```python
In [5]:  import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import seaborn as sns

         path=r'C:\Users\DELL\Documents\cognifyz Task\Data set\Dataset .csv'
         data_df=pd.read_csv(path)
         data_df
```

Out[5]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Century City Mall, Poblacion, Makati City
1	6304287	Izakaya Kikufuji	162	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...	Little Tokyo, Legaspi Village, Makati City
2	6300002	Heat - Edsa Shangri-La	162	Mandaluyong City	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...	Edsa Shangri-La, Ortigas, Mandaluyong City
3	6318506	Ooma	162	Mandaluyong City	Third Floor, Mega Fashion Hall, SM Megamall, O...	SM Megamall, Ortigas, Mandaluyong City
4	6314302	Sambo Kojin	162	Mandaluyong City	Third Floor, Mega Atrium, SM Megamall, Ortigas...	SM Megamall, Ortigas, Mandaluyong City
...	...	...	...	...	...	...
9546	5915730	Namlı Gurme	208	��stanbul	Kemanke�� Karamustafa Pa��a Mahallesi, Rıhtım ...	Karak�_y
9547	5908749	Ceviz A��acı	208	��stanbul	Ko��uyolu Mahallesi, Muhittin ��st�_nda�� Cadd...	Ko��uyolu
9548	5915807	Huqqa	208	��stanbul	Kuru�_e��me Mahallesi, Muallim Naci Caddesi, N...	Kuru�_e��me
9549	5916112	A���k Kahve	208	��stanbul	Kuru�_e��me Mahallesi, Muallim Naci Caddesi, N...	Kuru�_e��me
9550	5927402	Walter's Coffee Roastery	208	��stanbul	Cafea��a Mahallesi, Bademaltı Sokak, No 21/B, ...	Moda

9551 rows × 21 columns

Task-1: Top Cuisines

Determine the top three most common cuisines in the dataset

```
In [4]:  import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import seaborn as sns

         path=r'C:\Users\DELL\Documents\cognifyz Task\Data set\Dataset .csv'

         data_df=pd.read_csv(path)

         cuisine_counts = data_df['Cuisines'].str.split(',').explode().str.strip().value_

         top_three_cuisines = cuisine_counts.head(3)

         print("Top three most common cuisines:")
         pd.DataFrame(top_three_cuisines)
```

Top three most common cuisines:

Out[4]:

Cuisines	count
North Indian	3960
Chinese	2735
Fast Food	1986

Calculate the percentage of restaurants that serve each of the top cuisines.

```
In [3]:  import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import seaborn as sns

         path=r'C:\Users\DELL\Documents\cognifyz Task\Data set\Dataset .csv'

         data_df=pd.read_csv(path)

         top_cuisines= data_df['Cuisines'].str.split(',').explode().str.strip().value_cou

         percentages = (top_cuisines / len(data_df)) * 100

         print(percentages)
```

```
Cuisines
North Indian    41.461627
Chinese         28.635745
Fast Food       20.793634
Name: count, dtype: float64
```

Task-2: City Analysis

Identify the city with the highest number of restaurants in the dataset

```
In [4]:  import pandas as pd
         import numpy as np

         path=r'C:\Users\DELL\Documents\cognifyz Task\Data set\Dataset .csv'

         data_df=pd.read_csv(path)

         city_restaurant_counts = data_df.groupby('City')['Restaurant Name'].count()


         city_with_max_restaurants = city_restaurant_counts.idxmax()
         max_restaurant_count = city_restaurant_counts.max()

         print(f"The city with the highest number of restaurants is {city_with_max_restau
```

```
The city with the highest number of restaurants is New Delhi with 5473 restaurant
s.
```

Calculate the average rating fo restaurants in each city.

```
In [5]:  import pandas as pd
         import numpy as np

         path=r'C:\Users\DELL\Documents\cognifyz Task\Data set\Dataset .csv'

         data_df=pd.read_csv(path)

         average_ratings_per_city = data_df.groupby('City')['Aggregate rating'].mean().re


         average_ratings_per_city = average_ratings_per_city.sort_values(by='Aggregate ra


         pd.DataFrame(average_ratings_per_city)
```

	City	Aggregate rating
0	Inner City	4.900000
1	Quezon City	4.800000
2	Makati City	4.650000
3	Pasig City	4.633333
4	Mandaluyong City	4.625000
...	...	...
136	New Delhi	2.438845
137	Montville	2.400000
138	Mc Millan	2.400000
139	Noida	2.036204
140	Faridabad	1.866932

141 rows × 2 columns

Determine the city with the highest average rating

In [6]:
```python
import pandas as pd
import numpy as np

path=r'C:\Users\DELL\Documents\cognifyz Task\Data set\Dataset .csv'

data_df=pd.read_csv(path)

highest_average_ratings_city=data_df.groupby('City')['Aggregate rating'].mean().

highest_average_ratings_city= highest_average_ratings_city.sort_values(by='Aggre

print(highest_average_ratings_city)
```
```
City              Inner City
Aggregate rating         4.9
Name: 56, dtype: object
```
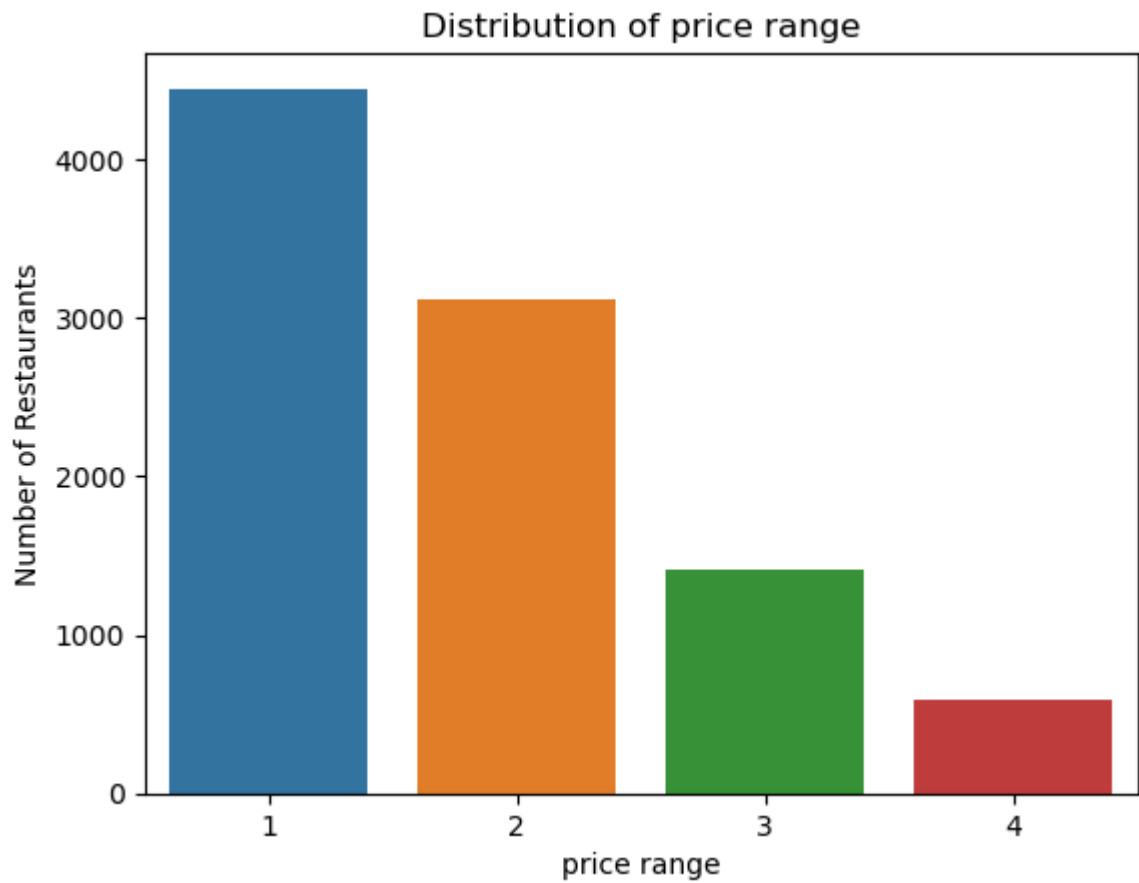
Task-3: Price Range Distribution

**Create a histogram or bar chart to visualize the distribution of price ranges among the restaurants**

In [7]:
```python
import seaborn as sns

sns.countplot(data=data_df,x='Price range')

plt.title('Distribution of price range')
plt.xlabel('price range')
plt.ylabel('Number of Restaurants')
plt.show()
```

Distribution of price range

```
In [8]:  plt.hist(data_df['Price range'])
         plt.title('Distrubution Price range')
         plt.xlabel('price range')
         plt.ylabel('Number of Restaurants')
         plt.show()
```

Distrubution Price range
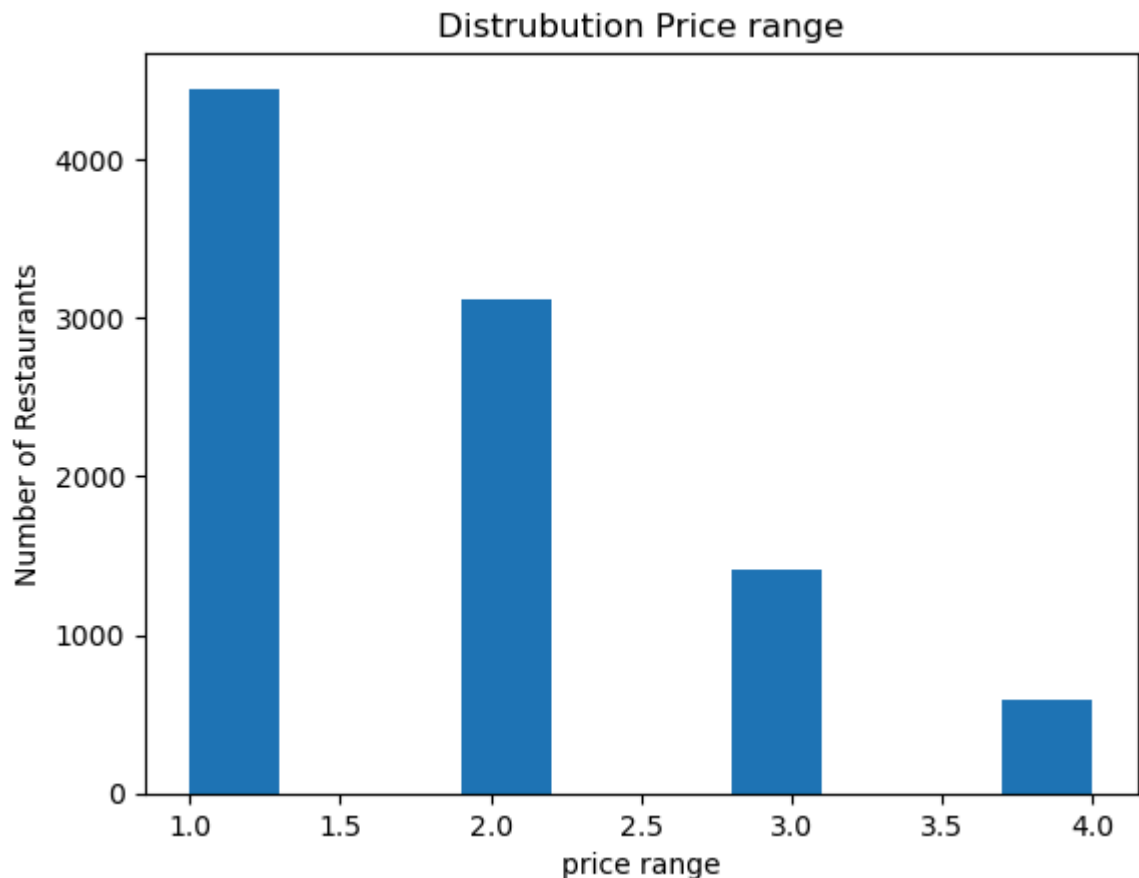
Calculate the percentage of restaurants in each price range category

```
In [9]:  import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import seaborn as sns

         path=r'C:\Users\DELL\Documents\cognifyz Task\Data set\Dataset .csv'

         data_df=pd.read_csv(path)

         price_range_percentage = data_df['Price range'].value_counts(normalize=True) * 1

         price_range_percentage_df = price_range_percentage.reset_index()
         price_range_percentage_df.columns = ['Price Range', 'Percentage']

         print(price_range_percentage_df)
```

```
   Price Range  Percentage
0            1   46.529159
1            2   32.593446
2            3   14.741912
3            4    6.135483
```

Task-4: Online Delivery

Determine the percentage of restaurants that offer online delivery

```
In [10]:  path=r'C:\Users\DELL\Documents\cognifyz Task\Data set\Dataset .csv'
```

```
data_df=pd.read_csv(path)

percentage_of_Online_delivery = data_df['Has Online delivery'].value_counts(norm

percentage_of_Online_delivery_df = percentage_of_Online_delivery.reset_index()
percentage_of_Online_delivery_df.columns = ['Has Online deliverye', 'Percentage'

print(percentage_of_Online_delivery_df)
```

```
  Has Online deliverye  Percentage
0                   No   74.337766
1                  Yes   25.662234
```

Compare the average ratings of restaurants with and without online delivery

In [11]:
```python
import pandas as pd


df = pd.read_csv(path)


average_ratings_comparison = df.groupby('Has Online delivery')['Aggregate rating


average_ratings_comparison.columns = ['Has Online Delivery', 'Average Rating']


print(average_ratings_comparison)
```

```
  Has Online Delivery  Average Rating
0                  No        2.465296
1                 Yes        3.248837
```

Level 2

Task-1: Restaurant Ratings

Analyze the distribution of aggregate ratings and determine the most common rating range

In [31]:
```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

df= pd.read_csv(path)

sns.histplot(data_df['Aggregate rating'],bins=20,kde=True)
plt.title('Aggregate rating')
plt.xlabel('Rating')
plt.ylabel('Frequency')
plt.show()

rating_bins = pd.cut(data_df['Aggregate rating'], bins=[0, 1, 2, 3, 4, 5], right
```
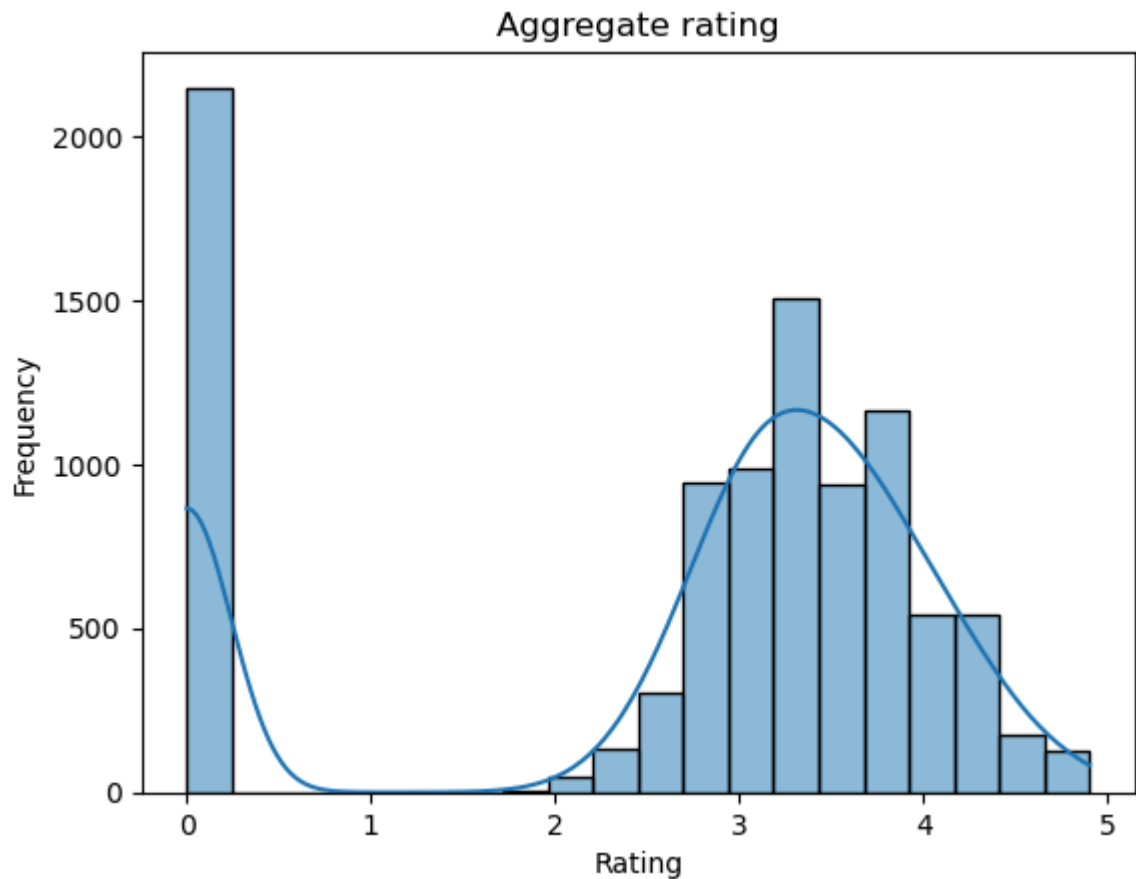
```
most_common_range = rating_bins.value_counts().idxmax()
print(f"The most common rating range is: {most_common_range}")
```

Aggregate rating

The most common rating range is: [3, 4)

Calculate the average number of votes received by restaurants

In [4]:
```python
import pandas as pd

path=r'C:\Users\DELL\Documents\cognifyz Task\Data set\Dataset .csv'

data_df=pd.read_csv(path)

df=pd.DataFrame(data_df)

avrages_vote=df['Votes'].mean()
avrages_vote
```

Out[4]: 156.909747670401

In [ ]: