Capstone Project

March 30, 2022

```python
[1]: import numpy as np
     import pandas as pd
```

```python
[2]: import seaborn as sns
```

```python
[3]: import matplotlib.pyplot as plt
```

```python
[4]: data=pd.read_excel('data.xlsx')
```

```python
[5]: pip install openpyxl
```

```
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: openpyxl in /usr/local/lib/python3.7/site-
packages (3.0.3)
Requirement already satisfied: jdcal in /usr/local/lib/python3.7/site-packages
(from openpyxl) (1.4.1)
Requirement already satisfied: et_xmlfile in /usr/local/lib/python3.7/site-
packages (from openpyxl) (1.0.1)
WARNING: You are using pip version 22.0.3; however, version 22.0.4 is

available.

You should consider upgrading via the '/usr/local/bin/python3.7 -m pip install

--upgrade pip' command.

Note: you may need to restart the kernel to use updated packages.
```

```python
[6]: cc=pd.read_excel('Country-Code.xlsx')
```

```python
[7]: data_restaurant=pd.merge(data,cc,on='Country Code' , how='left')
     data_restaurant.head()
```

```
[7]:    Restaurant ID             Restaurant Name  Country Code      City  \
     0        7402935                        Skye            94   Jakarta
     1        7410290     Satoo - Hotel Shangri-La           94   Jakarta
     2        7420899                  Sushi Masa            94   Jakarta
     3        7421967              3 Wise Monkeys            94   Jakarta
     4        7422489  Avec Moi Restaurant and Bar           94   Jakarta
```

```
                                      Address  \
0  Menara BCA, Lantai 56, Jl. MH. Thamrin, Thamri…
1                Hotel Shangri-La, Jl. Jend. Sudirman
2                    Jl. Tuna Raya No. 5, Penjaringan
3                    Jl. Suryo No. 26, Senopati, Jakarta
4  Gedung PIC, Jl. Teluk Betung 43, Thamrin, Jakarta


                      Locality                         Locality Verbose  \
0  Grand Indonesia Mall, Thamrin  Grand Indonesia Mall, Thamrin, Jakarta
1      Hotel Shangri-La, Sudirman      Hotel Shangri-La, Sudirman, Jakarta
2                    Penjaringan                       Penjaringan, Jakarta
3                      Senopati                          Senopati, Jakarta
4                       Thamrin                           Thamrin, Jakarta


    Longitude  Latitude                 Cuisines  Average Cost for two  \
0  106.821999 -6.196778       Italian, Continental                800000
1  106.818961 -6.203292  Asian, Indonesian, Western                800000
2  106.800144 -6.101298           Sushi, Japanese                500000
3  106.813400 -6.235241                  Japanese                450000
4  106.821023 -6.196270           French, Western                350000


                  Currency Has Table booking Has Online delivery  Price range  \
0  Indonesian Rupiah(IDR)               No                   No            3
1  Indonesian Rupiah(IDR)               No                   No            3
2  Indonesian Rupiah(IDR)               No                   No            3
3  Indonesian Rupiah(IDR)               No                   No            3
4  Indonesian Rupiah(IDR)               No                   No            3


   Aggregate rating Rating color Rating text  Votes    Country
0              4.1        Green    Very Good   1498  Indonesia
1              4.6   Dark Green    Excellent    873  Indonesia
2              4.9   Dark Green    Excellent    605  Indonesia
3              4.2        Green    Very Good    395  Indonesia
4              4.3        Green    Very Good    243  Indonesia
```

[8]: `data_restaurant.info()`

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 9551 entries, 0 to 9550
Data columns (total 20 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   Restaurant ID          9551 non-null   int64
 1   Restaurant Name        9550 non-null   object
 2   Country Code           9551 non-null   int64
 3   City                   9551 non-null   object
 4   Address                9551 non-null   object
```

```
 5   Locality            9551 non-null   object
 6   Locality Verbose    9551 non-null   object
 7   Longitude           9551 non-null   float64
 8   Latitude            9551 non-null   float64
 9   Cuisines            9542 non-null   object
 10  Average Cost for two 9551 non-null  int64
 11  Currency            9551 non-null   object
 12  Has Table booking   9551 non-null   object
 13  Has Online delivery 9551 non-null   object
 14  Price range         9551 non-null   int64
 15  Aggregate rating    9551 non-null   float64
 16  Rating color        9551 non-null   object
 17  Rating text         9551 non-null   object
 18  Votes               9551 non-null   int64
 19  Country             9551 non-null   object
dtypes: float64(3), int64(5), object(12)
memory usage: 1.5+ MB
```

[9]: `data_restaurant.isnull().sum()`

```
[9]: Restaurant ID         0
     Restaurant Name       1
     Country Code          0
     City                  0
     Address               0
     Locality              0
     Locality Verbose      0
     Longitude             0
     Latitude              0
     Cuisines              9
     Average Cost for two  0
     Currency              0
     Has Table booking     0
     Has Online delivery   0
     Price range           0
     Aggregate rating      0
     Rating color          0
     Rating text           0
     Votes                 0
     Country               0
     dtype: int64
```

[10]: `data_restaurant[data_restaurant['Restaurant Name'].isnull()]`

```
[10]:       Restaurant ID Restaurant Name  Country Code      City  \
      1646         113702             NaN             1  Ahmedabad
```

```
                                         Address  Locality  \
1646  Opposite Sindhu Bhawan, Bodakdev, Ahmedabad  Bodakdev


           Locality Verbose  Longitude   Latitude  \
1646  Bodakdev, Ahmedabad  72.501764  23.040163


                                   Cuisines  Average Cost for two  \
1646  North Indian, Continental, Mexican, Italian                   800


                Currency Has Table booking Has Online delivery  Price range  \
1646  Indian Rupees(Rs.)               No                 No            3


      Aggregate rating Rating color Rating text  Votes Country
1646               4.1        Green   Very Good    769   India
```

```python
data_restaurant.dropna(axis=0,subset=['Restaurant Name'],inplace=True)
data_restaurant.reset_index(drop=True,inplace=True)
data_restaurant[data_restaurant['Cuisines'].isnull()]
```

```
[11]:       Restaurant ID              Restaurant Name  Country Code  \
      9082       17374552               Corkscrew Cafe           216
      9085       17501439                     Dovetail           216
      9093       17059060                    Hillstone           216
      9405       17284158            Jimmie's Hot Dogs           216
      9493       17142698              Leonard's Bakery           216
      9503       17616465       Tybee Island Social Club          216
      9532       17284105                 Cookie Shoppe           216
      9534       17284211  Pearly's Famous Country Cookng         216
      9538       17606621            HI Lite Bar & Lounge          216


                  City                                          Address  \
      9082    Gainesville                   51 W Main St, Dahlonega, GA 30533
      9085         Macon                      543 Cherry St, Macon, GA 31201
      9093        Orlando  215 South Orlando Avenue, Winter Park, FL 32789
      9405         Albany                204 S Jackson St, Albany, GA 31701
      9493  Rest of Hawaii          933 Kapahulu Ave, Honolulu, HI 96816
      9503       Savannah          1311 Butler Ave, Tybee Island, GA 31328
      9532         Albany                115 N Jackson St, Albany, GA 31701
      9534         Albany              814 N Slappey Blvd, Albany, GA 31701
      9538         Miller          109 N Broadway Ave, Miller, SD 57362


              Locality        Locality Verbose    Longitude    Latitude Cuisines  \
      9082    Dahlonega    Dahlonega, Gainesville  -83.985800  34.531800      NaN
      9085        Macon            Macon, Macon  -83.627979  32.836410      NaN
      9093  Winter Park     Winter Park, Orlando  -81.365260  28.596682      NaN
      9405       Albany            Albany, Albany  -84.153400  31.575100      NaN
      9493       Kaimuki  Kaimuki, Rest of Hawaii -157.813432  21.284586      NaN
```

```
9503   Tybee Island   Tybee Island, Savannah   -80.848297   31.995810   NaN
9532        Albany          Albany, Albany   -84.154000   31.577200   NaN
9534        Albany          Albany, Albany   -84.175900   31.588200   NaN
9538        Miller          Miller, Miller   -98.989100   44.515800   NaN
```

	Average Cost for two	Currency	Has Table booking	Has Online delivery
9082	40	Dollar($)	No	No
9085	40	Dollar($)	No	No
9093	40	Dollar($)	No	No
9405	10	Dollar($)	No	No
9493	10	Dollar($)	No	No
9503	10	Dollar($)	No	No
9532	0	Dollar($)	No	No
9534	0	Dollar($)	No	No
9538	0	Dollar($)	No	No

	Price range	Aggregate rating	Rating color	Rating text	Votes
9082	3	3.9	Yellow	Good	209
9085	3	3.8	Yellow	Good	102
9093	3	4.4	Green	Very Good	1158
9405	1	3.9	Yellow	Good	160
9493	1	4.7	Dark Green	Excellent	707
9503	1	3.9	Yellow	Good	309
9532	1	3.4	Orange	Average	34
9534	1	3.4	Orange	Average	36
9538	1	3.4	Orange	Average	11

	Country
9082	United States
9085	United States
9093	United States
9405	United States
9493	United States
9503	United States
9532	United States
9534	United States
9538	United States

0.0.1 Now we fix cuisines errors

```
[13]: data_restaurant['Cuisines'].fillna('Other',inplace=True)
```

0.0.2 Check for duplicates

```
[15]: data_restaurant.duplicated().sum()
```

```
[15]: 0
```

0.1 Exploratory Data Analysis

0.1.1 Geographical distribution of the restautants

```
[18]: geo_dis=data_restaurant.groupby(['Country Code','Country']).
      ↪agg(Count=('Restaurant ID','count'))
      geo_dis.sort_values(by='Count',ascending=False)
```

```
[18]:                        Count
      Country Code Country
      1            India            8651
      216          United States     434
      215          United Kingdom     80
      30           Brazil             60
      189          South Africa       60
      214          UAE                60
      148          New Zealand        40
      208          Turkey             34
      14           Australia          24
      162          Phillipines        22
      94           Indonesia          21
      166          Qatar              20
      184          Singapore          20
      191          Sri Lanka          20
      37           Canada              4
```

0.1.2 India has the highest number of restaurants

```
[20]: city_distr=data_restaurant.groupby(['Country','City']).agg(Count=('Restaurant␣
      ↪ID','count'))
      city_distr.describe()
```

```
[20]:              Count
      count   141.000000
      mean     67.730496
      std     476.723952
      min       1.000000
      25%       1.000000
```

```
50%       20.000000
75%       20.000000
max     5473.000000
```

[21]: `city_distr.sort_values(by='Count',ascending=False)`

[21]:
```
                            Count
Country       City
India         New Delhi      5473
              Gurgaon        1118
              Noida          1080
              Faridabad       251
              Ghaziabad        25
...                          ...
              Panchkula         1
Australia     Balingup          1
Indonesia     Bandung           1
Phillipines   Quezon City       1
United States Winchester Bay    1

[141 rows x 1 columns]
```

0.2 New Delhi has the max number of restaurants, many cities have only one

0.3 Further analysis on cities with lowest number of restaurants

[24]:
```
low_cit=city_distr[city_distr['Count']==1]
low_cit.info()
low_cit
```

```
<class 'pandas.core.frame.DataFrame'>
MultiIndex: 46 entries, ('Australia', 'Armidale') to ('United States',
'Winchester Bay')
Data columns (total 1 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Count   46 non-null     int64
dtypes: int64(1)
memory usage: 1.8+ KB
```

[24]:
```
                        Count
Country     City
Australia   Armidale       1
            Balingup       1
            Beechworth     1
            Dicky Beach    1
```

	East Ballina	1
	Flaxton	1
	Forrest	1
	Huskisson	1
	Inverloch	1
	Lakes Entrance	1
	Lorn	1
	Macedon	1
	Mayfield	1
	Middleton Beach	1
	Montville	1
	Palm Cove	1
	Paynesville	1
	Penola	1
	Phillip Island	1
	Tanunda	1
	Trentham East	1
	Victor Harbor	1
Canada	Chatham-Kent	1
	Consort	1
	Vineland Station	1
	Yorkton	1
India	Mohali	1
	Panchkula	1
Indonesia	Bandung	1
Phillipines	Quezon City	1
	Tagaytay City	1
South Africa	Randburg	1
United States	Clatskanie	1
	Cochrane	1
	Fernley	1
	Lakeview	1
	Lincoln	1
	Mc Millan	1
	Miller	1
	Monroe	1
	Ojo Caliente	1
	Potrero	1
	Princeton	1
	Vernonia	1
	Weirton	1
	Winchester Bay	1

0.4 Ratio between restaurants that allow table booking vs that do not allow table booking.

```
[26]: data_restaurant1=data_restaurant.copy()
      data_restaurant1.columns
```

```
[26]: Index(['Restaurant ID', 'Restaurant Name', 'Country Code', 'City', 'Address',
             'Locality', 'Locality Verbose', 'Longitude', 'Latitude', 'Cuisines',
             'Average Cost for two', 'Currency', 'Has Table booking',
             'Has Online delivery', 'Price range', 'Aggregate rating',
             'Rating color', 'Rating text', 'Votes', 'Country'],
            dtype='object')
```

```
[27]: dummy=['Has Table booking' ,'Has Online delivery' ]
      data_restaurant1=pd.get_dummies(data_restaurant1,columns=dummy,drop_first=True)
```

```
[ ]:
```

```
[28]: data_restaurant1.head(5)
```

```
[28]:    Restaurant ID              Restaurant Name  Country Code     City  \
      0        7402935                         Skye            94  Jakarta
      1        7410290     Satoo - Hotel Shangri-La           94  Jakarta
      2        7420899                   Sushi Masa           94  Jakarta
      3        7421967                3 Wise Monkeys           94  Jakarta
      4        7422489   Avec Moi Restaurant and Bar          94  Jakarta

                                                  Address  \
      0  Menara BCA, Lantai 56, Jl. MH. Thamrin, Thamri…
      1              Hotel Shangri-La, Jl. Jend. Sudirman
      2                  Jl. Tuna Raya No. 5, Penjaringan
      3                Jl. Suryo No. 26, Senopati, Jakarta
      4  Gedung PIC, Jl. Teluk Betung 43, Thamrin, Jakarta

                             Locality                         Locality Verbose  \
      0  Grand Indonesia Mall, Thamrin  Grand Indonesia Mall, Thamrin, Jakarta
      1     Hotel Shangri-La, Sudirman      Hotel Shangri-La, Sudirman, Jakarta
      2                    Penjaringan                      Penjaringan, Jakarta
      3                       Senopati                        Senopati, Jakarta
      4                        Thamrin                          Thamrin, Jakarta

          Longitude  Latitude                    Cuisines  Average Cost for two  \
      0  106.821999 -6.196778       Italian, Continental                 800000
      1  106.818961 -6.203292  Asian, Indonesian, Western                800000
      2  106.800144 -6.101298             Sushi, Japanese                 500000
      3  106.813400 -6.235241                    Japanese                 450000
      4  106.821023 -6.196270             French, Western                 350000
```

```
               Currency  Price range  Aggregate rating Rating color  \
0  Indonesian Rupiah(IDR)            3               4.1        Green
1  Indonesian Rupiah(IDR)            3               4.6   Dark Green
2  Indonesian Rupiah(IDR)            3               4.9   Dark Green
3  Indonesian Rupiah(IDR)            3               4.2        Green
4  Indonesian Rupiah(IDR)            3               4.3        Green

   Rating text  Votes     Country  Has Table booking_Yes  \
0    Very Good   1498   Indonesia                      0
1    Excellent    873   Indonesia                      0
2    Excellent    605   Indonesia                      0
3    Very Good    395   Indonesia                      0
4    Very Good    243   Indonesia                      0

   Has Online delivery_Yes
0                        0
1                        0
2                        0
3                        0
4                        0
```

[29]:
```python
table_booking=data_restaurant1[data_restaurant1['Has Table
 ↪booking_Yes']==1]['Restaurant ID'].count()
table_bookingg=data_restaurant1[data_restaurant1['Has Table
 ↪booking_Yes']==0]['Restaurant ID'].count()
```

[30]:
```python
round((table_booking/table_bookingg),2)
```

[30]: 0.14

[31]:
```python
#Percentage of restaurants with online delivery
restaurants_online=data_restaurant1[data_restaurant1['Has Online delivery_Yes']
 ↪==1]['Restaurant ID'].count()
restaurants_onlinee=data_restaurant1[data_restaurant1['Has Online
 ↪delivery_Yes']==0]['Restaurant ID'].count()
```

[32]:
```python
print('Percentage of restaurants providing online delivery : {} %'.
 ↪format((round(restaurants_online/len(data_restaurant1),3)*100)))
```

Percentage of restaurants providing online delivery : 25.7 %

[33]:
```python
#Calculate the difference in number of votes for the restaurants that deliver
 ↪and the restaurants that do not deliver
```

[34]:
```python
restaurant_deliver=data_restaurant1[data_restaurant1['Has Table
 ↪booking_Yes']==1]['Votes'].sum()
```

```
restaurant_delivern=data_restaurant1[data_restaurant1['Has Table␣
 ↪booking_Yes']==0]['Votes'].sum()
restaurant_deliver-restaurant_delivern
```

[34]: -680082

0.4.1 Therefore difference in number of votes for restautant that deliver and don't deliver are :6880082

0.5 What are the top 10 cuisines served across cities?What is the maximum and minimum number of cuisines that a restaurant serves? Also, which is the most served cuisine across the restaurant for each city?

[37]: `data_restaurant.columns`

```
[37]: Index(['Restaurant ID', 'Restaurant Name', 'Country Code', 'City', 'Address',
       'Locality', 'Locality Verbose', 'Longitude', 'Latitude', 'Cuisines',
       'Average Cost for two', 'Currency', 'Has Table booking',
       'Has Online delivery', 'Price range', 'Aggregate rating',
       'Rating color', 'Rating text', 'Votes', 'Country'],
      dtype='object')
```

[38]: `data_restaurant.Cuisines`

```
[38]: 0                               Italian, Continental
      1                          Asian, Indonesian, Western
      2                                     Sushi, Japanese
      3                                            Japanese
      4                                     French, Western
                            …
      9545                  Chinese, North Indian, Fast Food
      9546                    Indian, Chinese, Continental
      9547    Cafe, Continental, Desserts, Ice Cream, Italia…
      9548                                     Street Food
      9549                         Chinese, North Indian
      Name: Cuisines, Length: 9550, dtype: object
```

0.6 Top cusines served across cities

```
[40]: data_restaurant.columns
      cuisines = data_restaurant['Cuisines'].apply(lambda x: pd.Series(x.split(',')))
      cuisines.columns =␣
       ↪['Cuisine_1','Cuisine_2','Cuisine_3','Cuisine_4','Cuisine_5','Cuisine_6','Cuisine_7','Cuisi
      cuisines.tail()
```

11

```
[40]:            Cuisine_1       Cuisine_2       Cuisine_3   Cuisine_4 Cuisine_5  \
      9545        Chinese    North Indian       Fast Food         NaN       NaN
      9546         Indian         Chinese     Continental         NaN       NaN
      9547           Cafe     Continental        Desserts   Ice Cream    Italian
      9548    Street Food             NaN             NaN         NaN       NaN
      9549        Chinese    North Indian             NaN         NaN       NaN


            Cuisine_6 Cuisine_7 Cuisine_8
      9545       NaN       NaN       NaN
      9546       NaN       NaN       NaN
      9547  Beverages       NaN       NaN
      9548       NaN       NaN       NaN
      9549       NaN       NaN       NaN
```

```
[41]: df_cuisines = pd.concat([data_restaurant1,cuisines],axis=1)
      df_cuisines.head()
```

```
[41]:    Restaurant ID             Restaurant Name  Country Code     City  \
      0        7402935                        Skye            94  Jakarta
      1        7410290    Satoo - Hotel Shangri-La            94  Jakarta
      2        7420899                  Sushi Masa            94  Jakarta
      3        7421967               3 Wise Monkeys            94  Jakarta
      4        7422489  Avec Moi Restaurant and Bar            94  Jakarta


                                            Address  \
      0  Menara BCA, Lantai 56, Jl. MH. Thamrin, Thamri…
      1            Hotel Shangri-La, Jl. Jend. Sudirman
      2               Jl. Tuna Raya No. 5, Penjaringan
      3            Jl. Suryo No. 26, Senopati, Jakarta
      4  Gedung PIC, Jl. Teluk Betung 43, Thamrin, Jakarta


                              Locality                     Locality Verbose  \
      0  Grand Indonesia Mall, Thamrin  Grand Indonesia Mall, Thamrin, Jakarta
      1    Hotel Shangri-La, Sudirman      Hotel Shangri-La, Sudirman, Jakarta
      2                  Penjaringan                      Penjaringan, Jakarta
      3                     Senopati                         Senopati, Jakarta
      4                      Thamrin                          Thamrin, Jakarta


         Longitude  Latitude                    Cuisines  … \
      0  106.821999 -6.196778        Italian, Continental  …
      1  106.818961 -6.203292  Asian, Indonesian, Western  …
      2  106.800144 -6.101298             Sushi, Japanese  …
      3  106.813400 -6.235241                    Japanese  …
      4  106.821023 -6.196270             French, Western  …


         Has Table booking_Yes Has Online delivery_Yes  Cuisine_1     Cuisine_2  \
      0                      0                       0    Italian   Continental
```

```
   1                       0                        0        Asian     Indonesian
   2                       0                        0        Sushi       Japanese
   3                       0                        0     Japanese            NaN
   4                       0                        0       French        Western

      Cuisine_3 Cuisine_4 Cuisine_5 Cuisine_6  Cuisine_7  Cuisine_8
   0      NaN       NaN       NaN       NaN        NaN        NaN
   1  Western       NaN       NaN       NaN        NaN        NaN
   2      NaN       NaN       NaN       NaN        NaN        NaN
   3      NaN       NaN       NaN       NaN        NaN        NaN
   4      NaN       NaN       NaN       NaN        NaN        NaN

   [5 rows x 28 columns]
```

[42]:
```python
rest_cuisine = pd.DataFrame(df_cuisines[['Restaurant␣
 ↪Name','City','Cuisine_1','Cuisine_2','Cuisine_3','Cuisine_4',
                                    ␣
 ↪'Cuisine_5','Cuisine_6','Cuisine_7','Cuisine_8']])
rest_cuisine_stack=pd.DataFrame(rest_cuisine.stack()) #stacking the columns
rest_cuisine.head()
```

[42]:
```
                     Restaurant Name     City Cuisine_1     Cuisine_2 Cuisine_3  \
   0                            Skye  Jakarta    Italian   Continental       NaN
   1      Satoo - Hotel Shangri-La  Jakarta      Asian    Indonesian   Western
   2                      Sushi Masa  Jakarta      Sushi      Japanese       NaN
   3                  3 Wise Monkeys  Jakarta   Japanese           NaN       NaN
   4  Avec Moi Restaurant and Bar  Jakarta     French       Western       NaN

      Cuisine_4 Cuisine_5 Cuisine_6 Cuisine_7 Cuisine_8
   0      NaN       NaN       NaN       NaN       NaN
   1      NaN       NaN       NaN       NaN       NaN
   2      NaN       NaN       NaN       NaN       NaN
   3      NaN       NaN       NaN       NaN       NaN
   4      NaN       NaN       NaN       NaN       NaN
```

[43]:
```python
keys = [c for c in rest_cuisine  if c.startswith('Cuisine')]
b=pd.melt(rest_cuisine, id_vars='Restaurant Name', value_vars=keys,␣
 ↪value_name='Cuisines')
#melting the stack into one row
max_rate1=pd.DataFrame(b.groupby(by=['Restaurant Name','variable','Cuisines']).
 ↪size().reset_index())
#find the highest restuarant in the city
max_rate1
del max_rate1['variable']
max_rate1.columns=['Restaurant Name','Cuisines','Count']
max_rate1.head(20)
```

```
[43]:            Restaurant Name        Cuisines   Count
      0                    12212       Fast Food       1
      1              Let's Burrrp         Chinese       1
      2              Let's Burrrp   North Indian       1
      3                      #45            Cafe       1
      4             #Dilliwaala6   North Indian       1
      5             #InstaFreeze       Ice Cream       1
      6              #OFF Campus            Cafe       1
      7              #OFF Campus     Continental       1
      8              #OFF Campus         Italian       1
      9              #OFF Campus       Fast Food       1
      10            #Urban CafÌ©   North Indian       1
      11            #Urban CafÌ©         Chinese       1
      12            #Urban CafÌ©         Italian       1
      13                #hashtag            Cafe       1
      14                  'Ohana        Hawaiian       1
      15        10 Downing Street   North Indian       2
      16        10 Downing Street         Chinese       2
      17        10 To 10 In Delhi          Indian       1
      18        10 To 10 In Delhi            Cafe       1
      19   11th Avenue Cafe Bistro            Cafe       1
```

```
[ ]:
```

```
[44]:  max_rate1.sort_values('Count',ascending=False)
```

```
[44]:            Restaurant Name        Cuisines   Count
      2479        Cafe Coffee Day            Cafe      83
      4596         Domino's Pizza           Pizza      79
      4597         Domino's Pizza       Fast Food      78
      12984                Subway           Salad      63
      12985                Subway    Healthy Food      63
      …                        …               …       …
      5568    Gabbar's Bar & Kitchen         Chinese       1
      5569    Gabbar's Bar & Kitchen         Mexican       1
      5570    Gabbar's Bar & Kitchen         Italian       1
      5571            Gaga Manjero   World Cuisine       1
      15963     Ìàukura€Ùa Sofras€±          Izgara       1

      [15964 rows x 3 columns]
```

0.6.1 max is 83, min is 1

```
[46]: rating = data_restaurant1[['Restaurant ID','Restaurant␣
      ↪Name','Country','City','Aggregate rating','Average Cost for␣
      ↪two','Votes','Price range','Has Table booking_Yes','Has Online␣
      ↪delivery_Yes']]
```

```
[47]: rating = rating.merge(max_rate1,left_on='Restaurant Name',right_on='Restaurant␣
      ↪Name',how='left')
      rating
```

```
[47]:           Restaurant ID           Restaurant Name     Country        City  \
      0               7402935                      Skye   Indonesia     Jakarta
      1               7402935                      Skye   Indonesia     Jakarta
      2               7410290  Satoo - Hotel Shangri-La   Indonesia     Jakarta
      3               7410290  Satoo - Hotel Shangri-La   Indonesia     Jakarta
      4               7410290  Satoo - Hotel Shangri-La   Indonesia     Jakarta
      ...                 ...                       ...         ...         ...
      23810          18312106                UrbanCrave       India      Kanpur
      23811          18312106                UrbanCrave       India      Kanpur
      23812           3900245        Deena Chat Bhandar       India    Varanasi
      23813          18246202           VNS Live Studio       India    Varanasi
      23814          18246202           VNS Live Studio       India    Varanasi

             Aggregate rating  Average Cost for two  Votes  Price range  \
      0                   4.1                800000   1498            3
      1                   4.1                800000   1498            3
      2                   4.6                800000    873            3
      3                   4.6                800000    873            3
      4                   4.6                800000    873            3
      ...                 ...                   ...    ...          ...
      23810               3.9                     0    127            1
      23811               3.9                     0    127            1
      23812               3.8                     0     78            1
      23813               3.5                     0    109            1
      23814               3.5                     0    109            1

             Has Table booking_Yes  Has Online delivery_Yes      Cuisines  Count
      0                          0                        0       Italian      1
      1                          0                        0   Continental      1
      2                          0                        0         Asian      1
      3                          0                        0     Indonesian      1
      4                          0                        0       Western      1
      ...                      ...                      ...           ...    ...
      23810                      0                        0       Italian      1
      23811                      0                        0     Beverages      1
      23812                      0                        0    Street Food      1
```

```
23813                        0            0       Chinese      1
23814                        0            0   North Indian      1

[23815 rows x 12 columns]
```

```
[48]: ratinggs=data_restaurant1.groupby(['Cuisines','City']).
       ↪agg(Count=('Cuisines','count'))
       ratinggs.sort_values(by='Count',ascending=False)
```

```
[48]:                                            Count
      Cuisines                       City
      North Indian                   New Delhi     658
      North Indian, Chinese          New Delhi     284
      Fast Food                      New Delhi     242
      Chinese                        New Delhi     228
      North Indian, Mughlai          New Delhi     207
      …                                     …
      Filipino, Japanese, Asian      Abu Dhabi       1
      Filipino, Mexican              Pasig City      1
                                     Quezon City     1
      Finger Food                    Faridabad       1
      World Cuisine, Patisserie, Cafe €¡stanbul      1

      [3030 rows x 1 columns]
```

0.6.2 Sorted values for the most served cuisine across the restaurant for each city

1 Distribution cost across restaurants

```
[51]: #Do/ scatterplot
```

```
[52]: plt.figure(figsize=(12,6))
      rating.hist(column='Average Cost for two')
      plt.show()
```
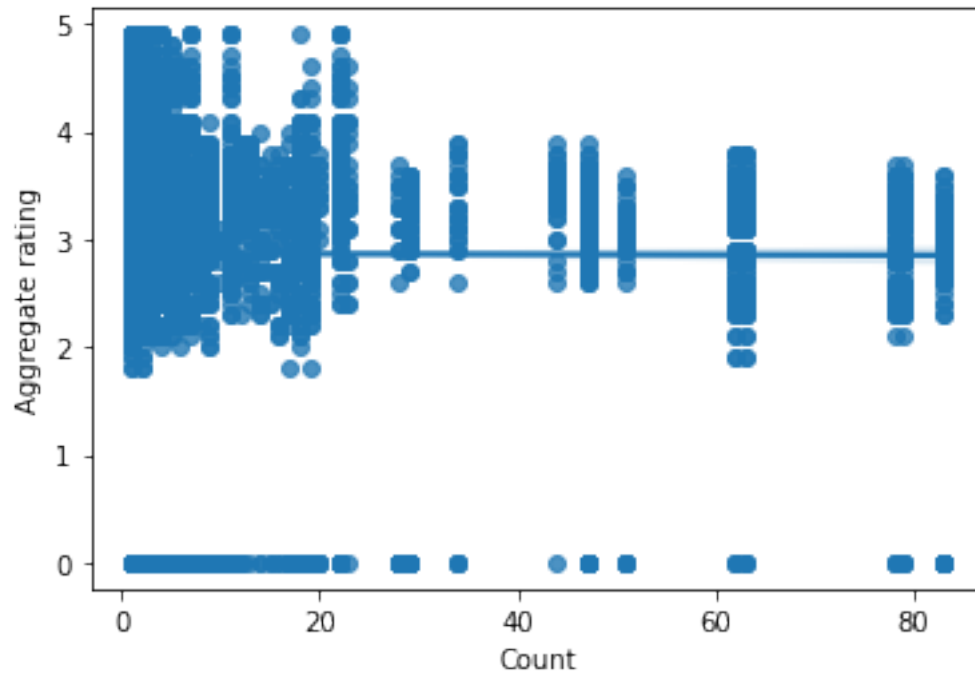
```
<Figure size 864x432 with 0 Axes>
```

Average Cost for two
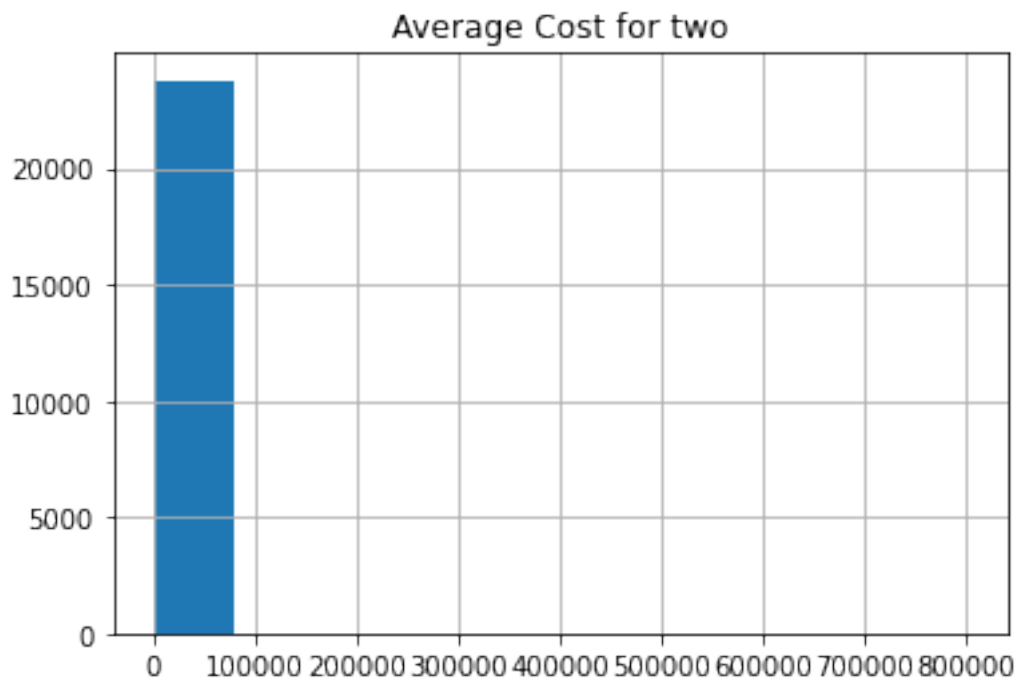

```
[53]:  sns.regplot(x='Count',y='Aggregate rating',data=rating)
       rating[["Count", "Aggregate rating"]].corr()
```

```
[53]:                       Count  Aggregate rating
       Count              1.000000         -0.001642
       Aggregate rating  -0.001642          1.000000
```
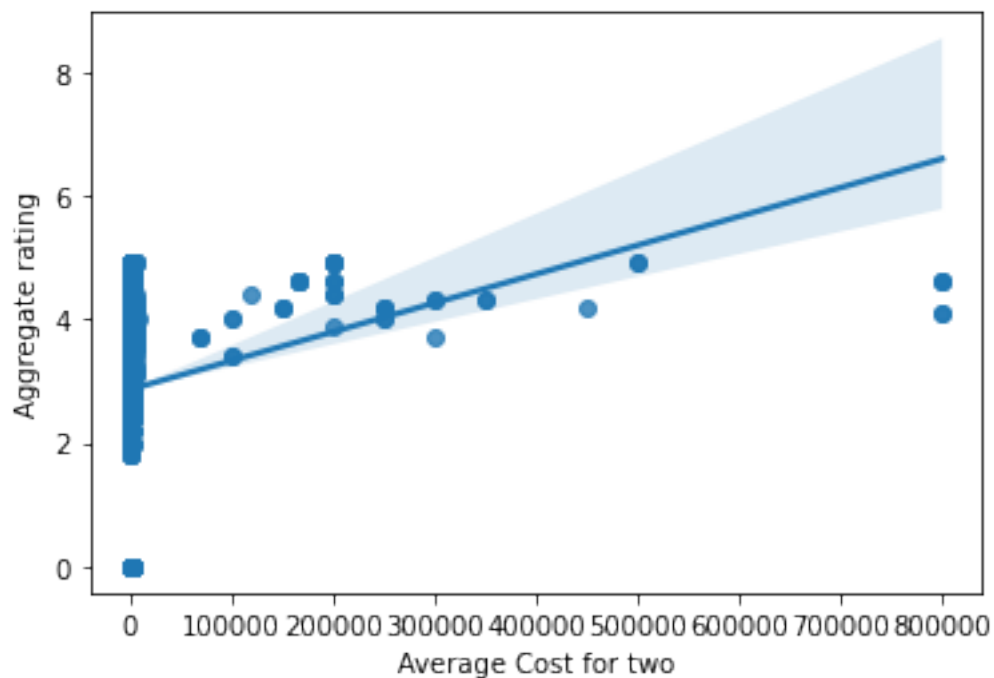
1.0.1 Number of cusines is not a good factor to decide the rating of a restaurant

```
[55]: sns.regplot(x='Average Cost for two',y='Aggregate rating',data=rating)
      rating[["Average Cost for two", "Aggregate rating"]].corr()
```

```
[55]:                       Average Cost for two   Aggregate rating
      Average Cost for two               1.00000            0.05011
      Aggregate rating                   0.05011            1.00000
```

1.0.2 Average cost for two is a suitable factor to decide the rating of a restaurant

```
[57]: onl = data_restaurant1[data_restaurant1['Has Online delivery_Yes'] ==␣
      ↪1]['Aggregate rating'].mean()
      onln = data_restaurant1[data_restaurant1['Has Online delivery_Yes'] ==␣
      ↪0]['Aggregate rating'].mean()
      sns.regplot(x='Price range',y='Aggregate rating',data=rating)
      rating[['Price range','Aggregate rating']].corr()
```

```
[57]:                   Price range  Aggregate rating
      Price range          1.000000          0.462983
      Aggregate rating     0.462983          1.000000
```
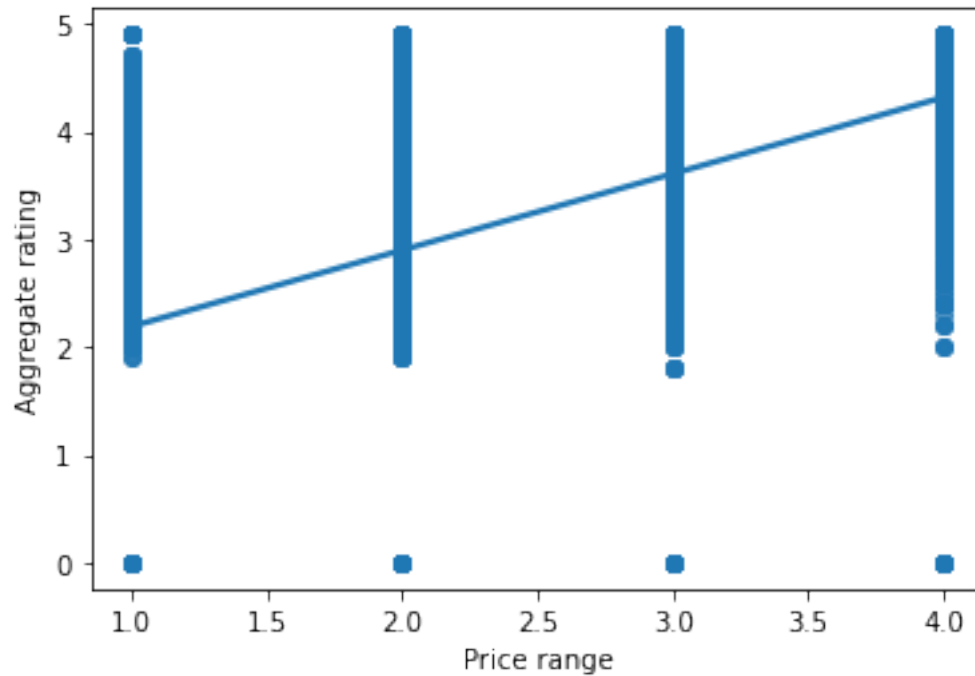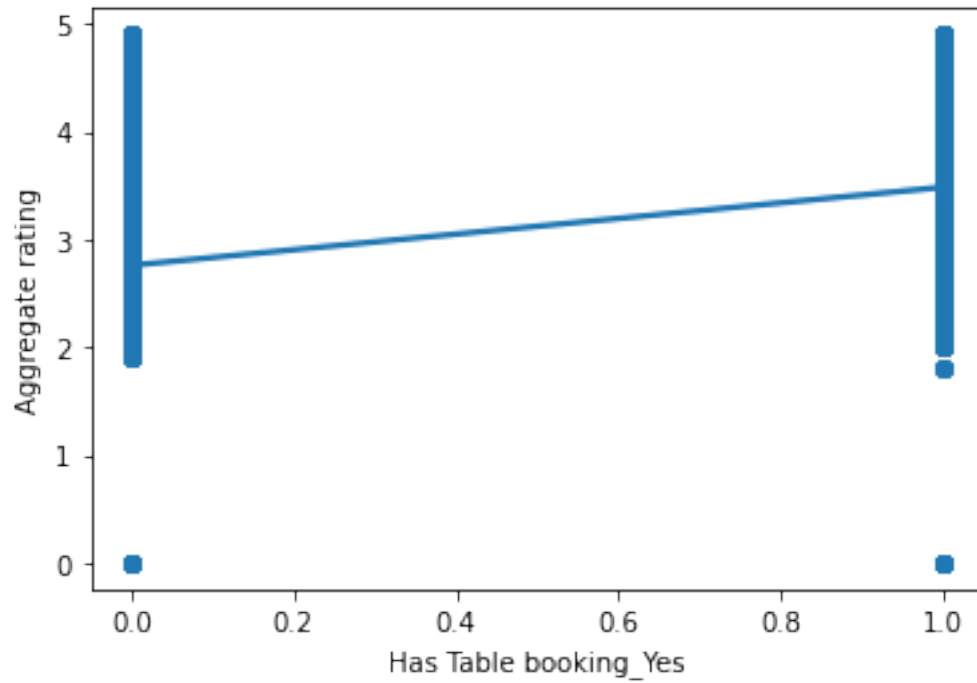
1.0.3 Price range can be a suitable factor to decide the rating of a restaurant

```
[59]: sns.regplot(x='Has Table booking_Yes',y='Aggregate rating',data=rating)
      rating[['Has Table booking_Yes','Aggregate rating']].corr()
```

```
[59]:                         Has Table booking_Yes  Aggregate rating
      Has Table booking_Yes               1.000000          0.181843
      Aggregate rating                    0.181843          1.000000
```

1.0.4 Table booking can be a factor to deduce the rating of a restaurant

[ ]:

[ ]:

[ ]: