

Final_Capstone_Project_Week_5

February 9, 2021

1 Opening a Restaurant in Colombo

1.1 Introduction / Business Problem

Colombo is the commercial capital and largest city of Sri Lanka by population. According to the Brookings Institution, Colombo metropolitan area has a population of 5.6 million, and 752,993 in the city proper. It is the financial center of the island and a tourist destination. Colombo is arguably Sri Lanka's most popular tourist destination. The coastal city is a beautiful place to spend a couple of days just relaxing, indulging in the local food, and leisure activities. In Colombo, most common restaurants are aimed at tourists, with a mix of Sri Lankan, seafood and Western dishes and South Indian-style places.

Our stakeholder is willing to open a restaurant in Colombo. Choosing a restaurant type and a good spot, relies on common sense and domain knowledge. Data science provides valuable insights regarding market trends and consumer lifestyles so that stakeholder can better address and meet public demand. In this project, we will implement the basic analysis and try to find the most optimal sub-urbs to open the restaurant according to those criteria. It's obvious, that there are many additional factors, such as distance from parking places or distance from the main streets, but this analysis mainly focuses on the type of restaurants

1.2 Data

The problem will be addressed in following steps:

Step 1. Using a table on '<https://en.wikipedia.org/wiki/Colombo#:~:text=Colombo%20>, collect information about Colombo suburbs.

Step 2. Use the Geopy and Folium library to get the coordinates of every locations and map geospatial data on a Colombo map.

Step 3. Using Foursquare API, collect the top 100 restaurants and their categories for each location within a radius 1000 meters.

Step 4. Group collected restaurants by location and by taking the mean of the frequency of occurrence of each type, preparing them for clustering.

Step 5. Cluster restaurants by k-means algorithm and analyze the top 5 most common restaurants in each cluster.

Step 6. Visualize clusters on the map, thus showing the best locations for opening the chosen restaurant.

1.3 Methodology

Before we get the data and start exploring it, let's import the packages that we will need.

Folium installed
Libraries imported.

1.3.1 Steps for data cleaning

The data was converted in to a pandas data frame.

```
<table class="wikitable">
<tbody><tr>
<td style="width:100px;">Postal number</td>
<td>City suburb
</td></tr>
<tr>
<td>Colombo 1</td>
<td><a href="/wiki/Fort_(Colombo)" title="Fort (Colombo)">Fort</a>
</td></tr>
<tr>
<td>Colombo 2</td>
<td><a href="/wiki/Slave_Island" title="Slave Island">Slave Island</a>, <a
href="/wiki/Union_Place" title="Union Place">Union Place</a>
</td></tr>
<tr>
<td>Colombo 3</td>
<td><a href="/wiki/Kollupitiya" title="Kollupitiya">Kollupitiya</a>
</td></tr>
<tr>
<td>Colombo 4</td>
<td><a href="/wiki/Bambalapitiya" title="Bambalapitiya">Bambalapitiya</a>
</td></tr>
<tr>
<td>Colombo 5</td>
<td><a href="/wiki/Havelock_Town" title="Havelock Town">Havelock Town</a>, <a
href="/wiki/Kirulapana" title="Kirulapana">Kirulapone</a>, <a
href="/wiki/Kirilapone_North" title="Kirilapone North">Kirulapone North</a>
</td></tr>
<tr>
<td>Colombo 6</td>
<td><a href="/wiki/Wellawatte" title="Wellawatte">Wellawatte</a>, <a
href="/wiki/Pamankada" title="Pamankada">Pamankada</a>,
</td></tr>
<tr>
<td>Colombo 7</td>
<td><a href="/wiki/Cinnamon_Gardens" title="Cinnamon Gardens">Cinnamon
Gardens</a>
</td></tr>
```

```

<tr>
<td>Colombo 8</td>
<td><a href="/wiki/Borella" title="Borella">Borella</a>
</td></tr>
<tr>
<td>Colombo 9</td>
<td><a href="/wiki/Dematagoda" title="Dematagoda">Dematagoda</a>
</td></tr>
<tr>
<td>Colombo 10</td>
<td><a href="/wiki/Maradana" title="Maradana">Maradana</a>, <a
href="/wiki/Panchikawatte" title="Panchikawatte">Panchikawatte</a>
</td></tr>
<tr>
<td>Colombo 11</td>
<td><a href="/wiki/Pettah,_Sri_Lanka" title="Pettah, Sri Lanka">Pettah</a>
</td></tr>
<tr>
<td>Colombo 12</td>
<td><a href="/wiki/Hulftsdorp" title="Hulftsdorp">Hulftsdorp</a>
</td></tr>
<tr>
<td>Colombo 13</td>
<td><a href="/wiki/Kotahena" title="Kotahena">Kotahena</a>, <a
href="/wiki/Bloemendhal" title="Bloemendhal">Bloemendhal</a>
</td></tr>
<tr>
<td>Colombo 14</td>
<td><a href="/wiki/Grandpass" title="Grandpass">Grandpass</a>
</td></tr>
<tr>
<td>Colombo 15</td>
<td><a href="/wiki/Modara" title="Modara">Modara/Mutwal</a>, <a
href="/wiki/Mattakkuliya" title="Mattakkuliya">Mattakkuliya</a>, <a
href="/wiki/Madampitiya" title="Madampitiya">Madampitiya</a>
</td></tr></tbody></table>

```

[123]:

	Postal number	City suburb
0	Postal number	City suburb
1	Colombo 1	Fort
2	Colombo 2	Slave Island, Union Place
3	Colombo 3	Kollupitiya
4	Colombo 4	Bambalapitiya

[124]:

	Postal number	City suburb
0	Colombo 1	Fort
1	Colombo 2	Slave Island, Union Place
2	Colombo 3	Kollupitiya

```

3      Colombo 4      Bambalapitiya
4      Colombo 5  Havelock Town, Kirulapone, Kirulapone North

```

1.3.2 Adding Longitude and Latitudes to the dataframe

The geograpical coordinate of Colombo City are 6.9349969, 79.8538463.

```

[131]:   Postal number      Neighborhood  Latitude \
0      Colombo 1      Fort 6.941480
1      Colombo 2      Slave Island, Union Place 6.922123
2      Colombo 3      Kollupitiya 6.912930
3      Colombo 4      Bambalapitiya 6.904710
4      Colombo 5  Havelock Town, Kirulapone, Kirulapone North 6.881350

      Longitude
0  79.846640
1  79.854221
2  79.851280
3  79.855070
4  79.875920

```

```

[138]: <folium.folium.Map at 0x1a2bb980910>

```

1.3.3 Let's explore a neighborhood in our dataframe

```

[139]: 'Fort'

```

Latitude and longitude values of suburb Fort are 6.941480000000007, 79.846640000000004.

1.3.4 Now, let's get the top 100 venues that are within a radius of 1000 meters.

```

[144]: {'meta': {'code': 200, 'requestId': '602221e05a01592844f42da5'},
      'response': {'headerLocation': 'Colombo',
                    'headerFullLocation': 'Colombo',
                    'headerLocationGranularity': 'city',
                    'query': 'restaurant',
                    'totalResults': 4,
                    'suggestedBounds': {'ne': {'lat': 6.9813000090000084,
                                                'lng': 79.88437013108789},
                                         'sw': {'lat': 6.9632999910000066, 'lng': 79.8662698689122}},
                    'groups': [{'type': 'Recommended Places',
                                'name': 'recommended',
                                'items': [{'reasons': {'count': 0,
                                                         'items': [{'summary': 'This spot is popular',
                                                                      'type': 'general',
                                                                      'reasonName': 'globalInteractionReason'}]}],
                                'venue': {'id': '4de24f537d8b2547eb002c07',

```

```

'name': 'Good Food',
'location': {'lat': 6.9715675775323245,
'lng': 79.87387934943084,
'labeledLatLngs': [{'label': 'display',
'lat': 6.9715675775323245,
'lng': 79.87387934943084}]},
'distance': 178,
'postalCode': '00150',
'cc': 'LK',
'city': ' ',
'state': 'Western Province',
'country': ' \u200d ',
'formattedAddress': [' 00150', ' \u200d ']],
'categories': [{'id': '4bf58dd8d48988d147941735',
'name': 'Diner',
'pluralName': 'Diners',
'shortName': 'Diner',
'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/food/diner_',
'suffix': '.png'},
'primary': True}],
'photos': {'count': 0, 'groups': []}},
'referralId': 'e-0-4de24f537d8b2547eb002c07-0'},
'reasons': {'count': 0,
'items': [{'summary': 'This spot is popular',
'type': 'general',
'reasonName': 'globalInteractionReason'}]},
'venue': {'id': '53482280498eae1591efc45d',
'name': 'Twoso',
'location': {'lat': 6.971693653580593,
'lng': 79.8754706801177,
'labeledLatLngs': [{'label': 'display',
'lat': 6.971693653580593,
'lng': 79.8754706801177}]},
'distance': 69,
'cc': 'LK',
'country': ' \u200d ',
'formattedAddress': [' \u200d ']],
'categories': [{'id': '52e81612bc57f1066b79ff',
'name': 'Halal Restaurant',
'pluralName': 'Halal Restaurants',
'shortName': 'Halal',
'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/food/halal_',
'suffix': '.png'},
'primary': True}],
'photos': {'count': 0, 'groups': []}},

```

```

    'referralId': 'e-0-53482280498eae1591efc45d-1'},
    {'reasons': {'count': 0,
        'items': [{'summary': 'This spot is popular',
            'type': 'general',
            'reasonName': 'globalInteractionReason'}]},
        'venue': {'id': '4fd75072af248dd23853ee1a',
            'name': 'Perera & Sons',
            'location': {'address': '29A Wistwike Road,',
                'lat': 6.97202703162331,
                'lng': 79.87346373942931,
                'labeledLatLngs': [{'label': 'display',
                    'lat': 6.97202703162331,
                    'lng': 79.87346373942931}]},
                'distance': 207,
                'cc': 'LK',
                'city': 'Colombo 15',
                'country': ' \u200d ',
                'formattedAddress': ['29A Wistwike Road,',
                    'Colombo 15',
                    ' \u200d ']},
                'categories': [{'id': '4bf58dd8d48988d16a941735',
                    'name': 'Bakery',
                    'pluralName': 'Bakeries',
                    'shortName': 'Bakery',
                    'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/food/bakery_',
                        'suffix': '.png'},
                    'primary': True}],
                'photos': {'count': 0, 'groups': []}},
            'referralId': 'e-0-4fd75072af248dd23853ee1a-2'},
    {'reasons': {'count': 0,
        'items': [{'summary': 'This spot is popular',
            'type': 'general',
            'reasonName': 'globalInteractionReason'}]},
        'venue': {'id': '52ed1ad2498e220cbe5b7575',
            'name': 'dine more',
            'location': {'address': 'Thurstan Road',
                'lat': 6.972405990423958,
                'lng': 79.86692856115866,
                'labeledLatLngs': [{'label': 'display',
                    'lat': 6.972405990423958,
                    'lng': 79.86692856115866}]},
                'distance': 927,
                'cc': 'LK',
                'country': ' \u200d ',
                'formattedAddress': ['Thurstan Road', ' \u200d ']},
                'categories': [{'id': '4bf58dd8d48988d147941735',

```

```

        'name': 'Diner',
        'pluralName': 'Diners',
        'shortName': 'Diner',
        'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/food/diner_',
        'suffix': '.png'},
        'primary': True}],
        'photos': {'count': 0, 'groups': []}},
        'referralId': 'e-0-52ed1ad2498e220cbe5b7575-3'}}]]]]}

```

1.3.5 Clean the json and structure it into a pandas dataframe

```

[147]:
      name      categories      lat      lng
0   Good Food           Diner  6.971568  79.873879
1    Twoso  Halal Restaurant  6.971694  79.875471
2  Perera & Sons          Bakery  6.972027  79.873464
3   dine more           Diner  6.972406  79.866929

```

1.3.6 Explore Neighborhoods in Colombo

Fort
 Slave Island, Union Place
 Kollupitiya
 Bambalapitiya
 Havelock Town, Kirulapone, Kirulapone North
 Wellawatte, Pamankada,
 Cinnamon Gardens
 Borella
 Dematagoda
 Maradana, Panchikawatte
 Pettah
 Hulftsdorp
 Kotahena, Bloemendhal
 Grandpass
 Modara/Mutwal, Mattakkuliya, Madampitiya
 (561, 7)

```

[152]:
Neighborhood Neighborhood Latitude Neighborhood Longitude \
0      Fort      6.94148      79.84664
1      Fort      6.94148      79.84664
2      Fort      6.94148      79.84664
3      Fort      6.94148      79.84664
4      Fort      6.94148      79.84664

```

```

      Venue Venue Latitude Venue Longitude Venue Category
0  Ministry Of Crab      6.933673      79.843499 Seafood Restaurant
1      & Co.      6.934104      79.842830 English Restaurant

```

2	Sri Suryas Hotel	6.942696	79.854711	Indian Restaurant
3	Pizza Hut	6.934417	79.844303	Pizza Place
4	Ikoi Tei	6.933578	79.843599	Japanese Restaurant

[153]:

Neighborhood	Latitude	Longitude	Venue
Bambalapitiya	100	100	
Borella	38	38	
Cinnamon Gardens	63	63	
Dematagoda	6	6	
Fort	26	26	
Grandpass	13	13	
Havelock Town, Kirulapone, Kirulapone North	30	30	
Hulftsdorp	25	25	
Kollupitiya	100	100	
Kotahena, Bloemendhal	8	8	
Maradana, Panchikawatte	9	9	
Modara/Mutwal, Mattakkuliya, Madampitiya	4	4	
Pettah	15	15	
Slave Island, Union Place	66	66	
Wellawatte, Pamankada,	58	58	

Neighborhood	Venue Latitude	Venue Longitude
Bambalapitiya	100	100
Borella	38	38
Cinnamon Gardens	63	63
Dematagoda	6	6
Fort	26	26

Grandpass	13	13
Havelock Town, Kirulapone, Kirulapone North	30	30
Hulftsdorp	25	25
Kollupitiya	100	100
Kotahena, Bloemendhal	8	8
Maradana, Panchikawatte	9	9
Modara/Mutwal, Mattakkuliya, Madampitiya	4	4
Pettah	15	15
Slave Island, Union Place	66	66
Wellawatte, Pamankada,	58	58

Venue Category

Neighborhood	
Bambalapitiya	100
Borella	38
Cinnamon Gardens	63
Dematagoda	6
Fort	26
Grandpass	13
Havelock Town, Kirulapone, Kirulapone North	30
Hulftsdorp	25
Kollupitiya	100
Kotahena, Bloemendhal	8
Maradana, Panchikawatte	9
Modara/Mutwal, Mattakkuliya, Madampitiya	4
Pettah	15
Slave Island, Union Place	66
Wellawatte, Pamankada,	58

1.3.7 Analyze Each Neighborhood

[154]:	Neighborhood	American Restaurant	Asian Restaurant	BBQ Joint	Bakery	\
0	Fort	0	0	0	0	
1	Fort	0	0	0	0	
2	Fort	0	0	0	0	
3	Fort	0	0	0	0	
4	Fort	0	0	0	0	
	Bistro	Breakfast Spot	Buffet	Burger Joint	Cafeteria	...
0	0	0	0	0	0	...
1	0	0	0	0	0	...
2	0	0	0	0	0	...
3	0	0	0	0	0	...
4	0	0	0	0	0	...
	Seafood Restaurant	Snack Place	South Indian Restaurant	\		
0	1	0	0			

1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0

	Sri Lankan Restaurant	Steakhouse	Sushi Restaurant	Taco Place \
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0

	Tapas Restaurant	Thai Restaurant	Vegetarian / Vegan Restaurant
0	0	0	0
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0

[5 rows x 56 columns]

There are 55 unique categories.

```
[157]: array(['Seafood Restaurant', 'English Restaurant', 'Indian Restaurant',
            'Pizza Place', 'Japanese Restaurant', 'Asian Restaurant',
            'Restaurant', 'Vegetarian / Vegan Restaurant',
            'American Restaurant', 'Mexican Restaurant',
            'Cajun / Creole Restaurant', 'Food Truck', 'Chinese Restaurant',
            'Sri Lankan Restaurant', 'Cafeteria', 'Food',
            'Fast Food Restaurant', 'Bakery', 'South Indian Restaurant',
            'German Restaurant', 'Café', 'Noodle House', 'Bistro', 'Diner',
            'French Restaurant', 'Gastropub', 'Sandwich Place',
            'Thai Restaurant', 'Steakhouse', 'Tapas Restaurant',
            'Italian Restaurant', 'Malay Restaurant', 'Snack Place',
            'Food Court', 'Cuban Restaurant', 'Breakfast Spot', 'Buffet',
            'Sushi Restaurant', 'Fondue Restaurant', 'North Indian Restaurant',
            'Middle Eastern Restaurant', 'Dim Sum Restaurant', 'Deli / Bodega',
            'Mediterranean Restaurant', 'Poke Place', 'Halal Restaurant',
            'BBQ Joint', 'Burger Joint', 'Donut Shop', 'Irish Pub',
            'Taco Place', 'Comfort Food Restaurant',
            'Indian Chinese Restaurant', 'Hot Dog Joint',
            'Eastern European Restaurant'], dtype=object)
```

Next, let's group rows by neighborhood and by taking the mean of the frequency of occurrence of each category

```
[158]:
```

	Neighborhood	American Restaurant \
0	Bambalapitiya	0.000000

1	Borella	0.000000
2	Cinnamon Gardens	0.031746
3	Dematagoda	0.000000
4	Fort	0.038462
5	Grandpass	0.000000
6	Havelock Town, Kirulapone, Kirulapone North	0.000000
7	Hulftsdorp	0.000000
8	Kollupitiya	0.000000
9	Kotahena, Bloemendhal	0.000000
10	Maradana, Panchikawatte	0.000000
11	Modara/Mutwal, Mattakkuliya, Madampitiya	0.000000
12	Pettah	0.000000
13	Slave Island, Union Place	0.000000
14	Wellawatte, Pamankada,	0.000000

	Asian Restaurant	BBQ Joint	Bakery	Bistro	Breakfast Spot	Buffet \
0	0.040000	0.01	0.060000	0.000000	0.000000	0.010000
1	0.026316	0.00	0.157895	0.000000	0.026316	0.026316
2	0.000000	0.00	0.079365	0.000000	0.031746	0.000000
3	0.166667	0.00	0.166667	0.000000	0.000000	0.000000
4	0.153846	0.00	0.000000	0.000000	0.000000	0.000000
5	0.153846	0.00	0.307692	0.000000	0.000000	0.000000
6	0.066667	0.00	0.066667	0.000000	0.000000	0.000000
7	0.240000	0.00	0.040000	0.000000	0.000000	0.000000
8	0.080000	0.00	0.050000	0.010000	0.010000	0.010000
9	0.125000	0.00	0.375000	0.000000	0.000000	0.000000
10	0.333333	0.00	0.111111	0.000000	0.000000	0.000000
11	0.000000	0.00	0.250000	0.000000	0.000000	0.000000
12	0.200000	0.00	0.000000	0.000000	0.000000	0.000000
13	0.045455	0.00	0.090909	0.015152	0.030303	0.015152
14	0.086207	0.00	0.120690	0.017241	0.034483	0.000000

	Burger Joint	Cafeteria	...	Seafood Restaurant	Snack Place \
0	0.000000	0.000000	...	0.030000	0.010000
1	0.000000	0.052632	...	0.000000	0.000000
2	0.000000	0.000000	...	0.000000	0.031746
3	0.000000	0.166667	...	0.000000	0.000000
4	0.000000	0.038462	...	0.076923	0.000000
5	0.000000	0.000000	...	0.000000	0.000000
6	0.033333	0.033333	...	0.033333	0.000000
7	0.000000	0.000000	...	0.000000	0.000000
8	0.000000	0.000000	...	0.050000	0.020000
9	0.000000	0.000000	...	0.000000	0.000000
10	0.000000	0.000000	...	0.000000	0.000000
11	0.000000	0.000000	...	0.000000	0.000000
12	0.000000	0.066667	...	0.000000	0.000000
13	0.000000	0.000000	...	0.060606	0.030303

14	0.000000	0.000000	...	0.017241	0.000000
----	----------	----------	-----	----------	----------

	South Indian Restaurant	Sri Lankan Restaurant	Steakhouse \
0	0.000000	0.070000	0.000000
1	0.000000	0.026316	0.000000
2	0.015873	0.031746	0.000000
3	0.000000	0.000000	0.000000
4	0.000000	0.038462	0.000000
5	0.000000	0.000000	0.000000
6	0.033333	0.000000	0.000000
7	0.000000	0.000000	0.000000
8	0.010000	0.050000	0.010000
9	0.000000	0.000000	0.000000
10	0.000000	0.000000	0.000000
11	0.000000	0.000000	0.000000
12	0.000000	0.066667	0.000000
13	0.015152	0.045455	0.015152
14	0.000000	0.051724	0.000000

	Sushi Restaurant	Taco Place	Tapas Restaurant	Thai Restaurant \
0	0.000000	0.000000	0.000000	0.040000
1	0.000000	0.000000	0.000000	0.000000
2	0.000000	0.015873	0.000000	0.015873
3	0.000000	0.000000	0.000000	0.000000
4	0.000000	0.000000	0.000000	0.000000
5	0.000000	0.000000	0.000000	0.000000
6	0.000000	0.000000	0.000000	0.033333
7	0.000000	0.000000	0.000000	0.000000
8	0.000000	0.000000	0.010000	0.010000
9	0.000000	0.000000	0.000000	0.000000
10	0.000000	0.000000	0.000000	0.000000
11	0.000000	0.000000	0.000000	0.000000
12	0.000000	0.000000	0.000000	0.000000
13	0.015152	0.000000	0.015152	0.015152
14	0.000000	0.000000	0.000000	0.000000

	Vegetarian / Vegan Restaurant
0	0.010000
1	0.052632
2	0.000000
3	0.000000
4	0.076923
5	0.076923
6	0.033333
7	0.120000
8	0.010000
9	0.000000

10	0.111111
11	0.000000
12	0.133333
13	0.000000
14	0.034483

[15 rows x 56 columns]

```
[164]:      Neighborhood 1st Most Common Venue 2nd Most Common Venue \
0      Bambalapitiya      Café Fast Food Restaurant
1      Borella      Restaurant      Bakery
2      Cinnamon Gardens      Café      Restaurant
3      Dematagoda      Asian Restaurant      Bakery
4      Fort      Asian Restaurant      Indian Restaurant

      3rd Most Common Venue      4th Most Common Venue \
0      Restaurant      Chinese Restaurant
1      Café Vegetarian / Vegan Restaurant
2      Fast Food Restaurant      Bakery
3      Food Court      Fast Food Restaurant
4      Vegetarian / Vegan Restaurant      Seafood Restaurant

      5th Most Common Venue 6th Most Common Venue      7th Most Common Venue \
0      Indian Restaurant Sri Lankan Restaurant      Bakery
1      Pizza Place      Cafeteria      Halal Restaurant
2      Indian Restaurant Japanese Restaurant      Chinese Restaurant
3      Cafeteria      Chinese Restaurant Vegetarian / Vegan Restaurant
4      Restaurant      Pizza Place      Mexican Restaurant

      8th Most Common Venue 9th Most Common Venue 10th Most Common Venue
0      Asian Restaurant      Thai Restaurant      Seafood Restaurant
1      Donut Shop      Indian Chinese Restaurant      Italian Restaurant
2      Breakfast Spot      American Restaurant Sri Lankan Restaurant
3      Dim Sum Restaurant      Food Truck      Food
4      Cafeteria Cajun / Creole Restaurant      Chinese Restaurant
```

Let's print each neighborhood along with the top 10 most common venues

----Bambalapitiya----

	venue	freq
0	Café	0.15
1	Fast Food Restaurant	0.11
2	Restaurant	0.11
3	Chinese Restaurant	0.10
4	Indian Restaurant	0.08
5	Sri Lankan Restaurant	0.07
6	Bakery	0.06
7	Thai Restaurant	0.04

8	Asian Restaurant	0.04
9	Japanese Restaurant	0.03

----Borella----

	venue	freq
0	Restaurant	0.18
1	Bakery	0.16
2	Café	0.16
3	Vegetarian / Vegan Restaurant	0.05
4	Cafeteria	0.05
5	Pizza Place	0.05
6	Sri Lankan Restaurant	0.03
7	Mediterranean Restaurant	0.03
8	Indian Chinese Restaurant	0.03
9	Mexican Restaurant	0.03

----Cinnamon Gardens----

	venue	freq
0	Café	0.22
1	Restaurant	0.14
2	Fast Food Restaurant	0.10
3	Bakery	0.08
4	Indian Restaurant	0.06
5	Japanese Restaurant	0.05
6	Snack Place	0.03
7	Chinese Restaurant	0.03
8	Sri Lankan Restaurant	0.03
9	American Restaurant	0.03

----Dematagoda----

	venue	freq
0	Bakery	0.17
1	Asian Restaurant	0.17
2	Cafeteria	0.17
3	Food Court	0.17
4	Chinese Restaurant	0.17
5	Fast Food Restaurant	0.17
6	American Restaurant	0.00
7	Pizza Place	0.00
8	Italian Restaurant	0.00
9	Japanese Restaurant	0.00

----Fort----

	venue	freq
--	-------	------

0	Indian Restaurant	0.15
1	Asian Restaurant	0.15
2	Vegetarian / Vegan Restaurant	0.08
3	Seafood Restaurant	0.08
4	Restaurant	0.08
5	Pizza Place	0.08
6	English Restaurant	0.04
7	Sri Lankan Restaurant	0.04
8	Mexican Restaurant	0.04
9	Japanese Restaurant	0.04

----Grandpass----

	venue	freq
0	Bakery	0.31
1	Indian Restaurant	0.15
2	Asian Restaurant	0.15
3	Fast Food Restaurant	0.15
4	Chinese Restaurant	0.08
5	Food Court	0.08
6	Vegetarian / Vegan Restaurant	0.08
7	Steakhouse	0.00
8	Middle Eastern Restaurant	0.00
9	Irish Pub	0.00

----Havelock Town, Kirulapone, Kirulapone North----

	venue	freq
0	Café	0.13
1	Indian Restaurant	0.13
2	Pizza Place	0.10
3	Restaurant	0.10
4	Asian Restaurant	0.07
5	Bakery	0.07
6	Fast Food Restaurant	0.07
7	Vegetarian / Vegan Restaurant	0.03
8	South Indian Restaurant	0.03
9	Middle Eastern Restaurant	0.03

----Hulftsdorp----

	venue	freq
0	Indian Restaurant	0.24
1	Asian Restaurant	0.24
2	Food Truck	0.12
3	Vegetarian / Vegan Restaurant	0.12
4	Eastern European Restaurant	0.04
5	Restaurant	0.04

6	Hot Dog Joint	0.04
7	Bakery	0.04
8	Food	0.04
9	Fast Food Restaurant	0.04

----Kollupitiya----

	venue	freq
0	Restaurant	0.12
1	Café	0.09
2	Fast Food Restaurant	0.08
3	Chinese Restaurant	0.08
4	Asian Restaurant	0.08
5	Indian Restaurant	0.06
6	Bakery	0.05
7	Sri Lankan Restaurant	0.05
8	Seafood Restaurant	0.05
9	Japanese Restaurant	0.03

----Kotahena, Bloemendhal----

	venue	freq
0	Bakery	0.38
1	German Restaurant	0.12
2	Fast Food Restaurant	0.12
3	Asian Restaurant	0.12
4	Food Court	0.12
5	Indian Restaurant	0.12
6	Sushi Restaurant	0.00
7	Middle Eastern Restaurant	0.00
8	Thai Restaurant	0.00
9	Irish Pub	0.00

----Maradana, Panchikawatte----

	venue	freq
0	Asian Restaurant	0.33
1	Restaurant	0.33
2	Vegetarian / Vegan Restaurant	0.11
3	Diner	0.11
4	Bakery	0.11
5	Indian Restaurant	0.00
6	Irish Pub	0.00
7	Italian Restaurant	0.00
8	Japanese Restaurant	0.00
9	Malay Restaurant	0.00

----Modara/Mutwal, Mattakkuliya, Madampitiya----

	venue	freq
0	Diner	0.50
1	Halal Restaurant	0.25
2	Bakery	0.25
3	Pizza Place	0.00
4	Indian Chinese Restaurant	0.00
5	Indian Restaurant	0.00
6	Irish Pub	0.00
7	Italian Restaurant	0.00
8	Japanese Restaurant	0.00
9	Malay Restaurant	0.00

----Pettah----

	venue	freq
0	Indian Restaurant	0.27
1	Asian Restaurant	0.20
2	Vegetarian / Vegan Restaurant	0.13
3	Sri Lankan Restaurant	0.07
4	Cafeteria	0.07
5	Food Truck	0.07
6	Food	0.07
7	Restaurant	0.07
8	Pizza Place	0.07
9	Tapas Restaurant	0.00

----Slave Island, Union Place----

	venue	freq
0	Restaurant	0.14
1	Bakery	0.09
2	Café	0.09
3	Fast Food Restaurant	0.06
4	Seafood Restaurant	0.06
5	Asian Restaurant	0.05
6	Sri Lankan Restaurant	0.05
7	Japanese Restaurant	0.03
8	Italian Restaurant	0.03
9	Breakfast Spot	0.03

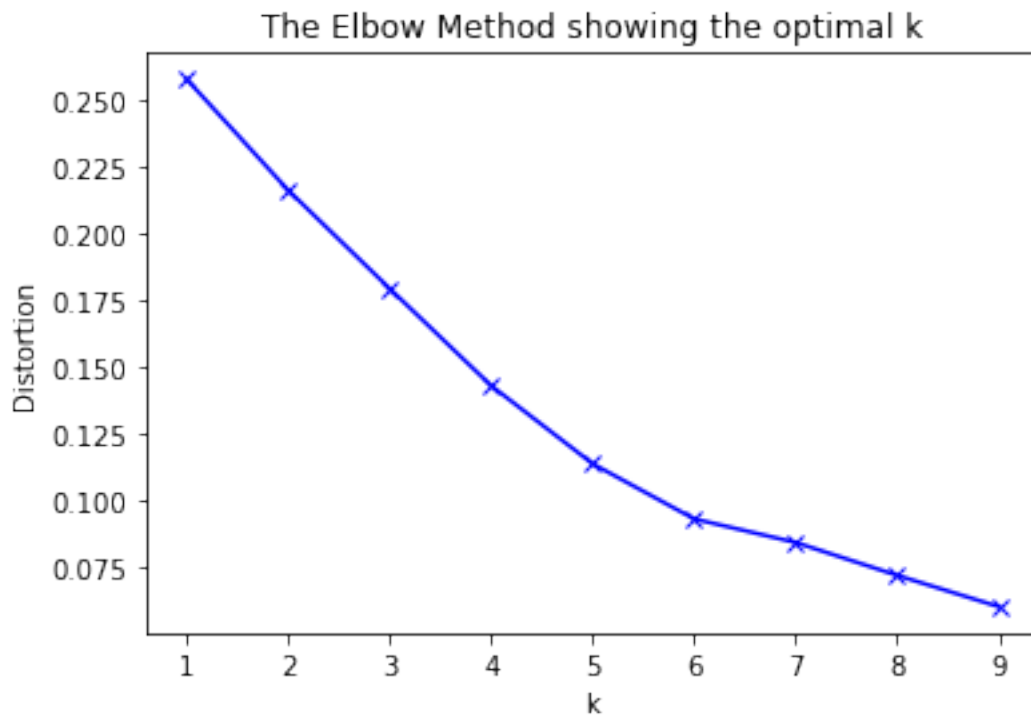
----Wellawatte, Pamankada,----

	venue	freq
0	Restaurant	0.19
1	Indian Restaurant	0.14
2	Bakery	0.12
3	Asian Restaurant	0.09

4	Fast Food Restaurant	0.07
5	Sri Lankan Restaurant	0.05
6	Café	0.05
7	Vegetarian / Vegan Restaurant	0.03
8	Chinese Restaurant	0.03
9	Pizza Place	0.03

1.3.8 Cluster Neighborhoods

To identify the optimal number of clusters, the Elbow method is used.



Based on the graph, we can see that 6 clusters are enough.

Run k-means to cluster the neighborhood into 6 clusters.

```
[168]: array([1, 1, 1, 5, 0, 3, 1, 0, 1, 3, 4, 2, 0, 1, 1])
```

Let's create a new dataframe that includes the cluster as well as the top 10 venues for each neighborhood.

```
[169]:
```

	Postal number	Neighborhood	Latitude	\
0	Colombo 1	Fort	6.941480	
1	Colombo 2	Slave Island, Union Place	6.922123	
2	Colombo 3	Kollupitiya	6.912930	

3	Colombo 4	Bambalapitiya	6.904710
4	Colombo 5	Havelock Town, Kirulapone, Kirulapone North	6.881350

	Longitude	Cluster	Labels	1st Most Common Venue	2nd Most Common Venue	\
0	79.846640		0	Asian Restaurant	Indian Restaurant	
1	79.854221		1	Restaurant	Café	
2	79.851280		1	Restaurant	Café	
3	79.855070		1	Café	Fast Food Restaurant	
4	79.875920		1	Indian Restaurant	Café	

	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	\
0	Vegetarian / Vegan Restaurant	Seafood Restaurant	Restaurant	
1	Bakery	Fast Food Restaurant	Seafood Restaurant	
2	Chinese Restaurant	Asian Restaurant	Fast Food Restaurant	
3	Restaurant	Chinese Restaurant	Indian Restaurant	
4	Pizza Place	Restaurant	Asian Restaurant	

	6th Most Common Venue	7th Most Common Venue	\
0	Pizza Place	Mexican Restaurant	
1	Asian Restaurant	Sri Lankan Restaurant	
2	Indian Restaurant	Bakery	
3	Sri Lankan Restaurant	Bakery	
4	Fast Food Restaurant	Bakery	

	8th Most Common Venue	9th Most Common Venue	\
0	Cafeteria	Cajun / Creole Restaurant	
1	Japanese Restaurant	Indian Restaurant	
2	Sri Lankan Restaurant	Seafood Restaurant	
3	Asian Restaurant	Thai Restaurant	
4	Vegetarian / Vegan Restaurant	Thai Restaurant	

	10th Most Common Venue
0	Chinese Restaurant
1	Diner
2	Japanese Restaurant
3	Seafood Restaurant
4	Donut Shop

Plotting Clusters in the Map

[179]: <folium.folium.Map at 0x1a2c10f3160>

MAP LEGEND

Cluster 1 - red dots

Cluster 2 - purple dots

Cluster 3 - blue dots

Cluster 5 - light blue dots

Cluster 5 - light green dots

Cluster 6 - orange dots

1.4 Results

And now we are ready to conclude our report.

1.4.1 Examine Clusters

Let's examine each cluster and the discriminating restaurant categories that distinguish a cluster.

Cluster 1

```
[180]: Neighborhood 1st Most Common Venue 2nd Most Common Venue \
0      Fort      Asian Restaurant      Indian Restaurant
10     Pettah     Indian Restaurant      Asian Restaurant
11  Hulftsdorp    Indian Restaurant      Asian Restaurant

      3rd Most Common Venue 4th Most Common Venue \
0  Vegetarian / Vegan Restaurant      Seafood Restaurant
10 Vegetarian / Vegan Restaurant      Food
11 Vegetarian / Vegan Restaurant      Food Truck

      5th Most Common Venue 6th Most Common Venue 7th Most Common Venue \
0      Restaurant      Pizza Place      Mexican Restaurant
10 Sri Lankan Restaurant      Cafeteria      Restaurant
11      Bakery      Food      Hot Dog Joint

      8th Most Common Venue      9th Most Common Venue \
0      Cafeteria      Cajun / Creole Restaurant
10      Pizza Place      Food Truck
11      Restaurant      Fast Food Restaurant

      10th Most Common Venue
0      Chinese Restaurant
10      Fondue Restaurant
11 Eastern European Restaurant
```

```
[181]: Neighborhood 1st Most Common Venue 2nd Most Common Venue \
count      3      3      3
unique      3      2      2
top      Pettah      Indian Restaurant      Asian Restaurant
freq      1      2      2

      3rd Most Common Venue 4th Most Common Venue \
count      3      3
unique      1      3
top      Vegetarian / Vegan Restaurant      Food
freq      3      1
```

	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	\
count	3	3	3	
unique	3	3	3	
top	Bakery	Pizza Place	Restaurant	
freq	1	1	1	

	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
count	3	3	3
unique	3	3	3
top	Pizza Place	Fast Food Restaurant	Fondue Restaurant
freq	1	1	1

Cluster 2

[182]:

	Neighborhood	1st Most Common Venue	\
1	Slave Island, Union Place	Restaurant	
2	Kollupitiya	Restaurant	
3	Bambalapitiya	Café	
4	Havelock Town, Kirulapone, Kirulapone North	Indian Restaurant	
5	Wellawatte, Pamankada,	Restaurant	
6	Cinnamon Gardens	Café	
7	Borella	Restaurant	

	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	\
1	Café	Bakery	Fast Food Restaurant	
2	Café	Chinese Restaurant	Asian Restaurant	
3	Fast Food Restaurant	Restaurant	Chinese Restaurant	
4	Café	Pizza Place	Restaurant	
5	Indian Restaurant	Bakery	Asian Restaurant	
6	Restaurant	Fast Food Restaurant	Bakery	
7	Bakery	Café	Vegetarian / Vegan Restaurant	

	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	\
1	Seafood Restaurant	Asian Restaurant	Sri Lankan Restaurant	
2	Fast Food Restaurant	Indian Restaurant	Bakery	
3	Indian Restaurant	Sri Lankan Restaurant	Bakery	
4	Asian Restaurant	Fast Food Restaurant	Bakery	
5	Fast Food Restaurant	Café	Sri Lankan Restaurant	
6	Indian Restaurant	Japanese Restaurant	Chinese Restaurant	
7	Pizza Place	Cafeteria	Halal Restaurant	

	8th Most Common Venue	9th Most Common Venue	\
1	Japanese Restaurant	Indian Restaurant	
2	Sri Lankan Restaurant	Seafood Restaurant	
3	Asian Restaurant	Thai Restaurant	
4	Vegetarian / Vegan Restaurant	Thai Restaurant	
5	Vegetarian / Vegan Restaurant	Breakfast Spot	

6	Breakfast Spot	American Restaurant
7	Donut Shop	Indian Chinese Restaurant

10th Most Common Venue	
1	Diner
2	Japanese Restaurant
3	Seafood Restaurant
4	Donut Shop
5	Diner
6	Sri Lankan Restaurant
7	Italian Restaurant

[183]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	\
count		7	7	7
unique		7	3	5
top	Borella	Restaurant		Café
freq		1	4	3

	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	\
count		7	7	7
unique		6	6	5
top	Bakery	Asian Restaurant	Fast Food Restaurant	
freq		2	2	2

	6th Most Common Venue	7th Most Common Venue	\
count		7	
unique		7	4
top	Fast Food Restaurant	Bakery	
freq		1	3

	8th Most Common Venue	9th Most Common Venue	\
count		7	7
unique		6	6
top	Vegetarian / Vegan Restaurant	Thai Restaurant	
freq		2	2

10th Most Common Venue	
count	7
unique	6
top	Diner
freq	2

Cluster 3

[184]:

	Neighborhood	1st Most Common Venue	\
14	Modara/Mutwal, Mattakkuliya, Madampitiya	Diner	

2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	\
-----------------------	-----------------------	-----------------------	---

14	Halal Restaurant	Bakery	Vegetarian / Vegan Restaurant
----	------------------	--------	-------------------------------

	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue \
14	Deli / Bodega	Food Truck	Food Court

	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
14	Food	Fondue Restaurant	Fast Food Restaurant

[185]:

	Neighborhood	1st Most Common Venue \
count		1
unique		1
top	Modara/Mutwal, Mattakkuliya, Madampitiya	Diner
freq		1

	2nd Most Common Venue	3rd Most Common Venue \
count	1	1
unique	1	1
top	Halal Restaurant	Bakery
freq	1	1

	4th Most Common Venue	5th Most Common Venue \
count	1	1
unique	1	1
top	Vegetarian / Vegan Restaurant	Deli / Bodega
freq	1	1

	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue \
count	1	1	1
unique	1	1	1
top	Food Truck	Food Court	Food
freq	1	1	1

	9th Most Common Venue	10th Most Common Venue
count	1	1
unique	1	1
top	Fondue Restaurant	Fast Food Restaurant
freq	1	1

Cluster 4

[186]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue \
12	Kotahena, Bloemendhal	Bakery	German Restaurant
13	Grandpass	Bakery	Asian Restaurant

	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue \
12	Asian Restaurant	Food Court	Indian Restaurant
13	Indian Restaurant	Fast Food Restaurant	Vegetarian / Vegan Restaurant

	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	\
12	Fast Food Restaurant	Deli / Bodega	Food Truck	
13	Food Court	Chinese Restaurant	Dim Sum Restaurant	

	9th Most Common Venue	10th Most Common Venue
12	Food	Fondue Restaurant
13	Food Truck	Food

[187]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	\
count		2	2	2
unique		2	1	2
top	Grandpass	Bakery	German Restaurant	
freq	1	2	1	

	3rd Most Common Venue	4th Most Common Venue	\
count	2	2	
unique	2	2	
top	Asian Restaurant	Fast Food Restaurant	
freq	1	1	

	5th Most Common Venue	6th Most Common Venue	\
count	2	2	
unique	2	2	
top	Vegetarian / Vegan Restaurant	Fast Food Restaurant	
freq	1	1	

	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	\
count	2	2	2	
unique	2	2	2	
top	Deli / Bodega	Dim Sum Restaurant	Food	
freq	1	1	1	

	10th Most Common Venue
count	2
unique	2
top	Fondue Restaurant
freq	1

Cluster 5

[188]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	\
9	Maradana, Panchikawatte	Asian Restaurant	Restaurant	

	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	\
9	Vegetarian / Vegan Restaurant	Bakery	Diner	

	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	\
9	Deli / Bodega	Food Truck	Food Court	

	9th Most Common Venue	10th Most Common Venue
9	Food	Fondue Restaurant

[189]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	\
count		1	1	1
unique		1	1	1
top	Maradana, Panchikawatte	Asian Restaurant		Restaurant
freq		1	1	1

	3rd Most Common Venue	4th Most Common Venue	\
count		1	1
unique		1	1
top	Vegetarian / Vegan Restaurant		Bakery
freq		1	1

	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	\
count		1	1	1
unique		1	1	1
top	Diner	Deli / Bodega		Food Truck
freq		1	1	1

	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
count		1	1
unique		1	1
top	Food Court	Food	Fondue Restaurant
freq		1	1

Cluster 6

[190]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	\
8	Dematagoda	Asian Restaurant		Bakery

	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	\
8	Food Court	Fast Food Restaurant		Cafeteria

	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	\
8	Chinese Restaurant	Vegetarian / Vegan Restaurant		Dim Sum Restaurant

	9th Most Common Venue	10th Most Common Venue
8	Food Truck	Food

[191]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	\
count		1	1	1
unique		1	1	1
top	Dematagoda	Asian Restaurant		Bakery
freq		1	1	1

	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	\
count	1	1	1	
unique	1	1	1	
top	Food Court	Fast Food Restaurant	Cafeteria	
freq	1	1	1	

	6th Most Common Venue	7th Most Common Venue	\
count	1	1	
unique	1	1	
top	Chinese Restaurant	Vegetarian / Vegan Restaurant	
freq	1	1	

	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
count	1	1	1
unique	1	1	1
top	Dim Sum Restaurant	Food Truck	Food
freq	1	1	1

1.5 Discussion

Analyzing the most popular restaurants in each cluster, it seems colombo is famous for spicy foods, especially Inian/Asian foods the stakeholder should prefer the popular types as a safe choice.

However stake holder can take a risk and can add novelty to open a restaurent related to mediterranean, american or european foods since Colombo is lacking oncusines other tha asian.

But bear in mind that descending on the popularity we might face an absence of demand for this type of food, and open a restaurant that is not needed in this particular location. Presence of interested customers is a must for a successful business.

Recommendations, based on description of each cluster: Cluster 1 Locations: Indian and or Pizza are popular Cluster 2 Locations: Asian and Fast Foods are populer Cluster 3 Locations: Halal or Vegetarian Foods are populer Cluster 4 Locations: Bakery and European Foods are populer Cluster 5 Locations: Asian Foods or Vegetarian Foods are populer Cluster 6 Locations: Asian Foods or Fast Foods are populer

After the type of restaurant is chosen, it is time to select a right place. Using the map created in section 3.8.3.

1.6 Conclusion

In this report we worked out a methodology to determine what the most promising type of restaurant is and where it should be opened.

We collected information about Colombo suburbs from Wikipedia, and using geospatial libraries mapped them. Using Foursquare API, we collected the top 100 restaurants and their types for each location within a radius 1000 meters from its central point. Then we grouped collected restaurants by location and by taking the mean of the frequency of occurrence of each type, preparing them for clustering. Finally we clustered restaurants by the k-means algorithm and analyze the top 10 most

common restaurants in each cluster, making useful observations. Eventually we visualized clusters on the map, thus showing the best locations for opening the chosen type of restaurant.

This type of analysis can be applied to any city of your choice that has available geospatial information.

This type of analysis can be applied to any type of venue (shopping, clubs, etc.) that is available in Foursquare database.