

Project Name - Indian Restaurants

Project Type - EDA

Project Summary -

"The zomato capstone project aims to provide data-driven insights for improving customer experience and restaurant success.

By analyzing ratings, locations, cuisines, and delivery options, the project identifies factors that influence restaurant

popularity and customer engagement."

Project Goals

- Knowing basic composition of data
- Removing duplicates
- Dealing with missing values
- Understanding features
- Plotting horizontal bar charts (multicolor)
- Using groupby, apply, and unique functions
- Scatter plot
- Word Cloud
- Box plot
- Density plot
- Bar Charts
- Drawing insights and conclusions from data

Data:

The project would require access to Indian Restaurants data. This data might include information such as

- res_id
- name
- establishment
- url
- address
- city
- city_id
- locality
- latitude
- longitude
- zipcode
- country_id
- locality_verbose
- cuisines
- timings
- average_cost_for_two
- price_range
- currency
- highlights
- aggregate_rating
- rating_text
- votes
- photo_count
- opentable_support
- delivery

- takeaway

Objective

To analyze zomato restaurant data to uncover insights related to:

- *Customer preferences
- *Restaurant performance
- *Pricing trends
- *Location-based demand
- *Online delivery & rating impact

Data Loading:

- Reading the sales data from a CSV file, Excel sheet, or a database into a Pandas

```
DataFrame import pandas as pd df = pd.read_csv('Downloads/Indian-Restaurants.csv')
```

Let's begin !

Know your data

Import Libraries

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

Load Dataset

```
df=pd.read_csv("Downloads/Indian-Restaurants.csv")
```

df

```
      res_id          name   establishment \
0    3400299      Bikanervala  ['Quick Bites']
1  3400005  Mama Chicken Mama Franky House  ['Quick Bites']
2  3401013        Bhagat Halwai  ['Quick Bites']
3  3400290        Bhagat Halwai  ['Quick Bites']
4  3401744  The Salt Cafe Kitchen & Bar  ['Casual Dining']
```

```
...     ...     ...     ...     ...
211939  3202251  Kali Mirch Cafe And Restaurant  ['Casual Dining']
211940  3200996          Raju Omlet    ['Quick Bites']
211941  18984164        The Grand Thakar  ['Casual Dining']
211942  3201138           Subway      ['Quick Bites']
211943  18879846  Freshco's - The Health Cafe  ['Café']
```

```
url \
0  https://www.zomato.com/agra/bikanervala-khanda...
1  https://www.zomato.com/agra/mama-chicken-mama-...
2  https://www.zomato.com/agra/bhagat-halwai-2-sh...
3  https://www.zomato.com/agra/bhagat-halwai-civi...
4  https://www.zomato.com/agra/the-salt-cafe-kitc...
...
211939 https://www.zomato.com/vadodara/kali-mirch-caf...
211940 https://www.zomato.com/vadodara/raju-omlet-kar...
211941 https://www.zomato.com/vadodara/the-grand-thak...
211942 https://www.zomato.com/vadodara/subway-1-akota...
211943 https://www.zomato.com/vadodara/freshcos-the-h...
```

		address	city
city_id \			
0	Kalyani Point, Near Tulsi Cinema, Bypass Road,...		Agra
34	Main Market, Sadar Bazaar, Agra Cantt, Agra		Agra
34	62/1, Near Easy Day, West Shivaji Nagar, Goalp...		Agra
34	Near Anjana Cinema, Nehru Nagar, Civil Lines, ...		Agra
34	1C,3rd Floor, Fatehabad Road, Tajganj, Agra		Agra
34
...
211939	Manu Smriti Complex, Near Navrachna School, GI...	Vadodara	
32	Mahalaxmi Apartment, Opposite B O B, Karoli Ba...	Vadodara	
32	3rd Floor, Shreem Shalini Mall, Opposite Conqu...	Vadodara	
32	G-2, Vedant Platina, Near Cosmos, Akota, Vadodara	Vadodara	
32	Shop 7, Ground Floor, Opposite Natubhai Circle...	Vadodara	

	locality	latitude	longitude	...	price_range
currency \					
0	Khandari	27.211450	78.002381	...	2 Rs.
1	Agra Cantt	27.160569	78.011583	...	2 Rs.

	name	lat	lon	rating	votes	open	support	delivery								
2	Shahganj	27.182938	77.979684	...		1	Rs.									
3	Civil Lines	27.205668	78.004799	...		1	Rs.									
4	Tajganj	27.157709	78.052421	...		3	Rs.									
...								
211939	Fatehgunj	22.336931	73.192356	...		2	Rs.									
211940	Karelibaug	22.322455	73.197203	...		1	Rs.									
211941	Alkapuri	22.310563	73.171163	...		2	Rs.									
211942	Akota	22.270027	73.143068	...		2	Rs.									
211943	Vadiwadi	22.309935	73.158768	...		2	Rs.									
highlights																
aggregate_rating \n0	['Lunch', 'Takeaway Available', 'Credit Card', ...															
4.4	['Delivery', 'No Alcohol Available', 'Dinner', ...															
4.4	['No Alcohol Available', 'Dinner', 'Takeaway A...']															
4.2	['Takeaway Available', 'Credit Card', 'Lunch', ...]															
3	['Lunch', 'Serves Alcohol', 'Cash', 'Credit Ca...']															
4.3	['Lunch', 'Serves Alcohol', 'Cash', 'Credit Ca...']															
4.9	[...]															
...	[...]															
...	[...]															
211939	['Dinner', 'Cash', 'Lunch', 'Delivery', 'Indoo...']															
4.1	['Dinner', 'Cash', 'Takeaway Available', 'Debi...']															
4.1	['Dinner', 'Cash', 'Debit Card', 'Lunch', 'Tak...']															
4.0	['Dinner', 'Delivery', 'Credit Card', 'Lunch', ...]															
3.7	['Dinner', 'Cash', 'Takeaway Available', 'Debi...']															
4.0	['Dinner', 'Cash', 'Takeaway Available', 'Debi...']															
rating_text votes photo_count opentable_support delivery																
takeaway																
0	Very Good	814		154		0.0	-1									
-1																
1	Very Good	1203		161		0.0	-1									

```

-1
2      Very Good    801       107        0.0      1
-1
3      Very Good    693       157        0.0      1
-1
4      Excellent    470       291        0.0      1
-1
...
...
211939  Very Good    243        40        0.0     -1
-1
211940  Very Good    187        40        0.0      1
-1
211941  Very Good    111        38        0.0     -1
-1
211942      Good     128        34        0.0      1
-1
211943  Very Good     93        53        0.0      1
-1

```

[211944 rows x 26 columns]

Use head function to look for first 10 rows

df.head(10)

	res_id	name	establishment	url
0	3400299	Bikanervala	['Quick Bites']	https://www.zomato.com/agra/bikanervala-khanda...
1	3400005	Mama Chicken Mama Franky House	['Quick Bites']	https://www.zomato.com/agra/mama-chicken-mama-...
2	3401013	Bhagat Halwai	['Quick Bites']	https://www.zomato.com/agra/bhagat-halwai-2-sh...
3	3400290	Bhagat Halwai	['Quick Bites']	https://www.zomato.com/agra/bhagat-halwai-civi...
4	3401744	The Salt Cafe Kitchen & Bar	['Casual Dining']	https://www.zomato.com/agra/the-salt-cafe-kitc...
5	3400275	Domino's Pizza	['Quick Bites']	https://www.zomato.com/agra/domino-s-pizza-civi...
6	3400296	Honeydew Restaurant	['Quick Bites']	https://www.zomato.com/agra/honeydew-restauran...
7	3400368	Domino's Pizza	['Quick Bites']	https://www.zomato.com/agra/domino-s-pizza-sika...
8	3401284	Cake House	['Bakery']	https://www.zomato.com/agra/cake-house-2-civil...
9	3400838	Sugar N Thyme	['Café']	https://www.zomato.com/agra/sugar-n-thyme-2-civil...

```
9 https://www.zomato.com/agra/sugar-n-thyme-tajg...
```

		address	city	city_id	\
0	Kalyani Point, Near Tulsi Cinema, Bypass Road,...	Agra	34		
1	Main Market, Sadar Bazaar, Agra Cantt, Agra	Agra	34		
2	62/1, Near Easy Day, West Shivaji Nagar, Goalp...	Agra	34		
3	Near Anjana Cinema, Nehru Nagar, Civil Lines, ...	Agra	34		
4	1C,3rd Floor, Fatehabad Road, Tajganj, Agra	Agra	34		
5	114/23 G, Deep Shikha Complex, Sanjay Place, C...	Agra	34		
6	Opposite Soami Bagh Temple, Dayal Bagh, Agra	Agra	34		
7	Plot C-1/6, Sector 13, Sikandra, Agra	Agra	34		
8	23/301, Wazirpura Rd, Judge Compound Chowraha,...	Agra	34		
9	1374 K/1375 K, Ground floor, Dinesh Nagar, Fat...	Agra	34		
	locality	latitude	longitude	...	price_range currency \
0	Khandari	27.211450	78.002381	...	2 Rs.
1	Agra Cantt	27.160569	78.011583	...	2 Rs.
2	Shahganj	27.182938	77.979684	...	1 Rs.
3	Civil Lines	27.205668	78.004799	...	1 Rs.
4	Tajganj	27.157709	78.052421	...	3 Rs.
5	Civil Lines	27.201516	78.007556	...	2 Rs.
6	Dayal Bagh	27.222175	78.010174	...	2 Rs.
7	Sikandra	27.203930	77.954260	...	2 Rs.
8	Civil Lines	27.204148	78.009025	...	2 Rs.
9	Tajganj	27.158243	78.045591	...	3 Rs.
			highlights	aggregate_rating	
\					
0	['Lunch', 'Takeaway Available', 'Credit Card',...]				4.4
1	['Delivery', 'No Alcohol Available', 'Dinner',...]				4.4
2	['No Alcohol Available', 'Dinner', 'Takeaway A...]				4.2
3	['Takeaway Available', 'Credit Card', 'Lunch',...]				4.3
4	['Lunch', 'Serves Alcohol', 'Cash', 'Credit Ca...]				4.9
5	['Credit Card', 'Lunch', 'Delivery', 'Dinner',...]				4.0
6	['Dinner', 'Delivery', 'Lunch', 'Cash', 'Takea...]				4.2
7	['Lunch', 'Delivery', 'Credit Card', 'No Alcoh...]				3.8
8	['Takeaway Available', 'Cash', 'Indoor Seating...]				3.4
9	['No Alcohol Available', 'Dinner', 'Delivery',...]				4.4
	rating_text	votes	photo_count	opentable_support	delivery takeaway

0	Very Good	814	154	0.0	-1	-1
1	Very Good	1203	161	0.0	-1	-1
2	Very Good	801	107	0.0	1	-1
3	Very Good	693	157	0.0	1	-1
4	Excellent	470	291	0.0	1	-1
5	Very Good	707	62	0.0	-1	-1
6	Very Good	647	46	0.0	1	-1
7	Good	617	18	0.0	-1	-1
8	Average	322	14	0.0	1	-1
9	Very Good	289	324	0.0	1	-1

[10 rows x 26 columns]

Dataset Rows and Columns count

```
df.shape
```

```
(211944, 26)
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 211944 entries, 0 to 211943
Data columns (total 26 columns):
 #   Column           Non-Null Count  Dtype  
 ---  --  
 0   res_id          211944 non-null   int64  
 1   name            211944 non-null   object 
 2   establishment   211944 non-null   object 
 3   url             211944 non-null   object 
 4   address         211810 non-null   object 
 5   city            211944 non-null   object 
 6   city_id         211944 non-null   int64  
 7   locality        211944 non-null   object 
 8   latitude        211944 non-null   float64 
 9   longitude       211944 non-null   float64 
 10  zipcode         48757 non-null   object 
 11  country_id     211944 non-null   int64
```

```

12 locality_verbose      211944 non-null object
13 cuisines              210553 non-null object
14 timings               208070 non-null object
15 average_cost_for_two 211944 non-null int64
16 price_range           211944 non-null int64
17 currency              211944 non-null object
18 highlights             211944 non-null object
19 aggregate_rating       211944 non-null float64
20 rating_text            211944 non-null object
21 votes                 211944 non-null int64
22 photo_count            211944 non-null int64
23 opentable_support      211896 non-null float64
24 delivery               211944 non-null int64
25 takeaway               211944 non-null int64
dtypes: float64(4), int64(9), object(13)
memory usage: 42.0+ MB

```

```
df.describe()
```

	res_id	city_id	latitude	longitude
country_id \ count	2.119440e+05	211944.000000	211944.000000	211944.000000
211944.0 mean	1.349411e+07	4746.785434	21.499758	77.615276
1.0 std	7.883722e+06	5568.766386	22.781331	7.500104
0.0 min	5.000000e+01	1.000000	0.000000	0.000000
1.0 25%	3.301027e+06	11.000000	15.496071	74.877961
1.0 50%	1.869573e+07	34.000000	22.514494	77.425971
1.0 75%	1.881297e+07	11306.000000	26.841667	80.219323
1.0 max	1.915979e+07	11354.000000	10000.000000	91.832769
	average_cost_for_two	price_range	aggregate_rating	-
votes \ count	211944.000000	211944.000000	211944.000000	
211944.000000 mean	595.812229	1.882535	3.395937	
378.001864 std	606.239363	0.892989	1.283642	
925.333370 min	0.000000	1.000000	0.000000	
18.000000				-

25%	250.000000	1.000000	3.300000
16.000000			
50%	400.000000	2.000000	3.800000
100.000000			
75%	700.000000	2.000000	4.100000
362.000000			
max	30000.000000	4.000000	4.900000
42539.000000			

	photo_count	opentable_support	delivery	takeaway
count	211944.000000	211896.0	211944.000000	211944.0
mean	256.971224	0.0	-0.255907	-1.0
std	867.668940	0.0	0.964172	0.0
min	0.000000	0.0	-1.000000	-1.0
25%	3.000000	0.0	-1.000000	-1.0
50%	18.000000	0.0	-1.000000	-1.0
75%	128.000000	0.0	1.000000	-1.0
max	17702.000000	0.0	1.000000	-1.0

Removing duplicates

Identify Duplicates

```
df.duplicated()      #Returns a boolean series (True/False) for
each row
```

0	False
1	False
2	False
3	False
4	False
...	
211939	True
211940	False
211941	True
211942	False
211943	True

Length: 211944, dtype: bool

```
df.duplicated().sum()    #Gives you the total count of duplicate row.
```

151527

```
df[df.duplicated()]    #View the actual duplicate rows
```

	res_id	name	establishment	\
101	3400059	Peshawri - ITC Mughal	['Fine Dining']	
116	3400060	Taj Bano - ITC Mughal	['Fine Dining']	
140	3400017	Pinch Of Spice	['Casual Dining']	
141	3400018	Pinch Of Spice	['Casual Dining']	
142	3400850	Urban Deck	['Casual Dining']	
...	
211937	18855810	Biryani aur Baatein	['Casual Dining']	
211938	18662583	Wok On Fire	['Casual Dining']	
211939	3202251	Kali Mirch Cafe And Restaurant	['Casual Dining']	
211941	18984164	The Grand Thakar	['Casual Dining']	
211943	18879846	Freshco's - The Health Cafe		['Café']
			url	\
101		https://www.zomato.com/agra/peshawri-itc-mugha...		
116		https://www.zomato.com/agra/taj-bano-itc-mugha...		
140		https://www.zomato.com/agra/pinch-of-spice-civ...		
141		https://www.zomato.com/agra/pinch-of-spice-taj...		
142		https://www.zomato.com/agra/urban-deck-2-civil...		
...		...		
211937		https://www.zomato.com/vadodara/biryani-aur-ba...		
211938		https://www.zomato.com/vadodara/wok-on-fire-fa...		
211939		https://www.zomato.com/vadodara/kali-mirch-caf...		
211941		https://www.zomato.com/vadodara/the-grand-thak...		
211943		https://www.zomato.com/vadodara/freshcos-the-h...		
			address	city
city_id	\			
101		ITC Mughal, Fatehabad Road, Tajganj, Agra		Agra
34				
116		ITC Mughal, Fatehabad Road, Tajganj, Agra		Agra
34				
140		23/453, Opposite Sanjay Cinema, Wazipura Road,...		Agra
34				
141		1076/2, Fatehabad Road, Tajganj, Agra		Agra
34				
142		5th Floor, The P L Palace Hotel, MG Road, Sanj...		Agra
34				
...		...		
...				
211937		Shop 14, Atlantis K-10, A Wing, Genda Circle R...		Vadodara
32				
211938		Ground Floor 1, Rossette Building, Opposite Se...		Vadodara
32				
211939		Manu Smriti Complex, Near Navrachna School, GI...		Vadodara
32				
211941		3rd Floor, Shreem Shalini Mall, Opposite Conqu...		Vadodara
32				
211943		Shop 7, Ground Floor, Opposite Natubhai Circle...		Vadodara
32				

```
          locality  latitude  longitude ... price_range
currency \
101      ITC Mughal, Tajganj  27.161150  78.043993 ...        4
Rs.
116      ITC Mughal, Tajganj  27.161132  78.044022 ...        4
Rs.
140          Civil Lines  27.201735  78.007625 ...        4
Rs.
141          Tajganj    27.159649  78.043304 ...        4
Rs.
142          Civil Lines  27.199573  78.003699 ...        4
Rs.
...
...
211937          Alkapuri  22.317746  73.168043 ...        2
Rs.
211938          Fatehgunj  22.323357  73.187461 ...        3
Rs.
211939          Fatehgunj  22.336931  73.192356 ...        2
Rs.
211941          Alkapuri  22.310563  73.171163 ...        2
Rs.
211943          Vadiwadi  22.309935  73.158768 ...        2
Rs.
```

```
                    highlights
aggregate_rating \
101      ['Lunch', 'Cash', 'Credit Card', 'Dinner', 'De... 4.4
4.4
116      ['Credit Card', 'Lunch', 'Cash', 'Debit Card',... 4.3
4.3
140      ['Lunch', 'Delivery', 'Credit Card', 'Dinner',... 4.6
4.6
141      ['Delivery', 'Dinner', 'Cash', 'Credit Card', ... 4.6
4.6
142      ['Dinner', 'Cash', 'Debit Card', 'Takeaway Ava... 4.3
4.3
...
...
211937  ['Dinner', 'Cash', 'Takeaway Available', 'Debi... 4.1
4.1
211938  ['Dinner', 'Cash', 'Debit Card', 'Lunch', 'Tak... 4.0
4.0
211939  ['Dinner', 'Cash', 'Lunch', 'Delivery', 'Indoo... 4.1
4.1
211941  ['Dinner', 'Cash', 'Debit Card', 'Lunch', 'Tak... 4.0
4.0
211943  ['Dinner', 'Cash', 'Takeaway Available', 'Debi... 4.0
4.0
```

	rating_text	votes	photo_count	opentable_support	delivery
takeaway					
101	Very Good	353	154	0.0	-1
-1					
116	Very Good	96	205	0.0	-1
-1					
140	Excellent	915	105	0.0	1
-1					
141	Excellent	965	690	0.0	1
-1					
142	Very Good	672	192	0.0	1
-1					
...
...					
211937	Very Good	154	96	0.0	-1
-1					
211938	Very Good	301	126	0.0	1
-1					
211939	Very Good	243	40	0.0	-1
-1					
211941	Very Good	111	38	0.0	-1
-1					
211943	Very Good	93	53	0.0	1
-1					

[151527 rows x 26 columns]

df.drop_duplicates()

	res_id		name	establishment
\				
0	3400299		Bikanervala	['Quick Bites']
1	3400005	Mama Chicken Mama Franky House		['Quick Bites']
2	3401013		Bhagat Halwai	['Quick Bites']
3	3400290		Bhagat Halwai	['Quick Bites']
4	3401744	The Salt Cafe Kitchen & Bar		['Casual Dining']
...
211882	19142822		Shree Janta Ice Cream	['Dessert Parlour']
211925	18984164		The Grand Thakar	['Casual Dining']
211926	18019952		Geeta lodge	['Casual Dining']

211940	3200996	Raju Omlet	['Quick Bites']
211942	3201138	Subway	['Quick Bites']

		url \
0		https://www.zomato.com/agra/bikanervala-khanda...
1		https://www.zomato.com/agra/mama-chicken-mama-...
2		https://www.zomato.com/agra/bhagat-halwai-2-sh...
3		https://www.zomato.com/agra/bhagat-halwai-civi...
4		https://www.zomato.com/agra/the-salt-cafe-kitc...
...		...
211882		https://www.zomato.com/vadodara/shree-janta-ic...
211925		https://www.zomato.com/vadodara/the-grand-thak...
211926		https://www.zomato.com/vadodara/geeta-lodge-al...
211940		https://www.zomato.com/vadodara/raju-omlet-kar...
211942		https://www.zomato.com/vadodara/subway-1-akota...

		address	city
city_id \			
0	Kalyani Point, Near Tulsi Cinema, Bypass Road,...		Agra
34	Main Market, Sadar Bazaar, Agra Cantt, Agra		Agra
1	62/1, Near Easy Day, West Shivaji Nagar, Goalp...		Agra
34	Near Anjana Cinema, Nehru Nagar, Civil Lines, ...		Agra
3	1C,3rd Floor, Fatehabad Road, Tajganj, Agra		Agra
34
4
32	Ground Floor, 5 Ronak Plaza, Tulsidham Char Ra...	Vadodara	
211882			
32	3rd Floor, Shreem Shalini Mall, Opposite Conqu...	Vadodara	
211925			
32	Shop 11, Ground Floor, Atlantis K-10, Tower A,...	Vadodara	
211926			
32	Mahalaxmi Apartment, Opposite B 0 B, Karoli Ba...	Vadodara	
211940			
32	G-2, Vedant Platina, Near Cosmos, Akota, Vadodara	Vadodara	
211942			
32			

	locality	latitude	longitude	...	price_range
currency \					
0	Khandari	27.211450	78.002381	...	2 Rs.
1	Agra Cantt	27.160569	78.011583	...	2 Rs.

2	Shahganj	27.182938	77.979684	...	1	Rs.
3	Civil Lines	27.205668	78.004799	...	1	Rs.
4	Tajganj	27.157709	78.052421	...	3	Rs.
...
211882	Manjalpur	22.270516	73.196408	...	1	Rs.
211925	Alkapuri	22.310563	73.171163	...	2	Rs.
211926	Alkapuri	22.317731	73.168107	...	1	Rs.
211940	Karelibaug	22.322455	73.197203	...	1	Rs.
211942	Akota	22.270027	73.143068	...	2	Rs.

```
                                highlights
aggregate_rating \
0      ['Lunch', 'Takeaway Available', 'Credit Card',...
4.4
1      ['Delivery', 'No Alcohol Available', 'Dinner',...
4.4
2      ['No Alcohol Available', 'Dinner', 'Takeaway A...
4.2
3      ['Takeaway Available', 'Credit Card', 'Lunch',...
4.3
4      ['Lunch', 'Serves Alcohol', 'Cash', 'Credit Ca...
4.9
...
...
211882 ['Cash', 'Takeaway Available', 'Delivery', 'In...
2.9
211925 ['Dinner', 'Cash', 'Debit Card', 'Lunch', 'Tak...
4.0
211926 ['Dinner', 'Cash', 'Credit Card', 'Lunch', 'Ta...
3.9
211940 ['Dinner', 'Cash', 'Takeaway Available', 'Debi...
4.1
211942 ['Dinner', 'Delivery', 'Credit Card', 'Lunch',...
3.7

           rating_text   votes  photo_count opentable_support delivery
takeaway
0      Very Good     814          154        0.0       -1
-1
1      Very Good    1203          161        0.0       -1
-1
```

```

2      Very Good    801        107        0.0       1
-1
3      Very Good    693        157        0.0       1
-1
4      Excellent    470        291        0.0       1
-1
...
...
211882    Average     4          1        0.0       1
-1
211925    Very Good   111        38        0.0      -1
-1
211926      Good     207        14        0.0      -1
-1
211940    Very Good   187        40        0.0       1
-1
211942      Good     128        34        0.0       1
-1

```

[60417 rows x 26 columns]

```
df_cleaned = df.drop_duplicates()      #Remove rows where all columns  
are identical
```

Verification : After dropping the rows, always check the shape of your dataframe to ensure the operation worked as expected.

```
print(f"Original shape:{df.shape}")

Original shape:(211944, 26)

print(f"New shape:{df_cleaned.shape}")

New shape:(60417, 26)
```

Dealing with missing values

```
print(df.isnull().sum())      #Check total missing values per column

res_id                  0
name                    0
establishment           0
url                     0
address                 134
city                    0
city_id                 0
locality                0
latitude                0
longitude               0
```

```
zipcode           163187
country_id        0
locality_verbose  0
cuisines          1391
timings           3874
average_cost_for_two  0
price_range       0
currency          0
highlights         0
aggregate_rating  0
rating_text       0
votes              0
photo_count       0
opentable_support 48
delivery           0
takeaway           0
dtype: int64
```

```
missing_pct = (df.isnull().sum() / len(df)) * 100      #Calculate the percentage of missing values
```

```
print(missing_pct)
```

```
res_id            0.000000
name              0.000000
establishment    0.000000
url               0.000000
address           0.063224
city              0.000000
city_id           0.000000
locality          0.000000
latitude          0.000000
longitude         0.000000
zipcode           76.995338
country_id        0.000000
locality_verbose  0.000000
cuisines          0.656305
timings           1.827841
average_cost_for_two  0.000000
price_range       0.000000
currency          0.000000
highlights         0.000000
aggregate_rating  0.000000
rating_text       0.000000
votes              0.000000
photo_count       0.000000
opentable_support 0.022647
delivery           0.000000
```

```
takeaway          0.000000
dtype: float64

df.drop(columns=['zipcode'], inplace=True)      #Drop Zipcode because
it has over 160k missing values

df['cuisines'] = df['cuisines'].fillna('Unknown')      #Fill
cuisines with 'Unknown'

df['timings'] = df['timings'].fillna(df['timings'].mode()[0])  #Fill
#Fill timings with the most common timing (Mode)

df.dropna(subset=['address', 'opentable_support'], inplace=True)  #Drop rows where address or opentable_support is missing

df.isnull().sum()

res_id           0
name             0
establishment   0
url              0
address          0
city              0
city_id          0
locality         0
latitude         0
longitude        0
country_id       0
locality_verbose 0
cuisines          0
timings          0
average_cost_for_two 0
price_range       0
currency          0
highlights         0
aggregate_rating  0
rating_text        0
votes              0
photo_count        0
opentable_support 0
delivery            0
takeaway           0
dtype: int64
```

Create Plots or visualizations

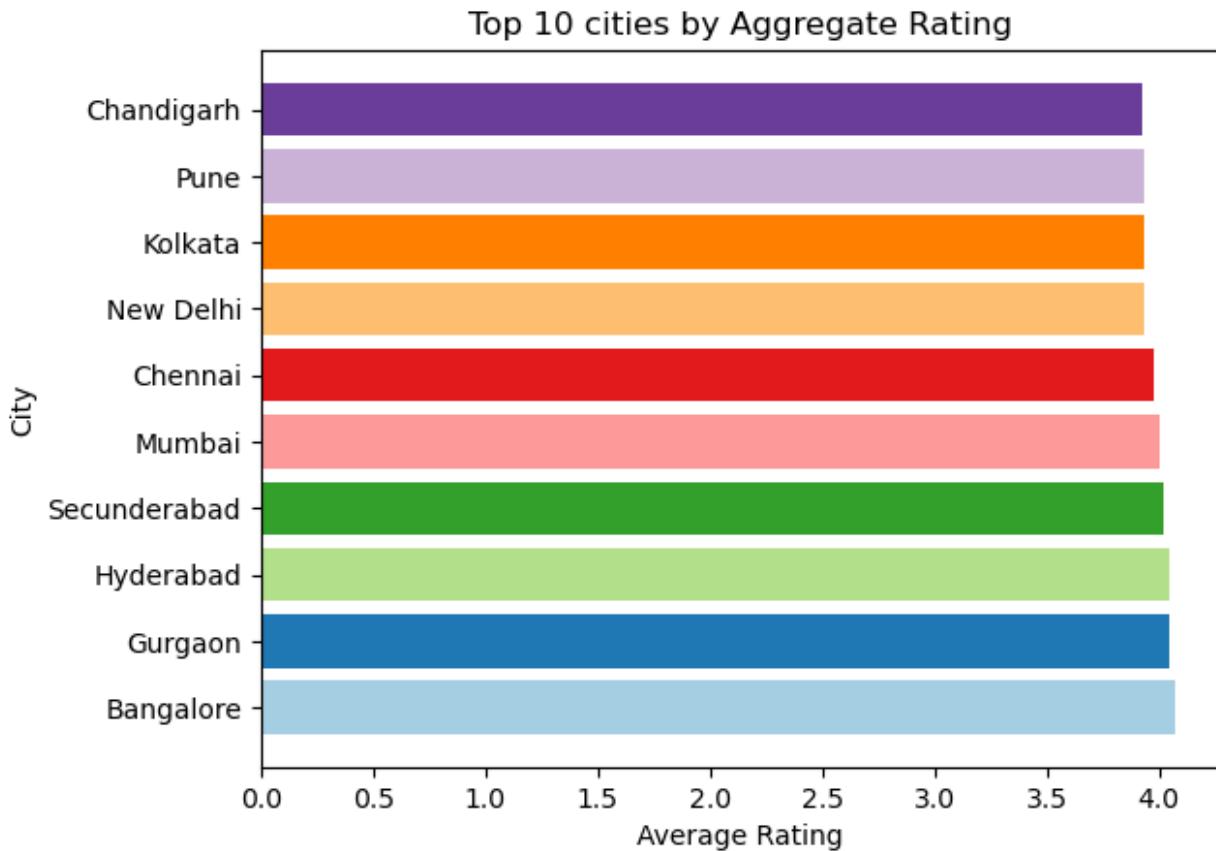
Plotting horizontal bar charts (multicolor)

Group by city, calculate mean rating, and take the top 10

```
top_cities = df.groupby('city')[['aggregate_rating']].mean().sort_values(ascending=False).head(10)
```

Plotting

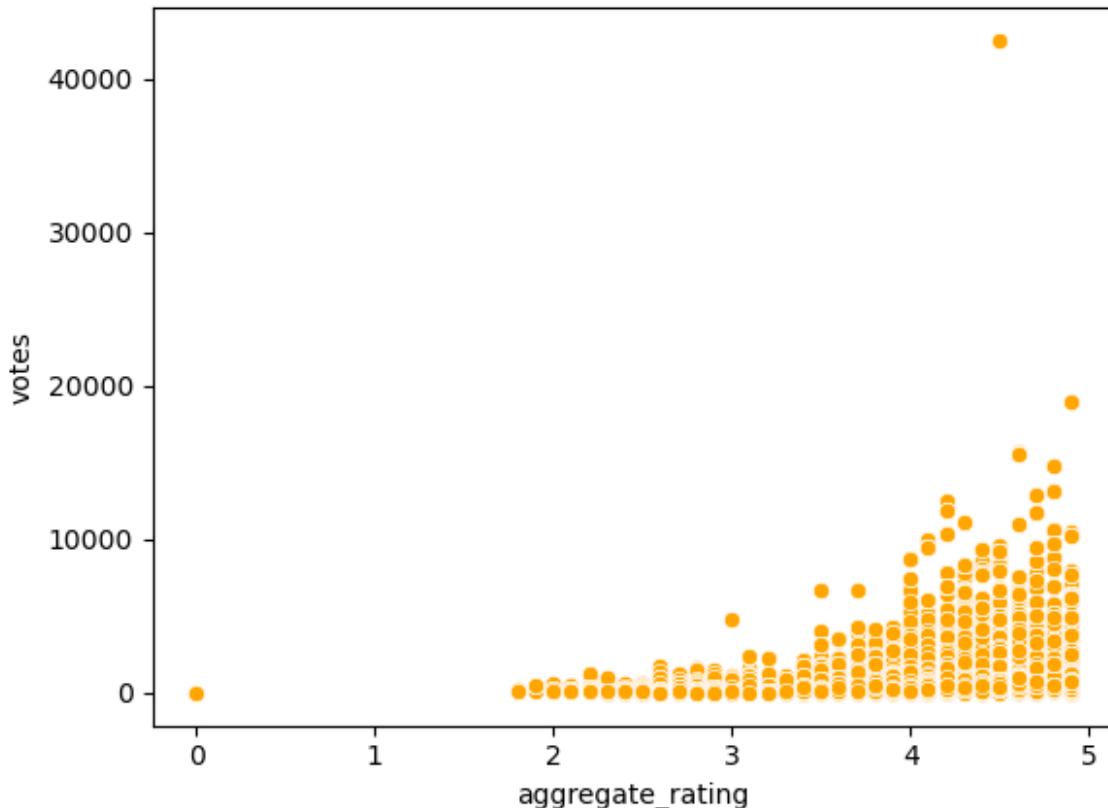
```
plt.barh(y=top_cities.index, width=top_cities.values,
          color=plt.cm.Paired(np.arange(len(top_cities))))
plt.xlabel('Average Rating')
plt.ylabel('City')
plt.title('Top 10 cities by Aggregate Rating')
plt.show()
```



Scatter plot

```
sns.scatterplot(x=df['aggregate_rating'], y=df['votes'],  
color='orange')
```

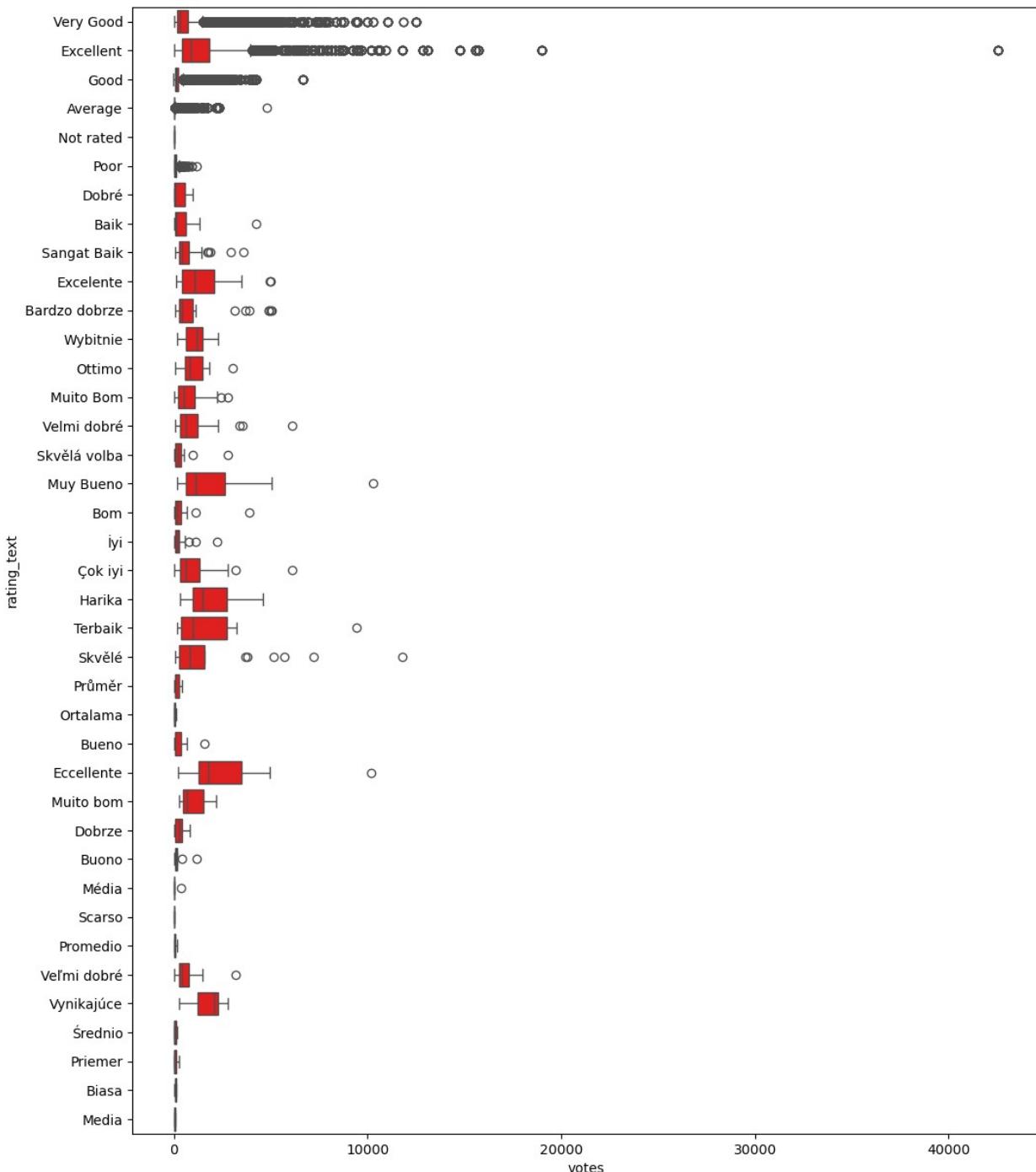
```
<Axes: xlabel='aggregate_rating', ylabel='votes'>
```



Word Cloud

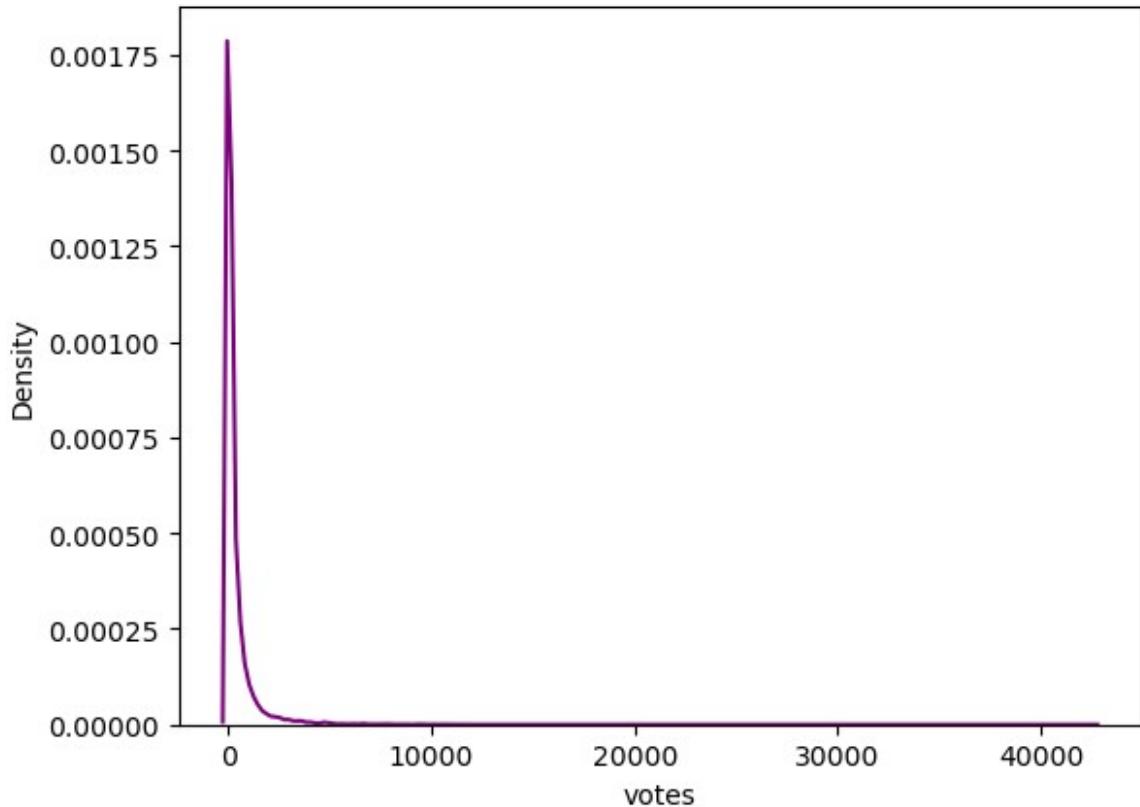
Box plot

```
plt.figure(figsize=(12,15))  
sns.boxplot(x=df['votes'], y=df['rating_text'], color='red')  
plt.show()
```



Density plot

```
sns.kdeplot(x=df['votes'], color='purple')  
<Axes: xlabel='votes', ylabel='Density'>
```



Conclusion

Improve restaurant recommendations

Optimize pricing strategy

Focus marketing on high-demand cuisines

Identify high-performing locations