In [34]:

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
z=pd.read_excel('zomato - merged.xlsx')
z
```

Out[34]:

	Restaurant ID	Restaurant Name	Country	City	Address	Locality	Loca Verb
0	6317637	Le Petit Souffle	Phillipines	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Century City Mall, Poblacion, Makati City	Century M Poblac Makati (Ma
1	6304287	Izakaya Kikufuji	Phillipines	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...	Little Tokyo, Legaspi Village, Makati City	Little Tok Lega Villa Makati (M
2	6300002	Heat - Edsa Shangri-La	Phillipines	Mandaluyong City	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...	Edsa Shangri-La, Ortigas, Mandaluyong City	Edsa Shar La, Ortig Mandaluy City, M
3	6318506	Ooma	Phillipines	Mandaluyong City	Third Floor, Mega Fashion Hall, SM Megamall, O...	SM Megamall, Ortigas, Mandaluyong City	SM Megan Ortig Mandaluy City, Mand
4	6314302	Sambo Kojin	Phillipines	Mandaluyong City	Third Floor, Mega Atrium, SM Megamall, Ortigas...	SM Megamall, Ortigas, Mandaluyong City	SM Megan Ortig Mandaluy City, Mand
...	...	...	...	...	...	...	
9546	5915730	NamlÛ± Gurme	Turkey	ÛÁstanbul	Kemanke□ô Karamustafa Pa□ôa Mahallesi, RÛ±htÛ±...	Karakí_y	Karak ÛÁstar
9547	5908749	Ceviz AÛôacÛ±	Turkey	ÛÁstanbul	Ko□ôuyolu Mahallesi, Muhittin íìstí_ndaÛô Cadd...	Ko□ôuyolu	Ko□ôuy ÛÁstar
9548	5915807	Huqqa	Turkey	ÛÁstanbul	Kuruí_e□ôme Mahallesi, Muallim Naci Caddesi, N...	Kuruí_e□ôme	Kuruí_e□ô ÛÁstar
9549	5916112	A□ô□ôk Kahve	Turkey	ÛÁstanbul	Kuruí_e□ôme Mahallesi, Muallim Naci Caddesi, N...	Kuruí_e□ôme	Kuruí_e□ô ÛÁstar
9550	5927402	Walter's Coffee Roastery	Turkey	ÛÁstanbul	CafeaÛôa Mahallesi, BademaltÛ± Sokak, No 21/B,...	Moda	Mc ÛÁstar

9551 rows × 22 columns

In [35]:

```
z.head(5)
```

Out[35]:

	Restaurant ID	Restaurant Name	Country	City	Address	Locality	Locality Verbose	Lo
0	6317637	Le Petit Souffle	Phillipines	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Century City Mall, Poblacion, Makati City	Century City Mall, Poblacion, Makati City, Mak...	12
1	6304287	Izakaya Kikufuji	Phillipines	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...	Little Tokyo, Legaspi Village, Makati City	Little Tokyo, Legaspi Village, Makati City, Ma...	12
2	6300002	Heat - Edsa Shangri-La	Phillipines	Mandaluyong City	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...	Edsa Shangri-La, Ortigas, Mandaluyong City	Edsa Shangri-La, Ortigas, Mandaluyong City, Ma...	12
3	6318506	Ooma	Phillipines	Mandaluyong City	Third Floor, Mega Fashion Hall, SM Megamall, O...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	12
4	6314302	Sambo Kojin	Phillipines	Mandaluyong City	Third Floor, Mega Atrium, SM Megamall, Ortigas...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	12

5 rows × 22 columns

In [36]:

```
z.shape
```

Out[36]:

```
(9551, 22)
```

In [37]:

```
z.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9551 entries, 0 to 9550
Data columns (total 22 columns):
 #   Column                Non-Null Count  Dtype  
---  ------                --------------  -----  
 0   Restaurant ID         9551 non-null   int64  
 1   Restaurant Name       9551 non-null   object 
 2   Country               9551 non-null   object 
 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  Is delivering now     9551 non-null   object 
 15  Switch to order menu  9551 non-null   object 
 16  Price range           9551 non-null   int64  
 17  Aggregate rating      9551 non-null   float64
 18  Rating color          9551 non-null   object 
 19  Rating text           9551 non-null   object 
 20  Votes                 9551 non-null   int64  
 21  Country.1             9551 non-null   object 
dtypes: float64(3), int64(4), object(15)
memory usage: 1.6+ MB
```

In [38]:

```
z.columns
```

Out[38]:

```
Index(['Restaurant ID', 'Restaurant Name', 'Country', 'City', 'Address',
       'Locality', 'Locality Verbose', 'Longitude', 'Latitude', 'Cuisine
s',
       'Average Cost for two', 'Currency', 'Has Table booking',
       'Has Online delivery', 'Is delivering now', 'Switch to order men
u',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
       'Votes', 'Country.1'],
      dtype='object')
```

In [39]:

```
# 1.Find total number of restaurant.
z['Restaurant ID'].value_counts().sum()
```

Out[39]:

```
9551
```

In [40]:

```
#2.Find total number of restaurant by restaurant .
z['Restaurant Name'].value_counts()
```

Out[40]:

```
Cafe Coffee Day             83
Domino's Pizza              79
Subway                      63
Green Chick Chop            51
McDonald's                  48
                            ..
Odeon Social                 1
Johnny Rockets               1
House of Commons             1
HotMess                      1
Walter's Coffee Roastery     1
Name: Restaurant Name, Length: 7446, dtype: int64
```

In [41]:

```
#3.find total number of country
z['Country'].value_counts()
```

Out[41]:

```
India             8652
United States      434
United Kingdom      80
Brazil              60
UAE                 60
South Africa        60
New Zealand         40
Turkey              34
Australia           24
Phillipines         22
Indonesia           21
Singapore           20
Qatar               20
Sri Lanka           20
Canada               4
Name: Country, dtype: int64
```

In [42]:

```
# 4.find tota number of City.
z['City'].value_counts()
```

Out[42]:

```
New Delhi           5473
Gurgaon             1118
Noida               1080
Faridabad            251
Ghaziabad             25
                    ...
Panchkula              1
Mc Millan              1
Mayfield               1
Macedon                1
Vineland Station       1
Name: City, Length: 141, dtype: int64
```

In [43]:

```
#5.find total number of address
z['Address'].value_counts()
```

Out[43]:

```
Dilli Haat, INA, New Delhi
11
Sector 41, Noida
11
Greater Kailash (GK) 1, New Delhi
10
The Imperial, Janpath, New Delhi
9
Cyber Hub, DLF Cyber City, Gurgaon
8

..
23-24, Defence Colony Market, Defence Colony, New Delhi
1
28, Main Market, Defence Colony, New Delhi
1
Daryaganj, New Delhi
1
Ground Floor, E-23, Netaji Subhash Marg, Opposite Golcha Cinema, Daryagan
j, New Delhi     1
CafeaÛôa Mahallesi, BademaltÛ± Sokak, No 21/B, KadÛ±kí_y, ÛÁstanbul
1
Name: Address, Length: 8918, dtype: int64
```

In [44]:

```
# 6.Find total number of Rating color
z['Rating color'].value_counts()
```

Out[44]:

```
Orange        3737
White         2148
Yellow        2100
Green         1079
Dark Green     301
Red            186
Name: Rating color, dtype: int64
```

In [45]:

```
# 7.Find total number of restaurant by Rating color
#z.groupby(['Restaurant Name','Rating color']).count()
```

In [46]:

```
pd.crosstab(z['Restaurant Name'],z['Rating color'])
```

Out[46]:

Rating color	Dark Green	Green	Orange	Red	White	Yellow
Restaurant Name						
12212	0	0	0	0	0	1
Let's Burrrp	0	0	0	0	1	0
#45	0	0	0	0	0	1
#Dilliwaala6	0	0	0	0	0	1
#InstaFreeze	0	0	0	0	1	0
...	...	...	...	...	...	...
t Lounge by Dilmah	0	0	0	0	0	1
tashas	0	1	0	0	0	0
wagamama	0	0	0	0	0	1
{Niche} - Cafe & Bar	0	1	0	0	0	0
íˆukuraÛôa SofrasÛ±	0	1	0	0	0	0

7446 rows × 6 columns

In [47]:

```
# 8.Find the mostly use Votes
z['Votes'].max()
```

Out[47]:

```
10934
```

In [48]:

```
# 9.Find the less use Votes
z['Votes'].min()
```

Out[48]:

```
0
```

In [51]:

```
# 10.Find the total number of restaurant by Rating color and city
z.groupby(['Rating color','City']).count()['Restaurant ID']
```

Out[51]:

```
Rating color  City
Dark Green    Abu Dhabi           7
              Agra                1
              Ahmedabad           3
              Ankara              5
              Athens              3
                                 ..
Yellow        Vizag              10
              Waterloo           19
              Weirton             1
              Wellington City     1
              ÛÁstanbul           2
Name: Restaurant ID, Length: 347, dtype: int64
```

In [52]:

```
z.columns
```

Out[52]:

```
Index(['Restaurant ID', 'Restaurant Name', 'Country', 'City', 'Address',
       'Locality', 'Locality Verbose', 'Longitude', 'Latitude', 'Cuisine
s',
       'Average Cost for two', 'Currency', 'Has Table booking',
       'Has Online delivery', 'Is delivering now', 'Switch to order men
u',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
       'Votes', 'Country.1'],
      dtype='object')
```

In [53]:

```
# 11.Find the unique number of City
u3=z['City'].unique()
len(np.unique(u3))
```

Out[53]:

```
141
```

In [54]:

```
# 12.Find the unique number of Rating color
u3=z['Rating color'].unique()
len(np.unique(u3))
```

Out[54]:

```
6
```

In [55]:

```
# 13.Find total number of Cuisines
z['Cuisines'].value_counts()
```

Out[55]:

```
North Indian                                            936
North Indian, Chinese                                   511
Chinese                                                 354
Fast Food                                               354
North Indian, Mughlai                                   334
                                                       ...
Bengali, Fast Food                                        1
North Indian, Rajasthani, Asian                           1
Chinese, Thai, Malaysian, Indonesian                      1
Bakery, Desserts, North Indian, Bengali, South Indian     1
Italian, World Cuisine                                    1
Name: Cuisines, Length: 1825, dtype: int64
```

In [57]:

```
# 14.Find total number of Rating text
z['Rating text'].value_counts()
```

Out[57]:

```
Average      3737
Not rated    2148
Good         2100
Very Good    1079
Excellent     301
Poor          186
Name: Rating text, dtype: int64
```

In [58]:

```
# 15.Find average of Aggregate rating
z['Aggregate rating'].mean()
```

Out[58]:

```
2.66637001361114
```

In [59]:

```
# 16.Find total of Aggregate rating
z['Aggregate rating'].sum()
```

Out[59]:

```
25466.499999999996
```

In [60]:

```
# 17.Find max of Aggregate rating
z['Aggregate rating'].max()
```

Out[60]:

```
4.9
```

In [61]:

```
# 18.Find total of Aggregate rating
z['Aggregate rating'].min()
```

Out[61]:

```
0.0
```

In [65]:

```
# 19.Find the total number of Restaurant by Country and City
z.groupby(['Country','City']).count()['Restaurant Name']
```

Out[65]:

```
Country        City
Australia      Armidale           1
               Balingup           1
               Beechworth         1
               Dicky Beach        1
               East Ballina       1
                                 ..
United States  Valdosta          20
               Vernonia           1
               Waterloo          20
               Weirton            1
               Winchester Bay     1
Name: Restaurant Name, Length: 141, dtype: int64
```

In [62]:

```
z.columns
```

Out[62]:

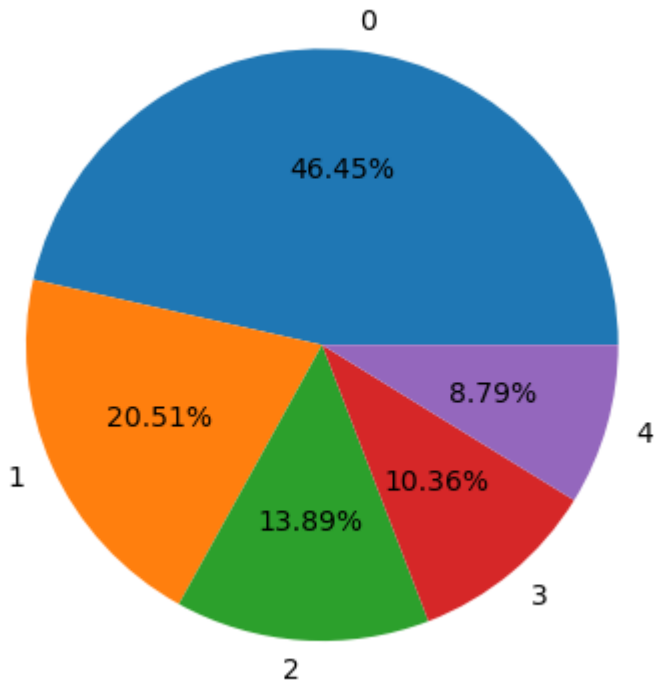
```
Index(['Restaurant ID', 'Restaurant Name', 'Country', 'City', 'Address',
       'Locality', 'Locality Verbose', 'Longitude', 'Latitude', 'Cuisine
s',
       'Average Cost for two', 'Currency', 'Has Table booking',
       'Has Online delivery', 'Is delivering now', 'Switch to order men
u',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
       'Votes', 'Country.1'],
      dtype='object')
```

In [66]:

```
# 20.Create a pie chart to show most Votes used in the restaurant.
pie_label=z.Votes.value_counts().index
pie_val=z.Votes.value_counts().values
plt.pie(pie_val[:5],labels=pie_label[:5],autopct='%1.2f%%')
```

Out[66]:

```
([<matplotlib.patches.Wedge at 0x20ab95683d0>,
  <matplotlib.patches.Wedge at 0x20ab95682e0>,
  <matplotlib.patches.Wedge at 0x20ab95690c0>,
  <matplotlib.patches.Wedge at 0x20ab9569750>,
  <matplotlib.patches.Wedge at 0x20ab9569de0>],
 [Text(0.12227556787221748, 1.0931828234570495, '0'),
  Text(-1.0037027331103823, -0.4500897949817888, '1'),
  Text(-0.07551220314841353, -1.0974050788909593, '2'),
  Text(0.7027480901125928, -0.8462535801065206, '3'),
  Text(1.0583265110866198, -0.29990831254105516, '4')],
 [Text(0.0666957642939368, 0.5962815400674816, '46.45%'),
  Text(-0.5474742180602085, -0.24550352453552113, '20.51%'),
  Text(-0.041188474444589195, -0.5985845884859777, '13.89%'),
  Text(0.3833171400614142, -0.4615928618762839, '10.36%'),
  Text(0.5772690060472471, -0.16358635229512097, '8.79%')])
```



In [70]:

```
# 21. Find the total number of restaurant and country
pd.crosstab(z['Restaurant Name'],z['Country'])
```

Out[70]:

Country	Australia	Brazil	Canada	India	Indonesia	New Zealand	Phillipines	Qatar	Singap
Restaurant Name									
12212	0	0	0	1	0	0	0	0	
Let's Burrrp	0	0	0	1	0	0	0	0	
#45	0	0	0	1	0	0	0	0	
#Dilliwaala6	0	0	0	1	0	0	0	0	
#InstaFreeze	0	0	0	1	0	0	0	0	
...	...	...	...	...	...	...	...	...	
t Lounge by Dilmah	0	0	0	1	0	0	0	0	
tashas	0	0	0	0	0	0	0	0	
wagamama	0	0	0	0	0	1	0	0	
{Niche} - Cafe & Bar	0	0	0	1	0	0	0	0	
íˆukuraÛôa SofrasÛ±	0	0	0	0	0	0	0	0	

7446 rows × 15 columns

In [71]:

```
# 22.check the descriptive statistics for Votes.
z.describe()
```

Out[71]:

	Restaurant ID	Longitude	Latitude	Average Cost for two	Price range	Aggregate rating	
count	9.551000e+03	9551.000000	9551.000000	9551.000000	9551.000000	9551.000000	9
mean	9.051128e+06	64.126574	25.854381	1199.210763	1.804837	2.666370	
std	8.791521e+06	41.467058	11.007935	16121.183073	0.905609	1.516378	
min	5.300000e+01	-157.948486	-41.330428	0.000000	1.000000	0.000000	
25%	3.019625e+05	77.081343	28.478713	250.000000	1.000000	2.500000	
50%	6.004089e+06	77.191964	28.570469	400.000000	2.000000	3.200000	
75%	1.835229e+07	77.282006	28.642758	700.000000	2.000000	3.700000	
max	1.850065e+07	174.832089	55.976980	800000.000000	4.000000	4.900000	10

In [72]:

```
# 23.Find total of Longitude
z['Longitude'].sum()
```

Out[72]:

612472.9126835731

In [73]:

```
# 24.Find Average of Longitude
z['Longitude'].mean()
```

Out[73]:

64.12657446168706

In [74]:

```
# 25. Find total of Latitude
z['Latitude'].sum()
```

Out[74]:

246935.19006641398

In [75]:

```
# 26.Find Average of Latitude
z['Latitude'].mean()
```

Out[75]:

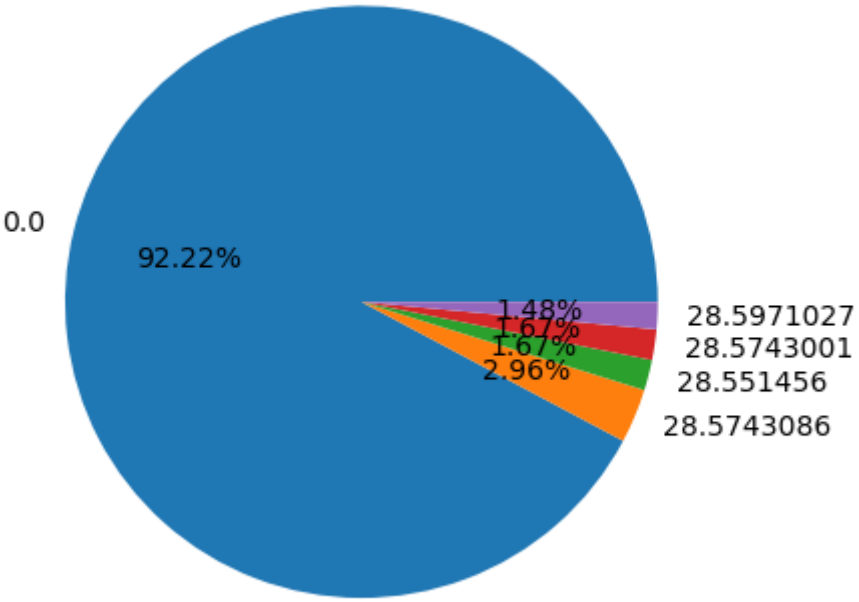
25.854380700074756

In [76]:

```
# 27.Create a pie chart to show most Latitude used in the restaurant.
pie_label=z.Latitude.value_counts().index
pie_val=z.Latitude.value_counts().values
plt.pie(pie_val[:5],labels=pie_label[:5],autopct='%1.2f%%')
```

Out[76]:

```
([<matplotlib.patches.Wedge at 0x20abe4310f0>,
  <matplotlib.patches.Wedge at 0x20abe431000>,
  <matplotlib.patches.Wedge at 0x20abe431d80>,
  <matplotlib.patches.Wedge at 0x20abe432410>,
  <matplotlib.patches.Wedge at 0x20abe432aa0>],
 [Text(-1.0673252778639226, 0.26611416954514167, '0.0'),
  Text(1.0150386210624949, -0.4239063549317812, '28.5743086'),
  Text(1.0657590129345347, -0.2723191626544239, '28.551456'),
  Text(1.088385778010785, -0.15942521200819565, '28.5743001'),
  Text(1.0988088147281376, -0.051178009689174915, '28.5971027')],
 [Text(-0.5821774242894123, 0.14515318338825908, '92.22%'),
  Text(0.5536574296704517, -0.23122164814460788, '2.96%'),
  Text(0.5813230979642915, -0.1485377250842312, '1.67%'),
  Text(0.5936649698240646, -0.08695920654992488, '1.67%'),
  Text(0.5993502625789842, -0.027915278012277223, '1.48%')])
```

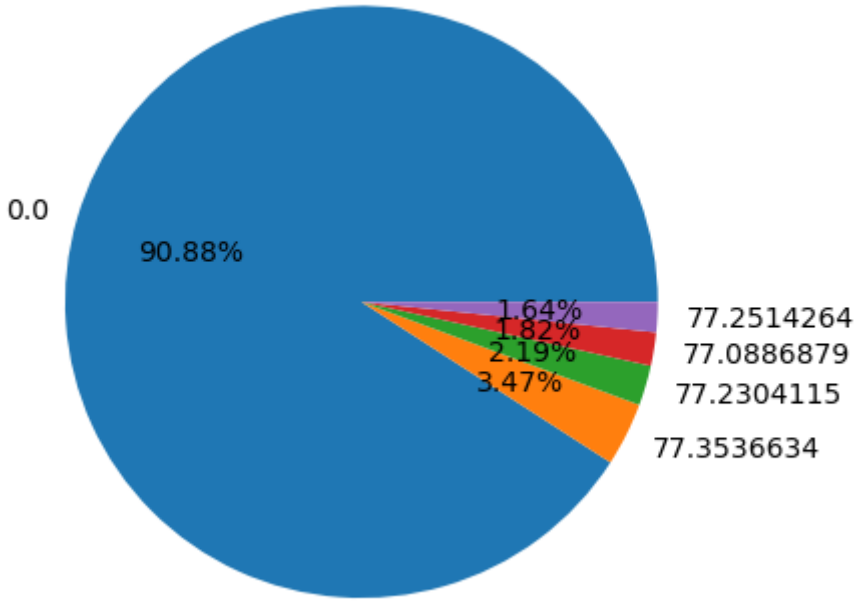


In [77]:

```
# 28.Create a pie chart to show most Longitude used in the restaurant.
pie_label=z.Longitude.value_counts().index
pie_val=z.Longitude.value_counts().values
plt.pie(pie_val[:5],labels=pie_label[:5],autopct='%1.2f%%')
```

Out[77]:

```
([<matplotlib.patches.Wedge at 0x20abe408c40>,
  <matplotlib.patches.Wedge at 0x20abe408b50>,
  <matplotlib.patches.Wedge at 0x20abe409810>,
  <matplotlib.patches.Wedge at 0x20abe409ea0>,
  <matplotlib.patches.Wedge at 0x20abe40a530>],
 [Text(-1.055118665728599, 0.3110057897067212, '0.0'),
  Text(0.983519437718095, -0.49263527647812877, '77.3536634'),
  Text(1.055118660268887, -0.3110058082293474, '77.2304115'),
  Text(1.0858588444598671, -0.17581402079550526, '77.0886879'),
  Text(1.0985361695484468, -0.05672992326653011, '77.2514264')],
 [Text(-0.5755192722155993, 0.16963952165821156, '90.88%'),
  Text(0.5364651478462336, -0.26871015080625205, '3.47%'),
  Text(0.5755192692375746, -0.1696395317614622, '2.19%'),
  Text(0.5922866424326547, -0.09589855679754831, '1.82%'),
  Text(0.5992015470264255, -0.03094359450901642, '1.64%')])
```



In [81]:

```
# 29.create a count plot for parental level of Restaurant
#z['Restaurant Name'].value_counts(normalize= True)
#z['Restaurant Name'].value_counts(dropna= False).plot.bar(color='green')
#plt.title('Restaurant')
#plt.xlabel('Restaurant Name')
#plt.ylabel('Count')
#plt.show()
```

In [79]:

```
z.columns
```

Out[79]:

```
Index(['Restaurant ID', 'Restaurant Name', 'Country', 'City', 'Address',
       'Locality', 'Locality Verbose', 'Longitude', 'Latitude', 'Cuisine
s',
       'Average Cost for two', 'Currency', 'Has Table booking',
       'Has Online delivery', 'Is delivering now', 'Switch to order men
u',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
       'Votes', 'Country.1'],
      dtype='object')
```

In [ ]:

```

```