if(to_cluster.loc[i,'Barangay']==dvobrgys_df.loc[j,'Barangay']):
                     to_cluster.loc[i,'Latitude'] = dvobrgys_df.loc[j,'Latitude']
                     to_cluster.loc[i,'Longitude'] = dvobrgys_df.loc[j,'Longitude']
         to_cluster.head()
```

Out[36]:

irport	Asian Restaurant	Athletics & Sports	Australian Restaurant	BBQ Joint	Bakery	Bar	Basketball Court	Beach	...	Theme Park	Theme Park Ride / Attraction	Thrift / Vintage Store	University	Vacation Rental	Vegetarian / Vegan Restaurant	Wine Bar	Wings Joint	Latitude	Longitu
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.192088	125.602
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.088696	125.627
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.060412	125.471
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.264411	125.663
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.109070	125.637

giving us the ***to_cluster*** dataframe to work on.

3.5 Clustering the barangays in Davao City

To group together barangays that are more similar in terms of the most frequent venue categories that exist nearby, two machine learning algorithms were used: k-Means clustering and DBSCAN.

3.5.1 k-Means clustering algorithm

The k-Means clustering algorithm was applied to the data containing the proportion of venue categories per barangay as well as the geographical coordinates of each barangay. Three clusters were formed as follows:

```
In [36]: from sklearn.cluster import KMeans

         # set number of clusters
         kclusters = 3

         dvobrgys_vcfreq_clustering = to_cluster.drop(['Barangay'], 1)

         # run k-means clustering
         kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(dvobrgys_vcfreq_clustering)

         # check cluster labels generated for each row in the dataframe
         kmeans.labels_

Out[36]: array([0, 0, 1, 1, 0, 0, 1, 2, 0, 0, 1, 0, 0, 0, 1, 1, 2, 1, 1, 1, 1, 1,
                0, 0, 0, 0, 0, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0,
                1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 0,
                0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 1, 1, 2, 1, 0, 1, 0, 2, 1,
                0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1, 0, 0, 2, 1, 1, 1, 0, 0, 0, 0,
                0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 0, 2, 1,
                0, 1, 0, 1, 2, 1], dtype=int32)
```

This clustering algorithm works by grouping together data points that are near each centroid and far from the other centroids. It is widely used by data scientists to segment and cluster data because of its relative simplicity in terms of implementation, its scalability, guarantee of convergence, and its ability to generalize to clusters of different shapes and sizes. However, for this algorithm, the number of clusters must be set manually, which may lead to an non-optimal choice. Also, the results are dependent on the initial centroids. Most of all, outliers are being dragged into a cluster when they should have been ignored, since all data points should belong to a cluster after performing the entire process. Because of this, another clustering algorithm was considered, namely the DBSCAN.

3.5.2 DBSCAN Clustering

DBSCAN stands for Density-Based Spatial Clustering for Applications with Noise. It is one of the unsupervised machine learning algorithms used to cluster observations especially when separating clusters of high density and low density within a given dataset. Its special attention to outliers, unlike k-Means clustering, is an advantage for datasets where outliers are inevitable. In the case of our dataset, some data points are likely to be outliers geographically since some barangays might be located in remote areas. This is not to mention yet the possibility of being outliers due to difference in venue categories. As such, DBSCAN was also applied as a

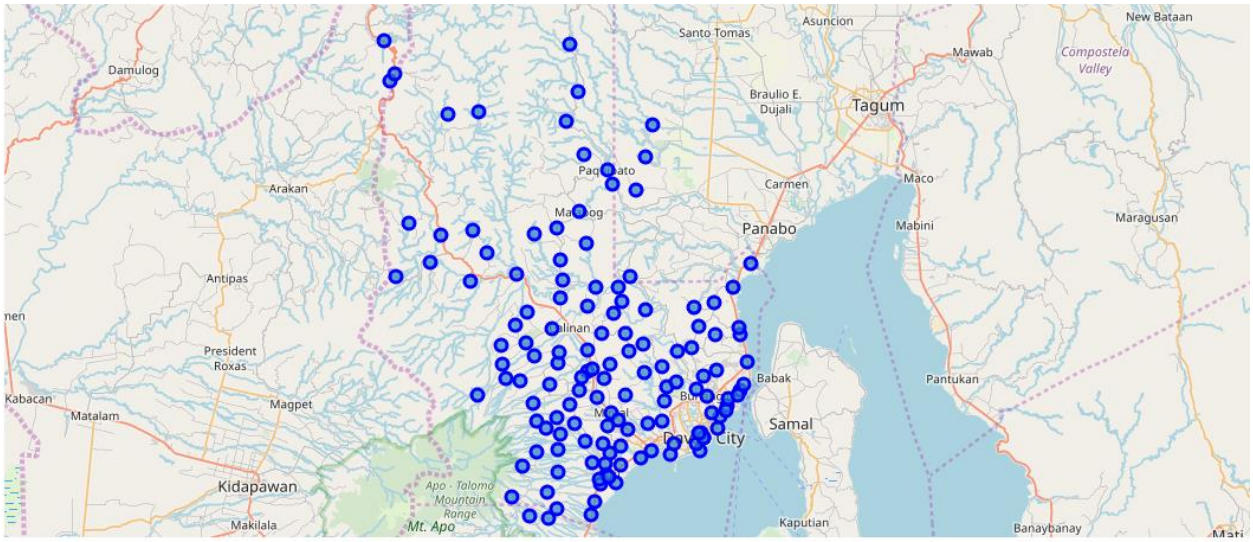
clustering method for the barangays in Davao City:

```
In [65]: from sklearn.cluster import DBSCAN
         import sklearn.utils
         from sklearn.preprocessing import StandardScaler
         sklearn.utils.check_random_state(1000)


         # Compute DBSCAN
         db = DBSCAN(eps=0.5, min_samples=10).fit(to_cluster)
         core_samples_mask = np.zeros_like(db.labels_, dtype=bool)
         core_samples_mask[db.core_sample_indices_] = True
         labels = db.labels_
         to_cluster["Clus_Db"]=labels

         realClusterNum=len(set(labels)) - (1 if -1 in labels else 0)
         clusterNum = len(set(labels))
```

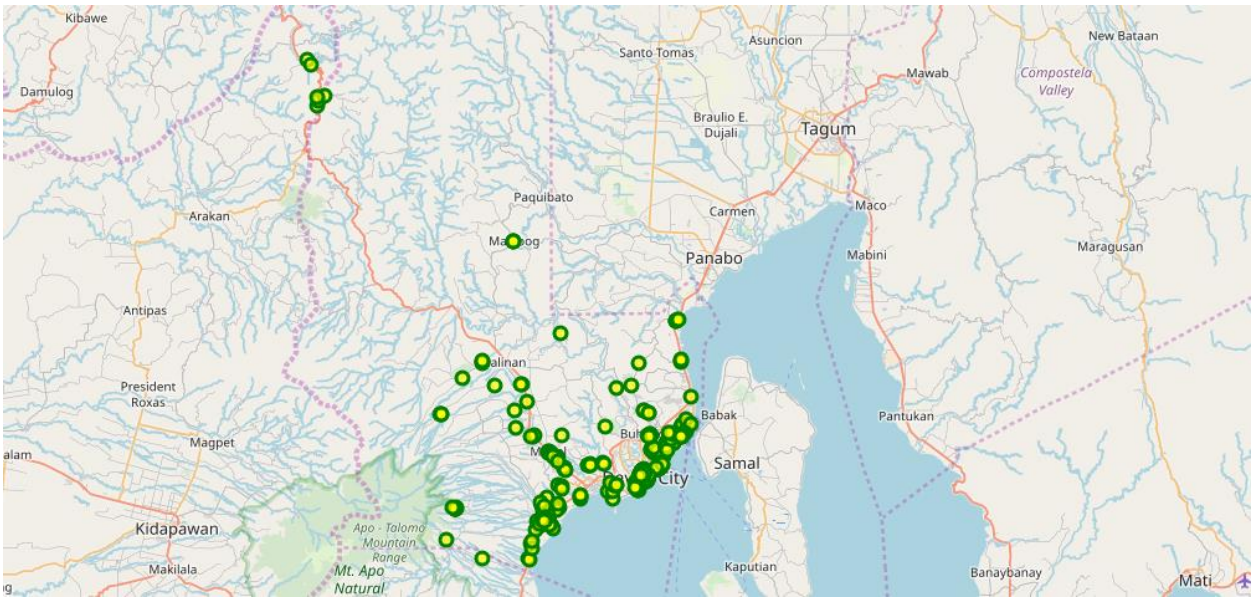
4. Results

4.1 Map of barangays and venues in Davao city

The following figure shows the distribution of the 138 barangays in Davao city:



Note how most barangays are denser near the coast and sparser in the north where mountains are located. As for the venues recorded in foursquare, the following shows their geospatial distribution among barangays:



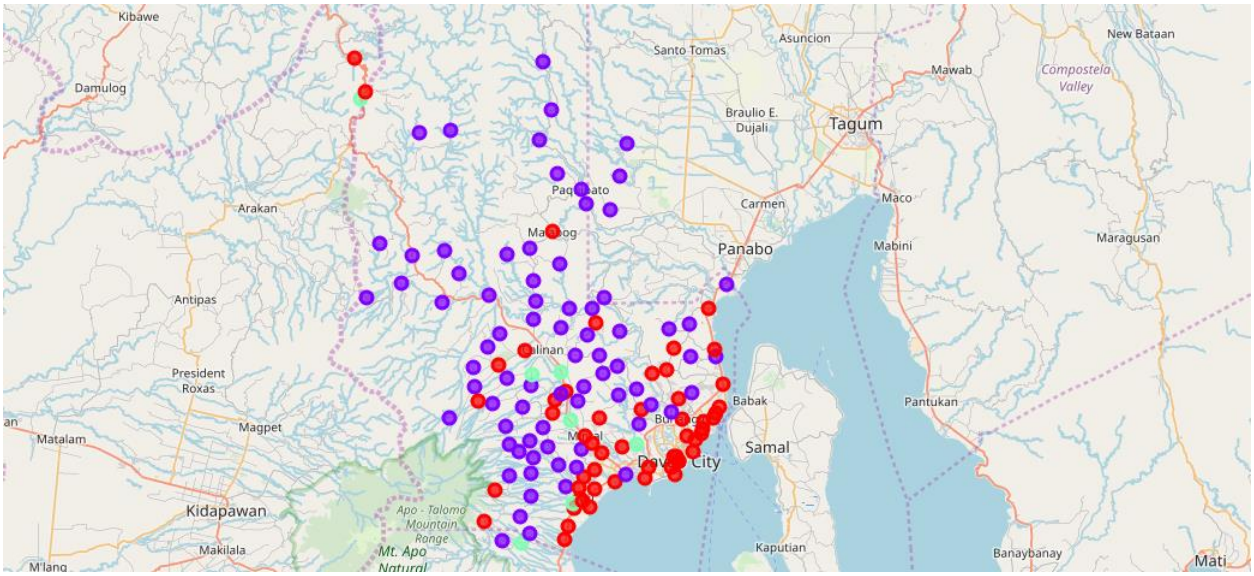
This figure shows the clustering of venues near the coast and only few can be found in mountainous areas; there are no venues in between. Those barangays in between have no recorded venues as per data from foursquare.com. Using the function ***unique()***

```
In [26]: print('There are {} unique venue categories covered in the entire Davao City.'.format(len(dvobrgys_venues['Venue Category'].unique())))
         There are 116 unique venue categories covered in the entire Davao City.
```

we that there are a total of 116 unique venue categories recorded in the entire Davao city.

4.2 k-Means cluster map

The aim of this project is to analyse the clustering of barangays in terms of the presence of absence as well as the variation of venue categories nearby. Using k-Means clustering algorithm, the following map shows the three(3) clusters formed:



In this map, barangays with a red color belong to cluster 0 while the violet and cyan belong to clusters 1 and 2, respectively.

Cluster 0 has 55 barangays and they are scattered along the coast and a few are found in mountainous areas, as shown in the map above. These barangays are:

Out[50]:

	Cluster Labels	Barangay	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	...	14th Most Common Venue	15th Most Common Venue	16th Most Common Venue	17th Most Common Venue	18th Most Common Venue	19th Most Common Venue
0	0	Acacia	Nudist Beach, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---
1	0	Agdao Proper	Gas Station, 0.25	Bakery, 0.25	Convenience Store, 0.25	Farmers Market, 0.25	---	---	---	---	...	---	---	---	---	---	---
2	0	Alfonso Angliongto Sr	Comfort Food Restaurant, 0.17	Cosmetics Shop, 0.17	Recreation Center, 0.17	Resort, 0.17	Hotel, 0.17	Farm, 0.17	---	---	...	---	---	---	---	---	---
3	0	Angalan	Department Store, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---
4	0	Bago Aplaya	Massage Studio, 0.25	Buffet, 0.25	Snack Place, 0.25	Beach, 0.25	---	---	---	---	...	---	---	---	---	---	---

5 rows × 25 columns

```
[53]: kmeans_cluster0.shape
```

Out[53]: (55, 25)

```
In [55]: np.array(kmeans_cluster0['Barangay'])

Out[55]: array(['Acacia', 'Agdao Proper', 'Alfonso Angliongto Sr', 'Angalan',
                'Bago Aplaya', 'Bago Gallera', 'Baguio Proper', 'Balengaeng',
                'Baliok', 'Binugao', 'Bucana', 'Buda', 'Buhangin Proper',
                'Bunawan Proper', 'Cabantian', 'Calinan Proper',
                'Catalunan Grande', 'Catalunan Pequeño', 'Crossing Bayabas',
                'Daliao', 'Datu Salumay', 'Dumoy', 'Eden', 'Gov. Paciano Bangoy',
                'Gov. Vicente Duterte', 'Indangan', 'Lampianao', 'Lapu-Lapu',
                'Leon Garcia', 'Lizada', 'Los Amigos', 'Lubogan', 'Ma-a',
                'Magtuod', 'Malabog', 'Malagos', 'Mandug', 'Matina Aplaya',
                'Matina Crossing', 'Mintal', 'Pampanga', 'Panacan', 'Poblacion ',
                'Rafael Castillo', 'San Antonio', 'Santo Niño', 'Sasa', 'Sibulan',
                'Sirawan', 'Tacunan', 'Tamayong', 'Tibungco', 'Toril Proper',
                'Ubalde', 'Vicente Hizon Sr.'], dtype=object)
```

Furthermore, there are 109 venue categories that are most frequent nearby.

```
There are  109 most frequent venue categories nearby barangays in cluster 0.

These venue categories are:
 ['Afghan Restaurant' 'Airport' 'Asian Restaurant' 'Athletics & Sports'
 'Australian Restaurant' 'Auto Workshop' 'BBQ Joint' 'Bakery' 'Bar'
 'Basketball Court' 'Beach' 'Beach Bar' 'Bed & Breakfast' 'Beer Garden'
 'Boat or Ferry' 'Border Crossing' 'Boutique' 'Breakfast Spot' 'Brewery'
 'Buffet' 'Burger Joint' 'Bus Station' 'Business Service' 'Cafeteria'
 'Café' 'Cajun / Creole Restaurant' 'Chinese Restaurant' 'Coffee Shop'
 'Comfort Food Restaurant' 'Construction & Landscaping'
 'Convenience Store' 'Cosmetics Shop' 'Cupcake Shop' 'Deli / Bodega'
 'Department Store' 'Dessert Shop' 'Dim Sum Restaurant' 'Diner'
 'Donut Shop' 'Electronics Store' 'Entertainment Service' 'Farm'
 'Farmers Market' 'Fast Food Restaurant' 'Filipino Restaurant'
 'Fishing Spot' 'Flea Market' 'Food Court' 'Food Service' 'Food Truck'
 'Fried Chicken Joint' 'Fruit & Vegetable Store' 'Furniture / Home Store'
 'Garden' 'Garden Center' 'Gas Station' 'Gourmet Shop' 'Gun Range' 'Gym'
 'Gym / Fitness Center' 'Harbor / Marina' 'History Museum' 'Hostel'
 'Hotel' 'Hotel Bar' 'Ice Cream Shop' 'Italian Restaurant'
 'Japanese Restaurant' 'Jewelry Store' 'Karaoke Bar' 'Korean Restaurant'
 'Lake' 'Market' 'Massage Studio' 'Mexican Restaurant'
 'Middle Eastern Restaurant' 'Mountain' 'Museum' 'Nightclub'
 'Nudist Beach' 'Outlet Store' 'Park' 'Persian Restaurant' 'Pharmacy'
 'Pizza Place' 'Pool' 'Pub' 'Recreation Center' 'Resort' 'Restaurant'
 'Road' 'Sandwich Place' 'Scenic Lookout' 'Science Museum'
 'Seafood Restaurant' 'Shopping Mall' 'Snack Place' 'Soup Place' 'Spa'
 'Sports Club' 'Steakhouse' 'Supermarket' 'Theme Park'
 'Theme Park Ride / Attraction' 'Thrift / Vintage Store' 'University'
 'Vacation Rental' 'Vegetarian / Vegan Restaurant' 'Wine Bar']
```

Notice that these categories are varied and represent how developed the barangays in this cluster are. For instance, there are simple restaurants such as fast food restaurants, but there are also restaurants with different cuisines like Afghan, Asian, Filipino, and so on. There are also accommodation venues such as hotels and bed & breakfast, as well as areas for socializing such

as bars and pubs. Physical fitness as well as educational facilities are also found nearby barangays in this cluster.

Cluster 1, on the other hand, is comprised of barangays with no recorded venues as per foursquare.com. There are 76 barangays in this cluster. The map reveals that these barangays are found between the coast and mountainous areas.

```
In [124]: kmeans_cluster1 = dvobrgys_venues_sorted.loc[dvobrgys_venues_sorted['Cluster Labels'] == 1].reset_index(drop=True)
          kmeans_cluster1.head()
```

Out[124]:

	Cluster Labels	Barangay	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	...	14th Most Common Venue	15th Most Common Venue	16th Most Common Venue	17th Most Common Venue	18th Most Common Venue	19th Most Common Venue	20th Most Common Venue
0	1	Alambre	None, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---	---
1	1	Alejandra Navarro (Lasang)	None, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---	---
2	1	Atan-Awe	None, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---	---
3	1	Bago Oshiro	None, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---	---
4	1	Bangkas Heights	None, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---	---

5 rows × 25 columns

```
In [126]: kmeans_cluster1.shape
```

Out[126]: (76, 25)

```
In [127]: np.array(kmeans_cluster1['Barangay'])
```

```
Out[127]: array(['Alambre', 'Alejandra Navarro (Lasang)', 'Atan-Awe', 'Bago Oshiro',
                 'Bangkas Heights', 'Bantol', 'Bato', 'Bayabas', 'Biao Escuela',
                 'Biao Guianga', 'Biao Joaquin', 'Cadalian', 'Callawa', 'Camansi',
                 'Carmen', 'Catigan', 'Cawayan', 'Colosas', 'Communal', 'Dacudao',
                 'Dalag', 'Dalagdag', 'Daliaon Plantation', 'Dominga',
                 'Fatima (Benowang)', 'Gatungan', 'Gumalang', 'Gumitan', 'Ilang',
                 'Inayangan', 'Kilate', 'Lacson', 'Lamanan', 'Langub', 'Lumiad',
                 'Mabuhay', 'Magsaysay', 'Mahayag', 'Malamba', 'Manambulan',
                 'Manuel Guianga', 'Mapula', 'Marilog Proper', 'Matina Biao',
                 'Megkawayan', 'Mudiang', 'Mulig', 'New Carmen', 'New Valencia',
                 'Panalum', 'Pandaitan', 'Pangyan', 'Paquibato Proper',
                 'Paradise Embak', 'Salapawan', 'Salaysay', 'Saloy', 'Sirib',
                 'Suawan (Tuli)', 'Subasta', 'Sumimao', 'Tagakpan', 'Tagluno',
                 'Tagurano', 'Talandang', 'Talomo Proper', 'Talomo River',
                 'Tamugan', 'Tapak', 'Tawan-Tawan', 'Tibuloy', 'Tigatto',
                 'Tungkalan', 'Ula', 'Waan', 'Wines'], dtype=object)
```

```
In [139]: print('There is ',len(np.unique(vc_list))-1,'most frequent venue category nearby barangays in cluster 1. \n')
          print('These venue categories are: \n',np.unique(vc_list)[1:])
```

```
There is  1 most frequent venue category nearby barangays in cluster 1.

These venue categories are:
 ['None']
```

Finally, cluster 2 is composed of 7 barangays with 6 most frequent venue categories:

```
In [132]: kmeans_cluster2 = dvobrgys_venues_sorted.loc[dvobrgys_venues_sorted['Cluster Labels'] == 2].reset_index(drop=True)
          kmeans_cluster2.head()
```

Out[132]:

	Cluster Labels	Barangay	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	...	14th Most Common Venue	15th Most Common Venue	16th Most Common Venue	17th Most Common Venue	18th Most Common Venue	19th Most Common Venue	20 Mc Comm Ven
0	2	Baganihan	Farm, 0.5	Café, 0.5	---	---	---	---	---	---	...	---	---	---	---	---	---	
1	2	Baracatan	Farm, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---	
2	2	Marapangi	Farm, 0.5	Fast Food Restaurant, 0.5	---	---	---	---	---	---	...	---	---	---	---	---	---	
3	2	Matina Pangi	River, 0.5	Farm, 0.5	---	---	---	---	---	---	...	---	---	---	---	---	---	
4	2	Riverside	Farm, 0.5	Grocery Store, 0.5	---	---	---	---	---	---	...	---	---	---	---	---	---	

5 rows × 25 columns

```
In [133]: kmeans_cluster2.shape
```

Out[133]: (7, 25)

```
In [134]: np.array(kmeans_cluster2['Barangay'])
```

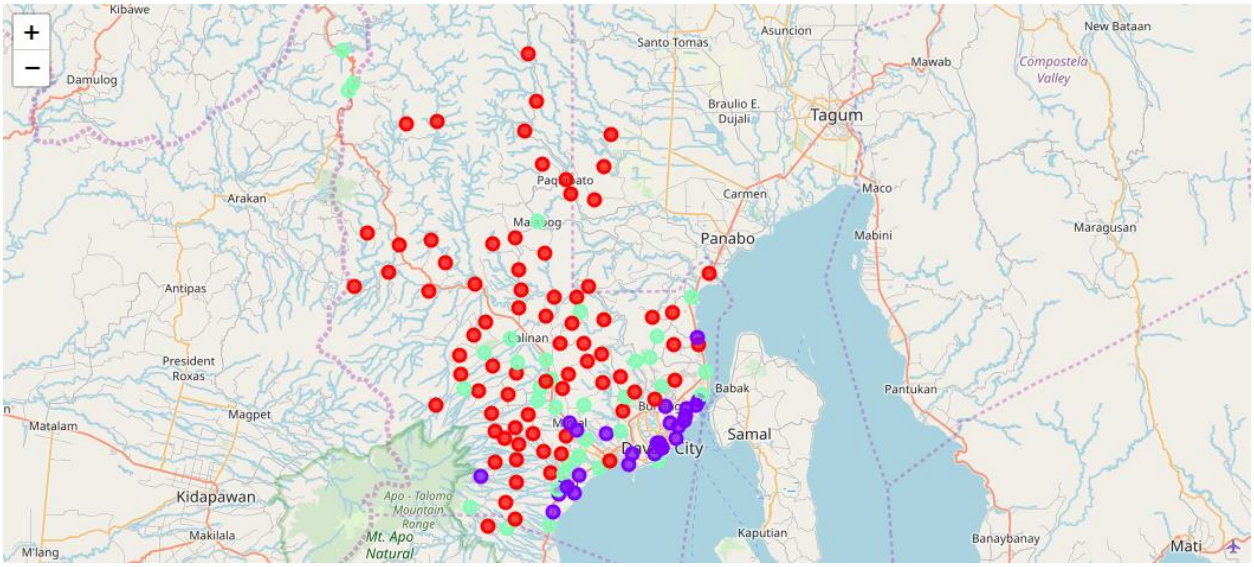
```
Out[134]: array(['Baganihan', 'Baracatan', 'Marapangi', 'Matina Pangi', 'Riverside',
                 'Tugbok Proper', 'Wangan'], dtype=object)
```

```
There are  6 most frequent venue categories nearby barangays in cluster 2.

These venue categories are:
 ['Bakery' 'Café' 'Farm' 'Fast Food Restaurant' 'Grocery Store' 'River']
```

Note how practical the venue categories are. The map shows that these barangays are located between the coast and mountainous areas, interspersed with barangays that have no venue records and some are interspersed with cluster 0 barangays.

4.3 DBSCAN

The following figure shows the grouping of barangays after applying DBSCAN. Barangays colored in cyan belong to cluster -1 which are outliers, while those in red and violet belong to clusters 0 and 1, respectively.



The barangays that are considered outliers belong to cluster -1, and there are 34 barangays in this cluster. The variability of venue categories in this cluster is at most 4 venue categories per barangay.

```
In [158]: dbscan_cluster_outlier = dvobrgys_venues_sorted_2.loc[dvobrgys_venues_sorted_2['Cluster Labels'] == -1].reset_index(drop=True)
          dbscan_cluster_outlier.head()
```

Out[158]:

	Cluster Labels	Barangay	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	...	14th Most Common Venue	15th Most Common Venue	16th Most Common Venue	17th Most Common Venue	18th Most Common Venue	19th Most Common Venue
0	-1	Acacia	Nudist Beach, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---
1	-1	Angalan	Department Store, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---
2	-1	Baganihan	Farm, 0.5	Café, 0.5	---	---	---	---	---	---	...	---	---	---	---	---	---
3	-1	Bago Aplaya	Massage Studio, 0.25	Buffet, 0.25	Snack Place, 0.25	Beach, 0.25	---	---	---	---	...	---	---	---	---	---	---
4	-1	Bago Gallera	Pool, 0.67	Construction & Landscaping, 0.33	---	---	---	---	---	---	...	---	---	---	---	---	---

```
In [159]: dbscan_cluster_outlier.shape
```

Out[159]: (34, 25)

```
In [160]: np.array(dbscan_cluster_outlier['Barangay'])
```

Out[160]:
```
array(['Acacia', 'Angalan', 'Baganihan', 'Bago Aplaya', 'Bago Gallera',
       'Baguio Proper', 'Balengaeng', 'Baliok', 'Baracatan', 'Binugao',
       'Bucana', 'Buda', 'Bunawan Proper', 'Cabantian',
       'Catalunan Pequeño', 'Datu Salumay', 'Indangan', 'Lampianao',
       'Los Amigos', 'Lubogan', 'Magtuod', 'Malabog', 'Malagos', 'Mandug',
       'Marapangi', 'Matina Pangi', 'Panacan', 'Riverside', 'Sasa',
       'Sibulan', 'Tacunan', 'Tamayong', 'Tugbok Proper', 'Wangan'],
      dtype=object)
```

```
There are  49 most frequent venue categories nearby barangays in cluster outlier.

These venue categories are:
 ['Afghan Restaurant' 'Airport' 'Auto Workshop' 'BBQ Joint' 'Bakery'
 'Beach' 'Beach Bar' 'Border Crossing' 'Buffet' 'Business Service'
 'Cafeteria' 'Café' 'Coffee Shop' 'Comfort Food Restaurant'
 'Construction & Landscaping' 'Convenience Store' 'Department Store'
 'Diner' 'Donut Shop' 'Electronics Store' 'Farm' 'Fast Food Restaurant'
 'Food Service' 'Fried Chicken Joint' 'Garden' 'Garden Center'
 'Grocery Store' 'Hotel Bar' 'Italian Restaurant' 'Jewelry Store' 'Lake'
 'Market' 'Massage Studio' 'Mountain' 'Museum' 'Nudist Beach'
 'Outlet Store' 'Pizza Place' 'Pool' 'Recreation Center' 'Restaurant'
 'River' 'Sandwich Place' 'Scenic Lookout' 'Shopping Mall' 'Snack Place'
 'Spa' 'Thrift / Vintage Store' 'Vacation Rental']
```

Cluster 1 represents the group of barangays with no recorded venues, and there are 76 barangays in this cluster, similar to cluster 1 of the k-means clustering algorithm.

```
In [164]: dbscan_cluster0 = dvobrgys_venues_sorted_2.loc[dvobrgys_venues_sorted_2['Cluster Labels'] == 0].reset_index(drop=True)
          dbscan_cluster0.head()
```

Out[164]:

	Cluster Labels	Barangay	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	...	14th Most Common Venue	15th Most Common Venue	16th Most Common Venue	17th Most Common Venue	18th Most Common Venue	19th Most Common Venue	20th Most Common Venue
0	0	Alambre	None, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---	---
1	0	Alejandra Navarro (Lasang)	None, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---	---
2	0	Atan-Awe	None, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---	---
3	0	Bago Oshiro	None, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---	---
4	0	Bangkas Heights	None, 1.0	---	---	---	---	---	---	---	...	---	---	---	---	---	---	---

5 rows × 25 columns

```
In [165]: dbscan_cluster0.shape
```

Out[165]: (76, 25)

```
In [166]: np.array(dbscan_cluster0['Barangay'])
```

```
Out[166]: array(['Alambre', 'Alejandra Navarro (Lasang)', 'Atan-Awe', 'Bago Oshiro',
                 'Bangkas Heights', 'Bantol', 'Bato', 'Bayabas', 'Biao Escuela',
                 'Biao Guianga', 'Biao Joaquin', 'Cadalian', 'Callawa', 'Camansi',
                 'Carmen', 'Catigan', 'Cawayan', 'Colosas', 'Communal', 'Dacudao',
                 'Dalag', 'Dalagdag', 'Daliaon Plantation', 'Dominga',
                 'Fatima (Benowang)', 'Gatungan', 'Gumalang', 'Gumitan', 'Ilang',
                 'Inayangan', 'Kilate', 'Lacson', 'Lamanan', 'Langub', 'Lumiad',
                 'Mabuhay', 'Magsaysay', 'Mahayag', 'Malamba', 'Manambulan',
                 'Manuel Guianga', 'Mapula', 'Marilog Proper', 'Matina Biao',
                 'Megkawayan', 'Mudiang', 'Mulig', 'New Carmen', 'New Valencia',
                 'Panalum', 'Pandaitan', 'Pangyan', 'Paquibato Proper',
                 'Paradise Embak', 'Salapawan', 'Salaysay', 'Saloy', 'Sirib',
                 'Suawan (Tuli)', 'Subasta', 'Sumimao', 'Tagakpan', 'Tagluno',
                 'Tagurano', 'Talandang', 'Talomo Proper', 'Talomo River',
                 'Tamugan', 'Tapak', 'Tawan-Tawan', 'Tibuloy', 'Tigatto',
                 'Tungkalan', 'Ula', 'Waan', 'Wines'], dtype=object)
```

```
There are  1 most frequent venue categories nearby barangays in cluster 0.

These venue categories are:
 ['None']
```

Lastly, cluster 1 represents barangays with very varied venue categories. In particular, there are at least 3 venue categories per barangay. This cluster is composed of 28 barangays, most of which are along the coast.

```
In [169]: dbscan_cluster1 = dvobrgys_venues_sorted_2.loc[dvobrgys_venues_sorted_2['Cluster Labels'] == 1].reset_index(drop=True)
          dbscan_cluster1.head()
```

Out[169]:

	Cluster Labels	Barangay	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	...	14th Most Common Venue	15th Most Common Venue	16th Most Common Venue	17th Most Common Venue	18th Most Common Venue
0	1	Agdao Proper	Gas Station, 0.25	Bakery, 0.25	Convenience Store, 0.25	Farmers Market, 0.25	---	---	---	---	...	---	---	---	---	---
1	1	Alfonso Angliongto Sr	Comfort Food Restaurant, 0.17	Cosmetics Shop, 0.17	Recreation Center, 0.17	Resort, 0.17	Hotel, 0.17	Farm, 0.17	---	---	...	---	---	---	---	---
2	1	Buhangin Proper	Chinese Restaurant, 0.1	BBQ Joint, 0.1	Café, 0.1	Market, 0.1	Burger Joint, 0.1	Donut Shop, 0.1	Bed & Breakfast, 0.1	Fast Food Restaurant, 0.1	...	---	---	---	---	---
3	1	Calinan Proper	Asian Restaurant, 0.12	Convenience Store, 0.09	Fast Food Restaurant, 0.09	Café, 0.09	Bar, 0.06	Pharmacy, 0.06	Athletics & Sports, 0.03	Donut Shop, 0.03	...	University, 0.03	Filipino Restaurant, 0.03	Australian Restaurant, 0.03	Farmers Market, 0.03	Diner, 0.03

```
In [170]: dbscan_cluster1.shape
Out[170]: (28, 25)
```

```
In [171]: np.array(dbscan_cluster1['Barangay'])
Out[171]: array(['Agdao Proper', 'Alfonso Angliongto Sr', 'Buhangin Proper',
                 'Calinan Proper', 'Catalunan Grande', 'Crossing Bayabas', 'Daliao',
                 'Dumoy', 'Eden', 'Gov. Paciano Bangoy', 'Gov. Vicente Duterte',
                 'Lapu-Lapu', 'Leon Garcia', 'Lizada', 'Ma-a', 'Matina Aplaya',
                 'Matina Crossing', 'Mintal', 'Pampanga', 'Poblacion ',
                 'Rafael Castillo', 'San Antonio', 'Santo Niño', 'Sirawan',
                 'Tibungco', 'Toril Proper', 'Ubalde', 'Vicente Hizon Sr.'],
                dtype=object)
```

```
There are  88 most frequent venue categories nearby barangays in cluster 1.

These venue categories are:
 ['Asian Restaurant' 'Athletics & Sports' 'Australian Restaurant'
 'BBQ Joint' 'Bakery' 'Bar' 'Basketball Court' 'Bed & Breakfast'
 'Beer Garden' 'Boat or Ferry' 'Boutique' 'Breakfast Spot' 'Brewery'
 'Buffet' 'Burger Joint' 'Bus Station' 'Cafeteria' 'Café'
 'Cajun / Creole Restaurant' 'Chinese Restaurant' 'Coffee Shop'
 'Comfort Food Restaurant' 'Construction & Landscaping'
 'Convenience Store' 'Cosmetics Shop' 'Cupcake Shop' 'Deli / Bodega'
 'Department Store' 'Dessert Shop' 'Dim Sum Restaurant' 'Diner'
 'Donut Shop' 'Entertainment Service' 'Farm' 'Farmers Market'
 'Fast Food Restaurant' 'Filipino Restaurant' 'Fishing Spot' 'Flea Market'
 'Food Court' 'Food Service' 'Food Truck' 'Fried Chicken Joint'
 'Fruit & Vegetable Store' 'Furniture / Home Store' 'Gas Station'
 'Gourmet Shop' 'Gun Range' 'Gym' 'Gym / Fitness Center' 'Harbor / Marina'
 'History Museum' 'Hostel' 'Hotel' 'Ice Cream Shop' 'Italian Restaurant'
 'Japanese Restaurant' 'Karaoke Bar' 'Korean Restaurant' 'Market'
 'Mexican Restaurant' 'Middle Eastern Restaurant' 'Nightclub' 'Park'
 'Persian Restaurant' 'Pharmacy' 'Pizza Place' 'Pool' 'Pub'
 'Recreation Center' 'Resort' 'Restaurant' 'Road' 'Sandwich Place'
 'Science Museum' 'Seafood Restaurant' 'Shopping Mall' 'Snack Place'
 'Soup Place' 'Spa' 'Sports Club' 'Steakhouse' 'Supermarket' 'Theme Park'
 'Theme Park Ride / Attraction' 'University'
 'Vegetarian / Vegan Restaurant' 'Wine Bar']
```

5. Discussion

The following table summarizes the results of the two clustering algorithms implemented:

Cluster	k-Means		DBSCAN	
	Barangays	Venue Categories	Barangays	Venue Categories
1	Alambre, Alejandra Navarro (Lasang), Atan-Awe, Bago Oshiro, Bangkas Heights, Bantol, Bato, Bayabas, Biao Escuela, Biao Guianga, Biao Joaquin, Cadalian, Callawa, Camansi, Carmen, Catigan, Cawayan, Colosas, Communal, Dacudao, Dalag, Dalagdag,	None	Alambre, Alejandra Navarro (Lasang), Atan-Awe, Bago Oshiro, Bangkas Heights, Bantol, Bato, Bayabas, Biao Escuela, Biao Guianga, Biao Joaquin, Cadalian, Callawa, Camansi, Carmen, Catigan, Cawayan, Colosas, Communal, Dacudao, Dalag, Dalagdag,	None

	Daliaon Plantation, Dominga, Fatima (Benowang), Gatungan, Gumalang, Gumitan, Ilang, Inayangan, Kilate, Lacson, Lamanan, Langub, Lumiad, Mabuhay, Magsaysay, Mahayag, Malamba, Manambulan, Manuel Guianga, Mapula, Marilog Proper, Matina Biao, Megkawayan, Mudiang, Mulig, New Carmen, New Valencia, Panalum, Pandaitan, Pangyan, Paguibato Proper, Paradise Embak, Salapawan, Salaysay, Saloy, Sirib, Suawan (Tuli), Subasta, Sumimao, Tagakpan, Tagluno, Tagurano, Talandang, Talomo Proper, Talomo River, Tamugan, Tapak, Tawan-Tawan, Tibuloy, Tigatto, Tungkalan, Ula, Waan, Wines		Daliaon Plantation, Dominga, Fatima (Benowang), Gatungan, Gumalang, Gumitan, Ilang, Inayangan, Kilate, Lacson, Lamanan, Langub, Lumiad, Mabuhay, Magsaysay, Mahayag, Malamba, Manambulan, Manuel Guianga, Mapula, Marilog Proper, Matina Biao, Megkawayan, Mudiang, Mulig, New Carmen, New Valencia, Panalum, Pandaitan, Pangyan, Paguibato Proper, Paradise Embak, Salapawan, Salaysay, Saloy, Sirib, Suawan (Tuli), Subasta, Sumimao, Tagakpan, Tagluno, Tagurano, Talandang, Talomo Proper, Talomo River, Tamugan, Tapak, Tawan-Tawan, Tibuloy, Tigatto, Tungkalan, Ula, Waan, Wines	
2	Baganihan, Baracatan, Marapangi, Matina Pangi, Riverside, Tugbok Proper, Wangan	Bakery, Café, Farm, Fast Food Restaurant, Grocery Store, River	Acacia, Angalan, Baganihan, Bago Aplaya, Bago Gallera, Baguio Proper, Balengaeng, Baliok, Baracatan, Binugao, Bucana, Buda, Bunawan Proper, Cabantian, Catalunan Pequeño, Datu Salumay, Indangan, Lampianao, Los	Afghan Restaurant, Airport, Auto Workshop, BBQ Joint, Bakery, Beach, Beach Bar, Border Crossing, Buffet, Business Service, Cafeteria, Café, Coffee Shop, Comfort Food Restaurant, Construction & Landscaping,

			Amigos, Lubogan, Magtuod, Malabog, Malagos, Mandug, Marapangi, Matina Pangi, Panacan, Riverside, Sasa, Sibulan, Tacunan, Tamayong, Tugbok Proper, Wangan	Convenience Store, Department Store, Diner, Donut Shop, Electronics Store, Farm, Fast Food Restaurant, Food Service, Fried Chicken Joint, Garden, Garden Center, Grocery Store, Hotel Bar, Italian Restaurant, Jewelry Store, Lake, Market, Massage Studio, Mountain, Museum, Nudist Beach, Outlet Store, Pizza Place, Pool, Recreation Center, Restaurant, River, Sandwich Place, Scenic Lookout, Shopping Mall, Snack Place, Spa, Thrift/Vintage Store, Vacation Rental
3	Acacia, Agdao Proper, Alfonso Angliongto Sr., Angalan, Bago Aplaya, Bago Gallera, Baguio Proper, Balengaeng, Baliok, Binugao, Bucana, Buda, Buhangin Proper, Bunawan Proper, Cabantian, Calinan Proper, Catalunan Grande, Catalunan Pequeño, Crossing Bayabas, Daliao, Datu Salumay, Dumoy, Eden, Gov. Paciano Bangoy, Gov. Vicente Duterte, Indangan,	Afghan Restaurant, Airport, Asian Restaurant, Athletics & Sports, Australian Restaurant, Auto Workshop, BBQ Joint, Bakery, Bar, Basketball Court, Beach, Beach Bar, Bed & Breakfast, Beer Garden, Boat or Ferry, Border Crossing, Boutique, Breakfast Spot, Brewery, Buffet, Burger Joint, Bus Station, Business Service, Cafeteria, Café, Cajun/Creole Restaurant, Chinese	Agdao Proper, Alfonso Angliongto Sr., Buhangin Proper, Calinan Proper, Catalunan Grande, Crossing Bayabas, Daliao, Dumoy, Eden, Gov. Paciano Bangoy, Gov. Vicente Duterte, Lapu-Lapu, Leon Garcia, Lizada, Ma-a, Matina Aplaya, Matina Crossing, Mintal, Pampanga, Poblacion, Rafael Castillo, San Antonio, Santo Niño, Sirawan, Tibungco, Toril Proper, Ubalde, Vicente Hizon Sr.	Asian Restaurant, Athletics & Sports, Australian Restaurant, BBQ Joint, Bakery, Bar, Basketball Court, Bed & Breakfast, Beer Garden, Boat or Ferry, Boutique, Breakfast Spot, Brewery, Buffet, Burger Joint, Bus Station, Cafeteria, Café, Cajun/Creole Restaurant, Chinese Restaurant, Coffee Shop, Comfort Food Restaurant, Construction & Landscaping, Convenience Store,

	Lampianao, Lapu-Lapu, Leon Garcia, Lizada, Los Amigos, Lubogan, Ma-a, Magtuod, Malabog, Malagos, Mandug, Matina Aplaya, Matina Crossing, Mintal, Pampanga, Panacan, Poblacion, Rafael Castillo, San Antonio, Santo Niño, Sasa, Sibulan, Sirawan, Tacunan, Tamayong, Tibungco, Toril Proper, Ubalde, Vicente Hizon Sr.	Restaurant, Coffee Shop, Comfort Food Restaurant, Construction & Landscaping, Convenience Store, Cosmetics Shop, Cupcake Shop, Deli/Bodega, Department Store, Dessert Shop, Dim Sum Restaurant, Diner, Donut Shop, Electronics Store, Entertainment Service, Farm, Farmers Market, Fast Food Restaurant, Filipino Restaurant, Fishing Spot, Flea Market, Food Court, Food Service, Food Truck, Fried Chicken Joint, Fruit & Vegetable Store, Furniture/Home Store, Garden, Garden Center, Gas Station, Gourmet Shop, Gun Range, Gym, Gym/Fitness Center, Harbor/Marina, History Museum, Hostel, Hotel, Hotel Bar, Ice Cream Shop, Italian Restaurant, Japanese Restaurant, Jewelry Store, Karaoke Bar, Korean Restaurant, Lake, Market, Massage Studio, Mexican Restaurant, Middle Eastern Restaurant, Mountain, Museum, Nightclub, Nudist		Cosmetics Shop, Cupcake Shop, Deli/Bodega, Department Store, Dessert Shop, Dim Sum Restaurant, Diner, Donut Shop, Entertainment Service, Farm, Farmers Market, Fast Food Restaurant, Filipino Restaurant, Fishing Spot, Flea Market, Food Court, Food Service, Food Truck, Fried Chicken Joint, Fruit & Vegetable Store, Furniture/Home Store, Gas Station, Gourmet Shop, Gun Range, Gym, Gym/Fitness Center, Harbor/Marina, History Museum, Hostel, Hotel, Ice Cream Shop, Italian Restaurant, Japanese Restaurant, Karaoke Bar, Korean Restaurant, Market, Mexican Restaurant, Middle Eastern Restaurant, Nightclub, Park, Persian Restaurant, Pharmacy, Pizza Place, Pool, Pub, Recreation Center, Resort, Restaurant, Road, Sandwich Place, Science Museum, Seafood Restaurant, Shopping Mall, Snack Place, Soup Place, Spa, Sports Club,

		Beach, Outlet Store, Park, Persian Restaurant, Pharmacy, Pizza Place, Pool, Pub, Recreation Center, Resort, Restaurant, Road, Sandwich Place, Scenic Lookout, Science Museum, Seafood Restaurant, Shopping Mall, Snack Place, Soup Place, Spa, Sports Club, Steakhouse, Supermarket, Theme Park, Theme Park Ride/Attraction, Thrift/Vintage Store, University, Vacation Rental, Vegetarian/Vegan Restaurant, Wine Bar		Steakhouse, Supermarket, Theme Park, Theme Park Ride/Attraction, University, Vegetarian/Vegan Restaurant, Wine Bar

Between the two algorithms, DBSCAN showed a more realistic clustering in the sense that those barangays located in the mountainous areas are clustered as outliers. There is less variation in the venue categories in these barangays, which can imply that those areas are either remote or not yet as developed as those in cluster 3. On the contrary, those barangays near the coast, belonging to cluster 3, have more similarities in terms of nearby venues, and the categories are more varied, indicating more developed areas.

Businessmen can take advantage of the results of this project by identifying areas where they can potentially put up their business, as well as identifying the kind of business they can put up in that location. For those barangays belonging to the cluster where there are no recorded venues yet, businessmen can locate nearby barangays belonging either in the outlier cluster or in the 3rd cluster and study which venue categories are most likely to gain much profit, either because it has been patronized in nearby barangays or there is no venue similar to that in the chosen barangay as well as in the surrounding barangays. For the outlier cluster, there is still room for more venues to be established, and businessmen can analyse how to enter the community since they already know which barangays to study through this project. Finally, businessmen can also think of ways to enrich barangays in the 3rd cluster by looking into barangays within the cluster that are more varied than the others and potentially mimic the variation to barangays with less varied venue categories.

6. Conclusion

The DBSCAN Clustering results returned a more realistic picture of the similarities and differences among barangays in Davao City. The first cluster includes barangays whose features composed of the venue categories nearby and the geospatial location are less varied and are far from the downtown and coastal areas. The second cluster is composed of barangays with no recorded venues yet as per data from foursquare.com. Finally, the third cluster consists of barangays that have varied venue categories and are located in the downtown and coastal areas. The results of this project will be very beneficial to entrepreneurs who would like to find out strategic locations for their potential business, as well as the kind of business that will be profitable in those areas.