Battle_of_neighbourhoods

January 20, 2021

[1]:
```
import numpy as np

import matplotlib as mpl
import matplotlib.pyplot as plt

import pandas as pd
pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)

import json

import requests
from pandas.io.json import json_normalize

import matplotlib.cm as cm
import matplotlib.colors as colors

from sklearn.cluster import KMeans

#!conda install -c conda-forge folium=0.5.0 --yes # uncomment this line if␣
↪needed
import folium
```

[2]:
```
#setting the needed information to call Foursquare API

CLIENT_ID = 'xxx' # your Foursquare ID
CLIENT_SECRET = 'xxx' # your Foursquare Secret
VERSION = '20180605' # Foursquare API version
```

[3]:
```
#setting the location to be searched, search word, and limits on results

toronto_lat = 43.653963
toronto_lon = -79.387207
search_query = 'Hotel'
radius = 500
limit_hotels = 50
```

[4]:
```
#url to be used forr calling the API

url = 'https://api.foursquare.com/v2/venues/search?
 ↪client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.
 ↪format(CLIENT_ID, CLIENT_SECRET, toronto_lat, toronto_lon, VERSION,␣
 ↪search_query, radius, limit_hotels)
url
```

[4]:
```
'https://api.foursquare.com/v2/venues/search?client_id=CNTR5AX4ENVQG0F5PJDPKWWSO
DAXDOYMZ10CBMSG03KJEAOZ&client_secret=CXZCYJ321QOXL3WFR5PH1WZVUGSZXPPDSANPIYJEUQ
4P1HGT&ll=43.653963,-79.387207&v=20180605&query=Hotel&radius=500&limit=50'
```

[5]:
```
results = requests.get(url).json()
```

[6]:
```
# assign relevant part of JSON to venues
hotels = results['response']['venues']

# tranform venues into a dataframe
hotels_df = json_normalize(hotels)
hotels_df.head()
```

```
/home/jupyterlab/conda/envs/python/lib/python3.6/site-
packages/ipykernel_launcher.py:5: FutureWarning: pandas.io.json.json_normalize
is deprecated, use pandas.json_normalize instead
  """
```

[6]:

	id	name	categories	referralId	hasPerk
0	4b68aed1f964a520de862be3	The Rex Hotel Jazz & Blues Bar	[{'id': '4bf58dd8d48988d1e7931735', 'name': 'J…	v-1611159359	False
1	4ab2d511f964a5209b6c20e3	Sheraton Centre Toronto Hotel	[{'id': '4bf58dd8d48988d1fa931735', 'name': 'H…	v-1611159359	False
2	4f343a31e4b0230a3b337a90	VFM Test Hotel	[{'id': '4bf58dd8d48988d1fa931735', 'name': 'H…	v-1611159359	False
3	52ce14b0498e50457ce11780	DoubleTree by Hilton Hotel Toronto Downtown	[{'id': '4bf58dd8d48988d1fa931735', 'name': 'H…	v-1611159359	False
4	4f53fb2ee4b036244bea152f	Stathcona Hotel	[{'id': '4bf58dd8d48988d1fa931735', 'name': 'H…	v-1611159359	False

	location.address	location.crossStreet	location.lat	location.lng
0	194 Queen St W	Queen & St. Patrick	43.650505	-79.388577
1	123 Queen Street West	at York St.	43.650817	-79.384534
2	123 Test Drive	at somewhere St	43.658434	-79.387894
3	108 Chestnut Street	Dundas St W	43.654608	-79.385942
4	NaN	NaN	43.654947	-79.386359

```
                              location.labeledLatLngs  location.distance  \
0  [{'label': 'display', 'lat': 43.65050475544005…                400
1  [{'label': 'display', 'lat': 43.6508169, 'lng'…                411
2  [{'label': 'display', 'lat': 43.658434, 'lng':…                500
3  [{'label': 'display', 'lat': 43.6546083, 'lng'…                124
4  [{'label': 'display', 'lat': 43.654947, 'lng':…                129

  location.postalCode location.cc location.city location.state  \
0             M5V 1Z1          CA       Toronto             ON
1             M5H 2M9          CA       Toronto             ON
2             M2M 2M2          CA       Toronto             ON
3             M5G 1R3          CA       Toronto             ON
4                 NaN          CA       Toronto             ON

  location.country                          location.formattedAddress  \
0           Canada  [194 Queen St W (Queen & St. Patrick), Toronto…
1           Canada  [123 Queen Street West (at York St.), Toronto …
2           Canada  [123 Test Drive (at somewhere St), Toronto ON …
3           Canada  [108 Chestnut Street (Dundas St W), Toronto ON…
4           Canada                                [Toronto ON, Canada]

  venuePage.id location.neighborhood
0     62225795                  NaN
1          NaN                  NaN
2          NaN                  NaN
3          NaN                  NaN
4          NaN                  NaN
```

[7]:
```
# keep only columns that include venue name, and anything that is associated␣
 ↪with location
filtered_columns = ['name', 'categories'] + [col for col in hotels_df.columns␣
 ↪if col.startswith('location.')] + ['id']
hotels_df_filtered = hotels_df.loc[:, filtered_columns]

# 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']

# filter the category for each row
```

```
hotels_df_filtered['categories'] = hotels_df_filtered.apply(get_category_type,␣
↪axis=1)

# clean column names by keeping only last term
hotels_df_filtered.columns = [column.split('.')[-1] for column in␣
↪hotels_df_filtered.columns]

hotels_df_filtered
```

[7]:

	name	categories	address	crossStreet	lat	lng
0	The Rex Hotel Jazz & Blues Bar	Jazz Club	194 Queen St W	Queen & St. Patrick	43.650505	-79.388577
1	Sheraton Centre Toronto Hotel	Hotel	123 Queen Street West	at York St.	43.650817	-79.384534
2	VFM Test Hotel	Hotel	123 Test Drive	at somewhere St	43.658434	-79.387894
3	DoubleTree by Hilton Hotel Toronto Downtown	Hotel	108 Chestnut Street	Dundas St W	43.654608	-79.385942
4	Stathcona Hotel	Hotel	NaN	NaN	43.654947	-79.386359
5	Sheraton Centre Toronto Hotel - Grand Ballroom	Ballroom	123 Queen Street West	at Bay Street	43.651200	-79.384520
6	Sheraton Centre Hotel Club Lounge	Lounge	NaN	NaN	43.651063	-79.384527
7	Be SixFifty Hotel	Hotel	650 Bay Street	Elm Street	43.657120	-79.384560
8	650 Hotel	Hotel	650 Bay Street	Elm Street	43.657046	-79.384411
9	89 Chestnut Residence	College Residence Hall	89 Chestnut St.	at Armoury St.	43.654160	-79.385291
10	Op mijn hotelkamer In Toronto	Hotel	NaN	NaN	43.654235	-79.386208
11	Marriott Downtown at CF Toronto Eaton Centre	Hotel	525 Bay Street	NaN	43.654728	-79.382422
12	Hilton	Hotel	145 Richmond St W	at University Ave	43.649946	-79.385479
13	Lobby Lounge at the Shangri-La Toronto	Lounge	188 University Ave.	NaN	43.649155	-79.386546
14	Grand Ballroom	Event Space				
15	Fitness Centre	Gym				
16	Bistro on Two	Breakfast Spot				
17	Fitness Centre	Gym				
18	Pool	Pool				
19	Shangri-La Toronto	Hotel				
20	Kumar For Men	Men's Store				

```
14        123 Queen St. W    in Sheraton Centre  43.651217 -79.383771
15            525 Bay St.  in Marriott Downtown  43.654690 -79.381739
16        123 Queen St. W    at Sheraton Centre  43.651011 -79.383658
17        Sheraton Centre        123 Queen St W  43.650985 -79.384002
18       Intercontinental                   NaN  43.650975 -79.384053
19    188 University Ave.     at Adelaide St. W  43.649129 -79.386557
20        Sheraton Centre        123 Queen St W  43.651044 -79.383302

                                        labeledLatLngs  distance postalCode  \
0   [{'label': 'display', 'lat': 43.65050475544005…       400    M5V 1Z1
1   [{'label': 'display', 'lat': 43.6508169, 'lng'…       411    M5H 2M9
2   [{'label': 'display', 'lat': 43.658434, 'lng':…       500    M2M 2M2
3   [{'label': 'display', 'lat': 43.6546083, 'lng'…       124    M5G 1R3
4   [{'label': 'display', 'lat': 43.654947, 'lng':…       129        NaN
5   [{'label': 'display', 'lat': 43.65119976208405…       376        NaN
6   [{'label': 'display', 'lat': 43.651063, 'lng':…       388    M5H 2M9
7   [{'label': 'display', 'lat': 43.65711975097656…       411        NaN
8   [{'label': 'display', 'lat': 43.65704642307161…       410        NaN
9   [{'label': 'display', 'lat': 43.65415952834927…       155    M5G 1R1
10  [{'label': 'display', 'lat': 43.654235, 'lng':…        85        NaN
11  [{'label': 'display', 'lat': 43.65472844428402…       394    M5G 2L2
12  [{'label': 'display', 'lat': 43.6499459, 'lng'…       468    M5H 2L2
13  [{'label': 'display', 'lat': 43.64915499986854…       537    M5H 0A3
14  [{'label': 'display', 'lat': 43.65121729803052…       412    M5H 2M9
15  [{'label': 'display', 'lat': 43.65469045595554…       447    M5G 2L2
16  [{'label': 'display', 'lat': 43.65101097139271…       435    M5H 2M9
17  [{'label': 'display', 'lat': 43.65098505166592…       420    M5H 2M9
18  [{'label': 'display', 'lat': 43.65097492328021…       418        NaN
19  [{'label': 'display', 'lat': 43.64912919417502…       540    M5H 0A3
20  [{'label': 'display', 'lat': 43.65104394334362…       452    M5H 2M9

    cc     city    state country  \
0   CA  Toronto       ON  Canada
1   CA  Toronto       ON  Canada
2   CA  Toronto       ON  Canada
3   CA  Toronto       ON  Canada
4   CA  Toronto       ON  Canada
5   CA  Toronto       ON  Canada
6   CA  Toronto       ON  Canada
7   CA  Toronto       ON  Canada
8   CA  Toronto       ON  Canada
9   CA  Toronto       ON  Canada
10  CA  Toronto       ON  Canada
11  CA  Toronto       ON  Canada
12  CA  Toronto       ON  Canada
13  CA  Toronto       ON  Canada
14  CA  Toronto       ON  Canada
```

```
15  CA  Toronto       ON  Canada
16  CA  Toronto       ON  Canada
17  CA  Toronto       ON  Canada
18  CA      NaN  Ontario  Canada
19  CA  Toronto       ON  Canada
20  CA  Toronto       ON  Canada
```

```
                                      formattedAddress        neighborhood  \
0   [194 Queen St W (Queen & St. Patrick), Toronto…                    NaN
1   [123 Queen Street West (at York St.), Toronto …                    NaN
2   [123 Test Drive (at somewhere St), Toronto ON …                    NaN
3   [108 Chestnut Street (Dundas St W), Toronto ON…                    NaN
4                                [Toronto ON, Canada]                    NaN
5   [123 Queen Street West (at Bay Street), Toront…  Financial District
6                        [Toronto ON M5H 2M9, Canada]                    NaN
7   [650 Bay Street (Elm Street), Toronto ON, Canada]                    NaN
8   [650 Bay Street (Elm Street), Toronto ON, Canada]                    NaN
9   [89 Chestnut St. (at Armoury St.), Toronto ON …                    NaN
10                               [Toronto ON, Canada]                    NaN
11       [525 Bay Street, Toronto ON M5G 2L2, Canada]                    NaN
12  [145 Richmond St W (at University Ave), Toront…                    NaN
13  [188 University Ave., Toronto ON M5H 0A3, Canada]                    NaN
14  [123 Queen St. W (in Sheraton Centre), Toronto…                    NaN
15  [525 Bay St. (in Marriott Downtown), Toronto O…                    NaN
16  [123 Queen St. W (at Sheraton Centre), Toronto…                    NaN
17  [Sheraton Centre (123 Queen St W), Toronto ON …                    NaN
18                [Intercontinental, Ontario, Canada]                    NaN
19  [188 University Ave. (at Adelaide St. W), Toro…                    NaN
20  [Sheraton Centre (123 Queen St W), Toronto ON …                    NaN
```

```
                          id
0   4b68aed1f964a520de862be3
1   4ab2d511f964a5209b6c20e3
2   4f343a31e4b0230a3b337a90
3   52ce14b0498e50457ce11780
4   4f53fb2ee4b036244bea152f
5   5545d07e498e2facac03f666
6   5d59d01867339e000897bc88
7   53e3f70c498e8221abc0db4c
8   55de0cf9498e74d841ab26c1
9   4be18ce28dd062b548533e3c
10  4e1632f31f6eb653a5e4aa66
11  4b0563c0f964a5200e5822e3
12  4a9d0a53f964a520d23720e3
13  53856fa411d2061fc84a3d0a
14  4b958a8ef964a5205ca834e3
15  4ff8e46fe4b03705cefc7075
```

```
16  4b99106ef964a520f95e35e3
17  4b5243a7f964a5209d7327e3
18  515e0d16e4b0e99f5f4764a8
19  4e31b74252b131dcebb08743
20  4b0d95f7f964a520a14b23e3
```

[8]:
```
#keeping only venues with the category "Hotel"

hotels_df_filtered = hotels_df_filtered[hotels_df_filtered.categories ==␣
↪'Hotel']
hotels_df_filtered.head(10)
```

[8]:
```
                                        name categories  \
1                  Sheraton Centre Toronto Hotel      Hotel
2                                VFM Test Hotel      Hotel
3    DoubleTree by Hilton Hotel Toronto Downtown      Hotel
4                               Stathcona Hotel      Hotel
7                            Be SixFifty Hotel      Hotel
8                                    650 Hotel      Hotel
10                Op mijn hotelkamer In Toronto      Hotel
11  Marriott Downtown at CF Toronto Eaton Centre      Hotel
12                                      Hilton      Hotel
19                           Shangri-La Toronto      Hotel

                  address        crossStreet        lat        lng  \
1   123 Queen Street West        at York St.  43.650817 -79.384534
2         123 Test Drive    at somewhere St  43.658434 -79.387894
3     108 Chestnut Street        Dundas St W  43.654608 -79.385942
4                     NaN                NaN  43.654947 -79.386359
7         650 Bay Street         Elm Street  43.657120 -79.384560
8         650 Bay Street         Elm Street  43.657046 -79.384411
10                    NaN                NaN  43.654235 -79.386208
11        525 Bay Street                NaN  43.654728 -79.382422
12      145 Richmond St W  at University Ave  43.649946 -79.385479
19    188 University Ave.  at Adelaide St. W  43.649129 -79.386557

                                      labeledLatLngs  distance postalCode  \
1   [{'label': 'display', 'lat': 43.6508169, 'lng'…       411    M5H 2M9
2   [{'label': 'display', 'lat': 43.658434, 'lng':…       500    M2M 2M2
3   [{'label': 'display', 'lat': 43.6546083, 'lng'…       124    M5G 1R3
4   [{'label': 'display', 'lat': 43.654947, 'lng':…       129        NaN
7   [{'label': 'display', 'lat': 43.65711975097656…       411        NaN
8   [{'label': 'display', 'lat': 43.65704642307161…       410        NaN
10  [{'label': 'display', 'lat': 43.654235, 'lng':…        85        NaN
11  [{'label': 'display', 'lat': 43.65472844428402…       394    M5G 2L2
12  [{'label': 'display', 'lat': 43.6499459, 'lng'…       468    M5H 2L2
19  [{'label': 'display', 'lat': 43.64912919417502…       540    M5H 0A3
```

```
    cc     city state country  \
1   CA  Toronto    ON  Canada
2   CA  Toronto    ON  Canada
3   CA  Toronto    ON  Canada
4   CA  Toronto    ON  Canada
7   CA  Toronto    ON  Canada
8   CA  Toronto    ON  Canada
10  CA  Toronto    ON  Canada
11  CA  Toronto    ON  Canada
12  CA  Toronto    ON  Canada
19  CA  Toronto    ON  Canada

                                    formattedAddress neighborhood  \
1   [123 Queen Street West (at York St.), Toronto …          NaN
2   [123 Test Drive (at somewhere St), Toronto ON …          NaN
3   [108 Chestnut Street (Dundas St W), Toronto ON…          NaN
4                                [Toronto ON, Canada]          NaN
7   [650 Bay Street (Elm Street), Toronto ON, Canada]          NaN
8   [650 Bay Street (Elm Street), Toronto ON, Canada]          NaN
10                               [Toronto ON, Canada]          NaN
11      [525 Bay Street, Toronto ON M5G 2L2, Canada]          NaN
12  [145 Richmond St W (at University Ave), Toront…          NaN
19  [188 University Ave. (at Adelaide St. W), Toro…          NaN

                          id
1   4ab2d511f964a5209b6c20e3
2   4f343a31e4b0230a3b337a90
3   52ce14b0498e50457ce11780
4   4f53fb2ee4b036244bea152f
7   53e3f70c498e8221abc0db4c
8   55de0cf9498e74d841ab26c1
10  4e1632f31f6eb653a5e4aa66
11  4b0563c0f964a5200e5822e3
12  4a9d0a53f964a520d23720e3
19  4e31b74252b131dcebb08743
```

[9]:
```
#resetting index

hotels_df_filtered.reset_index(drop=True, inplace=True)
hotels_df_filtered.head(10)
```

[9]:
```
                                        name categories  \
0                 Sheraton Centre Toronto Hotel      Hotel
1                              VFM Test Hotel      Hotel
2   DoubleTree by Hilton Hotel Toronto Downtown      Hotel
3                             Stathcona Hotel      Hotel
```

```
4                            Be SixFifty Hotel      Hotel
5                                  650 Hotel      Hotel
6                Op mijn hotelkamer In Toronto      Hotel
7  Marriott Downtown at CF Toronto Eaton Centre      Hotel
8                                     Hilton      Hotel
9                           Shangri-La Toronto      Hotel

                address        crossStreet        lat        lng  \
0  123 Queen Street West        at York St.  43.650817 -79.384534
1         123 Test Drive    at somewhere St  43.658434 -79.387894
2    108 Chestnut Street        Dundas St W  43.654608 -79.385942
3                    NaN                NaN  43.654947 -79.386359
4         650 Bay Street         Elm Street  43.657120 -79.384560
5         650 Bay Street         Elm Street  43.657046 -79.384411
6                    NaN                NaN  43.654235 -79.386208
7         525 Bay Street                NaN  43.654728 -79.382422
8      145 Richmond St W  at University Ave  43.649946 -79.385479
9    188 University Ave.  at Adelaide St. W  43.649129 -79.386557

                                     labeledLatLngs  distance postalCode  cc  \
0  [{'label': 'display', 'lat': 43.6508169, 'lng'…       411    M5H 2M9  CA
1  [{'label': 'display', 'lat': 43.658434, 'lng':…       500    M2M 2M2  CA
2  [{'label': 'display', 'lat': 43.6546083, 'lng'…       124    M5G 1R3  CA
3  [{'label': 'display', 'lat': 43.654947, 'lng':…       129        NaN  CA
4  [{'label': 'display', 'lat': 43.65711975097656…       411        NaN  CA
5  [{'label': 'display', 'lat': 43.65704642307161…       410        NaN  CA
6  [{'label': 'display', 'lat': 43.654235, 'lng':…        85        NaN  CA
7  [{'label': 'display', 'lat': 43.65472844428402…       394    M5G 2L2  CA
8  [{'label': 'display', 'lat': 43.6499459, 'lng'…       468    M5H 2L2  CA
9  [{'label': 'display', 'lat': 43.64912919417502…       540    M5H 0A3  CA

      city state country                                  formattedAddress  \
0  Toronto    ON  Canada  [123 Queen Street West (at York St.), Toronto …
1  Toronto    ON  Canada  [123 Test Drive (at somewhere St), Toronto ON …
2  Toronto    ON  Canada  [108 Chestnut Street (Dundas St W), Toronto ON…
3  Toronto    ON  Canada                                [Toronto ON, Canada]
4  Toronto    ON  Canada  [650 Bay Street (Elm Street), Toronto ON, Canada]
5  Toronto    ON  Canada  [650 Bay Street (Elm Street), Toronto ON, Canada]
6  Toronto    ON  Canada                                [Toronto ON, Canada]
7  Toronto    ON  Canada       [525 Bay Street, Toronto ON M5G 2L2, Canada]
8  Toronto    ON  Canada  [145 Richmond St W (at University Ave), Toront…
9  Toronto    ON  Canada  [188 University Ave. (at Adelaide St. W), Toro…

  neighborhood                        id
0          NaN  4ab2d511f964a5209b6c20e3
1          NaN  4f343a31e4b0230a3b337a90
2          NaN  52ce14b0498e50457ce11780
```

```
3          NaN  4f53fb2ee4b036244bea152f
4          NaN  53e3f70c498e8221abc0db4c
5          NaN  55de0cf9498e74d841ab26c1
6          NaN  4e1632f31f6eb653a5e4aa66
7          NaN  4b0563c0f964a5200e5822e3
8          NaN  4a9d0a53f964a520d23720e3
9          NaN  4e31b74252b131dcebb08743
```

[10]:
```
#creating a map to view the resulting hotels and their locations
hotels_map = folium.Map(location=[toronto_lat, toronto_lon], zoom_start=15) #␣
 ↪generate map centred around toronto

# add the hotels as blue circle markers
for lat, lng, label in zip(hotels_df_filtered.lat, hotels_df_filtered.lng,␣
 ↪hotels_df_filtered.categories):
    folium.features.CircleMarker(
        [lat, lng],
        radius=5,
        color='blue',
        popup=label,
        fill = True,
        fill_color='blue',
        fill_opacity=0.6
    ).add_to(hotels_map)

# display map
hotels_map
```

[10]:
```
<folium.folium.Map at 0x7fbfeebc1198>
```

[11]:
```
#How many hotels are there near toronto city center?
hotels_df_filtered.shape
```

[11]:
```
(10, 16)
```

[12]:
```
#prepare hotels dataframe for clustering

#adding an ID number to each venue will help in re-assembling the data after␣
 ↪clustering
id_n = list(range(1,len(hotels_df_filtered)+1))
hotels_df_filtered['id_n'] = id_n
hotels_df_filtered.head()
```

```
/home/jupyterlab/conda/envs/python/lib/python3.6/site-
packages/ipykernel_launcher.py:5: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  """
```

[12]:

```
                                      name categories  \
0               Sheraton Centre Toronto Hotel      Hotel
1                           VFM Test Hotel      Hotel
2  DoubleTree by Hilton Hotel Toronto Downtown      Hotel
3                            Stathcona Hotel      Hotel
4                         Be SixFifty Hotel      Hotel

                address      crossStreet        lat        lng  \
0  123 Queen Street West      at York St.  43.650817 -79.384534
1         123 Test Drive  at somewhere St  43.658434 -79.387894
2    108 Chestnut Street      Dundas St W  43.654608 -79.385942
3                    NaN              NaN  43.654947 -79.386359
4         650 Bay Street       Elm Street  43.657120 -79.384560

                                     labeledLatLngs  distance postalCode  cc  \
0  [{'label': 'display', 'lat': 43.6508169, 'lng'…       411    M5H 2M9  CA
1  [{'label': 'display', 'lat': 43.658434, 'lng':…       500    M2M 2M2  CA
2  [{'label': 'display', 'lat': 43.6546083, 'lng'…       124    M5G 1R3  CA
3  [{'label': 'display', 'lat': 43.654947, 'lng':…       129        NaN  CA
4  [{'label': 'display', 'lat': 43.65711975097656…       411        NaN  CA

      city state country                                  formattedAddress  \
0  Toronto    ON  Canada  [123 Queen Street West (at York St.), Toronto …
1  Toronto    ON  Canada  [123 Test Drive (at somewhere St), Toronto ON …
2  Toronto    ON  Canada  [108 Chestnut Street (Dundas St W), Toronto ON…
3  Toronto    ON  Canada                              [Toronto ON, Canada]
4  Toronto    ON  Canada  [650 Bay Street (Elm Street), Toronto ON, Canada]

  neighborhood                        id  id_n
0          NaN  4ab2d511f964a5209b6c20e3     1
1          NaN  4f343a31e4b0230a3b337a90     2
2          NaN  52ce14b0498e50457ce11780     3
3          NaN  4f53fb2ee4b036244bea152f     4
4          NaN  53e3f70c498e8221abc0db4c     5
```

[13]:

```
# The dataset that will be used in K-Means clustering
X=hotels_df_filtered.loc[:,['id_n','lat','lng']]
X.head(10)
```

[13]:

```
   id_n        lat        lng
0     1  43.650817 -79.384534
1     2  43.658434 -79.387894
```

```
2     3  43.654608 -79.385942
3     4  43.654947 -79.386359
4     5  43.657120 -79.384560
5     6  43.657046 -79.384411
6     7  43.654235 -79.386208
7     8  43.654728 -79.382422
8     9  43.649946 -79.385479
9    10  43.649129 -79.386557
```

[14]:
```
#clustering the hotels based on their locations
kmeans = KMeans(n_clusters = 3, init ='k-means++') # hotels will be clustered␣
 ↪in 3 groups
kmeans.fit(X[X.columns[1:3]]) # Compute k-means clustering.
X['cluster_label'] = kmeans.fit_predict(X[X.columns[1:3]])
centers = kmeans.cluster_centers_ # Coordinates of cluster centers.
labels = kmeans.predict(X[X.columns[1:3]]) # Labels of each point
centers = kmeans.cluster_centers_
print(centers) #to show coordinates of the center for each cluster
X.head(10)
```

```
[[ 43.658434   -79.387894  ]
 [ 43.649964   -79.3855231 ]
 [ 43.65544749 -79.38498343]]
```

[14]:

	id_n	lat	lng	cluster_label
0	1	43.650817	-79.384534	1
1	2	43.658434	-79.387894	0
2	3	43.654608	-79.385942	2
3	4	43.654947	-79.386359	2
4	5	43.657120	-79.384560	2
5	6	43.657046	-79.384411	2
6	7	43.654235	-79.386208	2
7	8	43.654728	-79.382422	2
8	9	43.649946	-79.385479	1
9	10	43.649129	-79.386557	1

[15]:
```
#adding the cluster label to each venue in the hotels data frame
X = X[['id_n','cluster_label']]
clustered_hotels = hotels_df_filtered.merge(X, left_on='id_n', right_on='id_n')
clustered_hotels.head()
```

[15]:

	name	categories	\
0	Sheraton Centre Toronto Hotel	Hotel	
1	VFM Test Hotel	Hotel	
2	DoubleTree by Hilton Hotel Toronto Downtown	Hotel	
3	Stathcona Hotel	Hotel	
4	Be SixFifty Hotel	Hotel	

```
                address      crossStreet        lat        lng  \
0  123 Queen Street West      at York St.  43.650817 -79.384534
1         123 Test Drive  at somewhere St  43.658434 -79.387894
2    108 Chestnut Street      Dundas St W  43.654608 -79.385942
3                    NaN              NaN  43.654947 -79.386359
4         650 Bay Street       Elm Street  43.657120 -79.384560

                                     labeledLatLngs  distance postalCode  cc  \
0  [{'label': 'display', 'lat': 43.6508169, 'lng'…       411    M5H 2M9  CA
1  [{'label': 'display', 'lat': 43.658434, 'lng':…       500    M2M 2M2  CA
2  [{'label': 'display', 'lat': 43.6546083, 'lng'…       124    M5G 1R3  CA
3  [{'label': 'display', 'lat': 43.654947, 'lng':…       129        NaN  CA
4  [{'label': 'display', 'lat': 43.65711975097656…       411        NaN  CA

      city state country                                  formattedAddress  \
0  Toronto    ON  Canada  [123 Queen Street West (at York St.), Toronto …
1  Toronto    ON  Canada  [123 Test Drive (at somewhere St), Toronto ON …
2  Toronto    ON  Canada  [108 Chestnut Street (Dundas St W), Toronto ON…
3  Toronto    ON  Canada                          [Toronto ON, Canada]
4  Toronto    ON  Canada  [650 Bay Street (Elm Street), Toronto ON, Canada]

  neighborhood                        id  id_n  cluster_label
0          NaN  4ab2d511f964a5209b6c20e3     1              1
1          NaN  4f343a31e4b0230a3b337a90     2              0
2          NaN  52ce14b0498e50457ce11780     3              2
3          NaN  4f53fb2ee4b036244bea152f     4              2
4          NaN  53e3f70c498e8221abc0db4c     5              2
```

[16]:
```
#map showing each hotel with different color based on the cluster
map_clusters = folium.Map(location=[toronto_lat, toronto_lon], zoom_start=15)

# add markers to the map
marker_colors = ['red','green','blue']
for lat, lon, poi, cluster in zip(clustered_hotels['lat'],␣
 ↪clustered_hotels['lng'], clustered_hotels['name'],␣
 ↪clustered_hotels['cluster_label']):
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
        color=marker_colors[cluster],
        fill=True,
        fill_color=marker_colors[cluster],
        fill_opacity=0.7).add_to(map_clusters)
```

```
map_clusters #show map
```

[16]:
```
<folium.folium.Map at 0x7fbfeeb72390>
```

[17]:
```
#putting the center of each cluster in a dataframe for later analysis
centers_df = pd.DataFrame(centers, columns =['c_lat', 'c_lon'])
cluster_label = list(range(0,3))
centers_df['cluster_label'] = cluster_label
centers_df
```

[17]:

	c_lat	c_lon	cluster_label
0	43.658434	-79.387894	0
1	43.649964	-79.385523	1
2	43.655447	-79.384983	2

[18]:
```
#map of centers
map_centers = folium.Map(location=[toronto_lat, toronto_lon], zoom_start=15)

# add markers to the map in different colors according to cluster
marker_colors = ['red','green','blue']
for lat, lon, poi, cluster in zip(centers_df['c_lat'], centers_df['c_lon'],␣
 ↪centers_df['cluster_label'], centers_df['cluster_label']):
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=10,
        popup=label,
        color=marker_colors[cluster],
        fill=True,
        fill_color=marker_colors[cluster],
        fill_opacity=0.7).add_to(map_centers)

map_centers #display the map showing 3 markers, one for each cluster
```

[18]:
```
<folium.folium.Map at 0x7fbfeeb05588>
```

[19]:
```
#creating the url to call on foursquare API

limit_restaurants = 500

url2 = 'https://api.foursquare.com/v2/venues/explore?
 ↪&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
    CLIENT_ID,
    CLIENT_SECRET,
    VERSION,
    toronto_lat,
    toronto_lon,
```

```
    radius,
    limit_restaurants)
url2
```

[19]:
```
'https://api.foursquare.com/v2/venues/explore?&client_id=CNTR5AX4ENVQG0F5PJDPKWW
SODAXDOYMZ10CBMSG03KJEAOZ&client_secret=CXZCYJ321QOXL3WFR5PH1WZVUGSZXPPDSANPIYJE
UQ4P1HGT&v=20180605&ll=43.653963,-79.387207&radius=500&limit=500'
```

[20]:
```
#sving the results
results2 = requests.get(url2).json()
```

[21]:
```
#getting the needed data and put it in a pandas dataframe
restaurants = results2['response']['groups'][0]['items']

restaurants_df = json_normalize(restaurants) # flatten JSON

# filter columns
filtered_columns = ['venue.name', 'venue.categories', 'venue.location.lat',␣
 ↪'venue.location.lng']
restaurants_df = restaurants_df.loc[:, filtered_columns]

# filter the category for each row
restaurants_df['venue.categories'] = restaurants_df.apply(get_category_type,␣
 ↪axis=1)

# clean columns
restaurants_df.columns = [col.split(".")[-1] for col in restaurants_df.columns]

restaurants_df
```

```
/home/jupyterlab/conda/envs/python/lib/python3.6/site-
packages/ipykernel_launcher.py:4: FutureWarning: pandas.io.json.json_normalize
is deprecated, use pandas.json_normalize instead
  after removing the cwd from sys.path.
```

[21]:

	name
0	Downtown Toronto
1	Japango
2	Textile Museum of Canada
3	Cafe Plenty
4	Poke Guys
5	Karine's
6	Chatime
7	The Library Specialty Coffee
8	Nathan Phillips Square
9	Aboveground Art Supplies
10	Ontario College of Art and Design University (…

11	Four Seasons Centre for the Performing Arts
12	Friendly Stranger - Cannabis Culture Shop
13	Art Gallery of Ontario
14	Canadian Opera Company
15	MUJI
16	Rosalinda
17	JaBistro
18	The Rex Hotel Jazz & Blues Bar
19	Daily Press Juicery
20	HotBlack Coffee
21	M Square Coffee Co
22	The Queen and Beaver Public House
23	The Elm Tree Restaurant
24	Jimmy's Coffee
25	Grange Park
26	Hakata Ikkousha Ramen
27	Norma Ridley Members' Lounge
28	Bulldog On The Block
29	Civello Salon & Spa
30	Old City Hall
31	Sin and Redemption
32	Assembly Chef's Hall
33	Jackman Hall
34	Art Square Gallery & Cafe
35	Midi Bistro
36	Krispy Kreme Doughnut Cafe
37	AGO Espresso Bar
38	Gift Shop
39	Omai
40	The Village Idiot Pub
41	Mo'Ramyun
42	Mangiacake

	categories	lat	lng
0	Neighborhood	43.653232	-79.385296
1	Sushi Restaurant	43.655268	-79.385165
2	Art Museum	43.654396	-79.386500
3	Café	43.654571	-79.389450
4	Poke Place	43.654895	-79.385052
5	Breakfast Spot	43.653699	-79.390743
6	Bubble Tea Shop	43.655542	-79.384684
7	Coffee Shop	43.654413	-79.390902
8	Plaza	43.652270	-79.383516
9	Arts & Crafts Store	43.652646	-79.390925
10	University	43.652803	-79.391074
11	Concert Hall	43.650592	-79.385806
12	Smoke Shop	43.650387	-79.388523

```
13                   Art Gallery  43.654003 -79.392922
14                   Opera House  43.650660 -79.386242
15             Miscellaneous Shop  43.656024 -79.383284
16  Vegetarian / Vegan Restaurant  43.650252 -79.385156
17                Sushi Restaurant  43.649687 -79.388090
18                       Jazz Club  43.650505 -79.388577
19                       Juice Bar  43.650388 -79.388792
20                    Coffee Shop  43.650364 -79.388669
21                    Coffee Shop  43.651218 -79.383555
22                      Gastropub  43.657472 -79.383524
23     Modern European Restaurant  43.657397 -79.383761
24                            Café  43.655827 -79.392042
25                            Park  43.652488 -79.392053
26                Ramen Restaurant  43.650299 -79.388753
27                    Art Gallery  43.652786 -79.392293
28                    Coffee Shop  43.650652 -79.384141
29              Salon / Barbershop  43.650020 -79.389400
30             Monument / Landmark  43.652009 -79.381744
31                             Bar  43.654661 -79.391737
32                     Food Court  43.650579 -79.383412
33                  Movie Theater  43.654090 -79.391705
34                    Art Gallery  43.654227 -79.392536
35              French Restaurant  43.655871 -79.392091
36                     Donut Shop  43.655772 -79.391993
37                            Café  43.654146 -79.392311
38                     Gift Shop  43.654151 -79.392729
39             Japanese Restaurant  43.656006 -79.392494
40                             Pub  43.654394 -79.391622
41              Korean Restaurant  43.656148 -79.392282
42                 Sandwich Place  43.655650 -79.391905
```

[22]:
```
#number of venues recieved
restaurants_df.shape
```

[22]:
```
(43, 4)
```

[23]:
```
#keeping only the venues with categories containing "Restaurant"
restaurants_df_cleaned = restaurants_df[restaurants_df['categories'].str.
 ↪contains('Restaurant')].reset_index(drop=True)
print(restaurants_df_cleaned.shape) #number of restaurants
restaurants_df_cleaned.head()
```

```
(8, 4)
```

[23]:
```
                    name                    categories        lat  \
0                Japango                Sushi Restaurant  43.655268
1              Rosalinda  Vegetarian / Vegan Restaurant  43.650252
```

```
2                 JaBistro               Sushi Restaurant  43.649687
3  The Elm Tree Restaurant     Modern European Restaurant  43.657397
4    Hakata Ikkousha Ramen               Ramen Restaurant  43.650299

         lng
0 -79.385165
1 -79.385156
2 -79.388090
3 -79.383761
4 -79.388753
```

[24]:
```
#creating a map of restaurants
restaurants_map = folium.Map(location=[toronto_lat, toronto_lon],␣
 ↪zoom_start=15) # generate map centred around toronto

# add the hotels as yellow circle markers
for lat, lng, label in zip(restaurants_df_cleaned.lat, restaurants_df_cleaned.
 ↪lng, restaurants_df_cleaned.categories):
    folium.features.CircleMarker(
        [lat, lng],
        radius=5,
        color='yellow',
        popup=label,
        fill = True,
        fill_color='yellow',
        fill_opacity=0.6
    ).add_to(restaurants_map)

# display map
restaurants_map
```

[24]:
```
<folium.folium.Map at 0x7fbfeea75080>
```

[25]:
```
#preparing the data that will be used in KNN classification
Xhat=restaurants_df_cleaned.loc[:,['lat','lng']] #classification based on the␣
 ↪location coordinates
print(Xhat.shape) #number of restaurants
Xhat.head(10) #show dataframe
```

```
(8, 2)
```

[25]:
```
         lat        lng
0  43.655268 -79.385165
1  43.650252 -79.385156
2  43.649687 -79.388090
3  43.657397 -79.383761
4  43.650299 -79.388753
```

```
5  43.655871 -79.392091
6  43.656006 -79.392494
7  43.656148 -79.392282
```

[26]:
```
#the centers dataframe to be used for training KNN
centers_df
```

[26]:
```
       c_lat      c_lon  cluster_label
0  43.658434 -79.387894              0
1  43.649964 -79.385523              1
2  43.655447 -79.384983              2
```

[27]:
```
#selecting classification features and target variable
X_train = centers_df[['c_lat', 'c_lon']].values
Y_train = centers_df['cluster_label']
```

[28]:
```
#importing KNN library
from sklearn.neighbors import KNeighborsClassifier
```

[29]:
```
k = 1 #classification based on the nearest center
#Train Model and Predict
neigh = KNeighborsClassifier(n_neighbors = k).fit(X_train,Y_train)
neigh
```

[29]:
```
KNeighborsClassifier(algorithm='auto', leaf_size=30, metric='minkowski',
           metric_params=None, n_jobs=None, n_neighbors=1, p=2,
           weights='uniform')
```

[30]:
```
#classify the restaurants based on the model created
yhat = neigh.predict(Xhat)
yhat[0:5]
```

[30]:
```
array([2, 1, 1, 2, 1])
```

[31]:
```
#putting the results in a dataframe
KNN_results = pd.DataFrame(yhat, columns=['class'])
KNN_results
```

[31]:
```
   class
0      2
1      1
2      1
3      2
4      1
5      0
6      0
7      0
```

[32]:
```
#joining the results with the restaurants dataframe
restaurants_classified = pd.concat([restaurants_df_cleaned, KNN_results],␣
↪axis=1, sort=False)
```

[33]:
```
#view resulting dataframe
restaurants_classified
```

[33]:

	name	categories	lat	lng	class
0	Japango	Sushi Restaurant	43.655268	-79.385165	2
1	Rosalinda	Vegetarian / Vegan Restaurant	43.650252	-79.385156	1
2	JaBistro	Sushi Restaurant	43.649687	-79.388090	1
3	The Elm Tree Restaurant	Modern European Restaurant	43.657397	-79.383761	2
4	Hakata Ikkousha Ramen	Ramen Restaurant	43.650299	-79.388753	1
5	Midi Bistro	French Restaurant	43.655871	-79.392091	0
6	Omai	Japanese Restaurant	43.656006	-79.392494	0
7	Mo'Ramyun	Korean Restaurant	43.656148	-79.392282	0

[34]:
```
#show restaurante in each area in different color
map_rest_class = folium.Map(location=[toronto_lat, toronto_lon], zoom_start=15)

# add markers to the map
marker_colors = ['red','green','blue']
for lat, lon, poi, cluster in zip(restaurants_classified['lat'],␣
↪restaurants_classified['lng'], restaurants_classified['name'],␣
↪restaurants_classified['class']):
    label = folium.Popup(str(poi) + ' Class ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
        color=marker_colors[cluster],
        fill=True,
        fill_color=marker_colors[cluster],
        fill_opacity=0.7).add_to(map_rest_class)

map_rest_class
```

[34]:
```
<folium.folium.Map at 0x7fbfeea462b0>
```

[35]:
```
#count number of restaurants in each area
restaurants_count = restaurants_classified['class'].value_counts().
 ↪rename_axis('class').reset_index(name='counts')
restaurants_count
```

[35]:

	class	counts
0	1	3
1	0	3
2	2	2

[36]:
```
#counting number of trending restaurants of each unique category
trending_restaurants = restaurants_classified['categories'].value_counts().
 ↪rename_axis('categories').reset_index(name='counts')
trending_restaurants
```

[36]:

	categories	counts
0	Sushi Restaurant	2
1	Japanese Restaurant	1
2	Korean Restaurant	1
3	French Restaurant	1
4	Vegetarian / Vegan Restaurant	1
5	Ramen Restaurant	1
6	Modern European Restaurant	1

[37]:
```
#keeping the five most frequent categories
trending_restaurants = trending_restaurants.head()
```

[38]:
```
#show top five categories
trending_restaurants
```

[38]:

	categories	counts
0	Sushi Restaurant	2
1	Japanese Restaurant	1
2	Korean Restaurant	1
3	French Restaurant	1
4	Vegetarian / Vegan Restaurant	1

[39]:
```
#make "categories" as index for visiualization
trending_restaurants = trending_restaurants.set_index('categories')
trending_restaurants
```

[39]:

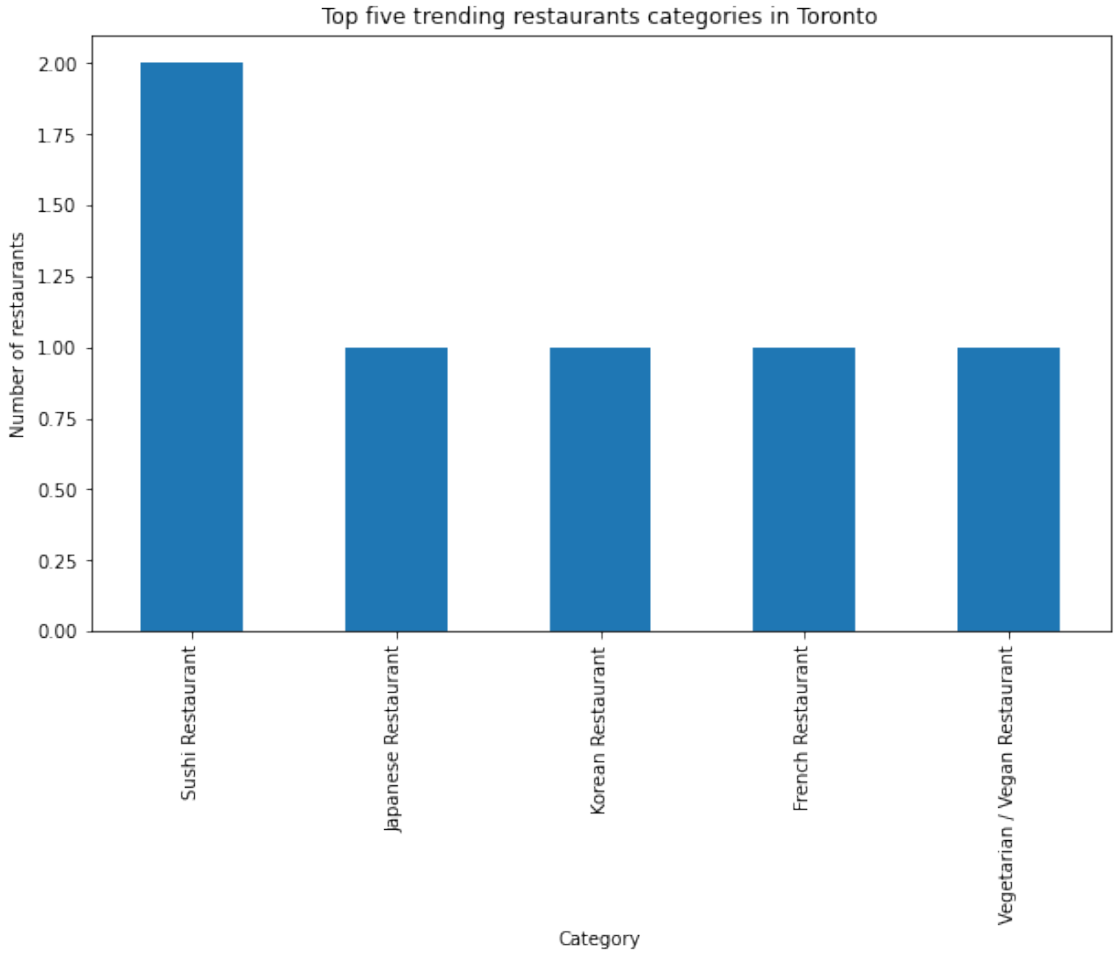
categories	counts
Sushi Restaurant	2
Japanese Restaurant	1

```
Korean Restaurant                   1
French Restaurant                   1
Vegetarian / Vegan Restaurant       1
```

[40]:
```
#plotting the top five restaurant categories
trending_restaurants.plot(kind='bar', figsize=(10, 6), legend=None)

plt.xlabel('Category') # add to x-label to the plot
plt.ylabel('Number of restaurants') # add y-label to the plot
plt.title('Top five trending restaurants categories in Toronto') # add title to␣
 ↪the plot

plt.show()
```



[41]:
```
#counting hotels in each area to evaluate potential
hotels_count = clustered_hotels['cluster_label'].value_counts().
 ↪rename_axis('cluster').reset_index(name='counts')
```

```
hotels_count
```

[41]:

	cluster	counts
0	2	6
1	1	3
2	0	1

[ ]: