

```
In [1]: import pandas as pd

In [2]: import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
import plotly.io as pio

In [3]: dataset=pd.read_csv('Zomato_Mumbai_Dataset.csv', delimiter='|')

In [4]: dataset
```

Out[4]:

	NAME	PRICE	CUSINE_CATEGORY	CITY	REGION	URL	PAGE NO	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
0	Hitchki	1200	Modern Indian,North Indian,Chinese,Momos,Birya...	Mumbai	First International Financial Centre--Bandra ...	https://www.zomato.com/mumbai/hitchki-bandra-k...	1	Casual Dining	12noon to 130am(Mon-Sun)	Excellent	4.9	3529
1	Baba Falooda	400	Desserts,Ice Cream,Beverages	Mumbai	Mahim	https://www.zomato.com/mumbai/baba-falooda-mah...	1	Dessert Parlor	2pm to 1am(Mon-Sun)	Very Good	4.4	1723
2	Chin Chin Chu	1800	Asian,Chinese	Mumbai	Juhu	https://www.zomato.com/mumbai/chin-chin-chu-ju...	1	Casual Dining	12noon to 1am(Mon-Sun)	Very Good	4.2	337
3	Butterfly High	1000	Modern Indian	Mumbai	Bandra Kurla Complex	https://www.zomato.com/mumbai/butterfly-high-b...	1	Bar	12noon to 130am(Mon-Sun)	Very Good	4.3	1200
4	BKC DIVE	1200	North Indian,Chinese,Continental	Mumbai	Bandra Kurla Complex	https://www.zomato.com/mumbai/bkc-dive-bandra-...	1	Bar	1130am to 1am(Mon-Sun)	Vefmi dobré	4.4	5995
...
15076	Hari Om Snack Bar	350	Fast Food,South Indian,Chinese	Mumbai	Kandivali West	https://www.zomato.com/mumbai/hari-om-snack-ba...	99	Quick Bites	11am to 230am(Mon-Sun)	Good	3.7	64
15077	PitaBurg	400	Fast Food,Lebanese	Mumbai	Lower Parel	https://www.zomato.com/mumbai/pitaburg-lower-p...	99	none	11am to 11pm(Mon,Tue,Wed,Thu,Sun),11am to ...	Average	3.4	99
15078	Uncha Otlawala	300	Desserts,Ice Cream	Mumbai	Kandivali West	https://www.zomato.com/mumbai/uncha-otlawala-1...	99	Dessert Parlor	9am to 1230AM(Mon-Sun)	Good	3.5	29
15079	Mandarin Panda	400	Desserts,Chinese,Thai	Mumbai	Malad West	https://www.zomato.com/mumbai/mandarin-panda-m...	99	none	12noon to 330pm,7pm to 1am(Mon-Sun)	Good	3.7	121
15080		NaN		NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

15081 rows × 12 columns

```
In [5]: dataset.shape

Out[5]: (15081, 12)
```

```
In [6]: dataset.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 15081 entries, 0 to 15080
Data columns (total 12 columns):
NAME                15081 non-null object
PRICE              15080 non-null object
CUSINE_CATEGORY     15079 non-null object
CITY                15080 non-null object
REGION             15080 non-null object
URL                15080 non-null object
PAGE_NO            15080 non-null object
CUSINE_TYPE        15080 non-null object
TIMING             15015 non-null object
RATING_TYPE        15080 non-null object
RATING             15080 non-null object
VOTES              15080 non-null object
dtypes: object(12)
memory usage: 1.4+ MB
```

```
In [7]: dataset.describe()

Out[7]:
```

	NAME	PRICE	CUSINE_CATEGORY	CITY	REGION	URL	PAGE NO	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
count	15081	15080	15079	15080	15080	15080	15080	15080	15015	15080	15080	15080
unique	12720	67	3183	2	241	13823	944	23	2551	32	35	1124
top	NAME	400	CUSINE_CATEGORY	Mumbai	REGION	URL	PAGE NO	Quick Bites	11am to 11pm(Mon-Sun)	Average	-	-
freq	942	2042	942	14138	942	942	942	5262	1192	5112	2360	2360

```
In [8]: wrongdata=dataset['PAGE_NO']!='PAGE_NO'

In [9]: wrongdata

Out[9]:
0      False
1      False
2      False
3      False
4      False
...
15076   False
15077   False
15078   False
15079   False
15080   False
Name: PAGE_NO, Length: 15081, dtype: bool

In [10]: dataset[wrongdata]
```

Out[10]:

	NAME	PRICE	CUSINE_CATEGORY	CITY	REGION	URL	PAGE NO	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
15	NAME	PRICE	CUSINE_CATEGORY	CITY	REGION	URL	PAGE NO	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
31	NAME	PRICE	CUSINE_CATEGORY	CITY	REGION	URL	PAGE NO	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
47	NAME	PRICE	CUSINE_CATEGORY	CITY	REGION	URL	PAGE NO	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
63	NAME	PRICE	CUSINE_CATEGORY	CITY	REGION	URL	PAGE NO	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
79	NAME	PRICE	CUSINE_CATEGORY	CITY	REGION	URL	PAGE NO	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
...
15000	NAME	PRICE	CUSINE_CATEGORY	CITY	REGION	URL	PAGE NO	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
15016	NAME	PRICE	CUSINE_CATEGORY	CITY	REGION	URL	PAGE NO	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
15032	NAME	PRICE	CUSINE_CATEGORY	CITY	REGION	URL	PAGE NO	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
15048	NAME	PRICE	CUSINE_CATEGORY	CITY	REGION	URL	PAGE NO	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
15064	NAME	PRICE	CUSINE_CATEGORY	CITY	REGION	URL	PAGE NO	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES

942 rows × 12 columns

In [11]:

dataset=dataset[~wrongdata]

In [12]:

dataset

Out[12]:

	NAME	PRICE	CUSINE_CATEGORY	CITY	REGION	URL	PAGE NO	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
0	Hitchki	1200	Modern Indian,North Indian,Chinese,Momos,Birya...	Mumbai	First International Financial Centre--Bandra ...	https://www.zomato.com/mumbai/hitchki-bandra-k...	1	Casual Dining	12noon to 130am(Mon-Sun)	Excellent	4.9	3529
1	Baba Falooda	400	Desserts,Ice Cream,Beverages	Mumbai	Mahim	https://www.zomato.com/mumbai/baba-falooda-mah...	1	Dessert Parlor	2pm to 1am(Mon-Sun)	Very Good	4.4	1723
2	Chin Chin Chu	1800	Asian,Chinese	Mumbai	Juhu	https://www.zomato.com/mumbai/chin-chin-chu-ju...	1	Casual Dining	12noon to 1am(Mon-Sun)	Very Good	4.2	337
3	Butterfly High	1000	Modern Indian	Mumbai	Bandra Kurla Complex	https://www.zomato.com/mumbai/butterfly-high-b...	1	Bar	12noon to 130am(Mon-Sun)	Very Good	4.3	1200
4	BKC DIVE	1200	North Indian,Chinese,Continental	Mumbai	Bandra Kurla Complex	https://www.zomato.com/mumbai/bkc-dive-bandra-...	1	Bar	1130am to 1am(Mon-Sun)	Velmi dobré	4.4	5995
...
15076	Hari Om Snack Bar	350	Fast Food,South Indian,Chinese	Mumbai	Kandivali West	https://www.zomato.com/mumbai/hari-om-snack-ba...	99	Quick Bites	11am to 230am(Mon-Sun)	Good	3.7	64
15077	PitaBurg	400	Fast Food,Lebanese	Mumbai	Lower Parel	https://www.zomato.com/mumbai/pitaburg-lower-p...	99	none	11pm(Mon,Tue,Wed,Thu,Sun), 11am to ...	Average	3.4	99
15078	Uncha Otlawala	300	Desserts,Ice Cream	Mumbai	Kandivali West	https://www.zomato.com/mumbai/uncha-otlawala-1...	99	Dessert Parlor	9am to 1230AM(Mon-Sun)	Good	3.5	29
15079	Mandarin Panda	400	Desserts,Chinese,Thai	Mumbai	Malad West	https://www.zomato.com/mumbai/mandarin-panda-m...	99	none	12noon to 330pm,7pm to 1am(Mon-Sun)	Good	3.7	121
15080		NaN		NaN	NaN		NaN	NaN	NaN	NaN	NaN	NaN

14139 rows × 12 columns

In [13]:

dataset.drop(['URL','CITY','PAGE NO'],axis=1,inplace=True)

C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\pandas\core\frame.py:4117: SettingWithCopyWarning: A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
errors=errors,

In [14]:

dataset.head()

Out[14]:

	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
0	Hitchki	1200	Modern Indian,North Indian,Chinese,Momos,Birya...	First International Financial Centre--Bandra ...	Casual Dining	12noon to 130am(Mon-Sun)	Excellent	4.9	3529
1	Baba Falooda	400	Desserts,Ice Cream,Beverages	Mahim	Dessert Parlor	2pm to 1am(Mon-Sun)	Very Good	4.4	1723
2	Chin Chin Chu	1800	Asian,Chinese	Juhu	Casual Dining	12noon to 1am(Mon-Sun)	Very Good	4.2	337
3	Butterfly High	1000	Modern Indian	Bandra Kurla Complex	Bar	12noon to 130am(Mon-Sun)	Very Good	4.3	1200
4	BKC DIVE	1200	North Indian,Chinese,Continental	Bandra Kurla Complex	Bar	1130am to 1am(Mon-Sun)	Velmi dobré	4.4	5995

In [15]:

dataset.isnull().sum()

Out[15]:

NAME 0
PRICE 1
CUSINE_CATEGORY 2
REGION 1
CUSINE TYPE 1
TIMING 66
RATING_TYPE 1
RATING 1
VOTES 1
dtype: int64

In [16]:

dataset[dataset['PRICE'].isnull()]

Out[16]:

NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES
15080	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

In [17]:

dataset.drop(labels=15080,axis=0,inplace=True)

In [18]:

dataset.isnull().sum()

Out[18]:

NAME 0
PRICE 0
CUSINE_CATEGORY 1
REGION 0
CUSINE TYPE 0
TIMING 65
RATING_TYPE 0
RATING 0
VOTES 0
dtype: int64

In [19]:

dataset.fillna('NA',inplace=True)

```
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\pandas\core\frame.py:4259: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy
**kwargs
```

```
In [20]: dataset.isnull().sum()
```

```
Out[20]: NAME          0
PRICE          0
Cuisine_Catgory  0
REGION         0
Cuisine_Type   0
TIMING         0
RATING_Type    0
RATING         0
VOTES          0
dtype: int64
```

```
In [21]: dataset['RATING'].value_counts()
```

```
Out[21]: -          2360
3.5       1094
3.4       1036
3.6        960
NEW        953
3.3        926
3.7        917
3.2        801
3.8        782
3.1        734
3.0        622
3.9        596
2.9        409
4.0        408
2.8        309
4.1        298
4.2        199
2.7        170
4.3        148
4.4         99
2.6         77
Opening     57
4.5         46
2.5         39
4.6         32
2.4         26
4.7         13
2.3         10
2.1          5
2.2          4
4.8          4
4.9          2
2.0          1
1.8          1
Name: RATING, dtype: int64
```

```
In [22]: dataset['RATING'].replace(to_replace=['-', 'NEW', 'Opening'], value=0, inplace=True)
```

```
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\pandas\core\generic.py:6786: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy
self._update_inplace(new_data)
```

```
In [23]: dataset['RATING'].value_counts()
```

```
Out[23]: 0          3370
3.5       1094
3.4       1036
3.6        960
3.3        926
3.7        917
3.2        801
3.8        782
3.1        734
3.0        622
3.9        596
2.9        409
4.0        408
2.8        309
4.1        298
4.2        199
2.7        170
4.3        148
4.4         99
2.6         77
4.5         46
2.5         39
4.6         32
2.4         26
4.7         13
2.3         10
2.1          5
4.8          4
2.2          4
4.9          2
2.0          1
1.8          1
Name: RATING, dtype: int64
```

```
In [24]: dataset['VOTES'].value_counts()
```

```
Out[24]: -          2360
NEW         953
4           364
5           320
6           288
...
822          1
672          1
1844         1
748          1
2157         1
Name: VOTES, Length: 1123, dtype: int64
```

```
In [25]: dataset['VOTES'].replace(to_replace=['-', 'NEW', 'Opening'], value=0, inplace=True)
```

```
In [26]: dataset['VOTES'].value_counts()
```

```
Out[26]: 0      3370
         4      364
         5      320
         6      288
         7      277
         ...
        805      1
        748      1
        822      1
        672      1
        1085     1
        Name: VOTES, Length: 1121, dtype: int64
```

```
In [27]: dataset.columns
```

```
Out[27]: Index(['NAME', 'PRICE', 'CUSINE_CATEGORY', 'REGION', 'CUSINE TYPE', 'TIMING',
              'RATING_TYPE', 'RATING', 'VOTES'],
              dtype='object')
```

```
In [28]: dataset['PRICE']=dataset['PRICE'].astype('int64')
         dataset['RATING']=dataset['RATING'].astype('float64')
         dataset['VOTES']=dataset['VOTES'].astype('int64')
```

```
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:1: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
"""Entry point for launching an IPython kernel.
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:2: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:3: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
This is separate from the ipykernel package so we can avoid doing imports until
```

```
In [29]: dataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 14138 entries, 0 to 15079
Data columns (total 9 columns):
 NAME                14138 non-null object
 PRICE              14138 non-null int64
 CUSINE_CATEGORY     14138 non-null object
 REGION             14138 non-null object
 CUSINE TYPE         14138 non-null object
 TIMING             14138 non-null object
 RATING_TYPE         14138 non-null object
 RATING             14138 non-null float64
 VOTES              14138 non-null int64
 dtypes: float64(1), int64(2), object(6)
memory usage: 1.1+ MB
```

```
In [30]: dataset['TIMING'].value_counts()
```

```
Out[30]: 11am to 11pm(Mon-Sun)          1192
         11am to 12midnight(Mon-Sun)    632
         12noon to 12midnight(Mon-Sun)  467
         11am to 1130pm(Mon-Sun)       309
         10am to 10pm(Mon-Sun)         267
         ...
         7am to 1145pm(Mon,Tue,Wed,Thu,Sat,Sun),7am... 1
         12midnight to 3am,6pm to 12midnight(Mon,Wed... 1
         1130am to 330pm,7pm to 12midnight(Mon,Tue... 1
         11am to 1230AM(Mon,Tue,Wed,Thu,Sat,Sun)... 1
         6am to 10pm(Mon,Tue,Wed,Fri,Sat,Sun),Closed... 1
         Name: TIMING, Length: 2551, dtype: int64
```

```
In [31]: temp=dataset['TIMING'].str.split('(',n=1,expand=True)
```

```
In [32]: temp
```

```
Out[32]:
```

	0	1
0	12noon to 130am	Mon-Sun)
1	2pm to 1am	Mon-Sun)
2	12noon to 1am	Mon-Sun)
3	12noon to 130am	Mon-Sun)
4	1130am to 1am	Mon-Sun)
...
15075	8am to 11pm,12midnight to 115am	Mon-Sun)
15076	11am to 230am	Mon-Sun)
15077	11am to 11pm	Mon,Tue,Wed,Thu,Sun),11am to ...
15078	9am to 1230AM	Mon-Sun)
15079	12noon to 330pm,7pm to 1am	Mon-Sun)

14138 rows × 2 columns

```
In [33]: dataset['TIMING']=temp[0]
         dataset['OpenDate']=temp[1]
         dataset
```

```
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:1: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
"""Entry point for launching an IPython kernel.
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:2: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
```

Out[33]:

	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE_TYPE	TIMING	RATING_TYPE	RATING	VOTES	OpenDate
0	Hitchki	1200	Modern Indian,North Indian,Chinese,Momos,Birya...	First International Financial Centre--Bandra ...	Casual Dining	12noon to 130am	Excellent	4.9	3529	Mon-Sun)
1	Baba Falooda	400	Desserts,Ice Cream,Beverages	Mahim	Dessert Parlor	2pm to 1am	Very Good	4.4	1723	Mon-Sun)
2	Chin Chin Chu	1800	Asian,Chinese	Juhu	Casual Dining	12noon to 1am	Very Good	4.2	337	Mon-Sun)
3	Butterfly High	1000	Modern Indian	Bandra Kurla Complex	Bar	12noon to 130am	Very Good	4.3	1200	Mon-Sun)
4	BKC DIVE	1200	North Indian,Chinese,Continental	Bandra Kurla Complex	Bar	1130am to 1am	Velfmi dobré	4.4	5995	Mon-Sun)
...
15075	Tirupati Balaji	500	Chinese,Fast Food,North Indian	Oshiwara-- Andheri West	Casual Dining	8am to 11pm,12midnight to 115am	Good	3.5	267	Mon-Sun)
15076	Hari Om Snack Bar	350	Fast Food,South Indian,Chinese	Kandivali West	Quick Bites	11am to 230am	Good	3.7	64	Mon-Sun)
15077	PitaBurg	400	Fast Food,Lebanese	Lower Parel	none	11am to 11pm	Average	3.4	99	Mon,Tue,Wed,Thu,Sun),11am to ...
15078	Uncha Otlawala	300	Desserts,Ice Cream	Kandivali West	Dessert Parlor	9am to 1230AM	Good	3.5	29	Mon-Sun)
15079	Mandarin Panda	400	Desserts,Chinese,Thai	Malad West	none	12noon to 330pm,7pm to 1am	Good	3.7	121	Mon-Sun)

14138 rows × 10 columns

In [34]:

```
dataset['OpenDate']=dataset['OpenDate'].str.replace(')','',regex=True)
dataset
```

C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:1: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
""Entry point for launching an IPython kernel.

Out[34]:

	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE_TYPE	TIMING	RATING_TYPE	RATING	VOTES	OpenDate
0	Hitchki	1200	Modern Indian,North Indian,Chinese,Momos,Birya...	First International Financial Centre--Bandra ...	Casual Dining	12noon to 130am	Excellent	4.9	3529	Mon-Sun
1	Baba Falooda	400	Desserts,Ice Cream,Beverages	Mahim	Dessert Parlor	2pm to 1am	Very Good	4.4	1723	Mon-Sun
2	Chin Chin Chu	1800	Asian,Chinese	Juhu	Casual Dining	12noon to 1am	Very Good	4.2	337	Mon-Sun
3	Butterfly High	1000	Modern Indian	Bandra Kurla Complex	Bar	12noon to 130am	Very Good	4.3	1200	Mon-Sun
4	BKC DIVE	1200	North Indian,Chinese,Continental	Bandra Kurla Complex	Bar	1130am to 1am	Velfmi dobré	4.4	5995	Mon-Sun
...
15075	Tirupati Balaji	500	Chinese,Fast Food,North Indian	Oshiwara-- Andheri West	Casual Dining	8am to 11pm,12midnight to 115am	Good	3.5	267	Mon-Sun
15076	Hari Om Snack Bar	350	Fast Food,South Indian,Chinese	Kandivali West	Quick Bites	11am to 230am	Good	3.7	64	Mon-Sun
15077	PitaBurg	400	Fast Food,Lebanese	Lower Parel	none	11am to 11pm	Average	3.4	99	Mon,Tue,Wed,Thu,Sun,11am to ...
15078	Uncha Otlawala	300	Desserts,Ice Cream	Kandivali West	Dessert Parlor	9am to 1230AM	Good	3.5	29	Mon-Sun
15079	Mandarin Panda	400	Desserts,Chinese,Thai	Malad West	none	12noon to 330pm,7pm to 1am	Good	3.7	121	Mon-Sun

14138 rows × 10 columns

In [35]:

```
dataset.isnull().sum()
```

Out[35]:

```
NAME          0
PRICE         0
CUSINE_CATEGORY  0
REGION        0
CUSINE_TYPE   0
TIMING        0
RATING_TYPE   0
RATING        0
VOTES         0
OpenDate      160
dtype: int64
```

In [36]:

```
dataset.fillna('NA',inplace=True)
```

In [37]:

```
dataset.isnull().sum()
```

Out[37]:

```
NAME          0
PRICE         0
CUSINE_CATEGORY  0
REGION        0
CUSINE_TYPE   0
TIMING        0
RATING_TYPE   0
RATING        0
VOTES         0
OpenDate      0
dtype: int64
```

In [38]:

```
dataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 14138 entries, 0 to 15079
Data columns (total 10 columns):
NAME          14138 non-null object
PRICE         14138 non-null int64
CUSINE_CATEGORY 14138 non-null object
REGION        14138 non-null object
CUSINE_TYPE   14138 non-null object
TIMING        14138 non-null object
RATING_TYPE   14138 non-null object
RATING        14138 non-null float64
VOTES         14138 non-null int64
OpenDate      14138 non-null object
dtypes: float64(1), int64(2), object(7)
memory usage: 1.2+ MB
```

In [39]:

```
waste=(dataset['RATING']==0.0) | (dataset['VOTES']==0)
dataset[waste]
```

Out[39]:

	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES	OpenDate
32	Hotel Annapoorna Refreshments	400	Maharashtrian,Mughlai,Chinese	Ghansoli	Quick Bites	1030am to 1230AM	Not rated	0.0	0	Mon-Sun
34	Biryani 9	600	Biryani,North Indian	Near Andheri East Station	none	11am to 3am	None	0.0	0	Mon-Sun
36	D Fusion Flavours	350	Chinese	Goregaon East	none	12noon to 330pm,7pm to 3am	None	0.0	0	Mon-Sun
39	Nation Tadka	400	North Indian,South Indian,Chinese,Fast Food	Worli	none	12noon to 1230AM	Not rated	0.0	0	Mon-Sun
83	Link Way Restaurant	500	North Indian,Chinese	Jogeshwari	Quick Bites	12noon to 4pm,8pm to 1am	Not rated	0.0	0	Mon-Sun
...
14998	Foodies House	0	Chinese	Goregaon East	none	12noon to 4am	None	0.0	0	Mon-Sun
14999	Khansama	0	Biryani	Lower Parel	none	12noon to 3am	None	0.0	0	Mon-Sun
15010	Earth Cafe @ Waterfield	800	Cafe,Healthy Food,Italian,Pizza,Beverages	Linking Road-- Bandra West	Café	10am to 10pm	None	0.0	0	Mon-Thu,10am to 11pm(Fri-Sun
15023	How About Some Cream	200	Beverages	Mumbai Central	Beverage Shop	12noon to 3am	None	0.0	0	Mon-Sun
15046	Food And Taste Theory	800	Continental,Italian	Phoenix Marketcity-- Kurla	Casual Dining	9am to 12midnight	None	0.0	0	Mon-Fri,9am to 1230AM(Sat-Sun

3371 rows × 10 columns

In [40]:

dataset=dataset[~waste]

In [41]:

dataset

Out[41]:

	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES	OpenDate
0	Hitchki	1200	Modern Indian,North Indian,Chinese,Momos,Birya...	First International Financial Centre-- Bandra ...	Casual Dining	12noon to 130am	Excellent	4.9	3529	Mon-Sun
1	Baba Falooda	400	Desserts,Ice Cream,Beverages	Mahim	Dessert Parlor	2pm to 1am	Very Good	4.4	1723	Mon-Sun
2	Chin Chin Chu	1800	Asian,Chinese	Juhu	Casual Dining	12noon to 1am	Very Good	4.2	337	Mon-Sun
3	Butterfly High	1000	Modern Indian	Bandra Kurla Complex	Bar	12noon to 130am	Very Good	4.3	1200	Mon-Sun
4	BKC DIVE	1200	North Indian,Chinese,Continental	Bandra Kurla Complex	Bar	1130am to 1am	Velmi dobré	4.4	5995	Mon-Sun
...
15075	Tirupati Balaji	500	Chinese,Fast Food,North Indian	Oshiwara-- Andheri West	Casual Dining	8am to 11pm,12midnight to 115am	Good	3.5	267	Mon-Sun
15076	Hari Om Snack Bar	350	Fast Food,South Indian,Chinese	Kandivali West	Quick Bites	11am to 230am	Good	3.7	64	Mon-Sun
15077	PitaBurg	400	Fast Food,Lebanese	Lower Parel	none	11am to 11pm	Average	3.4	99	Mon,Tue,Wed,Thu,Sun,11am to ...
15078	Uncha Otlawala	300	Desserts,Ice Cream	Kandivali West	Dessert Parlor	9am to 1230AM	Good	3.5	29	Mon-Sun
15079	Mandarin Panda	400	Desserts,Chinese,Thai	Malad West	none	12noon to 330pm,7pm to 1am	Good	3.7	121	Mon-Sun

10767 rows × 10 columns

In [42]:

dataset['RATING_TYPE'].value_counts()

Out[42]:

Average	5111
Good	4330
Very Good	1137
Excellent	95
Poor	47
VeImi dobré	6
Skvělá volba	4
Dobrze	4
Bardzo dobrze	3
Buono	2
Promedio	2
Iyi	2
Bom	2
Ortalama	2
Dobré	2
Bueno	2
Priemer	2
Průměr	2
Muito Bom	2
Velmi dobré	1
Sangat Baik	1
Média	1
Media	1
Biasa	1
Muy Bueno	1
Skvělé	1
Çok iyi	1
Excelente	1
Baik	1

Name: RATING_TYPE, dtype: int64

In [43]:

dataset['RATING_TYPE'].replace(to_replace='Excelente', value='Excellent', inplace=True)
dataset['RATING_TYPE'].replace(to_replace=['VeImi dobré','Bardzo dobrze','Muy Bueno','Velmi dobré'], value='Very Good',inplace=True)
dataset['RATING_TYPE'].replace(to_replace=['Skvělá volba','Dobrze','Bueno','Buono','Dobré','Bom','Skvělé'], value='Good',inplace=True)
dataset['RATING_TYPE'].replace(to_replace=['Priemer','Média','Çok iyi'], value='Average', inplace=True)
dataset['RATING_TYPE'].replace(to_replace=['Průměr','Promedio','Ortalama','Muito Bom','Iyi'], value='Poor', inplace=True)
dataset['RATING_TYPE'].replace(to_replace=['Baik','Biasa','Media','Sangat Baik'], value='Very Poor', inplace=True)

In [44]:

dataset['RATING_TYPE'].value_counts()

Out[44]:

Average	5115
Good	4347
Very Good	1148
Excellent	96
Poor	57
Very Poor	4

Name: RATING_TYPE, dtype: int64

In [45]:

dataset['REGION'].value_counts()

```
Out[45]: Mira Road          405
         Malad West       308
         Chembur          277
         Kharghar         268
         Borivali West    264
         ...
         Girgaon Chowpatty 1
         Aureole Hotel-- Andheri East 1
         Sea Princess-- Juhu 1
         Atria Mall-- Worli 1
         Royal Challenge Complex-- Goregaon 1
         Name: REGION, Length: 237, dtype: int64
```

```
In [46]: dataset['REGION']=dataset['REGION'].str.replace('[a-zA-Z].+--','',regex=True)
```

C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:1: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
""Entry point for launching an IPython kernel.

```
In [47]: dataset['REGION'].value_counts()
```

```
Out[47]: Thane West          712
         Andheri West       407
         Mira Road          405
         Malad West         308
         Bandra West        282
         ...
         Worli              1
         Turbhe             1
         Dadar              1
         Girgaon Chowpatty  1
         Goregaon           1
         Name: REGION, Length: 164, dtype: int64
```

```
In [48]: dataset['REGION']=dataset['REGION'].str.replace('East|east|West|west','',regex=True)
```

C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:1: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
""Entry point for launching an IPython kernel.

```
In [49]: dataset['REGION'].value_counts()
```

```
Out[49]: Thane          712
         Andheri        409
         Mira Road       405
         Malad           370
         Kandivali       361
         ...
         Turbhe          1
         Goregaon        1
         Parel           1
         Churchgate      1
         Worli           1
         Name: REGION, Length: 148, dtype: int64
```

```
In [50]: dataset['REGION'] = dataset['REGION'].str.replace('4 Bungalows|7 Andheri|Azad Nagar|Near Andheri Station|Veera Desai Area','Andheri',regex=True)
         dataset['REGION'] = dataset['REGION'].str.replace('Bandra Kurla Complex','Bandra',regex=True)
         dataset['REGION'] = dataset['REGION'].str.replace('CBD-Belapur','CBD Belapur',regex=True)
         dataset['REGION'] = dataset['REGION'].str.replace('Girgaon Chowpatty','Chowpatty',regex=True)
         dataset['REGION'] = dataset['REGION'].str.replace('Dadar Shivaji Park','Dadar',regex=True)
         dataset['REGION'] = dataset['REGION'].str.replace('Flea Bazaar Café|Kamala Mills Compound','Lower Parel',regex=True)
         dataset['REGION'] = dataset['REGION'].str.replace('Runwal Green','Mulund',regex=True)
         dataset['REGION'] = dataset['REGION'].str.replace('Mumbai CST Area','Mumbai Central',regex=True)
         dataset['REGION'] = dataset['REGION'].str.replace('Kopar Khairane|Seawoods|Turbhe|Ulwe','Navi Mumbai',regex=True)
         dataset['REGION'] = dataset['REGION'].str.replace('New Panvel|Old Panvel','Panvel',regex=True)
         dataset['REGION'] = dataset['REGION'].str.replace('Kamothe','Sion',regex=True)
         dataset['REGION'] = dataset['REGION'].str.replace('Ghodbunder Road|Majiwada','Thane',regex=True)
```

```
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:1: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
"""Entry point for launching an IPython kernel.
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:2: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:3: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
This is separate from the ipykernel package so we can avoid doing imports until
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:4: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
after removing the cwd from sys.path.
C:\Users\Lenovo\.conda\envs\hope\lib\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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
"""
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:6: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:7: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
import sys
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:8: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:9: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
if __name__ == '__main__':
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:10: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
# Remove the CWD from sys.path while we load stuff.
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:11: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
# This is added back by InteractiveShellApp.init_path()
C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\ipykernel_launcher.py:12: 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: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
if sys.path[0] == '':
```

In [51]: dataset['REGION'].value_counts()

Out[51]:

Thane	712
Andheri	409
Mira Road	405
Malad	370
Kandivali	361
...	
Dadar	1
Ulhasnagar	1
Churchgate	1
Parel	1
Chakala	1

Name: REGION, Length: 135, dtype: int64

In [52]: dataset[dataset.duplicated()]

Out[52]:

	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES	OpenDate
4064	Sai Sannidhi Restaurant & Bar	1000	North Indian,Konkan	Dahisar	Casual Dining	11am to 12midnight	Good	3.7	99	Mon-Sun
4068	Konkan Katta	400	Seafood,Maharashtrian,Malwani	Mahakali	Quick Bites	11am to 330pm,630pm to 1130pm	Good	3.5	181	Mon-Sun
4082	Usmaniya Hotel	600	Mughlai	Fort	Casual Dining	1030am to 1130pm	Average	3.2	8	Mon-Sun
4083	Gina's Cakes	450	Bakery	Dombivali	none	11am to 11pm	Good	3.5	49	Mon-Sun
4084	Konkanastha Lunch Home	400	Seafood,Malwani	Chakala	Casual Dining	12noon to 3pm,730pm to 1030pm	Good	3.5	44	Mon-Sun
...
14200	Mezbaan Family Restaurant	350	Chinese,Mughlai	Mumbra	Dhaba	12noon to 1230AM	Average	2.8	97	Mon-Sun
14204	Jyoti Lunch Home	650	Chinese,North Indian,Seafood,Mughlai	Mulund	Casual Dining	11am to 1230AM	Good	3.5	49	Mon-Sun
14253	On Toes	900	Italian,North Indian,Chinese	Malad	Casual Dining	12noon to 3pm,7pm to 1230AM	Good	3.6	76	Mon-Sun
14761	Frosty Farm	400	Ice Cream,Desserts,Fast Food	Malad	Dessert Parlor	1pm to 1215AM	Good	3.6	120	Mon-Sun
14928	Shree Manu Sagar	300	North Indian,Chinese,Indian	Ghansoli	Quick Bites	1130am to 415pm,7pm to 1215AM	Average	3.3	51	Mon-Sun

220 rows × 10 columns

In [53]: dataset=dataset.drop_duplicates()

In [54]: dataset

Out[54]:

	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES	OpenDate
0	Hitchki	1200	Modern Indian,North Indian,Chinese,Momos,Birya...	Bandra	Casual Dining	12noon to 130am	Excellent	4.9	3529	Mon-Sun
1	Baba Falooda	400	Desserts,Ice Cream,Beverages	Mahim	Dessert Parlor	2pm to 1am	Very Good	4.4	1723	Mon-Sun
2	Chin Chin Chu	1800	Asian,Chinese	Juhu	Casual Dining	12noon to 1am	Very Good	4.2	337	Mon-Sun
3	Butterfly High	1000	Modern Indian	Bandra	Bar	12noon to 130am	Very Good	4.3	1200	Mon-Sun
4	BKC DIVE	1200	North Indian,Chinese,Continental	Bandra	Bar	1130am to 1am	Very Good	4.4	5995	Mon-Sun
...
15075	Tirupati Balaji	500	Chinese,Fast Food,North Indian	Andheri	Casual Dining	8am to 11pm,12midnight to 115am	Good	3.5	267	Mon-Sun
15076	Hari Om Snack Bar	350	Fast Food,South Indian,Chinese	Kandivali	Quick Bites	11am to 230am	Good	3.7	64	Mon-Sun
15077	PitaBurg	400	Fast Food,Lebanese	Lower Parel	none	11am to 11pm	Average	3.4	99	Mon,Tue,Wed,Thu,Sun,11am to ...
15078	Uncha Otiawala	300	Desserts,Ice Cream	Kandivali	Dessert Parlor	9am to 1230AM	Good	3.5	29	Mon-Sun
15079	Mandarin Panda	400	Desserts,Chinese,Thai	Malad	none	12noon to 330pm,7pm to 1am	Good	3.7	121	Mon-Sun

10547 rows × 10 columns

1. How many restaurants are in mumbai for each type of cusine?

In [55]:

```
import plotly.express as px
graph1=px.histogram(dataset,x='CUSINE TYPE',color='CUSINE TYPE', title='No. of restaurants by cusine type', labels={'CUSINE TYPE':'cusine type'})
graph1.show()
```

1. What are the percentage of restaurants by Rating Type in Mumbai?

In [56]:

```
dataset.columns
```

Out[56]:

```
Index(['NAME', 'PRICE', 'CUSINE_CATEGORY', 'REGION', 'CUSINE TYPE', 'TIMING',
      'RATING_TYPE', 'RATING', 'VOTES', 'OpenDate'],
      dtype='object')
```

In [57]:

```
ratingtype=dataset['RATING_TYPE'].value_counts().reset_index()
```

In [58]:

```
ratingtype.rename(columns={'index':'Rating_type','RATING_TYPE':'Number of restaurants'},inplace=True)
```

In [59]:

```
ratingtype
```

Out[59]:

	Rating_type	Number of restaurants
0	Average	4983
1	Good	4263
2	Very Good	1145
3	Excellent	96
4	Poor	56
5	Very Poor	4

In [60]:

```
graph2=px.pie(ratingtype,names='Rating_type',values='Number of restaurants',color='Rating_type',title='percentage of restaurants by rating type').update_traces(textposition='inside',t
graph2.show()
```

1. Which are the Top 10 highest rated Seafood Restaurant in Mumbai?

In [61]: `#dataset`

In [62]: `seafood=dataset[dataset['CUSINE_CATEGORY'].str.contains('Seafood')]
seafood.sort_values(by='RATING',ascending=False).head(10)`

Out[62]:

	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE_TYPE	TIMING	RATING_TYPE	RATING	VOTES	OpenDate
7104	Thangabali	1000	Seafood,South Indian,Mangalorean,Andhra,Kerala	Khar	Bar	12noon to 4pm,7pm to 3am	Excellent	4.7	564	Mon-Sun
76	Ceremonial Kitchen & Co	1000	Seafood,Maharashtrian,North Indian,Chinese	Thane	Casual Dining	1130am to 1130pm	Excellent	4.6	350	Mon-Sun
13685	Maharashtra Lunch Home	600	Maharashtrian,Malwani,Konkan,Seafood	Kharghar	Casual Dining	11am to 345pm,7pm to 1145pm	Excellent	4.6	209	Mon-Sun
12433	Quarter Canteen	1100	North Indian,Seafood,Chinese	Bandra	Casual Dining	12noon to 330pm,7pm to 1am	Excellent	4.5	573	Mon-Sun
902	The Harbour Bay - SeaFood Kitchen & Bar	2400	Seafood,Beverages	Bandra	Casual Dining	12noon to 1am	Excellent	4.5	100	Mon-Sun
884	Rajmanya- Seafood family restaurant	800	Maharashtrian,Konkan,Seafood	Vashi	Casual Dining	11am to 11pm	Excellent	4.5	178	Mon-Sun
3380	Peco Peco	700	Chinese,Seafood,Asian	Powai	none	12noon to 330pm,7pm to 1230AM	Excellent	4.5	497	Mon-Sun
9954	Pi Bar and Kitchen	1600	Continental,European,Italian,Seafood,Pizza,Des...	Andheri	Bar	12noon to 6pm,7pm to 12midnight	Excellent	4.5	2068	Mon-Sun
903	Ferry Wharf	1500	Seafood,Mangalorean	Bandra	Casual Dining	11am to 330pm,7pm to 1230AM	Very Good	4.4	459	Mon-Sun
915	Monis Bar and Restaurant	1000	North Indian,Chinese,Continental,Seafood,Bever...	Thane	Casual Dining	1130am to 330pm,6pm to 1130pm	Very Good	4.4	662	Mon-Sun

1. Which is the best Food Truck in Mumbai?

In [63]: `truck=dataset[dataset['CUSINE_TYPE']=='Food Truck']
truck.sort_values(by='RATING',ascending=False).head(2)`

Out[63]:

	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE_TYPE	TIMING	RATING_TYPE	RATING	VOTES	OpenDate
262	Dumpling Delights	200	Momos	Matunga	Food Truck	430pm to 930pm	Very Good	4.3	212	Mon-Sun
1017	Street Food Co.	250	Fast Food,Chinese	Virar	Food Truck	6pm to 3am	Very Good	4.1	274	Mon-Sun

1. Which places have the highest rated restaurant for each Cuisine Type in Mumbai?

In [64]: `area=dataset[dataset['RATING']>=4.5]
area`

Out[64]:

	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE_TYPE	TIMING	RATING_TYPE	RATING	VOTES	OpenDate
0	Hitchki	1200	Modern Indian,North Indian,Chinese,Momos,Birya...	Bandra	Casual Dining	12noon to 130am	Excellent	4.9	3529	Mon-Sun
6	Persian Darbar	1300	Biryani,North Indian,Chinese,Mughlai	Marol	Casual Dining	10am to 3am	Excellent	4.5	3058	Mon-Sun
7	Tanatan	1500	Modern Indian	Juhu	Casual Dining	12noon to 130am	Excellent	4.7	1842	Mon-Sun
9	Plum by Bent Chair	1800	Asian	Lower Parel	Casual Dining	12noon to 1am	Excellent	4.7	1876	Mon-Sun
10	Angrezi Dhaba	1500	North Indian,Chinese,Thai,European	Dadar	Bar	12noon to 1am	Excellent	4.5	2092	Mon-Sun
...
14228	Zaika Crave - Club Aquaria	1300	North Indian,Continental,Chinese,Desserts	Borivali	Casual Dining	11am to 330pm,7pm to 1130pm	Excellent	4.5	1302	Mon,Tue,Wed...
14234	Cone Culture	250	European	Kharghar	Casual Dining	Closed	Excellent	4.6	492	Mon,12noon to 11pm(Tue-Sun
15007	Dessertino	300	Desserts,Ice Cream	Kandivali	Dessert Parlor	11am to 12midnight	Excellent	4.8	184	Mon-Sun
15051	Tick-eat	800	North Indian,Italian,Chinese,Mexican,Lebanese	Mulund	Casual Dining	1130am to 330pm,7pm to 1130pm	Excellent	4.5	754	Mon-Sun
15056	Daftar Goregaon	750	Pizza,Chinese,North Indian,Beverages	Goregaon	Casual Dining	12noon to 1130pm	Excellent	4.6	388	Mon-Sun

97 rows × 10 columns

In [65]: `graph3=px.histogram(area,x='REGION',color='CUSINE_TYPE',title='highest number of restaurant by cuisine type').update_xaxes(categoryorder='total descending')`

In [66]: `graph3.show()`

1. What is the Avg Price Distribution of highest rated restaurant for each Cuisine Type in Mumbai?

```
In [67]: highprice=area.groupby(by=['REGION', 'Cuisine Type'])['PRICE'].mean().reset_index()
highprice
```

```
Out[67]:
```

	REGION	Cuisine Type	PRICE
0	Andheri	Casual Dining	1000.0
1	Andheri	Bar	1600.0
2	Andheri	Café	800.0
3	Andheri	Casual Dining	1212.5
4	Andheri	Dessert Parlor	300.0
...
57	Thane	Café	800.0
58	Thane	Casual Dining	1000.0
59	Thane	Lounge	1400.0
60	Vashi	Casual Dining	700.0
61	Vile Parle	Casual Dining	1100.0

62 rows × 3 columns

```
In [68]: graph4=px.scatter(highprice,x='REGION',y='PRICE',color='Cuisine Type',symbol='Cuisine Type',title='Price Distribution of highest rated restaurant').update_traces(marker_size=10)
graph4.show()
```

1. Which areas have a large number of Chinese Restaurant Market?

```
In [69]: chinese=dataset[dataset['Cuisine_Category'].str.contains('Chinese')]
chinese
```

Out[69]:

	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE TYPE	TIMING	RATING_TYPE	RATING	VOTES	OpenDate
0	Hitchki	1200	Modern Indian,North Indian,Chinese,Momos,Birya...	Bandra	Casual Dining	12noon to 130am	Excellent	4.9	3529	Mon-Sun
2	Chin Chin Chu	1800	Asian,Chinese	Juhu	Casual Dining	12noon to 1am	Very Good	4.2	337	Mon-Sun
4	BKC DIVE	1200	North Indian,Chinese,Continental	Bandra	Bar	1130am to 1am	Very Good	4.4	5995	Mon-Sun
5	Flea Bazaar Café	800	American,Asian,Street Food,North Indian,Luckno...	Lower Parel	Café	12noon to 1am	Very Good	4.2	2042	Mon-Sun
6	Persian Darbar	1300	Biryani,North Indian,Chinese,Mughlai	Marol	Casual Dining	10am to 3am	Excellent	4.5	3058	Mon-Sun
...
15071	Lucknow Zaika	500	North Indian,Chinese	Kurla	Quick Bites	12noon to 2am	Average	2.6	36	Mon-Sun
15072	Zuha's Kitchen	400	Chinese,North Indian,Mughlai	Mumbai Central	Quick Bites	12noon to 4pm,730pm to 430am	Average	3.3	13	Mon-Sun
15075	Tirupati Balaji	500	Chinese,Fast Food,North Indian	Andheri	Casual Dining	8am to 11pm,12midnight to 115am	Good	3.5	267	Mon-Sun
15076	Hari Om Snack Bar	350	Fast Food,South Indian,Chinese	Kandivali	Quick Bites	11am to 230am	Good	3.7	64	Mon-Sun
15079	Mandarin Panda	400	Desserts,Chinese,Thai	Malad	none	12noon to 330pm,7pm to 1am	Good	3.7	121	Mon-Sun

5119 rows × 10 columns

In [84]:

```
chineseArea=chinese.groupby(by='REGION').agg({'NAME' : 'count', 'PRICE' : 'mean'}).rename(columns={'REGION':'Region','NAME':'COUNT OF RESTAURANTS'})
chineseArea=chineseArea.sort_values('COUNT OF RESTAURANTS',ascending=False).head(25)
chineseArea
```

Out[84]:

	COUNT OF RESTAURANTS	PRICE
REGION		
Thane	341	589.017595
Mira Road	220	545.681818
Malad	185	599.189189
Goregaon	160	540.937500
Dombivai	158	539.259494
Chembur	149	592.617450
Kandivali	139	574.244604
Andheri	137	658.394161
Kalyan	132	520.833333
Borivali	129	641.860465
Kharghar	124	535.887097
Mulund	121	581.818182
Vasai	111	552.252252
Sion	100	509.800000
Vashi	99	629.797980
Virar	95	500.000000
Navi Mumbai	94	562.765957
Ghatkopar	87	604.597701
Panvel	86	545.930233
Powai	86	679.069767
Near Andheri Station	85	567.058824
Santacruz	81	566.049383
Nalasopara	77	476.623377
Airoli	76	507.236842
Bhandup	73	522.602740

In [85]:

```
Region=chineseArea.index
```

In [86]:

```
graph5=px.bar(chineseArea, x=Region, y='COUNT OF RESTAURANTS',color='PRICE',title='No. of Chinese Restaurant by Places')
graph5.show()
```

1. Is there a relation between Price and Rating by each Cuisine Type?

In [87]:

```
priceRating=dataset.groupby(['CUSINE TYPE','RATING'])['PRICE'].mean().reset_index()
priceRating
```

```
Out[87]:
```

	CUSINE TYPE	RATING	PRICE
0	Bakery	2.7	400.000000
1	Bakery	2.8	285.714286
2	Bakery	2.9	328.571429
3	Bakery	3.0	300.000000
4	Bakery	3.1	369.117647
...
278	none	4.3	683.333333
279	none	4.4	555.000000
280	none	4.5	420.000000
281	none	4.6	687.500000
282	none	4.7	350.000000

283 rows × 3 columns

```
In [88]: graph6=px.line(priceRating, x='RATING',y='PRICE',color='CUSINE TYPE')
graph6.show()
```

1. Is there a relation between Region and Price?

```
In [97]: regionPrice=dataset.groupby(['REGION'])(['PRICE']).mean().reset_index()
regionPrice
```

```
Out[97]:
```

	REGION	PRICE
0	Andheri	1250.000000
1	Andheri	590.049383
2	Bandra	1675.000000
3	Bandra	718.273381
4	Bhayandar	150.000000
...
130	Vikhroli	480.434783
131	Vile Parle	464.457831
132	Virar	452.027027
133	Wadala	427.500000
134	Worli	517.763158

135 rows × 2 columns

```
In [98]: graph7=px.scatter(regionPrice,x='REGION',y='PRICE').update_traces(marker_size=8)
graph7.show()
```

1. Find the list of Affordable Restaurants?

```
In [99]: max_price = dataset['PRICE'].max()
Quater_price = max_price/4
Quater_price

Out[99]: 1250.0
```

```
In [115]: affordable=dataset[['NAME', 'PRICE', 'CUSINE_CATEGORY', 'REGION', 'CUSINE TYPE','RATING']]
affordable=affordable[affordable['PRICE']<=1250]
affordable.sort_values(by='PRICE',inplace=True)
affordable
```

	NAME	PRICE		CUSINE_CATEGORY	REGION	CUSINE TYPE	RATING
6137	Sanjog Wine N Dine	5		North Indian,Chinese	Thane	Casual Dining	3.5
2925	Jab We Eat	50	South Indian,North Indian,Maharashtrian,Fast Food	Girgaum	none		3.3
9598	Ho5 Store	50		Fast Food	Matunga	none	3.2
9589	Golden Butterfly	100		Bakery,Desserts	Mira Road	Bakery	3.0
5916	Madhuri Puranpoli	100		Maharashtrian	Vile Parle	none	3.2
...
2740	Peninsula Next	1200		North Indian,Mughlai,Chinese	Sion	Casual Dining	3.8
5528	The Thekka	1200	Finger Food,Continental,North Indian,Chinese	Vashi	Lounge		4.0
964	Bijoli Grill	1250		Bengali	Powai	Casual Dining	3.8
6045	Fabelle at The Chocolate Boutique - ITC Grand ...	1250		Desserts	Parel	Dessert Parlor	3.7
7301	SamBar Pub & Kitchen	1250	Finger Food,South Indian,North Indian	Khar	Pub		3.9

10190 rows × 6 columns

```
In [114]: highrating=dataset[['NAME', 'PRICE', 'CUSINE_CATEGORY', 'REGION', 'CUSINE TYPE','RATING']]
highrating=highrating[highrating['RATING']<=4.5]
highrating.sort_values(by='PRICE',inplace=True)
highrating
```

	NAME	PRICE		CUSINE_CATEGORY	REGION	CUSINE TYPE	RATING
6137	Sanjog Wine N Dine	5		North Indian,Chinese	Thane	Casual Dining	3.5
9598	Ho5 Store	50		Fast Food	Matunga	none	3.2
2925	Jab We Eat	50	South Indian,North Indian,Maharashtrian,Fast Food	Girgaum	none		3.3
11075	Majalkar Snacks	100		Fast Food	Vile Parle	Quick Bites	3.2
6957	Graduate Vadapav	100		Street Food	Byculla	Kiosk	3.5
...
15069	Uname	3000		Japanese	Churchgate	Fine Dining	4.2
1333	Haoochi - Hotel Sun N Sand	3000		Chinese,Thai	Juhu	Fine Dining	4.0
1740	The Oriental Blossom - Marine Plaza	3200		Chinese,Seafood	Marine Lines	Fine Dining	4.2
3561	Olive Bar & Kitchen	3500	American,Mediterranean,Italian,European	Khar	Fine Dining		4.2
4462	24/7 Restaurant - The Lalit Mumbai	5000	European,Italian,North Indian,Asian	Chakala	Fine Dining		4.0

10496 rows × 6 columns

```
In [122]: highrateAfford=pd.merge(affordable,highrating,how='inner',on=['NAME', 'REGION'])
highrateAfford
```

	NAME	PRICE_x		CUSINE_CATEGORY_x	REGION	CUSINE TYPE_x	RATING_x	PRICE_y		CUSINE_CATEGORY_y	CUSINE TYPE_y	RATING_y
0	Sanjog Wine N Dine	5		North Indian,Chinese	Thane	Casual Dining	3.5	5		North Indian,Chinese	Casual Dining	3.5
1	Jab We Eat	50	South Indian,North Indian,Maharashtrian,Fast Food	Girgaum	none		3.3	50	South Indian,North Indian,Maharashtrian,Fast Food	none		3.3
2	Ho5 Store	50		Fast Food	Matunga	none	3.2	50	Fast Food	none		3.2
3	Golden Butterfly	100		Bakery,Desserts	Mira Road	Bakery	3.0	100	Bakery,Desserts	Bakery		3.0
4	Madhuri Puranpoli	100		Maharashtrian	Vile Parle	none	3.2	100	Maharashtrian	none		3.2
...
10249	Peninsula Next	1200		North Indian,Mughlai,Chinese	Sion	Casual Dining	3.8	1200	North Indian,Mughlai,Chinese	Casual Dining		3.8
10250	The Thekka	1200	Finger Food,Continental,North Indian,Chinese	Vashi	Lounge		4.0	1200	Finger Food,Continental,North Indian,Chinese	Lounge		4.0
10251	Bijoli Grill	1250		Bengali	Powai	Casual Dining	3.8	1250	Bengali	Casual Dining		3.8
10252	Fabelle at The Chocolate Boutique - ITC Grand ...	1250		Desserts	Parel	Dessert Parlor	3.7	1250	Desserts	Dessert Parlor		3.7
10253	SamBar Pub & Kitchen	1250	Finger Food,South Indian,North Indian	Khar	Pub		3.9	1250	Finger Food,South Indian,North Indian	Pub		3.9

10254 rows × 10 columns

```
In [123]: highrateAfford=highrateAfford[['NAME', 'PRICE_x', 'CUSINE_CATEGORY_x', 'REGION', 'CUSINE TYPE_x', 'RATING_x']]
highrateAfford
```

	NAME	PRICE_x		CUSINE_CATEGORY_x	REGION	CUSINE TYPE_x	RATING_x
0	Sanjog Wine N Dine	5		North Indian,Chinese	Thane	Casual Dining	3.5
1	Jab We Eat	50	South Indian,North Indian,Maharashtrian,Fast Food	Girgaum	none		3.3
2	Ho5 Store	50		Fast Food	Matunga	none	3.2
3	Golden Butterfly	100		Bakery,Desserts	Mira Road	Bakery	3.0
4	Madhuri Puranpoli	100		Maharashtrian	Vile Parle	none	3.2
...
10249	Peninsula Next	1200		North Indian,Mughlai,Chinese	Sion	Casual Dining	3.8
10250	The Thekka	1200	Finger Food,Continental,North Indian,Chinese	Vashi	Lounge		4.0
10251	Bijoli Grill	1250		Bengali	Powai	Casual Dining	3.8
10252	Fabelle at The Chocolate Boutique - ITC Grand ...	1250		Desserts	Parel	Dessert Parlor	3.7
10253	SamBar Pub & Kitchen	1250	Finger Food,South Indian,North Indian	Khar	Pub		3.9

10254 rows × 6 columns

```
In [124]: highrateAfford.rename(columns={'NAME':'NAME', 'PRICE_x':'PRICE', 'CUSINE_CATEGORY_x':'CUSINE_CATEGORY',
```

```
'REGION': 'REGION', 'CUSINE_TYPE_x': 'CUSINE TYPE', 'RATING_x': 'RATING'), inplace=True)
highrateAfford
```

C:\Users\Lenovo\.conda\envs\hope\lib\site-packages\pandas\core\frame.py:4238: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame

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

Out[124]:

	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE TYPE	RATING
0	Sanjog Wine N Dine	5	North Indian,Chinese	Thane	Casual Dining	3.5
1	Jab We Eat	50	South Indian,North Indian,Maharashtrian,Fast Food	Girgaum	none	3.3
2	Ho5 Store	50	Fast Food	Matunga	none	3.2
3	Golden Butterfly	100	Bakery,Desserts	Mira Road	Bakery	3.0
4	Madhuri Puranpoli	100	Maharashtrian	Vile Parle	none	3.2
...
10249	Peninsula Next	1200	North Indian,Mughlai,Chinese	Sion	Casual Dining	3.8
10250	The Thekka	1200	Finger Food,Continental,North Indian,Chinese	Vashi	Lounge	4.0
10251	Bijoli Grill	1250	Bengali	Powai	Casual Dining	3.8
10252	Fabelle at The Chocolate Boutique - ITC Grand ...	1250	Desserts	Parel	Dessert Parlor	3.7
10253	SamBar Pub & Kitchen	1250	Finger Food,South Indian,North Indian	Khar	Pub	3.9

10254 rows × 6 columns

1. Find the list of most Reliable Restaurants?

In [125... dataset['VOTES'].mean()

Out[125]: 177.2656679624538

In [127... votes=dataset[['NAME', 'PRICE', 'CUSINE_CATEGORY', 'REGION', 'CUSINE TYPE', 'VOTES']]
votes=votes[votes['VOTES']>177]
votes.sort_values(by='VOTES', inplace=True)
votes

Out[127]:

	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE TYPE	VOTES
4194	Sai Sagar Veg Treat	500	North Indian,South Indian,Chinese,Fast Food,Be...	Kalyan	Casual Dining	178
884	Rajmanya- Seafood family restaurant	800	Maharashtrian,Konkan,Seafood	Vashi	Casual Dining	178
3914	Ice Cafe	500	Fast Food,Ice Cream,Beverages,Pizza	Borivali	Quick Bites	178
7897	Konkan Lajjattdar	500	Seafood,Biryani,Beverages,Chinese,Malwani,Konkan	Andheri	Casual Dining	178
3828	Frozen Delight -The Dessert Cafe	250	Desserts,Ice Cream	Airoli	Dessert Parlor	178
...
8539	Leopold Cafe & Bar	1600	American,Chinese,Mughlai,Italian	Colaba	Casual Dining	7327
1251	Joey's Pizza	800	Pizza	Malad	Quick Bites	7350
5337	Chilli's American Grill & Bar	1400	American,Mexican,Burger,Tex-Mex	Powai	Casual Dining	7377
3751	Prithvi Cafe	700	Cafe,Fast Food	Juhu	Café	8000
8897	Candies	700	Cafe,Italian,North Indian,Desserts	Bandra	Café	10217

2345 rows × 6 columns

In [136... reliable=pd.merge(highrateAfford,votes,how='inner',on=['NAME', 'REGION'])
reliable

Out[136]:

	NAME	PRICE_x	CUSINE_CATEGORY_x	REGION	CUSINE TYPE_x	RATING	PRICE_y	CUSINE_CATEGORY_y	CUSINE TYPE_y	VOTES
0	Laxman's Om Vada Pav	100	Street Food	Ghatkopar	Quick Bites	4.0	100	Street Food	Quick Bites	342
1	Ashok Vada Pav Stall	100	Street Food	Prabhadevi	Quick Bites	4.4	100	Street Food	Quick Bites	915
2	Manjeet Chole Puri Wala	100	North Indian,Street Food	Sion	Quick Bites	3.9	100	North Indian,Street Food	Quick Bites	248
3	Kunjvihar Vada Pav Shop	100	Street Food	Thane	Quick Bites	3.9	100	Street Food	Quick Bites	574
4	Eden Chaat Corner	150	Fast Food,Street Food	Powai	Quick Bites	3.5	150	Fast Food,Street Food	Quick Bites	180
...
2081	Vicious Circle Kitchen and Bar	1200	North Indian,Chinese,Mughlai	Vashi	Casual Dining	3.8	1200	North Indian,Chinese,Mughlai	Casual Dining	185
2082	Peninsula Next	1200	North Indian,Mughlai,Chinese	Sion	Casual Dining	3.8	1200	North Indian,Mughlai,Chinese	Casual Dining	354
2083	The Thekka	1200	Finger Food,Continental,North Indian,Chinese	Vashi	Lounge	4.0	1200	Finger Food,Continental,North Indian,Chinese	Lounge	854
2084	Bijoli Grill	1250	Bengali	Powai	Casual Dining	3.8	1250	Bengali	Casual Dining	813
2085	SamBar Pub & Kitchen	1250	Finger Food,South Indian,North Indian	Khar	Pub	3.9	1250	Finger Food,South Indian,North Indian	Pub	1090

2086 rows × 10 columns

In [137... reliable=reliable[['NAME', 'PRICE_x', 'CUSINE_CATEGORY_x', 'REGION', 'CUSINE TYPE_x', 'VOTES']]
reliable

Out[137]:

	NAME	PRICE_x	CUSINE_CATEGORY_x	REGION	CUSINE TYPE_x	VOTES
0	Laxman's Om Vada Pav	100	Street Food	Ghatkopar	Quick Bites	342
1	Ashok Vada Pav Stall	100	Street Food	Prabhadevi	Quick Bites	915
2	Manjeet Chole Puri Wala	100	North Indian,Street Food	Sion	Quick Bites	248
3	Kunjvihar Vada Pav Shop	100	Street Food	Thane	Quick Bites	574
4	Eden Chaat Corner	150	Fast Food,Street Food	Powai	Quick Bites	180
...
2081	Vicious Circle Kitchen and Bar	1200	North Indian,Chinese,Mughlai	Vashi	Casual Dining	185
2082	Peninsula Next	1200	North Indian,Mughlai,Chinese	Sion	Casual Dining	354
2083	The Thekka	1200	Finger Food,Continental,North Indian,Chinese	Vashi	Lounge	854
2084	Bijoli Grill	1250	Bengali	Powai	Casual Dining	813
2085	SamBar Pub & Kitchen	1250	Finger Food,South Indian,North Indian	Khar	Pub	1090

2086 rows × 6 columns

In [138... reliable.rename(columns={'PRICE_x': 'PRICE', 'CUSINE_CATEGORY_x': 'CUSINE_CATEGORY', 'CUSINE TYPE_x': 'CUSINE TYPE'}, inplace=True)
reliable

Out[138]:

	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE_TYPE	VOTES
0	Laxman's Om Vada Pav	100	Street Food	Ghatkopar	Quick Bites	342
1	Ashok Vada Pav Stall	100	Street Food	Prabhadevi	Quick Bites	915
2	Manjeet Chole Puri Wala	100	North Indian,Street Food	Sion	Quick Bites	248
3	Kunjvihar Vada Pav Shop	100	Street Food	Thane	Quick Bites	574
4	Eden Chaat Corner	150	Fast Food,Street Food	Powai	Quick Bites	180
...
2081	Vicious Circle Kitchen and Bar	1200	North Indian,Chinese,Mughlai	Vashi	Casual Dining	185
2082	Peninsula Next	1200	North Indian,Mughlai,Chinese	Sion	Casual Dining	354
2083	The Thekka	1200	Finger Food,Continental,North Indian,Chinese	Vashi	Lounge	854
2084	Bijoli Grill	1250	Bengali	Powai	Casual Dining	813
2085	SamBar Pub & Kitchen	1250	Finger Food,South Indian,North Indian	Khar	Pub	1090

2086 rows × 6 columns

In []:

In []: