Zomato Data Analysis by Python

```
Project By Ayush Kumar
```

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

Start coding or generate with AI.

dataframe = pd.read_csv("/content/Zomato data .csv")
```

dataframe.iloc[10:101, 0:4]

₹		name	online_order	book_table	rate	
	10	Village Café	Yes	No	4.1/5	11.
	11	Cafe Shuffle	Yes	Yes	4.2/5	
	12	The Coffee Shack	Yes	Yes	4.2/5	
	13	Caf-Eleven	No	No	4.0/5	
	14	San Churro Cafe	Yes	No	3.8/5	
	96	Kaggis	No	No	3.8/5	
	97	Ayda Persian Kitchen	No	No	3.7/5	
	98	Chatar Patar	No	No	3.7/5	
	99	Polar Bear	Yes	No	3.8/5	
	100	Kidambi's Kitchen	No	No	3.5/5	

91 rows × 4 columns

Start coding or generate with AI.

Create the dataframe

```
dataframe = pd.read_csv("Zomato data .csv")
```

print(dataframe)

_		name	online_order	book_table	rate	votes	١
_	0	Jalsa	Yes	Yes	4.1/5	775	
	1	Spice Elephant	Yes	No	4.1/5	787	
	2	San Churro Cafe	Yes	No	3.8/5	918	
	3	Addhuri Udupi Bhojana	No	No	3.7/5	88	
	4	Grand Village	No	No	3.8/5	166	
	143	Melting Melodies	No	No	3.3/5	0	
	144	New Indraprasta	No	No	3.3/5	0	
	145	Anna Kuteera	Yes	No	4.0/5	771	
	146	Darbar	No	No	3.0/5	98	
	147	Vijayalakshmi	Yes	No	3.9/5	47	

	approx_cost(for	two	people)	<pre>listed_in(type)</pre>
0			800	Buffet
1			800	Buffet
2			800	Buffet
3			300	Buffet
4			600	Buffet
143			100	Dining
144			150	Dining
145			450	Dining
146			800	Dining
147			200	Dining

П

ılı

[148 rows x 7 columns] dataframe ₹ name online_order book_table rate votes approx_cost(for two people) listed_in(type) 0 Jalsa Yes Yes 4.1/5 775 800 Buffet Spice Elephant 1 787 800 Buffet Yes No 4.1/5 2 San Churro Cafe Yes No 3.8/5 918 800 Buffet 3 Addhuri Udupi Bhojana 300 Buffet No 3.7/5 88 No Buffet 4 Grand Village No No 3.8/5 166 600 ... 143 3.3/5 0 100 Melting Melodies Nο Dining No 144 New Indraprasta No 3.3/5 0 150 Dining No 145 Anna Kuteera 4.0/5 771 Yes No 450 Dining 146 Darbar No No 3.0/5 98 800 Dining 147 Vijayalakshmi Yes No 3.9/5 47 200 Dining 148 rows × 7 columns Generate code with dataframe View recommended plots New interactive sheet Start coding or generate with AI. convert the data type of column-rate def handleRate(value): value = str(value).split('/') value = value[0]; return float(value) dataframe['rate']=dataframe['rate'].apply(handleRate) print(dataframe) _ name online_order book_table rate votes 0 Jalsa 4.1 775 Yes Yes Spice Elephant 1 Yes No 4.1 787 2 San Churro Cafe Yes No 3.8 918 3 Addhuri Udupi Bhojana 3.7 88 No No Grand Village 4 3.8 No 166 No 143 Melting Melodies 3.3 No No 144 New Indraprasta 0 No No 3.3 145 Anna Kuteera Yes No 4.0 771 146 Darbar No 3.0 98 147 Vijayalakshmi Yes No 3.9 47 approx_cost(for two people) listed_in(type) 0 800 Buffet 1 2 800 Buffet 3 300 Buffet 4 600 Buffet 143 100 Dining 144 150 Dining 145 450 Dining 146 800 Dining 147 Dining [148 rows x 7 columns]

This is formatted as code

dataframe.head() __ name online_order book_table rate votes approx_cost(for two people) listed_in(type) 0 Jalsa Yes 4.1 775 800 Buffet ıl. Spice Elephant Yes No 4.1 787 800 Buffet 2 San Churro Cafe No 3.8 918 800 Buffet Buffet 3 Addhuri Udupi Bhojana No No 3.7 88 300 Buffet 4 Grand Village No 3.8 166 600 Next steps: Generate code with dataframe View recommended plots New interactive sheet dataframe.info() <class 'pandas.core.frame.DataFrame'> RangeIndex: 148 entries, 0 to 147

object

object

Data columns (total 7 columns): # Column Non-Null Count Dtype -----0 name 148 non-null online_order 148 non-null 2 148 non-null book_table

object 3 rate 148 non-null float64 148 non-null int64 approx_cost(for two people) 148 non-null int64 listed_in(type) 148 non-null object

dtypes: float64(1), int64(2), object(4)

memory usage: 8.2+ KB

Start coding or generate with AI.

Type of Resturant

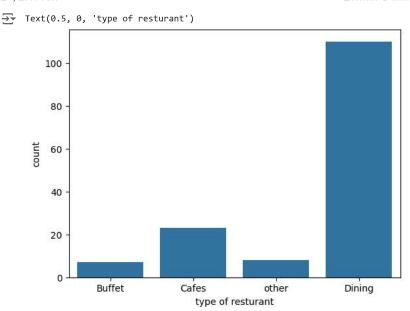
dataframe.head()

₹		name	online_order	book_table	rate	votes	approx_cost(for two people)	listed_in(type)	
	0	Jalsa	Yes	Yes	4.1	775	800	Buffet	ıl.
	1	Spice Elephant	Yes	No	4.1	787	800	Buffet	
	2	San Churro Cafe	Yes	No	3.8	918	800	Buffet	
	3	Addhuri Udupi Bhojana	No	No	3.7	88	300	Buffet	
	4	Grand Village	No	No	3.8	166	600	Buffet	

Next steps: Generate code with dataframe View recommended plots New interactive sheet

sns.countplot(x=dataframe['listed_in(type)']) plt.xlabel("type of resturant")

New interactive sheet



Start coding or generate with AI.

Conclusion: Majority of the resturant falls in dinning Category.

dataframe.head()

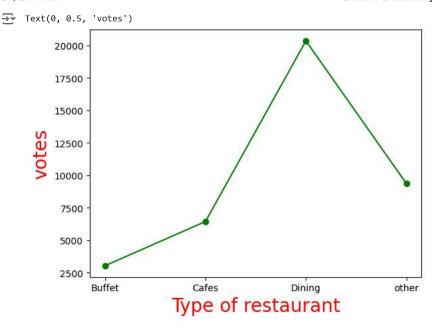
Next steps:

		name	online_order	book_table	rate	votes	approx_cost(for two people)	listed_in(type)	
	0	Jalsa	Yes	Yes	4.1	775	800	Buffet	ıl.
	1	Spice Elephant	Yes	No	4.1	787	800	Buffet	
	2	San Churro Cafe	Yes	No	3.8	918	800	Buffet	
	3	Addhuri Udupi Bhojana	No	No	3.7	88	300	Buffet	
	4	Grand Village	No	No	3.8	166	600	Buffet	

View recommended plots

grouped_data = dataframe.groupby('listed_in(type)')['votes'].sum()
result = pd.DataFrame({'votes': grouped_data})
plt.plot(result, c="green", marker="o")
plt.xlabel("Type of restaurant", c="red", size=20)
plt.ylabel("votes", c="red", size=20)

Generate code with dataframe



Start coding or $\underline{\text{generate}}$ with AI.

Conclusion: Dinning restaurants has recieved maximum votes.

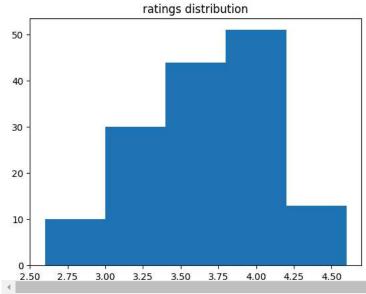
dataframe.head()

$\overline{\Rightarrow}$		name	online_order	book_table	rate	votes	approx_cost(for two people)	listed_in(type)	
	0	Jalsa	Yes	Yes	4.1	775	800	Buffet	ılı
	1	Spice Elephant	Yes	No	4.1	787	800	Buffet	
	2	San Churro Cafe	Yes	No	3.8	918	800	Buffet	
	3	Addhuri Udupi Bhojana	No	No	3.7	88	300	Buffet	
	4	Grand Village	No	No	3.8	166	600	Buffet	

Next steps: Generate code with dataframe View recommended plots New interactive sheet

plt.hist(dataframe['rate'],bins = 5)
plt.title("ratings distribution")

 \Rightarrow Text(0.5, 1.0, 'ratings distribution')



Start coding or generate with AI.

Conclusion: The majority restaurant recieved ratings from 3.6 to 6.

dataframe.head()

	name	online_order	book_table	rate	votes	approx_cost(for two people) listed_in(type)	
0	Jalsa	Yes	Yes	4.1	775	80	D Buffet	11.
1	Spice Elephant	Yes	No	4.1	787	80	D Buffet	
2	San Churro Cafe	Yes	No	3.8	918	80	D Buffet	
3	Addhuri Udupi Bhojana	No	No	3.7	88	30	D Buffet	
4	Grand Village	No	No	3.8	166	60	D Buffet	
	1 2 3	0 Jalsa1 Spice Elephant2 San Churro Cafe3 Addhuri Udupi Bhojana	0 Jalsa Yes 1 Spice Elephant Yes 2 San Churro Cafe Yes 3 Addhuri Udupi Bhojana No	0 Jalsa Yes Yes 1 Spice Elephant Yes No 2 San Churro Cafe Yes No 3 Addhuri Udupi Bhojana No No	0JalsaYesYes4.11Spice ElephantYesNo4.12San Churro CafeYesNo3.83Addhuri Udupi BhojanaNoNoNo	0 Jalsa Yes Yes 4.1 775 1 Spice Elephant Yes No 4.1 787 2 San Churro Cafe Yes No 3.8 918 3 Addhuri Udupi Bhojana No No 3.7 88	0 Jalsa Yes Yes 4.1 775 800 1 Spice Elephant Yes No 4.1 787 800 2 San Churro Cafe Yes No 3.8 918 800 3 Addhuri Udupi Bhojana No No 3.7 88 300	1 Spice Elephant Yes No 4.1 787 800 Buffet 2 San Churro Cafe Yes No 3.8 918 800 Buffet 3 Addhuri Udupi Bhojana No No 3.7 88 300 Buffet

Next steps:

Generate code with dataframe

View recommended plots

New interactive sheet

Start coding or generate with AI.

Average Order Spending by couples

dataframe.head()

_		name	online_order	book_table	rate	votes	approx_cost(for two people)) listed_in(type)	
	0	Jalsa	Yes	Yes	4.1	775	800	D Buffet	11.
	1	Spice Elephant	Yes	No	4.1	787	800	D Buffet	
	2	San Churro Cafe	Yes	No	3.8	918	800	D Buffet	
	3	Addhuri Udupi Bhojana	No	No	3.7	88	300	D Buffet	
	4	Grand Village	No	No	3.8	166	600	D Buffet	

Next steps:

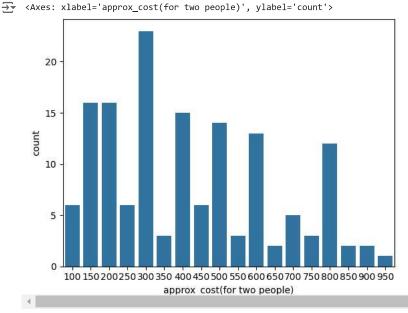
Generate code with dataframe

View recommended plots

New interactive sheet

couple_data=dataframe['approx_cost(for two people)']
sns.countplot(x=couple_data)

313. counteplot(x-couple_uata)



Start coding or generate with AI.

Conclusion: The majority of couples prefer restaurants with an approximate cost of 300 rupees.

8/29/24, 2:17 AM Zomato Data Analysis - Colab Start coding or $\underline{\text{generate}}$ with AI. Which mode recieves maximim rating dataframe.head() _ name online_order book_table rate votes approx_cost(for two people) listed_in(type) \blacksquare 0 Jalsa 4.1 775 800 Buffet Yes Yes ıl. 787 800 Buffet 1 Spice Elephant Yes No 4.1 2 San Churro Cafe Yes No 3.8 918 800 Buffet 3 Addhuri Udupi Bhojana No No 3.7 88 300 Buffet Grand Village No 3.8 600 Buffet Generate code with dataframe View recommended plots New interactive sheet Next steps: plt.figure(figsize = (6,6)) sns.boxplot(x='online_order', y = 'rate', data = dataframe) → <Axes: xlabel='online_order', ylabel='rate'> 4.50 4.25 4.00 3.75 3.50 3.25 0 3.00 0 0 2.75 0 2.50 Yes No online order Start coding or generate with AI. Conclusion: Offline order recieved lower rating in comparision to online order. dataframe.head() **₹** online_order book_table rate votes approx_cost(for two people) listed_in(type) name 0 Jalsa 775 800 Buffet Yes 4.1 Yes Buffet Spice Elephant 800 1 4.1 787 Yes No 2 Buffet San Churro Cafe 3.8 918 800 Yes No Addhuri Udupi Bhojana 88 300 Buffet No No 3.7

https://colab.res	search.google.co	m/drive/19FFjvXar0\	NAYFaHYrhwVas	s-14IFUgiQi#scrol	ITo=Ki6ezIRasJZ58	&printMode=true

No

3.8

View recommended plots

166

600

New interactive sheet

Buffet

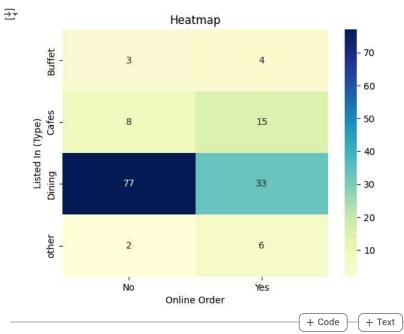
No

Grand Village

Next steps:

Generate code with dataframe

```
pivot_table = dataframe.pivot_table(index='listed_in(type)', columns='online_order', aggfunc='size', fill_value=0)
sns.heatmap(pivot_table, annot=True, cmap="YlGnBu", fmt='d')
plt.title("Heatmap")
plt.xlabel("Online Order")
plt.ylabel("Listed In (Type)")
plt.show()
```



Conclusion: Dining restaurants primarily accept offline orders, whereas cafes primarily receive online orders. This suggests that clients prefer to place orders in person at restaurants, but prefer online ordering at cafes.

Start coding or generate with AI.