

Level 2

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

df = pd.read_csv("Dataset .csv")
df
```

	Restaurant ID	Restaurant Name	Country Code
City \			
0	6317637	Le Petit Souffle	162
Makati City			
1	6304287	Izakaya Kikufuji	162
Makati City			
2	6300002	Heat - Edsa Shangri-La	162
Mandaluyong City			
3	6318506	Ooma	162
Mandaluyong City			
4	6314302	Sambo Kojin	162
Mandaluyong City			
...
...			
9546	5915730	Naml \ Gurme	208
00istanbul			
9547	5908749	Ceviz A00ac \	208
00istanbul			
9548	5915807	Huqqa	208
00istanbul			
9549	5916112	A000k Kahve	208
00istanbul			
9550	5927402	Walter's Coffee Roastery	208
00istanbul			

	Address \
0	Third Floor, Century City Mall, Kalayaan Avenu...
1	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...
2	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...
3	Third Floor, Mega Fashion Hall, SM Megamall, O...
4	Third Floor, Mega Atrium, SM Megamall, Ortigas...
...	...
9546	Kemanke00 Karamustafa Pa00a Mahallesi, R\ht\m ...
9547	Ko00uyolu Mahallesi, Muhittin 00st0_nda00 Cadd...
9548	Kuru0_e00me Mahallesi, Muallim Naci Caddesi, N...
9549	Kuru0_e00me Mahallesi, Muallim Naci Caddesi, N...
9550	Cafea00a Mahallesi, Bademalt\ Sokak, No 21/B, ...

Locality \

0	Century City Mall, Poblacion, Makati City
1	Little Tokyo, Legaspi Village, Makati City
2	Edsa Shangri-La, Ortigas, Mandaluyong City
3	SM Megamall, Ortigas, Mandaluyong City
4	SM Megamall, Ortigas, Mandaluyong City

...	...
9546	Karaköy
9547	Koşuyolu
9548	Kuruçeşme
9549	Kuruçeşme
9550	Moda

	Locality Verbose	Longitude \
0	Century City Mall, Poblacion, Makati City, Mak...	121.027535
1	Little Tokyo, Legaspi Village, Makati City, Ma...	121.014101
2	Edsa Shangri-La, Ortigas, Mandaluyong City, Ma...	121.056831
3	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.056475
4	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.057508
...
9546	Karaköy, İstanbul	28.977392
9547	Koşuyolu, İstanbul	29.041297
9548	Kuruçeşme, İstanbul	29.034640
9549	Kuruçeşme, İstanbul	29.036019
9550	Moda, İstanbul	29.026016

	Latitude	Cuisines ...	
Currency \			
0	14.565443	French, Japanese, Desserts ...	Botswana
Pula(P)			
1	14.553708	Japanese ...	Botswana
Pula(P)			
2	14.581404	Seafood, Asian, Filipino, Indian ...	Botswana
Pula(P)			
3	14.585318	Japanese, Sushi ...	Botswana
Pula(P)			
4	14.584450	Japanese, Korean ...	Botswana
Pula(P)			
...
..			
9546	41.022793	Turkish ...	Turkish
Lira(TL)			
9547	41.009847	World Cuisine, Patisserie, Cafe ...	Turkish
Lira(TL)			
9548	41.055817	Italian, World Cuisine ...	Turkish
Lira(TL)			
9549	41.057979	Restaurant Cafe ...	Turkish
Lira(TL)			
9550	40.984776	Cafe ...	Turkish
Lira(TL)			

	Has Table booking	Has Online delivery	Is delivering now	\
0	Yes	No	No	
1	Yes	No	No	
2	Yes	No	No	
3	No	No	No	
4	Yes	No	No	
...	
9546	No	No	No	
9547	No	No	No	
9548	No	No	No	
9549	No	No	No	
9550	No	No	No	

	Switch to order menu	Price range	Aggregate rating	Rating color
\				
0	No	3	4.8	Dark Green
1	No	3	4.5	Dark Green
2	No	4	4.4	Green
3	No	4	4.9	Dark Green
4	No	4	4.8	Dark Green
...
9546	No	3	4.1	Green
9547	No	3	4.2	Green
9548	No	4	3.7	Yellow
9549	No	4	4.0	Green
9550	No	2	4.0	Green

	Rating text	Votes
0	Excellent	314
1	Excellent	591
2	Very Good	270
3	Excellent	365
4	Excellent	229
...
9546	Very Good	788
9547	Very Good	1034
9548	Good	661
9549	Very Good	901
9550	Very Good	591

```
[9551 rows x 21 columns]
```

```
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
```

```
# Display the number of rows and columns
```

```
df.shape
```

```
(9551, 21)
```

```
#Check for missing values in each column
```

```
missing_values = df.isnull().sum()
```

```
missing_values
```

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

#Drop rows with missing values
df.dropna(inplace=True)

#After handling missing values
df.shape

(9542, 21)

```

Task 1: Table Booking and Online Delivery

```

# Determine the percentage of restaurants that offer table booking
table_booking_percentage = (df['Has Table
booking'].value_counts(normalize=True) * 100).round(2)

table_booking_percentage

No      87.86
Yes     12.14
Name: Has Table booking, dtype: float64

# Determine the percentage of restaurants that offer online delivery
online_delivery_percentage = (df['Has Online
delivery'].value_counts(normalize=True) * 100).round(2)
online_delivery_percentage

No      74.31
Yes     25.69
Name: Has Online delivery, dtype: float64

# Convert 'Aggregate rating' column to numeric (if not already)
df['Aggregate rating'] = pd.to_numeric(df['Aggregate rating'],
errors='coerce')

# Group the data based on 'Has Table booking' and calculate the
average ratings

```

```

average_ratings = df.groupby('Has Table booking')['Aggregate
rating'].mean()

average_ratings

Has Table booking
No      2.557956
Yes     3.441969
Name: Aggregate rating, dtype: float64

# Calculate the total number of restaurants for each price range
total_restaurants_per_price_range = df['Price range'].value_counts()

# Calculate the number of restaurants offering online delivery for
each price range
online_delivery_per_price_range = df[df['Has Online delivery'] ==
'Yes']['Price range'].value_counts()

# Calculate the percentage of restaurants offering online delivery for
each price range
online_delivery_percentage_per_price_range =
(online_delivery_per_price_range / total_restaurants_per_price_range)
* 100

online_delivery_percentage_per_price_range

1      15.795403
2      41.310633
3      29.252669
4       9.044369
Name: Price range, dtype: float64

```

Task 2: Price Range Analysis

```

# Determine the most common price range among all the restaurants
most_common_price_range = df['Price range'].value_counts().idxmax()
most_common_price_range

1

# Convert 'Aggregate rating' column to numeric (if not already)
df['Aggregate rating'] = pd.to_numeric(df['Aggregate rating'],
errors='coerce')

# Calculate the average rating for each price range
average_rating_per_price_range = df.groupby('Price range')['Aggregate
rating'].mean().round(2)

average_rating_per_price_range

```

```

Price range
1      2.00
2      2.94
3      3.68
4      3.82
Name: Aggregate rating, dtype: float64

# Calculate the average rating for each price range
average_rating_per_price_range = df.groupby('Price range')['Aggregate
rating'].mean()

# Find the price range with the highest average rating
highest_avg_rating_price_range =
average_rating_per_price_range.idxmax()

# Assuming that the color is represented by the 'Rating color' column
color_for_highest_avg_rating = df[df['Price range'] ==
highest_avg_rating_price_range]['Rating color'].iloc[0]

color_for_highest_avg_rating
'Green'

```

Task 3: Feature Engineering

```

# Extract additional features: length of restaurant name and address
df['Restaurant Name Length'] = df['Restaurant Name'].str.len()
df['Address Length'] = df['Address'].str.len()

df[['Restaurant Name', 'Restaurant Name Length']]

```

	Restaurant Name	Restaurant Name Length
0	Le Petit Souffle	16
1	Izakaya Kikufuji	16
2	Heat - Edsa Shangri-La	22
3	Ooma	4
4	Sambo Kojin	11
...
9546	Naml \ Gurme	11
9547	Ceviz A00ac \	12
9548	Huqqa	5
9549	A000k Kahve	11
9550	Walter's Coffee Roastery	24

```

[9542 rows x 2 columns]

df[['Address', 'Address Length']]

```

```

                                Address  Address
Length
0      Third Floor, Century City Mall, Kalayaan Avenu...
71
1      Little Tokyo, 2277 Chino Roces Avenue, Legaspi...
67
2      Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...
56
3      Third Floor, Mega Fashion Hall, SM Megamall, O...
70
4      Third Floor, Mega Atrium, SM Megamall, Ortigas...
64
...
.
9546  Kemanke00 Karamustafa Pa00a Mahallesi, R\ht\m ...
103
9547  Ko00uyolu Mahallesi, Muhittin 00st0_nda00 Cadd...
77
9548  Kuru0_e00me Mahallesi, Muallim Naci Caddesi, N...
73
9549  Kuru0_e00me Mahallesi, Muallim Naci Caddesi, N...
75
9550  Cafea00a Mahallesi, Bademalt\ Sokak, No 21/B, ...
65

```

```
[9542 rows x 2 columns]
```

```
# Perform one-hot encoding for 'Has Table booking' and 'Has Online
delivery' columns
```

```
df_encoded = pd.get_dummies(df, columns=['Has Table booking', 'Has
Online delivery'])
```

```
df_encoded.head()
```

```

    Restaurant ID      Restaurant Name  Country Code
City \
0      6317637      Le Petit Souffle          162      Makati
City
1      6304287      Izakaya Kikufuji          162      Makati
City
2      6300002      Heat - Edsa Shangri-La          162      Mandaluyong
City
3      6318506                      Ooma          162      Mandaluyong
City
4      6314302      Sambo Kojin          162      Mandaluyong
City

```

```

                                Address  \
0      Third Floor, Century City Mall, Kalayaan Avenu...
1      Little Tokyo, 2277 Chino Roces Avenue, Legaspi...

```



```

2 Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...
3 Third Floor, Mega Fashion Hall, SM Megamall, O...
4 Third Floor, Mega Atrium, SM Megamall, Ortigas...

```

```

                                Locality \
0 Century City Mall, Poblacion, Makati City
1 Little Tokyo, Legaspi Village, Makati City
2 Edsa Shangri-La, Ortigas, Mandaluyong City
3 SM Megamall, Ortigas, Mandaluyong City
4 SM Megamall, Ortigas, Mandaluyong City

```

```

                                Locality Verbose Longitude
Latitude \
0 Century City Mall, Poblacion, Makati City, Mak... 121.027535
14.565443
1 Little Tokyo, Legaspi Village, Makati City, Ma... 121.014101
14.553708
2 Edsa Shangri-La, Ortigas, Mandaluyong City, Ma... 121.056831
14.581404
3 SM Megamall, Ortigas, Mandaluyong City, Mandal... 121.056475
14.585318
4 SM Megamall, Ortigas, Mandaluyong City, Mandal... 121.057508
14.584450

```

```

                                Cuisines ... Aggregate rating Rating
color \
0 French, Japanese, Desserts ... 4.8 Dark
Green
1 Japanese ... 4.5 Dark
Green
2 Seafood, Asian, Filipino, Indian ... 4.4
Green
3 Japanese, Sushi ... 4.9 Dark
Green
4 Japanese, Korean ... 4.8 Dark
Green

```

```

Rating text Votes Restaurant Name Length Address Length \
0 Excellent 314 16 71
1 Excellent 591 16 67
2 Very Good 270 22 56
3 Excellent 365 4 70
4 Excellent 229 11 64

```

```

Has Table booking_No Has Table booking_Yes Has Online
delivery_No \
0 0 1 1
1 0 1 1

```

2	0	1	1
3	1	0	1
4	0	1	1

Has Online delivery_Yes	
0	0
1	0
2	0
3	0
4	0

[5 rows x 25 columns]