Neighborhood Restaurant

Capstone Project Data Analysis part

Description of the data and its sources that will be used to solve the problem

Description of data:

The following data is required to answer the issues of the problem:

- List of neighborhoods restaurant of USA with their geodata (latitud and longitud)
- List of City in USA where the restaurant situated.
- List of Cuisine of each neighborhood in metro city.
- Venue id for each Metro city neighborhood (than can be clustered)
- Rating of each restaurant in particular city.

How the data will be used to solve the problem¶

The data will be used as follows:

- Use Foursquare and geopy data to map top 10 venues for all Metro city neighborhoods and clustered in groups.
- Use Foursquare and geopy data to map the location of particular cuisine, in some form, linked to the subway locations.
- create a map that depicts, for instance, the average cuisine rating and feeback and tips.
- Addresses neighborhood from metro city will be converted to geodata(lat, long) using Geopy-distance and Nominatim.
- Data will be searched in open data sources if available, from food app.

The procesing of these DATA will allow to answer the key questions to make a decision:

- what is the expected rating and popularity of neighborhood at each city?
- what is the area of city with best cuisine and rating that meets criteria of customer?
- What is the distance from work place (Park Ave and 53 rd St) and the tentative future home?
- What are the venues of the two best places to live? How the prices compare?
- How venues distribute among cities neighborhoods and around metro stations?
- Are there tradeoffs between cuisine, Tips and location?
- Any other interesting statistical data findings of the real estate and overall data.

The Data analysis part has done with the help foursquare data. Data downloaded from kaggle.com and put it in Github so that it is for reviewer to check the data.

```
df= pd.read_csv('https://raw.githubusercontent.com/jaanu/Capstone_neighboring/master/us_restaurant
.csv')
```

Here is the below data with Restaurant name , city and cuisine.

	Restaurant Name	City	Cuisines	Latitude	Longitude
0	El Vaquero Mexican Restaurant	Albany	Mexican	40.094257	-83.085363
1	Chick-fil-A	Albany	Fast Food	38.573684	-75.286200
2	Guang Zhou Chinese Restaurant	Albany	Asian, Chinese, Vegetarian	14.068107	-60.955487
3	Harvest Moon	Albany	Pizza, Bar Food, Sandwich	54.312130	-1.690167
4	Hong Kong Cafe	Albany	Chinese, Seafood, Vegetarian	43.001832	-75.977458
5	Locos Grill & Pub	Albany	American, Burger, Sandwich	32.079725	-81.095565
6	Longhorn Steakhouse	Albany	American, Steak	40.540493	-105.076547
7	3 Squares Diner	Albany	American, Breakfast, Diner	31.039941	-84.876857
8	Mama's Boy Restaurant	Athens	Southern	41.890615	-87.629398
9	Sr. Sol 1	Athens	Mexican	42.425646	-2.078956
10	Big City Bread Cafe	Athens	Breakfast, Sandwich	33.959298	-83.384128
11	Taqueria Del Sol	Athens	Mexican, Spanish	36.126600	-86.789405
12	The National	Athens	International, Southern	47.611228	-122.339494
13	The Royal Peasant	Athens	Bar Food	33.938038	-83.387066
14	Transmetropolitan	Athens	Italian, Pizza, Sandwich	33.958487	-83.376544
15	Trappeze Pub	Athens	Burger, Bar Food	33.958494	-83.379016
16	The Bee's Knees	Augusta	International, Tapas, Vegetarian	41.001192	-111.910303
17	Boll Weevil Cafe	Augusta	Desserts, Sandwich, Southern	31.314453	-85.854016
18	Mellow Mushroom	Augusta	Italian, Pizza, Sandwich	33.921787	-84.379529
19	Rhinehart's Oyster Bar	Augusta	Bar Food, Sandwich, Seafood	33.513429	-82.050222

```
        location = geolocator.geocode(address)
        latitude = location.latitude
        longitude =location.longitude
    except:
        latitude=0
        longitude=0
    coordinates = coordinates.append({'Latitude':latitude,'Longitude':longitude},ignore_index=True
)
```

```
df= df.join(coordinates, how ='outer')
df
```

```
import requests
# tranforming json file into a pandas dataframe library
from pandas.io.json import json_normalize



venueid=pd.DataFrame(columns=['venue_id'])
for i,r in df.iterrows():
    search_query=r['Restaurant Name']
    latitude = r['Latitude']
    longitude = r['Longitude']
    try:
        url = 'http://cladiusfernando-eval-test.apigee.net/foursquare/v2/venues/search?client_id={
}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, lat
itude, longitude, VERSION, search_query, radius, LIMIT)
        venues = requests.get(url).json()['response']['venues']
        dataframe = json_normalize(venues)
        venue_id=dataframe['id'][0]
    except:
        print(search_query)
        venue_id=np.nan

    venueid =venueid.append({'venue_id': venue_id}, ignore_index = True)
    """url2 = 'https://api.foursquare.com/v2/venues/{}?client_id={}&client_secret={}&v={}'.format(
venue_id, CLIENT_ID, CLIENT_SECRET, VERSION)
    try:
        rating = requests.get(url2).json()['response']['venue']['rating']
    except:
        rating=0
    print(rating)
    Rating =Rating.append({'rating': rating}, ignore_index = True)"""
df= df.join(venueid, how ='outer')
```

	Restaurant Name	City	Cuisines	Latitude	Longitude	venue_id
0	Chick-fil-A	Albany	Fast Food	38.573684	-75.286200	5ab51bd3446ea6289e2bf8c6
1	Hong Kong Cafe	Albany	Chinese, Seafood, Vegetarian	43.001832	-75.977458	4ea4a5a5be7ba4918f303261
2	Locos Grill & Pub	Albany	American, Burger, Sandwich	32.079725	-81.095565	50c9385ffe1e45fe50f0e06f
3	Longhorn Steakhouse	Albany	American, Steak	40.540493	-105.076547	56d42849cd10d6b76f71a6d1
4	3 Squares Diner	Albany	American, Breakfast, Diner	31.039941	-84.876857	4c3f06100596c928dc0a8578
5	Mama's Boy Restaurant	Athens	Southern	41.890615	-87.629398	57cf6782498e577643efc1bd
6	Big City Bread Cafe	Athens	Breakfast, Sandwich	33.959298	-83.384128	4af18c02f964a5204de121e3
7	Taqueria Del Sol	Athens	Mexican, Spanish	36.126600	-86.789405	50f8a693e4b05404c154501e
8	The National	Athens	International, Southern	47.611228	-122.339494	4dc6d5311f6ef43b8a382bec
9	The Royal Peasant	Athens	Bar Food	33.938038	-83.387066	4b5b9ec3f964a520800b29e3
10	Transmetropolitan	Athens	Italian, Pizza, Sandwich	33.958487	-83.376544	4b0752a1f964a520fffb22e3
11	Trappeze Pub	Athens	Burger, Bar Food	33.958494	-83.379016	4e766e10ae60c3285192db22
12	Boll Weevil Cafe	Augusta	Desserts, Sandwich, Southern	31.314453	-85.854016	4bc0aaa62a89ef3b46def088

```
Rating =pd.DataFrame(columns=['rating'])
for i,r in df.iterrows():

    url2 = 'https://api.foursquare.com/v2/venues/{}?client_id={}&client_secret={}&v={}'.format(r['
venue_id'], CLIENT_ID, CLIENT_SECRET, VERSION)
    try:
        rating = requests.get(url2).json()['response']['venue']['rating']
    except:
        rating=np.nan
    print(rating)
    Rating =Rating.append({'rating': rating}, ignore_index = True)
```

	Restaurant Name	City	Cuisines	Latitude	Longitude	venue_id	rating
0	Chick-fil-A	Albany	Fast Food	38.573684	-75.286200	5ab51bd3446ea6289e2bf8c6	8.1
1	Longhorn Steakhouse	Albany	American, Steak	40.540493	-105.076547	56d42849cd10d6b76f71a6d1	6.9
2	Mama's Boy Restaurant	Athens	Southern	41.890615	-87.629398	57cf6782498e577643efc1bd	6.3
3	Big City Bread Cafe	Athens	Breakfast, Sandwich	33.959298	-83.384128	4af18c02f964a5204de121e3	8.1
4	Taqueria Del Sol	Athens	Mexican, Spanish	36.126600	-86.789405	50f8a693e4b05404c154501e	8.2
5	The National	Athens	International, Southern	47.611228	-122.339494	4dc6d5311f6ef43b8a382bec	6.5
6	The Royal Peasant	Athens	Bar Food	33.938038	-83.387066	4b5b9ec3f964a520800b29e3	9.2
7	Transmetropolitan	Athens	Italian, Pizza, Sandwich	33.958487	-83.376544	4b0752a1f964a520fffb22e3	8.4
8	Trappeze Pub	Athens	Burger, Bar Food	33.958494	-83.379016	4e766e10ae60c3285192db22	9.1
9	Boll Weevil Cafe	Augusta	Desserts, Sandwich, Southern	31.314453	-85.854016	4bc0aaa62a89ef3b46def088	7.3
10	Mellow Mushroom	Augusta	Italian, Pizza, Sandwich	33.921787	-84.379529	4a4797edf964a520dca91fe3	8.6
11	Rhinehart's Oyster Bar	Augusta	Bar Food, Sandwich, Seafood	33.513429	-82.050222	4bc25dbc2a89ef3b7fbcf388	8.3
12	Takosushi	Augusta	Mexican, Southwestern, Sushi	34.849926	-82.399637	4b44135ff964a52023f125e3	7.6

Now we find out all the coordinate of all restaurant using geopy.geocoder.
Lets find out all the venue and extract the venue id . Here we use foursquare data.

Using foursquare data, find out all rating for corresponding venue id. We need this rating to analyze the new restaurant in any area