Capstone Project - The Battle of the Neighborhoods (Week 2)¶

Applied Data Science Capstone by IBM/Coursera¶

Table of contents¶

Introduction: Business Problem ¶

In this project, I will try to find countries where you can eat what you usually eat in your country.
Let's assume that you are to live abroad for some reasons and you are wondering where to go. If you are a gourmet, 'lineup of restaurants' (the most common categories of restaurants) will probably affect your decision. Even if you are not, what kind of restaurants are available in the country will affect whether you will be comfortable there.
I will cluster all the countries in the world into 5 clusters based on their lineup of restaurants. By selecting one of the countries in the same cluster as yours, you will be satisfied with your life there.

Data ¶

Based on definition of our problem, factors that will influence our decision are: number of existing restaurants in the neighborhood category of each restaurant
Each country has a lot of cities. I use capital cities to evaluate lineup of restaurants of the county as a capital city is a typical city of the country.
Following data sources will be needed to extract/generate the required information:

- List of countries will be obtained from this site: https://geographyfieldwork.com/WorldCapitalCities.htm

- Location of the capital city of each country will be obtained by using *geopy.geocoders* library.
- List of restaurants near the center of capital cities and their categories will be obtained by using Foursquare API.

Download libraries¶

In [1]:

```
!conda install -c conda-forge geopy --yes # uncomment this line if you haven't
 completed the Foursquare API lab

!conda install -c conda-forge folium=0.5.0 --yes # uncomment this line if you
haven't completed the Foursquare API lab

!conda install -c conda-forge lxml html5lib beautifulsoup4 --yes
```

```
Solving environment: done
```

Import libraries¶

In [2]:

```
import numpy as np # library to handle data in a vectorized manner


import pandas as pd # library for data analsysis

pd.set_option('display.max_columns', None)

pd.set_option('display.max_rows', None)


import json # library to handle JSON files


#!conda install -c conda-forge geopy --yes # uncomment this line if you haven'
t completed the Foursquare API lab

from geopy.geocoders import Nominatim # convert an address into latitude and l
ongitude values


import requests # library to handle requests

from pandas.io.json import json_normalize # tranform JSON file into a pandas d
ataframe


# Matplotlib and associated plotting modules

import matplotlib.cm as cm

import matplotlib.colors as colors
```

```
# import k-means from clustering stage
from sklearn.cluster import KMeans

#!conda install -c conda-forge folium=0.5.0 --yes # uncomment this line if you
 haven't completed the Foursquare API lab
import folium # map rendering library

# Regular Expression
import re

print('Libraries imported.')
```

```
Libraries imported.
```

Get the list of capital cities¶

In [3]:

```
url = 'https://geographyfieldwork.com/WorldCapitalCities.htm'
df_cap0 = pd.read_html(url)[2]

# Remove the last line (total)
df_cap0 = df_cap0.drop(len(df_cap0)-1)

# Remove duplicated cities (London)
df_cap0 = df_cap0.groupby(['Capital City']).nth(0)
df_cap0.reset_index(inplace=True)
df_cap0 = df_cap0[['Country', 'Capital City']]

# Remove numbers
df_cap0['Country'] = df_cap0['Country'].str.translate(str.maketrans('Dummy', '
Dummy', '0123456789'))
df_cap0['Capital City'] = df_cap0['Capital City'].str.translate(str.maketrans
('Dummy', 'Dummy', '0123456789'))

df_cap0.head()
```

Out[3]:

	Country	Capital City
0	United Arab Emirates	Abu Dhabi
1	Nigeria	Abuja

	Country	Capital City
2	Ghana	Accra
3	Ethiopia	Addis Ababa
4	Algeria	Algiers

Get the latitude and longitude of each capital city by using geopy.geocoders¶

In [4]:

```
df_cap1 = df_cap0.copy()
geolocator = Nominatim(user_agent="ny_explorer")

for index, row in df_cap1.iterrows():
    country = row['Country']
    capital_city = row['Capital City']
    try:
        location = geolocator.geocode(capital_city + ", " + country)
        latitude = location.latitude
        longitude = location.longitude
    except:
        latitude = 0
        longitude = 0
    print('{}: The geographical coordinate of {}, {} are {}, {}.'.format(index,
 country, capital_city, latitude, longitude))
    df_cap1.loc[index, 'Latitude']=latitude
    df_cap1.loc[index, 'Longitude']=longitude

df_cap1

# Remove rows with Latitude 0.0 (geolocation does not support the capital city
 name)
df_cap1 = df_cap1[df_cap1['Latitude'] != 0.0]
df_cap1.reset_index(drop=True, inplace=True)

# Save the DataFrame
df_cap1.to_csv('capitals.csv')
df_cap1.head()
```

	Country	Capital City	Latitude	Longitude
0	United Arab Emirates	Abu Dhabi	23.997644	53.643910
1	Nigeria	Abuja	9.064331	7.489297
2	Ghana	Accra	5.560014	-0.205744
3	Ethiopia	Addis Ababa	9.010793	38.761252
4	Algeria	Algiers	28.000027	2.999983

Or, get the latitude and longitude of each capital city from the saved file.¶

In [5]:

```
df_cap1 = pd.read_csv('capitals.csv', index_col=0)
# df_cap1 = df_cap1[:20] # To reduce Foursquare API calls
df_cap1.head()
```

Out[5]:

	Country	Capital City	Latitude	Longitude
0	United Arab Emirates	Abu Dhabi	23.997644	53.643910
1	Nigeria	Abuja	9.064330	7.489297
2	Ghana	Accra	5.560014	-0.205744
3	Ethiopia	Addis Ababa	9.010793	38.761253
4	Algeria	Algiers	28.000027	2.999983

Display the world map with capital cities¶

In [6]:

```
# create world map using latitude and longitude values

# Center
latitude = 0.0
longitude = 0.0

map_world= folium.Map(location=[latitude, longitude], zoom_start=2)
world_data = df_cap1

# add markers to map
for lat, lng, country, city in zip(world_data['Latitude'], world_data['Longitu
de'], world_data['Country'], world_data['Capital City']):
    label = '{}, {}'.format(city, country)
    # print(type(label)) # str
    # print(label) # ex. Wakefield, Bronx
```

```
    label = folium.Popup(label, parse_html=True)
    # print(type(label)) # folium.map.Popup
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(map_world)

map_world
```

Out[6]:

Methodology ¶

In this project, I will cluster almost all the countries in the world into 5 clusters based on the lineup of restaurants.
I assume a capital city has a typical lineup of the restaurants in the country. The location of capital cities have been already collected above. Foursquare API is available to explore venues by specifying a location. We can get only the food-related venues by specifying section=food.
In the first step, I will find the restaurants and their category within a radius of 1000 meters from the center of each capital city by Foursquare API.
In the second step, I will get the top 10 categories for each capital city.
In the third step, I will run k-means to cluster all the capital cities into 5 clusters.

Analysis ¶

Define Foursquare Credentials and Version¶

In [7]:

```
CLIENT_ID = 'XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX' # your Foursquare ID

CLIENT_SECRET = 'XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX' # your Foursquare Secret

VERSION = '20180605' # Foursquare API version
```

```
print('Your credentails:')
print('CLIENT_ID: ' + CLIENT_ID)
print('CLIENT_SECRET:' + CLIENT_SECRET)
```

```
Your credentails:
CLIENT_ID: XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
CLIENT_SECRET:XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
```

Create a list of capital cities with venues in food section¶

From the Foursquare lab in the previous module, we know that all the information is in the *items* key. Before we proceed, let's borrow the **get_category_type** function from the Foursquare lab.

In [8]:

```
# function that extracts the category of the venue
def get_category_type(row):
    try:
        categories_list = row['categories']
    except:
        categories_list = row['venue.categories']

    if len(categories_list) == 0:
        return None
    else:
        return categories_list[0]['name']
```

Function to explore venues with specified section¶

In [9]:

```
LIMIT = 100 # limit of number of venues returned by Foursquare API

def getNearbyVenues(names, latitudes, longitudes, section='food', radius=100
0):

    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={}&clie
nt_secret={}&v={}&ll={},{}&radius={}&limit={}&section={}'.format(
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            lat,
            lng,
            radius,
            LIMIT,
            section
            )

        # make the GET request
        try:
            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])
        except:
            print('Error: ' + url)
            break



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

```
    return(nearby_venues)
```

Execute the function for all the capital cities in the world¶

In [10]:

```
world_venues = getNearbyVenues(names=world_data['Capital City'],
                                   latitudes=world_data['Latitude'],
                                   longitudes=world_data['Longitude'],
                                   section='food'
                              )
```

```
Abu Dhabi
Abuja
Accra
Addis Ababa
Algiers
…
```

Display the results.

In [11]:

```
print(world_venues.shape)
world_venues.to_csv('world_venues.csv')
world_venues.head()
```

```
(8276, 7)
```

Out[11]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Abuja	9.06433	7.489297	the secret garden	9.066731	7.490142	Italian Restaurant
1	Abuja	9.06433	7.489297	River Plate Garden	9.066260	7.490060	Pizza Place
2	Abuja	9.06433	7.489297	Papillion Restaurant	9.064327	7.484541	African Restaurant
3	Abuja	9.06433	7.489297	Yahuza Suya Spot	9.071558	7.485696	BBQ Joint
4	Abuja	9.06433	7.489297	Hatlab	9.071944	7.488913	Deli / Bodega

In [12]:

```
world_venues = pd.read_csv('world_venues.csv', index_col=0)
world_venues.head()
```

Out[12]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Abuja	9.064331	7.489297	the secret garden	9.066731	7.490142	Italian Restaurant
1	Abuja	9.064331	7.489297	River Plate Garden	9.066260	7.490060	Pizza Place
2	Abuja	9.064331	7.489297	Papillion Restaurant	9.064327	7.484541	African Restaurant
3	Abuja	9.064331	7.489297	Yahuza Suya Spot	9.071558	7.485696	BBQ Joint
4	Abuja	9.064331	7.489297	Hatlab	9.071944	7.488913	Deli / Bodega

Display the number of unique food categories.

In [13]:

```
print('There are {} uniques categories.'.format(len(world_venues['Venue Catego
ry'].unique())))
```

```
There are 187 uniques categories.
```

Analyze each capital city¶

In [14]:

```
# one hot encoding
world_onehot = pd.get_dummies(world_venues[['Venue Category']], prefix="", pre
fix_sep="")
#    Accessories Store   Adult Boutique  ...
# 0

# add neighborhood column back to dataframe
world_onehot['Neighborhood'] = world_venues['Neighborhood']
#    Accessories Store   Adult Boutique  ...  Neighborhood
# 0                                           Marbele Hill

# move neighborhood column to the first column from the last
fixed_columns = [world_onehot.columns[-1]] + list(world_onehot.columns[:-1])
world_onehot = world_onehot[fixed_columns]

world_onehot.head()
#   Neighborhood Accessories Store Adult Boutinue ...
# 0  Marble Hill                 0              0
```

Out[14]:

In [15]:

```
world_onehot.shape
```

Out[15]:

```
(8276, 188)
```

Group rows by neighborhood and by taking the mean of the frequency of occurrence of each category¶

In [16]:

```
world_grouped = world_onehot.groupby('Neighborhood').mean().reset_index()
world_grouped.head()
```

Out[16]:

Group rows by neighborhood and by taking the mean of the frequency of occurrence of each category¶

In [17]:

```
world_grouped.shape
```

Out[17]:

```
(173, 188)
```

Print each neighborhood along with the top 5 most common venues¶

In [18]:

```
num_top_venues = 5


for hood in world_grouped['Neighborhood']:
    print("----"+hood+"----")
    temp = world_grouped[world_grouped['Neighborhood'] == hood].T.reset_index
()
    temp.columns = ['venue','freq']
    temp = temp.iloc[1:]
    temp['freq'] = temp['freq'].astype(float)
    temp = temp.round({'freq': 2})
    print(temp.sort_values('freq', ascending=False).reset_index(drop=True).hea
d(num_top_venues))
    print('\n')
```

Put that into a *pandas* dataframe¶

Function to sort the venues in descending order¶

In [19]:

```
def return_most_common_venues(row, num_top_venues):
    row_categories = row.iloc[1:]
    row_categories_sorted = row_categories.sort_values(ascending=False)

    return row_categories_sorted.index.values[0:num_top_venues]
```

Create the new dataframe and display the top 10 venues for each neighborhood.¶

In [20]:

```
num_top_venues = 10

indicators = ['st', 'nd', 'rd']

# create columns according to number of top venues
columns = ['Neighborhood']
for ind in np.arange(num_top_venues):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[in
d]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

# create a new dataframe
neighborhoods_venues_sorted = pd.DataFrame(columns=columns)
neighborhoods_venues_sorted['Neighborhood'] = world_grouped['Neighborhood']

for ind in np.arange(world_grouped.shape[0]):
    neighborhoods_venues_sorted.iloc[ind, 1:] = return_most_common_venues(worl
d_grouped.iloc[ind, :], num_top_venues)

neighborhoods_venues_sorted.head()
```

Out[20]:

Cluster capital cities¶

Run *k*-means to cluster the neighborhood into 5 clusters.¶

In [21]:

```
# set number of clusters
kclusters = 5

world_grouped_clustering = world_grouped.drop('Neighborhood', 1)

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

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]
```

Out[21]:

```
array([0, 3, 0, 2, 1, 0, 0, 1, 0, 1], dtype=int32)
```

Create a new dataframe that includes the cluster as well as the top 10 venues for each neighborhood.¶

In [22]:

```
# add clustering labels
neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)
# print(type(neighborhoods_venues_sorted)) # DataFrame
# print(neighborhoods_venues_sorted.shape) # (99, 12)
# print(world_data.shape) # (103, 5)
```

In [23]:

```
world_merged = world_data
# print(world_merged)

# merge toronto_grouped with toronto_data to add latitude/longitude for each n
eighborhood
world_merged = world_merged.join(neighborhoods_venues_sorted.set_index('Neighb
orhood'), on='Capital City', how='inner')
# print(type(toronto_merged['Cluster Labels'][0])) # int32 if joined with how=
'inner'. float64 if how='inner'
```

```
world_merged.to_csv('world_merged.csv')
world_merged # check the last columns!
```

Out[23]:

	Country	Capital City	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
1	Nigeria	Abuja	9.064330	7.489297	0	Italian Restaurant	Pizza Place	African Restaurant
2	Ghana	Accra	5.560014	-0.205744	3	Restaurant	African Restaura nt	Breakfast Spot
3	Ethiopia	Addis Ababa	9.010793	38.761253	0	Ethiopian Restaurant	Restaura nt	American Restaurant
4	Algeria	Algiers	28.000027	2.999983	2	Café	Yunnan Restaura nt	Fast Food Restaurant
5	Jordan	Amman	31.951569	35.923963	1	Café	Middle Eastern Restaura nt	Italian Restaurant
6	Netherlands	Amsterdam	52.374540	4.897976	0	Café	Italian Restaura nt	Restaurant
7	Andorra	Andorra la Vella	42.506939	1.521247	0	Restaurant	Spanish Restaura nt	Tapas Restaurant
8	Turkey	Ankara	39.920777	32.854067	1	Café	Turkish Restaura nt	Kebab Restaurant
9	Madagascar	Antananarivo	-18.910012	47.525581	0	Restaurant	Burger Joint	Café
10	Samoa	Apia	-13.834369	-171.769279	1	Café	Restaura nt	Pizza Place
11	Turkmenistan	Ashgabat	37.940438	58.382279	1	Café	Restaura nt	Snack Place
12	Eritrea	Asmara	15.338967	38.932676	1	Café	Restaura nt	Asian Restaurant
13	Kazakhstan	Astana	51.150921	71.438860	0	Asian Restaurant	Fast Food Restaura nt	Restaurant
14	Paraguay	Asuncion	-25.280046	-57.634381	0	Restaurant	Café	Fast Food Restaurant
15	Greece	Athens	37.984149	23.727984	1	Café	Greek Restaura nt	Souvlaki Shop
16	Iraq	Baghdad	33.302431	44.378799	2	Café	Yunnan Restaura nt	Fast Food Restaurant
17	Azerbaijan	Baku	40.375443	49.832675	1	Café	Restaura nt	Turkish Restaurant

	Country	Capital City	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
19	Brunei	Bandar Seri Begawan	4.889545	114.941757	1	Asian Restaurant	Café	Food Court
20	Thailand	Bangkok	13.753893	100.816080	0	Buffet	Food Truck	Fast Food Restaurant
23	Saint Kitts and Nevis	Basseterre	17.296092	-62.722301	4	Fast Food Restaurant	Caribbean Restaurant	Restaurant
24	China	Beijing	39.906217	116.391276	0	Chinese Restaurant	French Restaurant	Asian Restaurant
25	Lebanon	Beirut	33.895920	35.478430	0	Café	Restaurant	Middle Eastern Restaurant
26	Northern Ireland	Belfast	54.596441	-5.930276	0	Restaurant	Café	Sandwich Place
27	Serbia	Belgrade	44.817813	20.456897	0	Restaurant	Café	Italian Restaurant
28	Belize	Belmopan	17.250199	-88.770018	1	Café	Deli / Bodega	Wings Joint
29	Germany	Berlin	52.517036	13.388860	0	German Restaurant	Café	Italian Restaurant
30	Switzerland	Bern	46.948271	7.451451	0	Swiss Restaurant	Café	Restaurant
31	Kyrgyzstan	Bishkek	42.876745	74.606995	0	Restaurant	Café	Turkish Restaurant
32	Guinea-Bissau	Bissau	11.861324	-15.583055	0	African Restaurant	French Restaurant	Restaurant
33	Colombia	Bogota	4.598080	-74.076044	0	Restaurant	Café	Italian Restaurant
35	Slovakia	Bratislava	48.151699	17.109306	1	Café	Vegetarian / Vegan Restaurant	Bistro
36	Barbados	Bridgetown	13.097783	-59.618418	4	Fast Food Restaurant	Caribbean Restaurant	Seafood Restaurant
37	Belgium	Brussels	50.846557	4.351697	0	Belgian Restaurant	Italian Restaurant	Gastropub
38	Romania	Bucharest	44.436141	26.102720	0	Café	Italian Restaurant	Romanian Restaurant
39	Hungary	Budapest	47.498382	19.040471	0	Restaurant	Hungarian Restaurant	Italian Restaurant
40	Argentina	Buenos Aires	-34.607562	-58.437076	0	Café	Pizza Place	Bakery

	Country	Capital City	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
41	Egypt	Cairo	30.048819	31.243666	1	Café	Egyptian Restaurant	Falafel Restaurant
42	Australia	Canberra	-35.297591	149.101268	3	Restaurant	Yunnan Restaurant	Falafel Restaurant
43	Venezuela	Caracas	10.506098	-66.914602	0	Bakery	Café	Fried Chicken Joint
44	Wales	Cardiff	51.481655	-3.179193	0	Restaurant	Café	Burger Joint
46	French Guiana	Cayenne	4.937114	-52.325831	4	Fast Food Restaurant	Japanese Restaurant	Chinese Restaurant
47	Moldova	Chisinau	47.024471	28.832253	0	Café	Pizza Place	Bakery
48	Sri Lanka	Colombo	6.921812	79.865561	0	Bakery	Restaurant	Café
49	Guinea	Conakry	9.517060	-13.699843	1	Italian Restaurant	Sandwich Place	Café
50	Denmark	Copenhagen	55.686724	12.570072	0	Café	Scandinavian Restaurant	Bakery
51	Senegal	Dakar	14.693425	-17.447938	0	Fast Food Restaurant	African Restaurant	Restaurant
52	Syria	Damascus	33.513070	36.309581	3	Restaurant	Diner	Café
53	Bangladesh	Dhaka	23.759357	90.378814	0	Asian Restaurant	Chinese Restaurant	Indian Restaurant
54	East Timor	Dili	-8.553681	125.578409	0	Indonesian Restaurant	Japanese Restaurant	Chinese Restaurant
56	Tanzania	Dodoma	-6.179118	35.746817	0	BBQ Joint	Burger Joint	Restaurant
57	Qatar	Doha	25.301327	51.495705	0	Restaurant	Asian Restaurant	Cafeteria
58	Ireland	Dublin	53.349764	-6.260273	1	Café	Restaurant	Italian Restaurant
59	Tajikistan	Dushanbe	38.576271	68.786357	1	Restaurant	Café	Burger Joint
60	Scotland	Edinburgh	55.953346	-3.188375	0	Café	Restaurant	Italian Restaurant
61	Sierra Leone	Freetown	8.479004	-13.267950	0	Bakery	African Restaurant	Restaurant
63	Botswana	Gaborone	-24.658136	25.908847	4	Fast Food Restaurant	Restaurant	Steakhouse
64	Guyana	Georgetown	6.802577	-58.162861	0	Burger Joint	Brazilian Restaurant	Café

	Country	Capital City	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
65	Burundi	Gitega	-3.428495	29.924972	3	American Restaurant	Restaurant	Yunnan Restaurant
66	Guatemala	Guatemala City	14.622233	-90.518519	0	Café	Pizza Place	Fast Food Restaurant
67	Vietnam	Hanoi	21.029450	105.854444	0	Vietnamese Restaurant	Café	Noodle House
68	Zimbabwe	Harare	-17.831773	31.045686	0	Breakfast Spot	African Restaurant	American Restaurant
69	Cuba	Havana	23.135305	-82.358963	0	Cuban Restaurant	Café	Italian Restaurant
70	Finland	Helsinki	60.167409	24.942568	0	Scandinavian Restaurant	Café	Restaurant
71	Solomon Islands	Honiara	-9.431297	159.955277	1	Café	Japanese Restaurant	Sushi Restaurant
72	Pakistan	Islamabad	33.635739	72.923047	0	BBQ Joint	Bakery	Pizza Place
73	Indonesia	Jakarta	-6.175394	106.827183	0	Indonesian Restaurant	Café	Fast Food Restaurant
74	South Sudan	Juba	4.847202	31.595166	3	Restaurant	Eastern European Restaurant	Indian Restaurant
75	Afghanistan	Kabul	34.526013	69.177648	0	Bakery	Afghan Restaurant	French Restaurant
76	Uganda	Kampala	0.317714	32.581354	1	Café	Fast Food Restaurant	Pizza Place
77	Nepal	Kathmandu	27.708796	85.320244	0	Restaurant	Asian Restaurant	Café
78	Sudan	Khartoum	15.593325	32.535650	0	Restaurant	Pizza Place	Burger Joint
79	Ukraine	Kiev	50.450034	30.524136	0	Café	Italian Restaurant	Bakery
81	Jamaica	Kingston	17.971215	-76.792813	4	Fast Food Restaurant	Restaurant	Bakery
82	Saint Vincent and the Grenadines	Kingstown	13.156186	-61.227962	0	Sandwich Place	Restaurant	Café
83	Malaysia	Kuala Lumpur	3.151664	101.694303	0	Malay Restaurant	Café	Asian Restaurant
84	Kuwait	Kuwait City	29.379709	47.973563	1	Café	Middle Eastern Restaurant	Restaurant

	Country	Capital City	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
85	Bolivia	La Paz	-16.495545	-68.133623	0	Café	Latin American Restaurant	Vegetarian / Vegan Restaurant
86	Gabon	Libreville	0.390002	9.454001	0	Bakery	Diner	American Restaurant
87	Malawi	Lilongwe	-13.987465	33.768056	4	Fast Food Restaurant	Mexican Restaurant	Italian Restaurant
88	Peru	Lima	-12.062107	-77.036526	0	Seafood Restaurant	Restaurant	Fried Chicken Joint
89	Portugal	Lisbon	38.707751	-9.136592	0	Portuguese Restaurant	Restaurant	Tapas Restaurant
90	Slovenia	Ljubljana	46.049815	14.506782	0	Restaurant	Café	Eastern European Restaurant
91	Togo	Lome	6.130419	1.215829	3	Restaurant	Japanese Restaurant	Spanish Restaurant
92	England	London	51.507322	-0.127647	0	Bakery	French Restaurant	Steakhouse
93	Angola	Luanda	-8.827270	13.243951	0	Pizza Place	Restaurant	Portuguese Restaurant
94	Zambia	Lusaka	-15.416449	28.282153	4	Fast Food Restaurant	American Restaurant	Café
95	Luxembourg	Luxembourg	49.815868	6.129675	0	Pizza Place	BBQ Joint	Yunnan Restaurant
96	Spain	Madrid	40.416705	-3.703582	0	Spanish Restaurant	Restaurant	Tapas Restaurant
97	Marshall Islands	Majuro	7.090992	171.381635	2	Café	Yunnan Restaurant	Fast Food Restaurant
98	Equatorial Guinea	Malabo	3.752828	8.780061	0	Pizza Place	French Restaurant	African Restaurant
99	Maldives	Male	4.177988	73.510739	1	Café	Restaurant	Pizza Place
100	Nicaragua	Managua	12.146124	-86.273717	0	Restaurant	Bakery	Sandwich Place
101	Bahrain	Manama	26.223504	50.582244	1	Café	Indian Restaurant	Asian Restaurant
102	Philippines	Manila	14.590622	120.979970	4	Fast Food Restaurant	Chinese Restaurant	Filipino Restaurant
103	Mozambique	Maputo	-25.966213	32.567450	0	Fast Food Restaurant	Breakfast Spot	Portuguese Restaurant

	Country	Capital City	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
104	Lesotho	Maseru	-29.310054	27.478222	1	Steakhouse	Café	Restaurant
105	Palau	Melekeok	7.495414	134.636753	2	Café	Yunnan Restaurant	Fast Food Restaurant
106	Mexico	Mexico City	19.432601	-99.133342	0	Mexican Restaurant	Taco Place	Café
107	Belarus	Minsk	53.902334	27.561879	0	Café	Restaurant	Eastern European Restaurant
109	Monaco	Monaco	43.732349	7.427683	0	Italian Restaurant	French Restaurant	Restaurant
110	Liberia	Monrovia	6.328034	-10.797788	4	Fast Food Restaurant	Yunnan Restaurant	French Restaurant
111	Uruguay	Montevideo	-34.905904	-56.191357	0	Restaurant	Bakery	Café
112	Comoros	Moroni	-11.693126	43.254304	1	Falafel Restaurant	Italian Restaurant	Café
113	Russia	Moscow	55.750446	37.617494	0	Italian Restaurant	Restaurant	Café
115	Chad	N'Djamena	12.119154	15.050276	0	Sandwich Place	Middle Eastern Restaurant	Yunnan Restaurant
116	Kenya	Nairobi	-1.283253	36.817245	0	African Restaurant	Fast Food Restaurant	Café
117	Bahamas	Nassau	25.078346	-77.338333	0	Caribbean Restaurant	Fast Food Restaurant	Greek Restaurant
118	India	New Delhi	28.614179	77.202266	0	Pizza Place	Indian Restaurant	BBQ Joint
119	Niger	Niamey	13.524834	2.109823	3	Restaurant	French Restaurant	African Restaurant
120	Cyprus	Nicosia	35.180892	33.373259	1	Café	Restaurant	Greek Restaurant
121	Mauritania	Nouakchott	18.079238	-15.978007	0	Breakfast Spot	Yunnan Restaurant	Fast Food Restaurant
122	Tonga	Nuku'alofa	-21.134340	-175.201808	1	Café	Italian Restaurant	Restaurant
123	Norway	Oslo	59.913330	10.738970	0	Scandinavian Restaurant	Café	Pizza Place

	Country	Capital City	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
124	Canada	Ottawa	45.421106	-75.690308	0	Café	Restaurant	New American Restaurant
125	Burkina Faso	Ouagadougou	12.368148	-1.527085	3	Restaurant	Bistro	Yunnan Restaurant
127	Panama	Panama City	8.971449	-79.534180	0	Latin American Restaurant	Restaurant	Mexican Restaurant
128	Suriname	Paramaribo	5.821609	-55.177043	1	Café	Fast Food Restaurant	Burger Joint
129	France	Paris	48.856610	2.351499	0	French Restaurant	Bakery	Café
130	Cambodia	Phnom Penh	11.568271	104.922443	0	Café	Asian Restaurant	French Restaurant
131	Montenegro	Podgorica	42.441524	19.262108	1	Café	Pizza Place	Restaurant
132	Mauritius	Port Louis	-20.163728	57.504533	0	Chinese Restaurant	Fast Food Restaurant	Pizza Place
133	Papua New Guinea	Port Moresby	-9.474330	147.159950	0	Diner	Asian Restaurant	Café
134	Vanuatu	Port Vila	-17.741497	168.315016	1	Café	Caribbean Restaurant	Restaurant
135	Haiti	Port au Prince	18.547327	-72.339593	3	Restaurant	Buffet	Fast Food Restaurant
136	Trinidad and Tobago	Port of Spain	10.657268	-61.518017	0	Caribbean Restaurant	Chinese Restaurant	Pizza Place
138	Czech Republic (Czechia)	Prague	50.099270	14.376683	0	Restaurant	Pizza Place	Café
139	Cape Verde	Praia	14.916017	-23.509613	3	Restaurant	Bakery	Café
140	Kosovo	Pristina	42.663877	21.164085	1	Restaurant	Café	Bakery
141	North Korea	Pyongyang	39.019474	125.753388	0	Korean Restaurant	Pizza Place	German Restaurant
142	Ecuador	Quito	-0.220164	-78.512327	0	Restaurant	Breakfast Spot	Diner
143	Morocco	Rabat	34.022405	-6.834543	1	Café	Moroccan Restaurant	French Restaurant
145	Latvia	Riga	56.949398	24.105185	0	Restaurant	Eastern European	Café

	Country	Capital City	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
							Restaurant	
146	Saudi Arabia	Riyadh	24.631969	46.715065	0	Middle Eastern Restaurant	Donut Shop	Bakery
147	Italy	Rome	41.894802	12.485338	0	Italian Restaurant	Pizza Place	Sandwich Place
148	Dominica	Roseau	15.299192	-61.387287	0	Restaurant	Pizza Place	Caribbean Restaurant
149	Grenada	Saint George's	12.053533	-61.751805	4	Caribbean Restaurant	Seafood Restaurant	Fast Food Restaurant
150	Antigua and Barbuda	Saint John's	17.118457	-61.844851	0	Fast Food Restaurant	Café	Middle Eastern Restaurant
151	Costa Rica	San Jose	9.932543	-84.079578	0	Latin American Restaurant	Sandwich Place	Café
152	San Marino	San Marino	43.945862	12.458306	0	Pizza Place	Fast Food Restaurant	Restaurant
153	El Salvador	San Salvador	13.698994	-89.191425	1	Café	Pizza Place	Fried Chicken Joint
154	Yemen	Sana'a	15.342101	44.200520	2	Café	Kebab Restaurant	Fast Food Restaurant
155	Chile	Santiago	-33.437797	-70.650445	0	Sandwich Place	Peruvian Restaurant	Café
156	Dominican Republic	Santo Domingo	18.480197	-69.942111	0	Pizza Place	Sandwich Place	Fast Food Restaurant
157	Sao Tome and Principe	Sao Tome	0.338924	6.731303	1	Café	Bakery	BBQ Joint
158	Bosnia and Herzegovina	Sarajevo	43.851977	18.386687	1	Café	Restaurant	Italian Restaurant
159	South Korea	Seoul	37.566679	126.978291	0	Korean Restaurant	Café	Chinese Restaurant
160	Singapore	Singapore	1.340863	103.830392	1	Japanese Restaurant	Sandwich Place	Restaurant
162	Bulgaria	Sofia	42.697863	23.322179	0	Bakery	Italian Restaurant	Restaurant
163	Sweden	Stockholm	59.325117	18.071093	0	Scandinavian Restaurant	Café	Bakery
164	Fiji	Suva	-18.141588	178.442166	4	Fast Food Restaurant	Japanese Restaurant	Asian Restaurant

	Country	Capital City	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
165	Taiwan	Taipei	25.037520	121.563680	0	Café	Japanese Restaurant	Chinese Restaurant
166	Estonia	Tallinn	59.437216	24.745369	1	Café	Restaurant	Italian Restaurant
167	Uzbekistan	Tashkent	41.312336	69.278708	0	Fast Food Restaurant	Pizza Place	Café
168	Georgia	Tbilisi	41.693459	44.801449	1	Café	Restaurant	Caucasian Restaurant
169	Honduras	Tegucigalpa	14.093192	-87.201263	0	Restaurant	Mexican Restaurant	Steakhouse
170	Iran	Tehran	35.700618	51.401378	2	Café	Persian Restaurant	Breakfast Spot
172	Albania	Tirana (Tirane)	41.327946	19.818532	1	Café	Italian Restaurant	Restaurant
173	Japan	Tokyo	35.682839	139.759455	0	Japanese Restaurant	Café	French Restaurant
174	Libya	Tripoli	32.896672	13.177792	1	Café	Italian Restaurant	Seafood Restaurant
177	Liechtenstein	Vaduz	47.139286	9.522796	1	Café	Swiss Restaurant	Italian Restaurant
178	Malta	Valletta	35.898982	14.513676	0	Mediterranean Restaurant	Café	Restaurant
179	Vatican City	Vatican City	41.903491	12.452835	0	Italian Restaurant	Café	Pizza Place
180	Seychelles	Victoria	-4.623208	55.452359	1	Café	Fast Food Restaurant	Asian Restaurant
181	Austria	Vienna	48.208354	16.372504	0	Restaurant	Austrian Restaurant	Italian Restaurant
182	Laos	Vientiane	17.964099	102.613371	1	Café	Asian Restaurant	Italian Restaurant
183	Lithuania	Vilnius	54.687046	25.282911	0	Restaurant	Café	Steakhouse
184	Poland	Warsaw	52.233717	21.071411	0	Italian Restaurant	Sushi Restaurant	Eastern European Restaurant
185	United States	Washington D.C.	38.895009	-77.036563	0	Sandwich Place	American Restaurant	Food Truck
186	New Zealand	Wellington	-41.288795	174.777211	0	Café	Restaurant	Vietnamese Restaurant
187	Namibia	Windhoek	-22.574418	17.079123	1	Restaurant	Café	Fast Food Restaurant

	Country	Capital City	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
188	Cameroon	Yaounde	3.868987	11.521334	0	Bakery	Restaurant	Breakfast Spot
189	Armenia	Yerevan	40.177612	44.512585	1	Café	Restaurant	Italian Restaurant
190	Croatia	Zagreb	45.813177	15.977048	1	Café	Restaurant	Mediterranean Restaurant

Visualize the resulting clusters¶

In [24]:

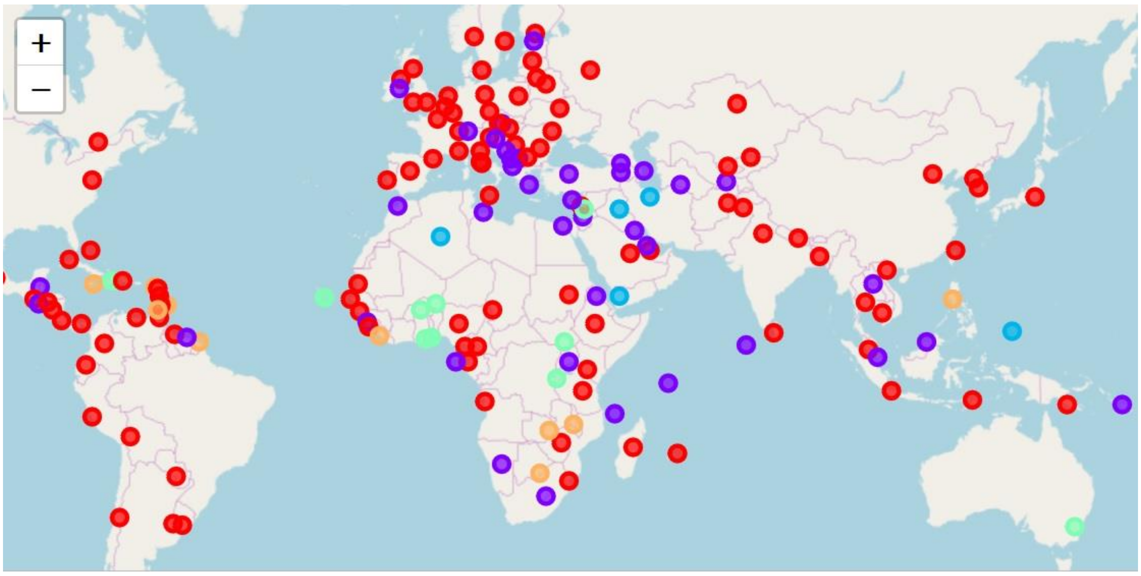
```
# create map
latitude = 0.0
longitude = 0.0
map_clusters = folium.Map(location=[latitude, longitude], zoom_start=2)

# set color scheme for the clusters
x = np.arange(kclusters)
ys = [i + x + (i*x)**2 for i in range(kclusters)]
colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
rainbow = [colors.rgb2hex(i) for i in colors_array]

# add markers to the map
markers_colors = []
for lat, lon, poi, cluster in zip(world_merged['Latitude'], world_merged['Longitude'], world_merged['Capital City'], world_merged['Cluster Labels']):
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
        color=rainbow[cluster-1],
        fill=True,
        fill_color=rainbow[cluster-1],
        fill_opacity=0.7).add_to(map_clusters)

map_clusters
```

Out[24]:



Cluster 1¶

In [25]:

```
cluster1 = world_merged.loc[world_merged['Cluster Labels'] == 0, world_merged.
columns[[1] + list(range(5, world_merged.shape[1]))]]

print(cluster1.shape)

cluster1
```

```
(104, 11)
```

Out[25]:

	Capital City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
1	Abuja	Italian Restaurant	Pizza Place	African Restaurant	Café	BBQ Joint	Bakery
3	Addis Ababa	Ethiopian Restaurant	Restaurant	American Restaurant	Italian Restaurant	Greek Restaurant	Café
6	Amsterdam	Café	Italian Restaurant	Restaurant	Chinese Restaurant	French Restaurant	Bakery
7	Andorra la Vella	Restaurant	Spanish Restaurant	Tapas Restaurant	Burger Joint	Café	Diner
9	Antananarivo	Restaurant	Burger Joint	Café	French Restaurant	African Restaurant	Italian Restaurant
13	Astana	Asian Restaurant	Fast Food Restaurant	Restaurant	Eastern European Restaurant	Café	Modern European Restaurant
14	Asuncion	Restaurant	Café	Fast Food Restaurant	Bakery	Pizza Place	Breakfast Spot
20	Bangkok	Buffet	Food Truck	Fast Food Restaurant	French Restaurant	Food Stand	Food Court

	Capital City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
24	Beijing	Chinese Restaurant	French Restaurant	Asian Restaurant	Café	Peking Duck Restaurant	Fast Food Restaurant
25	Beirut	Café	Restaurant	Middle Eastern Restaurant	Lebanese Restaurant	Fast Food Restaurant	Diner
26	Belfast	Restaurant	Café	Sandwich Place	Pizza Place	Indian Restaurant	Asian Restaurant
27	Belgrade	Restaurant	Café	Italian Restaurant	Bistro	Eastern European Restaurant	BBQ Joint
29	Berlin	German Restaurant	Café	Italian Restaurant	Restaurant	Vegetarian / Vegan Restaurant	Steakhouse
30	Bern	Swiss Restaurant	Café	Restaurant	Italian Restaurant	Bakery	Asian Restaurant
31	Bishkek	Restaurant	Café	Turkish Restaurant	Asian Restaurant	Chinese Restaurant	Italian Restaurant
32	Bissau	African Restaurant	French Restaurant	Restaurant	Portuguese Restaurant	Pizza Place	Bakery
33	Bogota	Restaurant	Café	Italian Restaurant	Latin American Restaurant	Mexican Restaurant	Burger Joint
37	Brussels	Belgian Restaurant	Italian Restaurant	Gastropub	Bakery	Seafood Restaurant	Thai Restaurant
38	Bucharest	Café	Italian Restaurant	Romanian Restaurant	Restaurant	Pizza Place	Gastropub
39	Budapest	Restaurant	Hungarian Restaurant	Italian Restaurant	Café	Eastern European Restaurant	Steakhouse
40	Buenos Aires	Café	Pizza Place	Bakery	Argentinian Restaurant	Restaurant	Burger Joint
43	Caracas	Bakery	Café	Fried Chicken Joint	Restaurant	Fast Food Restaurant	Breakfast Spot
44	Cardiff	Restaurant	Café	Burger Joint	Deli / Bodega	Bakery	Italian Restaurant
47	Chisinau	Café	Pizza Place	Bakery	Romanian Restaurant	Modern European Restaurant	Greek Restaurant
48	Colombo	Bakery	Restaurant	Café	Fast Food Restaurant	Indian Restaurant	Asian Restaurant
50	Copenhagen	Café	Scandinavian Restaurant	Bakery	Burger Joint	French Restaurant	Pizza Place
51	Dakar	Fast Food Restaurant	African Restaurant	Restaurant	Creperie	Food Court	Yunnan Restaurant
53	Dhaka	Asian Restaurant	Chinese Restaurant	Indian Restaurant	Fried Chicken Joint	Café	Thai Restaurant
54	Dili	Indonesian Restaurant	Japanese Restaurant	Chinese Restaurant	Steakhouse	Brazilian Restaurant	Café
56	Dodoma	BBQ Joint	Burger Joint	Restaurant	Chinese Restaurant	Bakery	Fast Food Restaurant

	Capital City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
57	Doha	Restaurant	Asian Restaurant	Cafeteria	Café	Fast Food Restaurant	Turkish Restaurant
60	Edinburgh	Café	Restaurant	Italian Restaurant	Indian Restaurant	Scottish Restaurant	Seafood Restaurant
61	Freetown	Bakery	African Restaurant	Restaurant	Food Court	Pizza Place	Yunnan Restaurant
64	Georgetown	Burger Joint	Brazilian Restaurant	Café	Asian Restaurant	Cajun / Creole Restaurant	Pizza Place
66	Guatemala City	Café	Pizza Place	Fast Food Restaurant	Burger Joint	Steakhouse	Restaurant
67	Hanoi	Vietnamese Restaurant	Café	Noodle House	Vegetarian / Vegan Restaurant	Sandwich Place	French Restaurant
68	Harare	Breakfast Spot	African Restaurant	American Restaurant	French Restaurant	Chinese Restaurant	Fast Food Restaurant
69	Havana	Cuban Restaurant	Café	Italian Restaurant	Bistro	Spanish Restaurant	Restaurant
70	Helsinki	Scandinavian Restaurant	Café	Restaurant	Pizza Place	Chinese Restaurant	Burger Joint
72	Islamabad	BBQ Joint	Bakery	Pizza Place	Yunnan Restaurant	Fast Food Restaurant	French Restaurant
73	Jakarta	Indonesian Restaurant	Café	Fast Food Restaurant	Asian Restaurant	Food Court	Bakery
75	Kabul	Bakery	Afghan Restaurant	French Restaurant	Pizza Place	Food Truck	Food Stand
77	Kathmandu	Restaurant	Asian Restaurant	Café	Italian Restaurant	Fast Food Restaurant	Bakery
78	Khartoum	Restaurant	Pizza Place	Burger Joint	Café	Fast Food Restaurant	Middle Eastern Restaurant
79	Kiev	Café	Italian Restaurant	Bakery	Caucasian Restaurant	Ukrainian Restaurant	Hot Dog Joint
82	Kingstown	Sandwich Place	Restaurant	Café	Chinese Restaurant	Pizza Place	Snack Place
83	Kuala Lumpur	Malay Restaurant	Café	Asian Restaurant	Indian Restaurant	Restaurant	Chinese Restaurant
85	La Paz	Café	Latin American Restaurant	Vegetarian / Vegan Restaurant	Breakfast Spot	Restaurant	Pizza Place
86	Libreville	Bakery	Diner	American Restaurant	Soup Place	French Restaurant	Food Truck
88	Lima	Seafood Restaurant	Restaurant	Fried Chicken Joint	Peruvian Restaurant	Chinese Restaurant	Bakery
89	Lisbon	Portuguese Restaurant	Restaurant	Tapas Restaurant	Bakery	Burger Joint	Café
90	Ljubljana	Restaurant	Café	Eastern European Restaurant	Pizza Place	Bistro	Thai Restaurant
92	London	Bakery	French Restaurant	Steakhouse	Italian Restaurant	Burger Joint	Seafood Restaurant

	Capital City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
93	Luanda	Pizza Place	Restaurant	Portuguese Restaurant	Indian Restaurant	Middle Eastern Restaurant	African Restaurant
95	Luxembourg	Pizza Place	BBQ Joint	Yunnan Restaurant	Fast Food Restaurant	French Restaurant	Food Truck
96	Madrid	Spanish Restaurant	Restaurant	Tapas Restaurant	Café	Pizza Place	Argentinian Restaurant
98	Malabo	Pizza Place	French Restaurant	African Restaurant	Café	Fast Food Restaurant	Food Truck
100	Managua	Restaurant	Bakery	Sandwich Place	Snack Place	Latin American Restaurant	Soup Place
103	Maputo	Fast Food Restaurant	Breakfast Spot	Portuguese Restaurant	Café	Fried Chicken Joint	Restaurant
106	Mexico City	Mexican Restaurant	Taco Place	Café	Restaurant	Bakery	Seafood Restaurant
107	Minsk	Café	Restaurant	Eastern European Restaurant	Gastropub	French Restaurant	Italian Restaurant
109	Monaco	Italian Restaurant	French Restaurant	Restaurant	Café	Mediterranean Restaurant	Sandwich Place
111	Montevideo	Restaurant	Bakery	Café	Pizza Place	BBQ Joint	Sandwich Place
113	Moscow	Italian Restaurant	Restaurant	Café	Bakery	Russian Restaurant	Caucasian Restaurant
115	N'Djamena	Sandwich Place	Middle Eastern Restaurant	Yunnan Restaurant	Fast Food Restaurant	Food Truck	Food Stand
116	Nairobi	African Restaurant	Fast Food Restaurant	Café	Fried Chicken Joint	Restaurant	Italian Restaurant
117	Nassau	Caribbean Restaurant	Fast Food Restaurant	Greek Restaurant	Bakery	Chinese Restaurant	Italian Restaurant
118	New Delhi	Pizza Place	Indian Restaurant	BBQ Joint	Falafel Restaurant	Food Truck	Food Stand
121	Nouakchott	Breakfast Spot	Yunnan Restaurant	Fast Food Restaurant	French Restaurant	Food Truck	Food Stand
123	Oslo	Scandinavian Restaurant	Café	Pizza Place	Restaurant	Italian Restaurant	Sushi Restaurant
124	Ottawa	Café	Restaurant	New American Restaurant	Pizza Place	Mexican Restaurant	Middle Eastern Restaurant
127	Panama City	Latin American Restaurant	Restaurant	Mexican Restaurant	Fast Food Restaurant	Café	Steakhouse
129	Paris	French Restaurant	Bakery	Café	Burger Joint	Restaurant	Pizza Place
130	Phnom Penh	Café	Asian Restaurant	French Restaurant	Malay Restaurant	Cambodian Restaurant	Restaurant
132	Port Louis	Chinese Restaurant	Fast Food Restaurant	Pizza Place	Asian Restaurant	Food	Café

	Capital City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
133	Port Moresby	Diner	Asian Restaurant	Café	Chinese Restaurant	Pizza Place	Filipino Restaurant
136	Port of Spain	Caribbean Restaurant	Chinese Restaurant	Pizza Place	Italian Restaurant	Food Truck	Sushi Restaurant
138	Prague	Restaurant	Pizza Place	Café	Breakfast Spot	Bakery	Chinese Restaurant
141	Pyongyang	Korean Restaurant	Pizza Place	German Restaurant	Café	Chinese Restaurant	Fast Food Restaurant
142	Quito	Restaurant	Breakfast Spot	Diner	Café	Pizza Place	South American Restaurant
145	Riga	Restaurant	Eastern European Restaurant	Café	Modern European Restaurant	Italian Restaurant	Bakery
146	Riyadh	Middle Eastern Restaurant	Donut Shop	Bakery	African Restaurant	Restaurant	North Indian Restaurant
147	Rome	Italian Restaurant	Pizza Place	Sandwich Place	Café	Restaurant	Bakery
148	Roseau	Restaurant	Pizza Place	Caribbean Restaurant	Falafel Restaurant	Fast Food Restaurant	Café
150	Saint John's	Fast Food Restaurant	Café	Middle Eastern Restaurant	Seafood Restaurant	Caribbean Restaurant	Bakery
151	San Jose	Latin American Restaurant	Sandwich Place	Café	Breakfast Spot	Diner	Restaurant
152	San Marino	Pizza Place	Fast Food Restaurant	Restaurant	Food	Café	Yunnan Restaurant
155	Santiago	Sandwich Place	Peruvian Restaurant	Café	Restaurant	Pizza Place	Burger Joint
156	Santo Domingo	Pizza Place	Sandwich Place	Fast Food Restaurant	Asian Restaurant	BBQ Joint	Restaurant
159	Seoul	Korean Restaurant	Café	Chinese Restaurant	Japanese Restaurant	BBQ Joint	Bakery
162	Sofia	Bakery	Italian Restaurant	Restaurant	Café	Vegetarian / Vegan Restaurant	Bistro
163	Stockholm	Scandinavian Restaurant	Café	Bakery	Restaurant	Asian Restaurant	Italian Restaurant
165	Taipei	Café	Japanese Restaurant	Chinese Restaurant	Hotpot Restaurant	Taiwanese Restaurant	Steakhouse
167	Tashkent	Fast Food Restaurant	Pizza Place	Café	Restaurant	Comfort Food Restaurant	Bakery
169	Tegucigalpa	Restaurant	Mexican Restaurant	Steakhouse	Burger Joint	Breakfast Spot	Café
173	Tokyo	Japanese Restaurant	Café	French Restaurant	Chinese Restaurant	Italian Restaurant	Ramen Restaurant
178	Valletta	Mediterranean Restaurant	Café	Restaurant	Italian Restaurant	Pizza Place	Bistro
179	Vatican City	Italian Restaurant	Café	Pizza Place	Sandwich Place	Restaurant	Burger Joint

	Capital City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
181	Vienna	Restaurant	Austrian Restaurant	Italian Restaurant	Café	Bakery	Asian Restaurant
183	Vilnius	Restaurant	Café	Steakhouse	Asian Restaurant	Chinese Restaurant	Burger Joint
184	Warsaw	Italian Restaurant	Sushi Restaurant	Eastern European Restaurant	Restaurant	Café	Greek Restaurant
185	Washington D.C.	Sandwich Place	American Restaurant	Food Truck	Steakhouse	Café	Seafood Restaurant
186	Wellington	Café	Restaurant	Vietnamese Restaurant	Malay Restaurant	Italian Restaurant	Burger Joint
188	Yaounde	Bakery	Restaurant	Breakfast Spot	Comfort Food Restaurant	Café	Diner

Cluster 2¶

In [26]:

```
cluster2 = world_merged.loc[world_merged['Cluster Labels'] == 1, world_merged.
columns[[1] + list(range(5, world_merged.shape[1]))]]

print(cluster2.shape)

cluster2
```

```
(42, 11)
```

Out[26]:

	Capital City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
5	Amman	Café	Middle Eastern Restaurant	Italian Restaurant	Burger Joint	Breakfast Spot	Falafel Restaurant
8	Ankara	Café	Turkish Restaurant	Kebab Restaurant	Sandwich Place	Restaurant	Doner Restaurant
10	Apia	Café	Restaurant	Pizza Place	Burger Joint	Fast Food Restaurant	Indian Restaurant
11	Ashgabat	Café	Restaurant	Snack Place	Fast Food Restaurant	Italian Restaurant	Gastropub
12	Asmara	Café	Restaurant	Asian Restaurant	BBQ Joint	Yunnan Restaurant	Filipino Restaurant
15	Athens	Café	Greek Restaurant	Souvlaki Shop	Falafel Restaurant	Bistro	Taverna
17	Baku	Café	Restaurant	Turkish Restaurant	Steakhouse	Italian Restaurant	Eastern European Restaurant
19	Bandar Seri Begawan	Asian Restaurant	Café	Food Court	Seafood Restaurant	Chinese Restaurant	Italian Restaurant
28	Belmopan	Café	Deli / Bodega	Wings Joint	Restaurant	Chinese Restaurant	Pizza Place

	Capital City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
35	Bratislava	Café	Vegetarian / Vegan Restaurant	Bistro	Bakery	Burger Joint	Indian Restaurant
41	Cairo	Café	Egyptian Restaurant	Falafel Restaurant	Fast Food Restaurant	Middle Eastern Restaurant	Kebab Restaurant
49	Conakry	Italian Restaurant	Sandwich Place	Café	Yunnan Restaurant	Falafel Restaurant	Food Stand
58	Dublin	Café	Restaurant	Italian Restaurant	Vietnamese Restaurant	Burrito Place	Bakery
59	Dushanbe	Restaurant	Café	Burger Joint	Asian Restaurant	Ukrainian Restaurant	Chinese Restaurant
71	Honiara	Café	Japanese Restaurant	Sushi Restaurant	Australian Restaurant	Yunnan Restaurant	Filipino Restaurant
76	Kampala	Café	Fast Food Restaurant	Pizza Place	Restaurant	African Restaurant	Indian Restaurant
84	Kuwait City	Café	Middle Eastern Restaurant	Restaurant	Burger Joint	American Restaurant	Breakfast Spot
99	Male	Café	Restaurant	Pizza Place	Fast Food Restaurant	Asian Restaurant	Thai Restaurant
101	Manama	Café	Indian Restaurant	Asian Restaurant	Middle Eastern Restaurant	Thai Restaurant	BBQ Joint
104	Maseru	Steakhouse	Café	Restaurant	Indian Restaurant	Portuguese Restaurant	Seafood Restaurant
112	Moroni	Falafel Restaurant	Italian Restaurant	Café	Fast Food Restaurant	Food Truck	Food Stand
120	Nicosia	Café	Restaurant	Greek Restaurant	Sandwich Place	Fast Food Restaurant	Cafeteria
122	Nuku'alofa	Café	Italian Restaurant	Restaurant	Asian Restaurant	Yunnan Restaurant	Fast Food Restaurant
128	Paramaribo	Café	Fast Food Restaurant	Burger Joint	Indonesian Restaurant	Fried Chicken Joint	Filipino Restaurant
131	Podgorica	Café	Pizza Place	Restaurant	Fast Food Restaurant	Italian Restaurant	BBQ Joint
134	Port Vila	Café	Caribbean Restaurant	Restaurant	Seafood Restaurant	Falafel Restaurant	Food Stand
140	Pristina	Restaurant	Café	Bakery	Burger Joint	Fast Food Restaurant	Pizza Place
143	Rabat	Café	Moroccan Restaurant	French Restaurant	Diner	Italian Restaurant	Restaurant
153	San Salvador	Café	Pizza Place	Fried Chicken Joint	Diner	Seafood Restaurant	Fast Food Restaurant
157	Sao Tome	Café	Bakery	BBQ Joint	Yunnan Restaurant	French Restaurant	Food Truck
158	Sarajevo	Café	Restaurant	Italian Restaurant	BBQ Joint	Eastern European Restaurant	Fast Food Restaurant
160	Singapore	Japanese Restaurant	Sandwich Place	Restaurant	Café	Falafel Restaurant	Food Stand

	Capital City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
166	Tallinn	Café	Restaurant	Italian Restaurant	Eastern European Restaurant	Asian Restaurant	Burger Joint
168	Tbilisi	Café	Restaurant	Caucasian Restaurant	Bakery	Eastern European Restaurant	Irish Pub
172	Tirana (Tirane)	Café	Italian Restaurant	Restaurant	Bistro	Seafood Restaurant	Fast Food Restaurant
174	Tripoli	Café	Italian Restaurant	Seafood Restaurant	Middle Eastern Restaurant	Bakery	Yunnan Restaurant
177	Vaduz	Café	Swiss Restaurant	Italian Restaurant	Indian Restaurant	German Restaurant	Pizza Place
180	Victoria	Café	Fast Food Restaurant	Asian Restaurant	Pizza Place	Yunnan Restaurant	Food Truck
182	Vientiane	Café	Asian Restaurant	Italian Restaurant	Noodle House	French Restaurant	Indian Restaurant
187	Windhoek	Restaurant	Café	Fast Food Restaurant	German Restaurant	Italian Restaurant	Yunnan Restaurant
189	Yerevan	Café	Restaurant	Italian Restaurant	Fast Food Restaurant	Eastern European Restaurant	Bakery
190	Zagreb	Café	Restaurant	Mediterranean Restaurant	Bistro	Pizza Place	Eastern European Restaurant

Cluster 3¶

In [27]:

```
cluster3 = world_merged.loc[world_merged['Cluster Labels'] == 2, world_merged.
columns[[1] + list(range(5, world_merged.shape[1]))]]

print(cluster3.shape)

cluster3
```

```
(6, 11)
```

Out[27]:

	Capital City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
4	Algiers	Café	Yunnan Restaurant	Fast Food Restaurant	Food Truck	Food Stand	Food Court
16	Baghdad	Café	Yunnan Restaurant	Fast Food Restaurant	Food Truck	Food Stand	Food Court
97	Majuro	Café	Yunnan Restaurant	Fast Food Restaurant	Food Truck	Food Stand	Food Court
105	Melekeok	Café	Yunnan Restaurant	Fast Food Restaurant	Food Truck	Food Stand	Food Court
154	Sana'a	Café	Kebab Restaurant	Fast Food Restaurant	Food Truck	Food Stand	Food Court

	Capital City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
170	Tehran	Café	Persian Restaurant	Breakfast Spot	Pizza Place	Falafel Restaurant	Restaurant

Cluster 4¶

In [28]:

```
cluster4 = world_merged.loc[world_merged['Cluster Labels'] == 3, world_merged.
columns[[1] + list(range(5, world_merged.shape[1]))]]

print(cluster4.shape)

cluster4
```

```
(10, 11)
```

Out[28]:

	Capital City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
2	Accra	Restaurant	African Restaurant	Breakfast Spot	Indian Restaurant	Swiss Restaurant	Donut Shop
42	Canberra	Restaurant	Yunnan Restaurant	Falafel Restaurant	Food Truck	Food Stand	Food Court
52	Damascus	Restaurant	Diner	Café	Persian Restaurant	Fast Food Restaurant	Food Stand
65	Gitega	American Restaurant	Restaurant	Yunnan Restaurant	Fast Food Restaurant	French Restaurant	Food Truck
74	Juba	Restaurant	Eastern European Restaurant	Indian Restaurant	Yunnan Restaurant	Fast Food Restaurant	Food Truck
91	Lome	Restaurant	Japanese Restaurant	Spanish Restaurant	Falafel Restaurant	Food Stand	Food Court
119	Niamey	Restaurant	French Restaurant	African Restaurant	Italian Restaurant	Fast Food Restaurant	Food Truck
125	Ouagadougou	Restaurant	Bistro	Yunnan Restaurant	Filipino Restaurant	French Restaurant	Food Truck
135	Port au Prince	Restaurant	Buffet	Fast Food Restaurant	French Restaurant	Food Truck	Food Stand
139	Praia	Restaurant	Bakery	Café	Yunnan Restaurant	Food Truck	Food Stand

Cluster 5¶

In [29]:

```
cluster5 = world_merged.loc[world_merged['Cluster Labels'] == 4, world_merged.
columns[[1] + list(range(5, world_merged.shape[1]))]]

print(cluster5.shape)

cluster5
```

```
(11, 11)
```

Out[29]:

	Capital City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
23	Basseterre	Fast Food Restaurant	Caribbean Restaurant	Restaurant	American Restaurant	Chinese Restaurant	Pizza Place
36	Bridgetown	Fast Food Restaurant	Caribbean Restaurant	Seafood Restaurant	Restaurant	Burger Joint	Fish & Chips Shop
46	Cayenne	Fast Food Restaurant	Japanese Restaurant	Chinese Restaurant	French Restaurant	Vietnamese Restaurant	Falafel Restaurant
63	Gaborone	Fast Food Restaurant	Restaurant	Steakhouse	Portuguese Restaurant	Bistro	Pizza Place
81	Kingston	Fast Food Restaurant	Restaurant	Bakery	Pizza Place	Café	Caribbean Restaurant
87	Lilongwe	Fast Food Restaurant	Mexican Restaurant	Italian Restaurant	Café	Pizza Place	Yunnan Restaurant
94	Lusaka	Fast Food Restaurant	American Restaurant	Café	Indian Restaurant	Bakery	Yunnan Restaurant
102	Manila	Fast Food Restaurant	Chinese Restaurant	Filipino Restaurant	Café	Pizza Place	Bakery
110	Monrovia	Fast Food Restaurant	Yunnan Restaurant	French Restaurant	Food Truck	Food Stand	Food Court
149	Saint George's	Caribbean Restaurant	Seafood Restaurant	Fast Food Restaurant	Pizza Place	Ethiopian Restaurant	Food Stand
164	Suva	Fast Food Restaurant	Japanese Restaurant	Asian Restaurant	Mexican Restaurant	Chinese Restaurant	Pizza Place

Results and Discussion ¶

173 out of 200 countries were clustered into 5 clusters based on their lineup of restaurants. 18 countries were not clustered because Foursquare API did not return venues. The other 9 countries were not clustered because the list of the countries and capital cities I used has extra characters and congeolocator.geocode failed. I could have cleaned up the list.

Cluster 1 104 countries, like China, Japan and United States, were clustered into Cluster 1.

Cluster 2 42 countries, like Greece, Egypt and Singapore, were clustered into Cluster 2.

Cluster 3 6 countries, like Algeria, Iraq and Iran, were clustered into Cluster 3.

Cluster 4 10 countries, like Ghana, Australia and Syria, were clustered into Cluster 4.

Cluster 5 11 countries, like Jamaica, Philippine and Grenada, were clustered into Cluster 5.

For improvements

- The granularity of categories is not consistent. For example, some restaurants are categorized as 'Restaurant' while others 'Sushi Restaurant'. It may be better to ignore the restaurants categorized as 'Restaurant'.
- I explored only within a radius of 1000 meters from the center of each capital city. Larger area will give us more information and accurate analysis.

- This analysis focuses on the rate of categories rather than the number. You may not be able to easily find a restaurant in some countries. Another analysis is required to avoid going to countries with few restaurants.

Conclusion ¶

This project successfully clustered most of the countries in the world into 5 clusters based on their lineup of restaurants. By selecting one of the countries in the same cluster as yours, you will be able to find similar restaurants there as in your country.
Technically, the geographical information obtained by using geopy.geocoders library and the venue information provided by Foursquare API can be applied not only to restaurants but other categories like arts, outdoors, sights, etc. with a very small modification to the source code above.