Exploratory data analysis on zomata dataset

In [10]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline

```

In [11]:
```
df=pd.read_csv(r"C:\Users\Antima\Downloads\Zomatodataset\zomato.csv",encoding='latin-1')
```

In [13]:
```
df.head()
```

Out[13]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	...	Currency	Has Table booking	Has Online delivery	Is delivering now	Switch to order menu	Price range	Aggregate rating	Rating color	Rating text	Votes
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Century City Mall, Poblacion, Makati City	Century City Mall, Poblacion, Makati City, Mak...	121.027535	14.565443	French, Japanese, Desserts	...	Botswana Pula(P)	Yes	No	No	No	3	4.8	Dark Green	Excellent	314
1	6304287	Izakaya Kikufuji	162	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...	Little Tokyo, Legaspi Village, Makati City	Little Tokyo, Legaspi Village, Makati City, Ma...	121.014101	14.553708	Japanese	...	Botswana Pula(P)	Yes	No	No	No	3	4.5	Dark Green	Excellent	591
2	6300002	Heat - Edsa Shangri-La	162	Mandaluyong City	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...	Edsa Shangri-La, Ortigas, Mandaluyong City	Edsa Shangri-La, Ortigas, Mandaluyong City, Ma...	121.056831	14.581404	Seafood, Asian, Filipino, Indian	...	Botswana Pula(P)	Yes	No	No	No	4	4.4	Green	Very Good	270
3	6318506	Ooma	162	Mandaluyong City	Third Floor, Mega Fashion Hall, SM Megamall, O...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.056475	14.585318	Japanese, Sushi	...	Botswana Pula(P)	No	No	No	No	4	4.9	Dark Green	Excellent	365
4	6314302	Sambo Kojin	162	Mandaluyong City	Third Floor, Mega Atrium, SM Megamall, Ortigas...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.057508	14.584450	Japanese, Korean	...	Botswana Pula(P)	Yes	No	No	No	4	4.8	Dark Green	Excellent	229

5 rows × 21 columns

In [14]:
```
df.columns
```

Out[14]:
```
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', 'Is delivering now', 'Switch to order menu',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
       'Votes'],
      dtype='object')
```

In [15]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9551 entries, 0 to 9550
Data columns (total 21 columns):
 #   Column                Non-Null Count  Dtype  
---  ------                --------------  -----  
 0   Restaurant ID         9551 non-null   int64  
 1   Restaurant Name       9551 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  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  
dtypes: float64(3), int64(5), object(13)
memory usage: 1.5+ MB
```

In data analysis things we do: 1.missing values 2.explore about numerical values 3.explore about categorical values 4.finding relationship between different features

In [18]:
```
df.isna().sum()
```

Out[18]:
```
Restaurant ID         0
Restaurant Name       0
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
Is delivering now     0
Switch to order menu  0
Price range           0
Aggregate rating      0
Rating color          0
Rating text           0
Votes                 0
dtype: int64
```

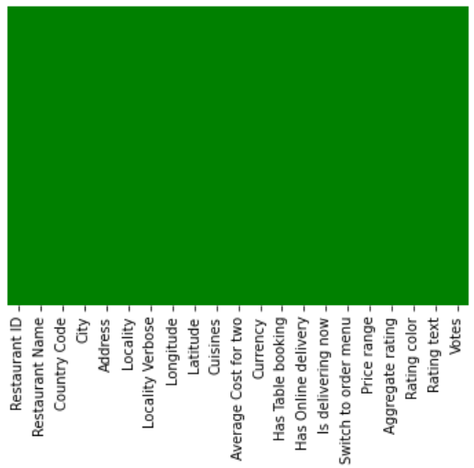
In [24]:
```
#other way to know the name of the specific columns contain null values:
[features for features in df.columns if df[features].isnull().sum()>0]
```

Out[24]:
```
['Cuisines']
```

In [31]:
```
sns.heatmap(df.isnull(),yticklabels=False,cbar=False,cmap='ocean')
```

```
<Figure size 1800x720 with 0 Axes>
```

Out[31]: <AxesSubplot:>



In [40]:
```
df_country=pd.read_excel(r"C:\Users\Antima\Downloads\Zomatodataset\Country-Code.xlsx")
```

In [41]:
```
df.columns
```

Out[41]:
```
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', 'Is delivering now', 'Switch to order menu',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
       'Votes'],
      dtype='object')
```

In [42]:
```
final_df=pd.merge(df,df_country,on='Country Code',how='left')
```

In [43]:
```
final_df
```

Out[43]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	...	Has Table booking	Has Online delivery	Is delivering now	Switch to order menu	Price range	Aggregate rating	Rating color	Rating text	Votes	Country
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Century City Mall, Poblacion, Makati City	Century City Mall, Poblacion, Makati City, Mak...	121.027535	14.565443	French, Japanese, Desserts	...	Yes	No	No	No	3	4.8	Dark Green	Excellent	314	Phillipines
1	6304287	Izakaya Kikufuji	162	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...	Little Tokyo, Legaspi Village, Makati City	Little Tokyo, Legaspi Village, Makati City, Ma...	121.014101	14.553708	Japanese	...	Yes	No	No	No	3	4.5	Dark Green	Excellent	591	Phillipines
2	6300002	Heat - Edsa Shangri-La	162	Mandaluyong City	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...	Edsa Shangri-La, Ortigas, Mandaluyong City	Edsa Shangri-La, Ortigas, Mandaluyong City, Ma...	121.056831	14.581404	Seafood, Asian, Filipino, Indian	...	Yes	No	No	No	4	4.4	Green	Very Good	270	Phillipines
3	6318506	Ooma	162	Mandaluyong City	Third Floor, Mega Fashion Hall, SM Megamall, O...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.056475	14.585318	Japanese, Sushi	...	No	No	No	No	4	4.9	Dark Green	Excellent	365	Phillipines
4	6314302	Sambo Kojin	162	Mandaluyong City	Third Floor, Mega Atrium, SM Megamall, Ortigas...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.057508	14.584450	Japanese, Korean	...	Yes	No	No	No	4	4.8	Dark Green	Excellent	229	Phillipines
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
9546	5915730	NamlÛ± Gurme	208	ÛÁstanbul	Kemanke ô Karamustafa Pa ôa Mahallesi, RÛ±htÛ±...	Karakí_y	Karakí_y, ÛÁstanbul	28.977392	41.022793	Turkish	...	No	No	No	No	3	4.1	Green	Very Good	788	Turkey
9547	5908749	Ceviz AÛôacÛ±	208	ÛÁstanbul	Ko ôuyolu Mahallesi, Muhittin îistí_ndaÛô Cadd...	Ko ôuyolu	Ko ôuyolu, ÛÁstanbul	29.041297	41.009847	World Cuisine, Patisserie, Cafe	...	No	No	No	No	3	4.2	Green	Very Good	1034	Turkey
9548	5915807	Huqqa	208	ÛÁstanbul	Kuruí_e ôme Mahallesi, Muallim Naci Caddesi, N...	Kuruí_e ôme	Kuruí_e ôme, ÛÃstanbul	29.034640	41.055817	Italian, World Cuisine	...	No	No	No	No	4	3.7	Yellow	Good	661	Turkey
9549	5916112	A ô ôk Kahve	208	ÛÁstanbul	Kuruí_e ôme Mahallesi, Muallim Naci Caddesi, N...	Kuruí_e ôme	Kuruí_e ôme, ÛÃstanbul	29.036019	41.057979	Restaurant Cafe	...	No	No	No	No	4	4.0	Green	Very Good	901	Turkey
9550	5927402	Walter's Coffee Roastery	208	ÛÁstanbul	CafeaÛôa Mahallesi, BademaltÛ± Sokak, No 21/B,...	Moda	Moda, ÛÁstanbul	29.026016	40.984776	Cafe	...	No	No	No	No	2	4.0	Green	Very Good	591	Turkey

9551 rows × 22 columns

In [45]:
```
final_df.head()
```

Out[45]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	...	Has Table booking	Has Online delivery	Is delivering now	Switch to order menu	Price range	Aggregate rating	Rating color	Rating text	Votes	Country
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Century City Mall, Poblacion, Makati City	Century City Mall, Poblacion, Makati City, Mak...	121.027535	14.565443	French, Japanese, Desserts	...	Yes	No	No	No	3	4.8	Dark Green	Excellent	314	Phillipines
1	6304287	Izakaya Kikufuji	162	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...	Little Tokyo, Legaspi Village, Makati City	Little Tokyo, Legaspi Village, Makati City, Ma...	121.014101	14.553708	Japanese	...	Yes	No	No	No	3	4.5	Dark Green	Excellent	591	Phillipines
2	6300002	Heat - Edsa Shangri-La	162	Mandaluyong City	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...	Edsa Shangri-La, Ortigas, Mandaluyong City	Edsa Shangri-La, Ortigas, Mandaluyong City, Ma...	121.056831	14.581404	Seafood, Asian, Filipino, Indian	...	Yes	No	No	No	4	4.4	Green	Very Good	270	Phillipines
3	6318506	Ooma	162	Mandaluyong City	Third Floor, Mega Fashion Hall, SM Megamall, O...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.056475	14.585318	Japanese, Sushi	...	No	No	No	No	4	4.9	Dark Green	Excellent	365	Phillipines
4	6314302	Sambo Kojin	162	Mandaluyong City	Third Floor, Mega Atrium, SM Megamall, Ortigas...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.057508	14.584450	Japanese, Korean	...	Yes	No	No	No	4	4.8	Dark Green	Excellent	229	Phillipines

5 rows × 22 columns

In [53]:
```
country_names=final_df.Country.value_counts().index
```

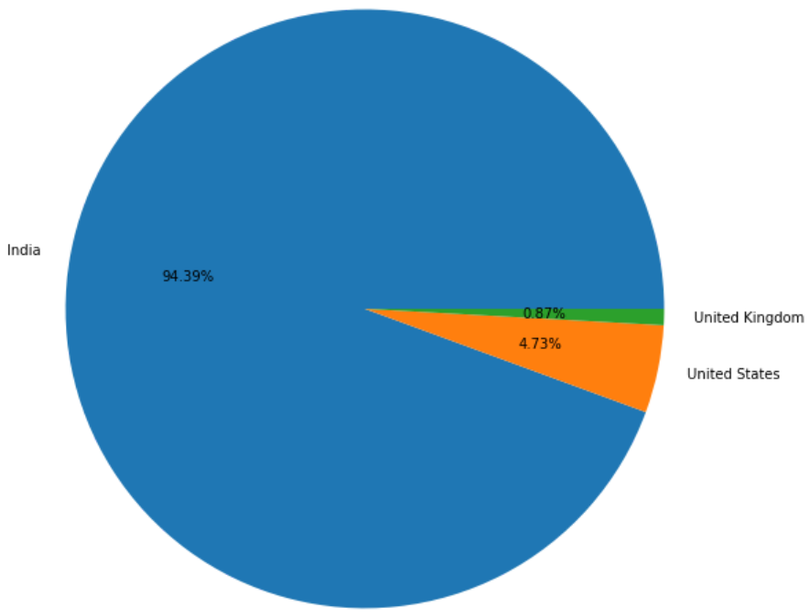
In [50]:
```
country_values=final_df.Country.value_counts().values
```

In [54]:
```
country_values
```

Out[54]:
```
array([8652,  434,   80,   60,   60,   60,   40,   34,   24,   22,   21,
         20,   20,   20,    4], dtype=int64)
```

In [62]:
```
#pie chart: top 3 country that uses Zomato most
from matplotlib import rcParams
plt.figure(figsize=(25,10))
plt.pie(country_values[:3],labels=country_names[:3],autopct='%1.2f%%')
```

Out[62]:
```
([<matplotlib.patches.Wedge at 0x27eb2070040>,
  <matplotlib.patches.Wedge at 0x27eb2070760>,
  <matplotlib.patches.Wedge at 0x27eb2070e80>],
 [Text(-1.0829742700952103, 0.19278674827836725, 'India'),
  Text(1.077281715838356, -0.22240527134123297, 'United States'),
  Text(1.0995865153823035, -0.03015783794312073, 'United Kingdom')],
 [Text(-0.590713238233751, 0.10515640815183668, '94.39%'),
  Text(0.5876082086391032, -0.12131196618612707, '4.73%'),
  Text(0.5997744629358018, -0.01644972978715676, '0.87%')])
```



Observation:

1. Zomato maximum transaction is done by India ...94.39% of transaction is done by India then 4.73% by united States and 0.87% by United Kingdom

In [69]:
```
Ratings=final_df.groupby(['Aggregate rating', 'Rating color', 'Rating text']).size().reset_index().rename(columns={0:"rating_count"})
```

Observation: 1.when rating is between 4.5 to 4.9 -->Excellent 2.when rating is between 4.0 to 4.4 -->Very Good 3.when rating is between 3.5 to 3.9 -->Good 4.when rating is between 2.5 to 3.4 -->Average 5.when rating is between 1.8 to 2.4 -->Poor

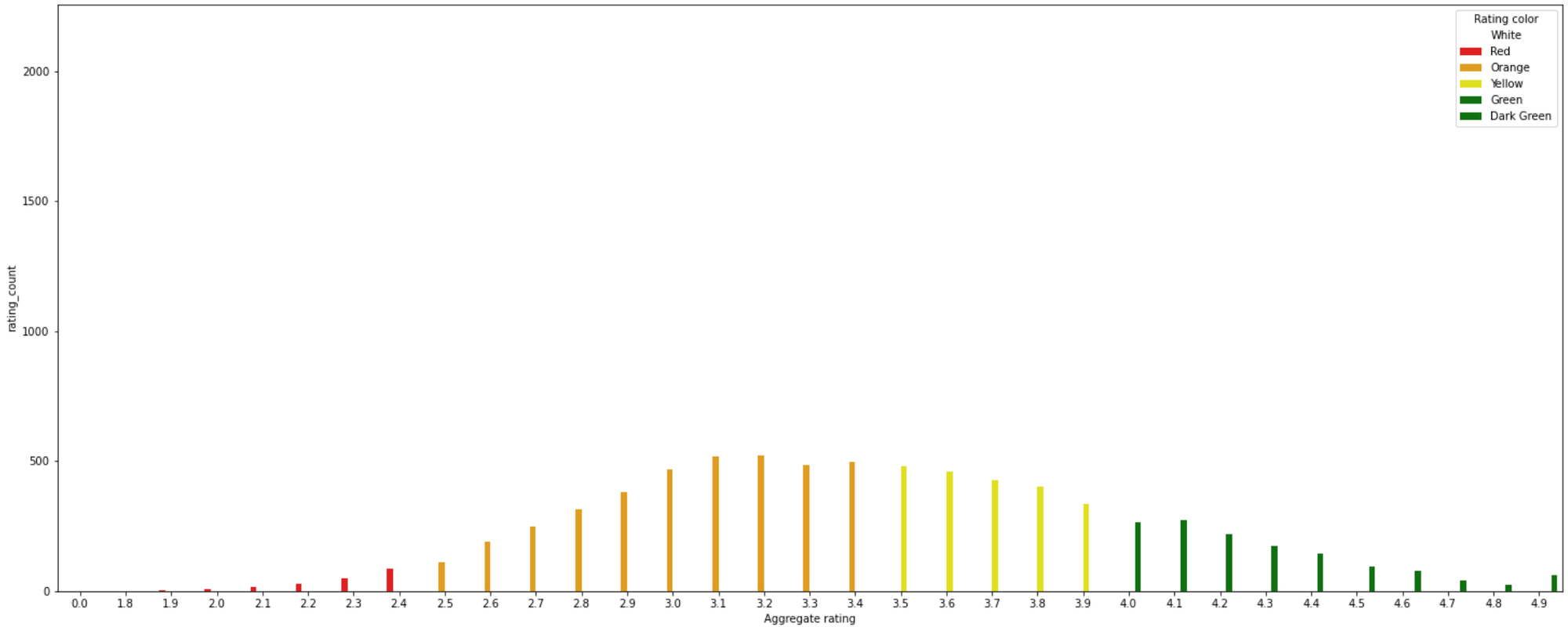
In [71]:
```
Ratings
```

Out[71]:

	Aggregate rating	Rating color	Rating text	rating_count
0	0.0	White	Not rated	2148
1	1.8	Red	Poor	1
2	1.9	Red	Poor	2
3	2.0	Red	Poor	7
4	2.1	Red	Poor	15
5	2.2	Red	Poor	27
6	2.3	Red	Poor	47
7	2.4	Red	Poor	87
8	2.5	Orange	Average	110
9	2.6	Orange	Average	191
10	2.7	Orange	Average	250
11	2.8	Orange	Average	315
12	2.9	Orange	Average	381
13	3.0	Orange	Average	468
14	3.1	Orange	Average	519
15	3.2	Orange	Average	522
16	3.3	Orange	Average	483
17	3.4	Orange	Average	498
18	3.5	Yellow	Good	480
19	3.6	Yellow	Good	458
20	3.7	Yellow	Good	427
21	3.8	Yellow	Good	400
22	3.9	Yellow	Good	335
23	4.0	Green	Very Good	266
24	4.1	Green	Very Good	274
25	4.2	Green	Very Good	221
26	4.3	Green	Very Good	174
27	4.4	Green	Very Good	144
28	4.5	Dark Green	Excellent	95
29	4.6	Dark Green	Excellent	78
30	4.7	Dark Green	Excellent	42
31	4.8	Dark Green	Excellent	25
32	4.9	Dark Green	Excellent	61

In [79]:
```
from matplotlib import rcParams
plt.figure(figsize=(25,10))
sns.barplot(data=Ratings,x='Aggregate rating',y='rating_count',hue='Rating color',palette=['white','red','orange','yellow',
            'green','green'])
```

Out[79]: <AxesSubplot:xlabel='Aggregate rating', ylabel='rating_count'>

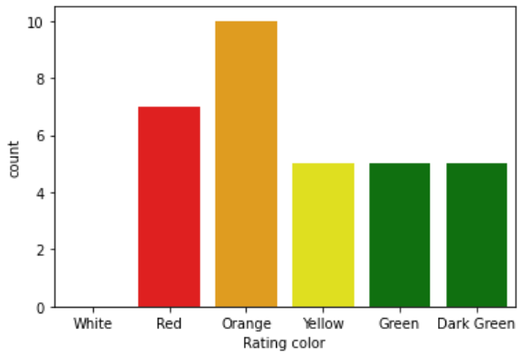


Observation:

1. Not rated count is very high.
2. Maximum number of rating are between 2.5 to 3.9

In [86]:
```
sns.countplot(data=Ratings,x='Rating color',palette=['white','red','orange','yellow','green','green'])
```

Out[86]: <AxesSubplot:xlabel='Rating color', ylabel='count'>



In [88]:
```
Ratings
```

Out[88]:

	Aggregate rating	Rating color	Rating text	rating_count
0	0.0	White	Not rated	2148
1	1.8	Red	Poor	1
2	1.9	Red	Poor	2
3	2.0	Red	Poor	7
4	2.1	Red	Poor	15
5	2.2	Red	Poor	27
6	2.3	Red	Poor	47
7	2.4	Red	Poor	87
8	2.5	Orange	Average	110
9	2.6	Orange	Average	191
10	2.7	Orange	Average	250
11	2.8	Orange	Average	315
12	2.9	Orange	Average	381
13	3.0	Orange	Average	468
14	3.1	Orange	Average	519
15	3.2	Orange	Average	522
16	3.3	Orange	Average	483
17	3.4	Orange	Average	498
18	3.5	Yellow	Good	480
19	3.6	Yellow	Good	458
20	3.7	Yellow	Good	427
21	3.8	Yellow	Good	400
22	3.9	Yellow	Good	335
23	4.0	Green	Very Good	266
24	4.1	Green	Very Good	274
25	4.2	Green	Very Good	221
26	4.3	Green	Very Good	174
27	4.4	Green	Very Good	144
28	4.5	Dark Green	Excellent	95
29	4.6	Dark Green	Excellent	78
30	4.7	Dark Green	Excellent	42
31	4.8	Dark Green	Excellent	25
32	4.9	Dark Green	Excellent	61

In [95]:
```
##country names that has given 0 rating
final_df[final_df['Aggregate rating']<1.8].groupby('Country').size().reset_index()
```

Out[95]:

	Country	0
0	Brazil	5
1	India	2139
2	United Kingdom	1
3	United States	3

observation:

1.maximum no of zero rating is given by Indian customers

In [97]:
```
final_df.columns
```

Out[97]:
```
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', 'Is delivering now', 'Switch to order menu',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
       'Votes', 'Country'],
      dtype='object')
```

In [101]:
```
final_df[['Country','Currency']].groupby(['Country','Currency']).size().reset_index()
```

Out[101]:

	Country	Currency	0
0	Australia	Dollar($)	24
1	Brazil	Brazilian Real(R$)	60
2	Canada	Dollar($)	4
3	India	Indian Rupees(Rs.)	8652
4	Indonesia	Indonesian Rupiah(IDR)	21
5	New Zealand	NewZealand($)	40
6	Phillipines	Botswana Pula(P)	22
7	Qatar	Qatari Rial(QR)	20
8	Singapore	Dollar($)	20
9	South Africa	Rand(R)	60
10	Sri Lanka	Sri Lankan Rupee(LKR)	20
11	Turkey	Turkish Lira(TL)	34
12	UAE	Emirati Diram(AED)	60
13	United Kingdom	Pounds(£)	80
14	United States	Dollar($)	434

In [108]:
```
#which countries do have online delivery option
final_df[final_df["Has Online delivery"]=="Yes"].Country.value_counts()
```

Out[108]:
```
India    2423
UAE        28
Name: Country, dtype: int64
```

Observation:

1. online delivery are available in India and UAE

In [109]:
```
#Create a piechart for cities distribution
city_names=final_df.City.value_counts().index
```

In [110]:
```
city_names
```

Out[110]:
```
Index(['New Delhi', 'Gurgaon', 'Noida', 'Faridabad', 'Ghaziabad',
       'Bhubaneshwar', 'Amritsar', 'Ahmedabad', 'Lucknow', 'Guwahati',
       ...
       'Ojo Caliente', 'Montville', 'Monroe', 'Miller', 'Middleton Beach',
       'Panchkula', 'Mc Millan', 'Mayfield', 'Macedon', 'Vineland Station'],
      dtype='object', length=141)
```

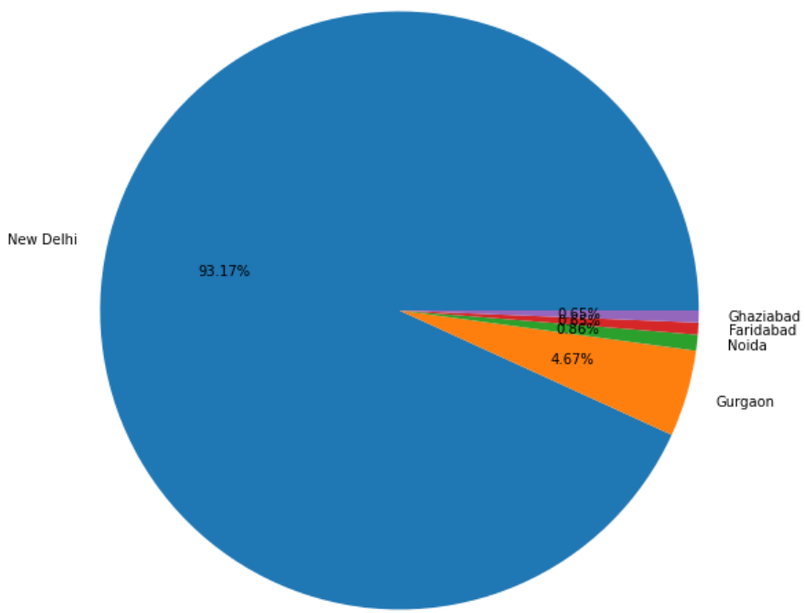
In [112]:
```
city_values=final_df.Country.value_counts().values
```

In [113]:
```
city_values
```

Out[113]:
```
array([8652,  434,   80,   60,   60,   60,   40,   34,   24,   22,   21,
         20,   20,   20,    4], dtype=int64)
```

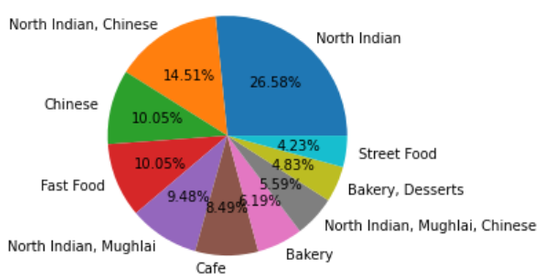
In [115]:
```
from matplotlib import rcParams
plt.figure(figsize=(25,10))
plt.pie(city_values[:5],labels=city_names[:5],autopct='%1.2f%%')
```

Out[115]:
```
([<matplotlib.patches.Wedge at 0x27eb960d220>,
  <matplotlib.patches.Wedge at 0x27eb960d940>,
  <matplotlib.patches.Wedge at 0x27eb96110d0>,
  <matplotlib.patches.Wedge at 0x27eb96117f0>,
  <matplotlib.patches.Wedge at 0x27eb9611f10>],
 [Text(-1.07479320772197, 0.23413577393196094, 'New Delhi'),
  Text(1.056503577535216, -0.3062681678779072, 'Gurgaon'),
  Text(1.093560093908513, -0.118854200644337, 'Noida'),
  Text(1.0979610170116392, -0.06694479160298568, 'Faridabad'),
  Text(1.0997733864424146, -0.022327079365287286, 'Ghaziabad')],
 [Text(-0.5862508405756199, 0.12771042214470596, '93.17%'),
  Text(0.5762746786555722, -0.1670553642970403, '4.67%'),
  Text(0.596487323950098, -0.06482956398782018, '0.86%'),
  Text(0.598887827460894, -0.03651534087435582, '0.65%'),
  Text(0.5998763926049533, -0.012178406926520337, '0.65%')])
```



In [128]:
```
## The top 10 Cuisines
cuisine_names = final_df["Cuisines"].value_counts().index
cuisine_vals = final_df["Cuisines"].value_counts().values
plt.pie(cuisine_vals[:10],labels=cuisine_names[:10],autopct="%1.2f%%")
```

Out[128]:
```
([<matplotlib.patches.Wedge at 0x27eb9810a60>,
  <matplotlib.patches.Wedge at 0x27eb981b1c0>,
  <matplotlib.patches.Wedge at 0x27eb981b8e0>,
  <matplotlib.patches.Wedge at 0x27eb9827040>,
  <matplotlib.patches.Wedge at 0x27eb9827760>,
  <matplotlib.patches.Wedge at 0x27eb9827e80>,
  <matplotlib.patches.Wedge at 0x27eb98335e0>,
  <matplotlib.patches.Wedge at 0x27eb9833d00>,
  <matplotlib.patches.Wedge at 0x27eb9801cd0>,
  <matplotlib.patches.Wedge at 0x27eb962a7f0>],
 [Text(0.7383739846958008, 0.8153550507137645, 'North Indian'),
  Text(-0.5794679314239953, 0.9349956772366362, 'North Indian, Chinese'),
  Text(-1.067309479615702, 0.26617752482593154, 'Chinese'),
  Text(-1.0185984499802057, -0.4152796620326146, 'Fast Food'),
  Text(-0.5935788454809928, -0.9261015895664211, 'North Indian, Mughlai'),
  Text(-0.005887079599915552, -1.0999842463843672, 'Cafe'),
  Text(0.4842062514572988, -0.9876964645323336, 'Bakery'),
  Text(0.808736477166136, -0.7456174022251013, 'North Indian, Mughlai, Chinese'),
  Text(1.0055375294202338, -0.44597564611473206, 'Bakery, Desserts'),
  Text(1.090298995560443, -0.14576728123927227, 'Street Food')],
 [Text(0.4027494461977095, 0.4447391185711442, '26.58%'),
  Text(-0.316073417140361, 0.5099976421290743, '14.51%'),
  Text(-0.5821688070631101, 0.14518774081414446, '10.05%'),
  Text(-0.5555991545346576, -0.22651617929051704, '10.05%'),
  Text(-0.32377027935326874, -0.5051463215816842, '9.48%'),
  Text(-0.003211134327226664, -0.5999914071187457, '8.49%'),
  Text(0.26411250079489024, -0.5387435261085456, '6.19%'),
  Text(0.441128987545165, -0.40670040121369155, '5.59%'),
  Text(0.5484750160474001, -0.24325944333530836, '4.83%'),
  Text(0.5947085430329688, -0.07950942613051214, '4.23%')])
```



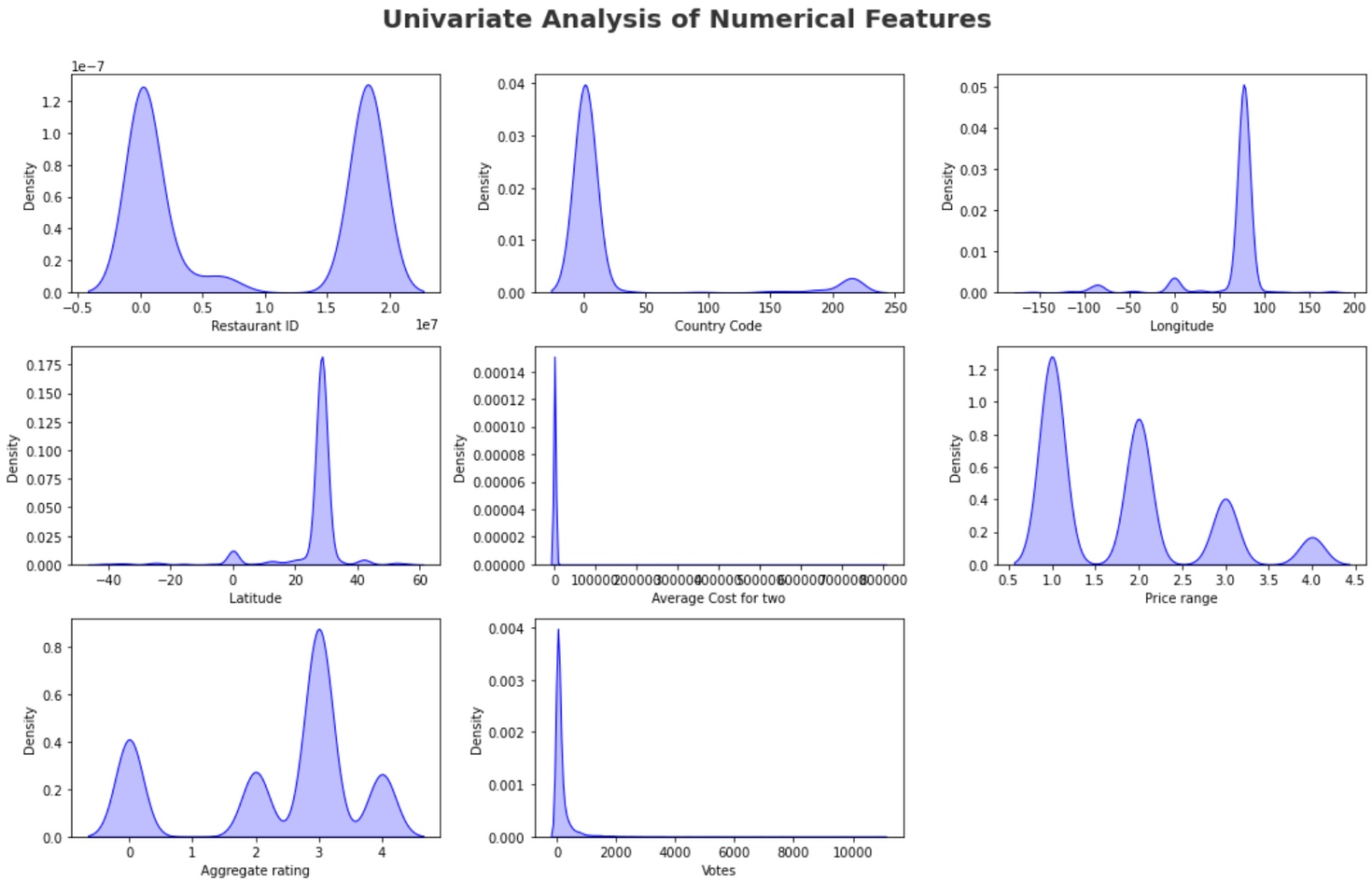
Numerical and Categorical features

In [124]:
```
# define numerical & categorical columns
categorical_features = [i for i in final_df.columns if final_df[i].dtypes == 'object']
numeric_features = [i for i in final_df.columns if final_df[i].dtypes != 'object']
print(categorical_features)
print(numeric_features)
```

```
['Restaurant Name', 'City', 'Address', 'Locality', 'Locality Verbose', 'Cuisines', 'Currency', 'Has Table booking', 'Has Online delivery', 'Is delivering now', 'Switch to order menu', 'Rating color', 'R
ating text', 'Country']
['Restaurant ID', 'Country Code', 'Longitude', 'Latitude', 'Average Cost for two', 'Price range', 'Aggregate rating', 'Votes']
```

In [129]:
```
plt.figure(figsize=(15, 15))
plt.suptitle('Univariate Analysis of Numerical Features', fontsize=20, fontweight='bold', alpha=0.8, y=1.)

for i in range(0, len(numeric_features)):
    plt.subplot(5, 3, i+1)
    sns.kdeplot(x=df[numeric_features[i]],shade=True, color='b')
    plt.xlabel(numeric_features[i])
    plt.tight_layout()
```



In [131]:
```
final_df.head()
```

Out[131]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	...	Has Table booking	Has Online delivery	Is delivering now	Switch to order menu	Price range	Aggregate rating	Rating color	Rating text	Votes	Country
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Century City Mall, Poblacion, Makati City	Century City Mall, Poblacion, Makati City, Mak...	121.027535	14.565443	French, Japanese, Desserts	...	Yes	No	No	No	3	4.8	Dark Green	Excellent	314	Phillipines
1	6304287	Izakaya Kikufuji	162	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...	Little Tokyo, Legaspi Village, Makati City	Little Tokyo, Legaspi Village, Makati City, Ma...	121.014101	14.553708	Japanese	...	Yes	No	No	No	3	4.5	Dark Green	Excellent	591	Phillipines
2	6300002	Heat - Edsa Shangri-La	162	Mandaluyong City	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...	Edsa Shangri-La, Ortigas, Mandaluyong City	Edsa Shangri-La, Ortigas, Mandaluyong City, Ma...	121.056831	14.581404	Seafood, Asian, Filipino, Indian	...	Yes	No	No	No	4	4.4	Green	Very Good	270	Phillipines
3	6318506	Ooma	162	Mandaluyong City	Third Floor, Mega Fashion Hall, SM Megamall, O...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.056475	14.585318	Japanese, Sushi	...	No	No	No	No	4	4.9	Dark Green	Excellent	365	Phillipines
4	6314302	Sambo Kojin	162	Mandaluyong City	Third Floor, Mega Atrium, SM Megamall, Ortigas...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.057508	14.584450	Japanese, Korean	...	Yes	No	No	No	4	4.8	Dark Green	Excellent	229	Phillipines

5 rows × 22 columns

In [ ]:
```
```