In [30]:
```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import geopandas as gpd
import contextily as ctx
import warnings
warnings.filterwarnings('ignore')
```

In [31]:
```python
#load and read dataset
df=pd.read_csv('Dataset.csv')
df
```

Out[31]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	...
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Century City Mall, Poblacion, Makati City	Century City Mall, Poblacion, Makati City, Mak...	121.027535	14.565443	French, Japanese, Desserts	...
1	6304287	Izakaya Kikufuji	162	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...	Little Tokyo, Legaspi Village, Makati City	Little Tokyo, Legaspi Village, Makati City, Ma...	121.014101	14.553708	Japanese	...
2	6300002	Heat - Edsa Shangri-La	162	Mandaluyong City	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...	Edsa Shangri-La, Ortigas, Mandaluyong City	Edsa Shangri-La, Ortigas, Mandaluyong City, Ma...	121.056831	14.581404	Seafood, Asian, Filipino, Indian	...
3	6318506	Ooma	162	Mandaluyong City	Third Floor, Mega Fashion Hall, SM Megamall, O...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.056475	14.585318	Japanese, Sushi	...
4	6314302	Sambo Kojin	162	Mandaluyong City	Third Floor, Mega Atrium, SM Megamall, Ortigas...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.057508	14.584450	Japanese, Korean	...
...	...	...	...	...	...	...	...	...	...	...	...
9546	5915730	Naml' Gurme	208	��stanbul	Kemanke�� Karamustafa Pa��a Mahallesi, R'ht'm ...	Karak�_y	Karak�_y, ��stanbul	28.977392	41.022793	Turkish	...
9547	5908749	Ceviz A��ac'	208	��stanbul	Ko��uyolu Mahallesi, Muhittin ��st�_nda�� Cadd...	Ko��uyolu	Ko��uyolu, ��stanbul	29.041297	41.009847	World Cuisine, Patisserie, Cafe	...
9548	5915807	Huqqa	208	��stanbul	Kuru�_e��me Mahallesi, Muallim Naci Caddesi, N...	Kuru�_e��me	Kuru�_e��me, ��stanbul	29.034640	41.055817	Italian, World Cuisine	...
9549	5916112	A���k Kahve	208	��stanbul	Kuru�_e��me Mahallesi, Muallim Naci Caddesi, N...	Kuru�_e��me	Kuru�_e��me, ��stanbul	29.036019	41.057979	Restaurant Cafe	...
9550	5927402	Walter's Coffee Roastery	208	��stanbul	Cafea��a Mahallesi, Bademalt' Sokak, No 21/B, ...	Moda	Moda, ��stanbul	29.026016	40.984776	Cafe	...

9551 rows × 21 columns

Level 2

Task 1

Cognifyz
Where Data Meets Intelligence

Task: Restaurant Ratings

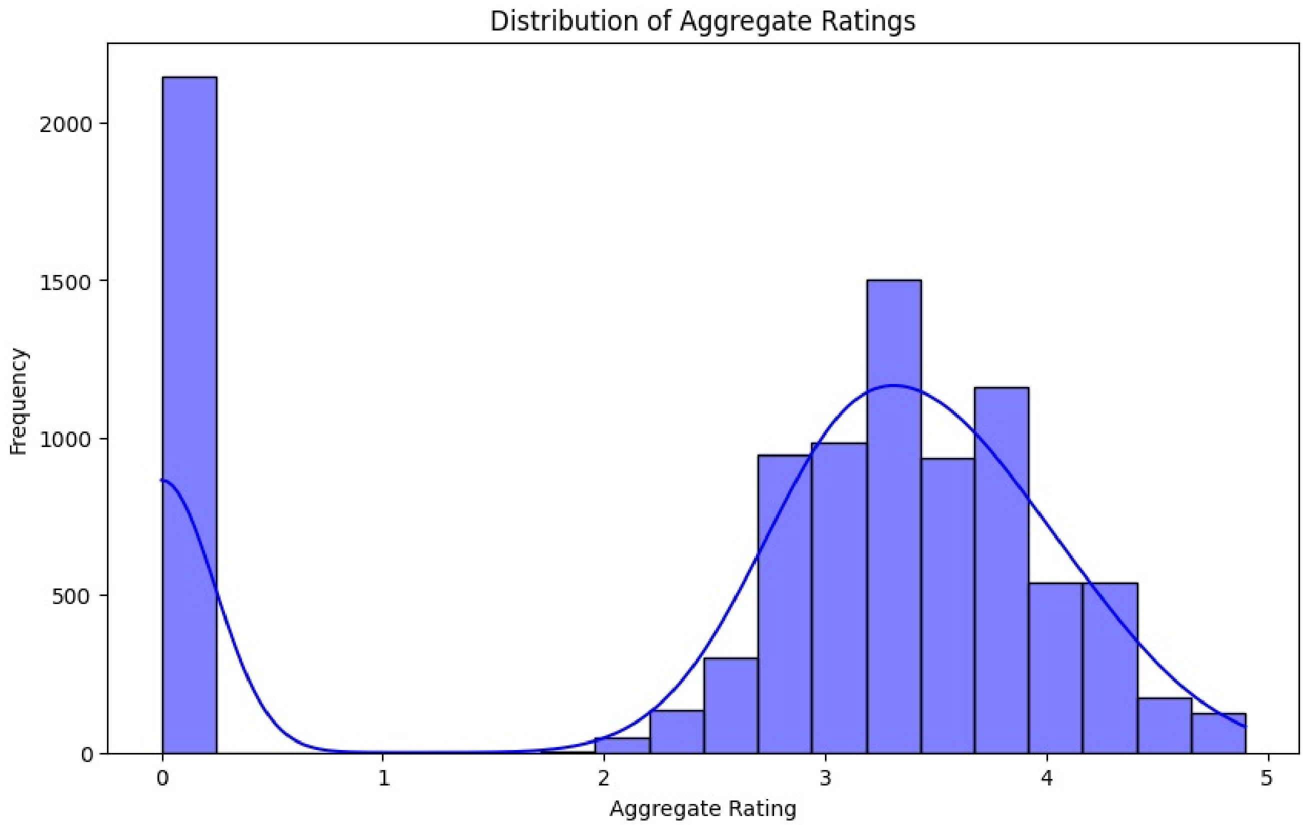
Analyze the distribution of aggregate ratings and determine the most common rating range.

Calculate the average number of votes received by restaurants.

Task

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

In [32]:
```
# Analyze the distribution of aggregate ratings
plt.figure(figsize=(10, 6))
sns.histplot(df['Aggregate rating'], bins=20, kde=True, color='blue')
plt.title('Distribution of Aggregate Ratings')
plt.xlabel('Aggregate Rating')
plt.ylabel('Frequency')
plt.show()
```



In [33]:
```
# Determine the most common rating range
rating_distribution = df['Aggregate rating'].value_counts().sort_index()
rating_distribution
```

Out[33]:
```
Aggregate rating
0.0    2148
1.8       1
1.9       2
2.0       7
2.1      15
2.2      27
2.3      47
2.4      87
2.5     110
2.6     191
2.7     250
2.8     315
2.9     381
3.0     468
3.1     519
3.2     522
3.3     483
3.4     498
3.5     480
3.6     458
3.7     427
3.8     400
3.9     335
4.0     266
4.1     274
4.2     221
4.3     174
4.4     144
4.5      95
4.6      78
4.7      42
4.8      25
4.9      61
Name: count, dtype: int64
```

In [34]:
```
most_common_rating=rating_distribution.idxmax()
print(most_common_rating)
```

```
0.0
```

Task

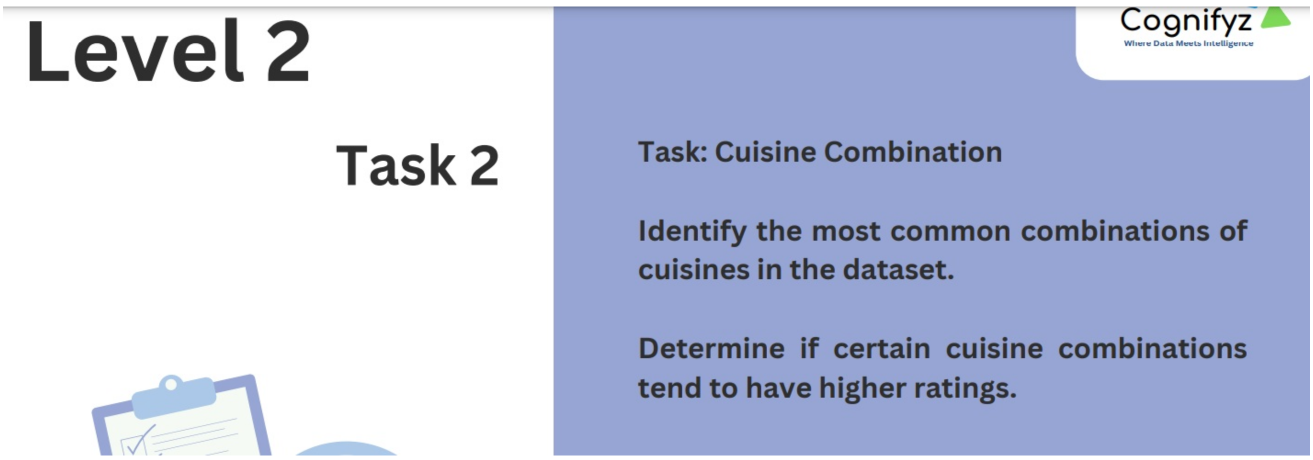
1. Calculate the average number of votes received by restaurants.

In [35]:
```
average_rating = df["Votes"].mean()
print(f"Average number of votes received by restaurants: {average_rating:.2f}")
```

```
Average number of votes received by restaurants: 156.91
```

Insight:

The average number of votes received by restaurants is 156.91



Task

1. Identify the most common combinations of cuisines in the dataset.

In [36]:
```
most_common_combinations=df["Cuisines"].str.split(', ').value_counts()
most_common_combinations.head(10)
```

```
Out[36]:
Cuisines
[North Indian]                          936
[North Indian, Chinese]                 511
[Chinese]                               354
[Fast Food]                             354
[North Indian, Mughlai]                 334
[Cafe]                                  299
[Bakery]                                218
[North Indian, Mughlai, Chinese]        197
[Bakery, Desserts]                      170
[Street Food]                           149
Name: count, dtype: int64
```

Insight

1. North Indian cuisine is the most popular This is evident from its high frequency in the dataset.
2. Diverse cuisine options: While North Indian cuisine dominates, there's a variety of other cuisines available, including Chinese, Fast Food, and various regional options.

Task

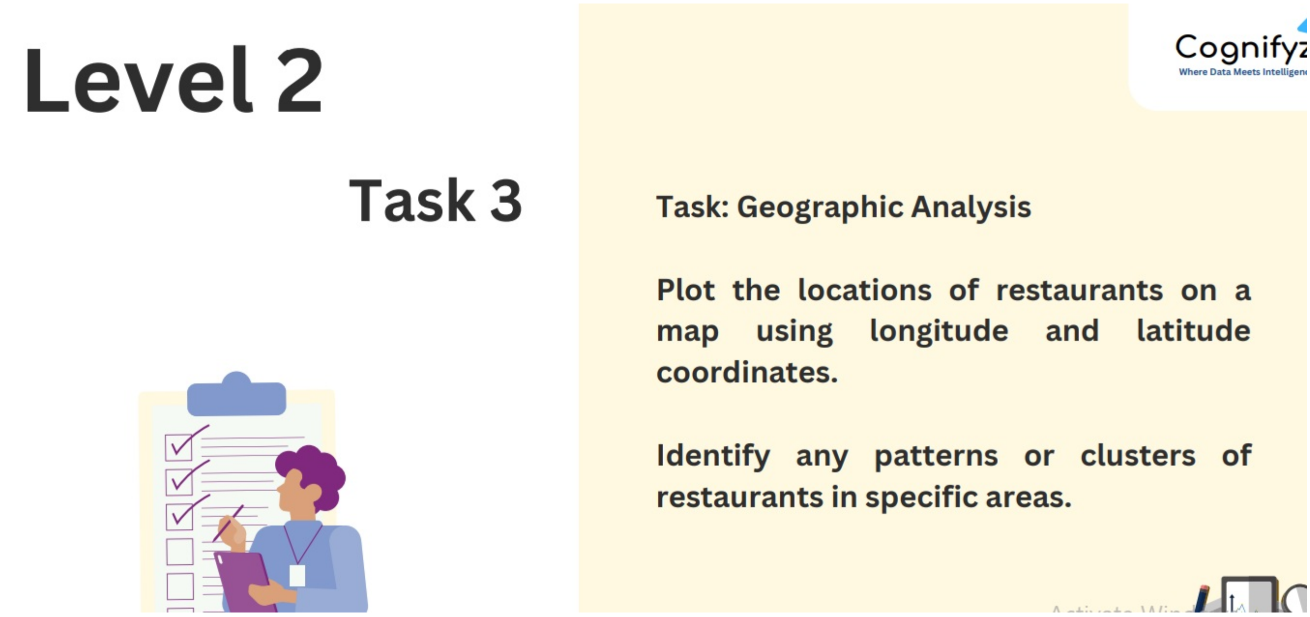
4.Determine if certain cuisine combinations tend to have higher ratings.

```python
In [37]:
# Split the cuisines into individual entries
df["Cuisines"]=df["Cuisines"].str.split(', ')
# # Explode the dataframe so each cuisine combination is a separate row
to_explode = df.explode("Cuisines")
```

```python
In [38]:
cuisines_rating = to_explode.groupby("Cuisines")["Aggregate rating"].mean().reset_index()
```

```python
In [39]:
# Sort by the mean rating in descending order
cuisines_rating = cuisines_rating.sort_values(by='Aggregate rating',ascending=False)
# Display the top 10 cuisines with the highest average ratings
print(cuisines_rating.head())
```

```
      Cuisines  Aggregate rating
130      Sunda          4.900000
26      B�_rek          4.700000
132  Taiwanese          4.650000
112      Ramen          4.500000
43     Dim Sum          4.466667
```



Task

5.Plot the locations of restaurants on a map using longitude and latitude coordinates.

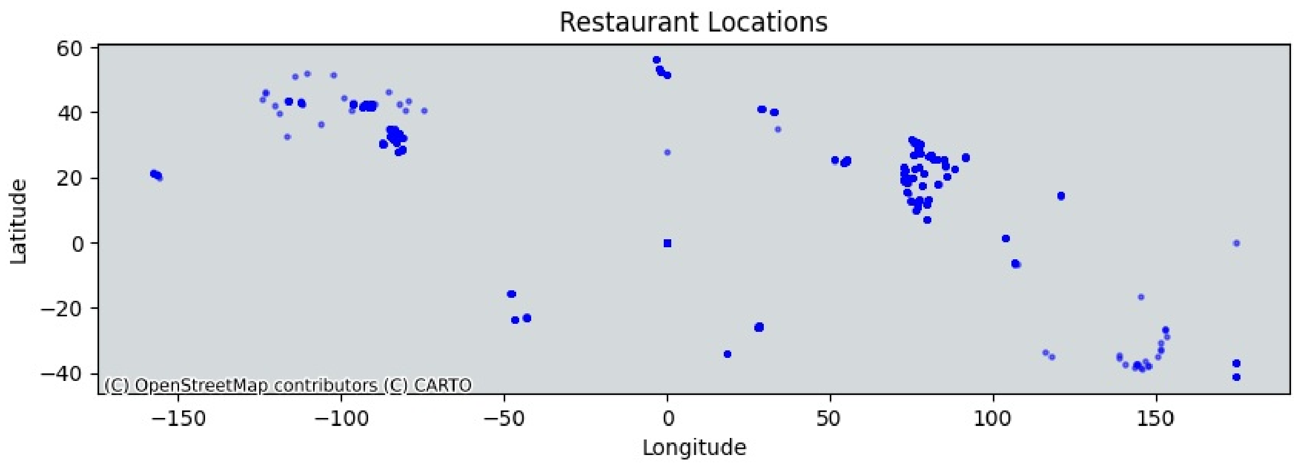
```python
In [40]:
# Assuming df is your DataFrame containing the Latitude and Longitude columns
# Create a GeoDataFrame from the DataFrame
geometry = gpd.points_from_xy(df["Longitude"], df["Latitude"])  # Note: Longitude first, then Latitude
gdf = gpd.GeoDataFrame(df, geometry=geometry)
geometry,gdf
```

```
Out[40]:
(<GeometryArray>
```

```
[<POINT (121.028 14.565)>, <POINT (121.014 14.554)>, <POINT (121.057 14.581)>,
 <POINT (121.056 14.585)>, <POINT (121.058 14.584)>, <POINT (121.056 14.584)>,
  <POINT (120.98 14.531)>,  <POINT (120.979 14.54)>,  <POINT (120.98 14.553)>,
 <POINT (121.057 14.572)>,
 ...
  <POINT (29.057 41.105)>,   <POINT (29.023 40.99)>,  <POINT (28.978 41.025)>,
  <POINT (28.978 41.023)>,  <POINT (28.981 41.026)>,  <POINT (28.977 41.023)>,
   <POINT (29.041 41.01)>,  <POINT (29.035 41.056)>,  <POINT (29.036 41.058)>,
  <POINT (29.026 40.985)>]
Length: 9551, dtype: geometry,
      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              Namlı Gurme           208         ��stanbul
9547        5908749             Ceviz A��acı           208         ��stanbul
9548        5915807                    Huqqa           208         ��stanbul
9549        5916112              A���k Kahve           208         ��stanbul
9550        5927402  Walter's Coffee Roastery           208         ��stanbul

                                                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  Kemanke�� Karamustafa Pa��a Mahallesi, Rıhtım ...
9547  Ko��uyolu Mahallesi, Muhittin ��st�_nda�� Cadd...
9548  Kuru�_e��me Mahallesi, Muallim Naci Caddesi, N...
9549  Kuru�_e��me Mahallesi, Muallim Naci Caddesi, N...
9550  Cafea��a Mahallesi, Bademaltı 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                                   Karak�_y
9547                                  Ko��uyolu
9548                                Kuru�_e��me
9549                                Kuru�_e��me
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                                Karak�_y, ��stanbul   28.977392
9547                               Ko��uyolu, ��stanbul   29.041297
9548                             Kuru�_e��me, ��stanbul   29.034640
9549                             Kuru�_e��me, ��stanbul   29.036019
9550                                    Moda, ��stanbul   29.026016

       Latitude                          Cuisines  ...  Has Table booking  \
0     14.565443        [French, Japanese, Desserts]  ...                Yes
1     14.553708                          [Japanese]  ...                Yes
2     14.581404  [Seafood, Asian, Filipino, Indian]  ...                Yes
3     14.585318                  [Japanese, Sushi]  ...                 No
4     14.584450                 [Japanese, Korean]  ...                Yes
...         ...                                 ...  ...                ...
9546  41.022793                           [Turkish]  ...                 No
9547  41.009847   [World Cuisine, Patisserie, Cafe]  ...                 No
9548  41.055817            [Italian, World Cuisine]  ...                 No
9549  41.057979                  [Restaurant Cafe]  ...                 No
9550  40.984776                              [Cafe]  ...                 No

     Has Online delivery Is delivering now Switch to order menu Price range  \
0                     No                No                   No           3
1                     No                No                   No           3
2                     No                No                   No           4
3                     No                No                   No           4
4                     No                No                   No           4
...                  ...               ...                  ...         ...
9546                  No                No                   No           3
9547                  No                No                   No           3
9548                  No                No                   No           4
9549                  No                No                   No           4
9550                  No                No                   No           2

     Aggregate rating  Rating color  Rating text Votes  \
```

```
0                 4.8    Dark Green    Excellent   314
1                 4.5    Dark Green    Excellent   591
2                 4.4         Green    Very Good   270
3                 4.9    Dark Green    Excellent   365
4                 4.8    Dark Green    Excellent   229
...               ...           ...          ...   ...
9546              4.1         Green    Very Good   788
9547              4.2         Green    Very Good  1034
9548              3.7        Yellow         Good   661
9549              4.0         Green    Very Good   901
9550              4.0         Green    Very Good   591

                        geometry
0     POINT (121.02754 14.56544)
1      POINT (121.0141 14.55371)
2      POINT (121.05683 14.5814)
3     POINT (121.05648 14.58532)
4     POINT (121.05751 14.58445)
...                          ...
9546   POINT (28.97739 41.02279)
9547    POINT (29.0413 41.00985)
9548   POINT (29.03464 41.05582)
9549   POINT (29.03602 41.05798)
9550   POINT (29.02602 40.98478)

[9551 rows x 22 columns])
```

In [41]:
```
# Set the coordinate reference system to WGS84
gdf.set_crs(epsg=4326, inplace=True)
# Plot the locations
fig, ax = plt.subplots(figsize=(10, 10))
gdf.plot(ax=ax, color='blue', markersize=5, alpha=0.5)
# Add a basemap
gdf = gdf.to_crs(epsg=3857)
ctx.add_basemap(ax, source=ctx.providers.CartoDB.Positron)
# Show the plot
plt.title('Restaurant Locations')
plt.xlabel('Longitude')
plt.ylabel('Latitude')
plt.show()
```



Task

1. Identify any patterns or clusters of restaurants in specific areas.

In [42]:
```
# Count the number of restaurants in each city
city_counts = df['City'].value_counts()
print("Number of restaurants in each city:")
print(city_counts)
```

```
Number of restaurants in each city:
City
New Delhi         5473
Gurgaon           1118
Noida             1080
Faridabad          251
Ghaziabad           25
                  ...
Lakes Entrance       1
Mohali               1
Panchkula            1
Bandung              1
Randburg             1
Name: count, Length: 141, dtype: int64
```

In [43]:
```
grouped_data = df.groupby(['City', 'Locality']).size().reset_index(name='Restaurant Count')
```

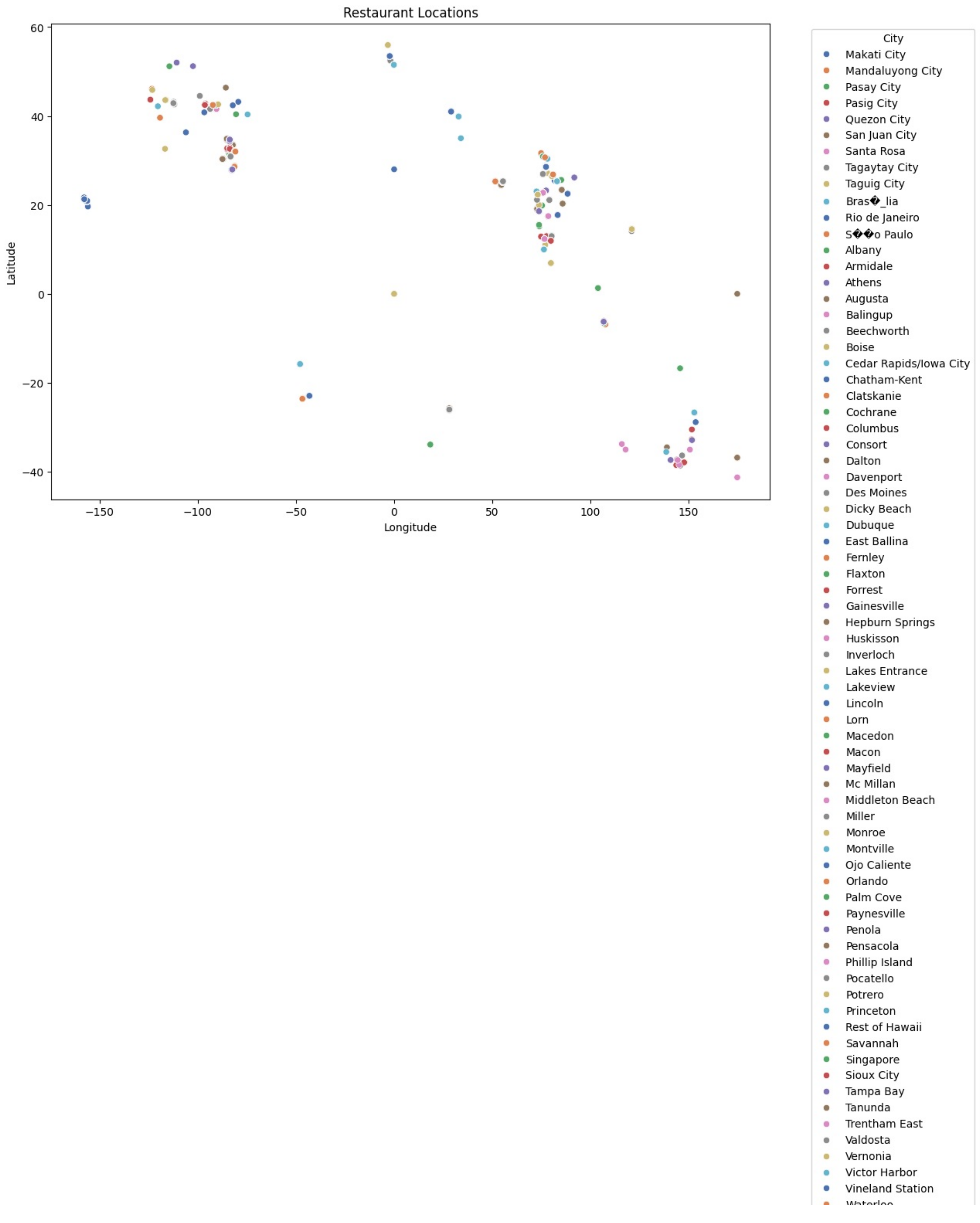
```
grouped_data.head()
```

Out[43]:

	City	Locality	Restaurant Count
0	Abu Dhabi	Abu Dhabi Mall, Tourist Club Area (Al Zahiyah)	2
1	Abu Dhabi	Al Dhafrah	2
2	Abu Dhabi	Al Mushrif	1
3	Abu Dhabi	Al Wahda Mall, Al Wahda	2
4	Abu Dhabi	Crowne Plaza Abu Dhabi, Al Markaziya	1

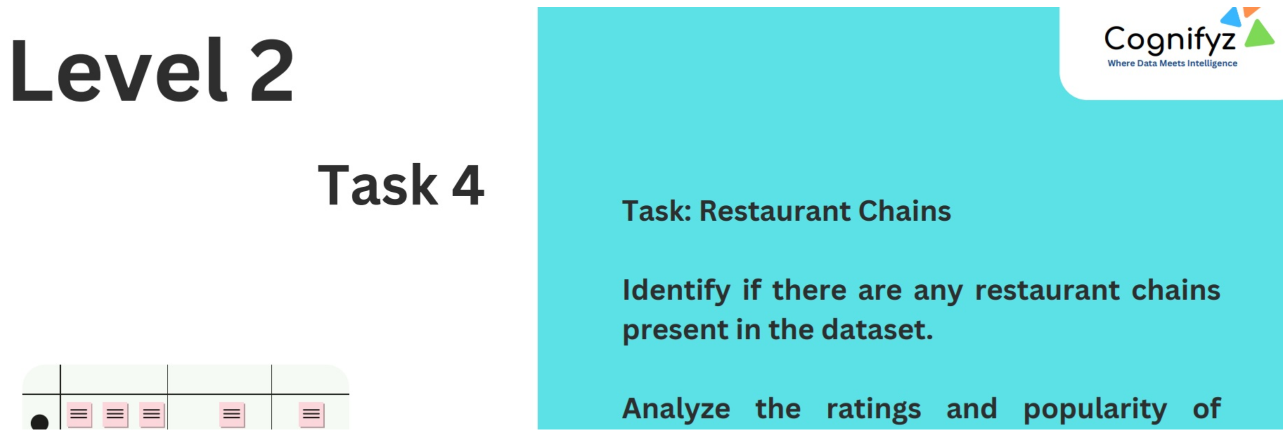
In [44]:

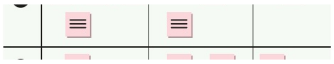
```
# Create a scatter plot of restaurant locations
plt.figure(figsize=(12, 8))
sns.scatterplot(data=df, x='Longitude', y='Latitude', hue='City', palette='deep')
plt.title('Restaurant Locations')
plt.xlabel('Longitude')
plt.ylabel('Latitude')
plt.legend(title='City', bbox_to_anchor=(1.05, 1), loc='upper left')
plt.show()
```



Observation

The scatter plot above shows the distribution of restaurants based on their longitude and latitude coordinates, with different colors representing different cities. We can observe some clear clusters of restaurants in specific areas.





different restaurant chains.

Task

1. Identify if there are any restaurant chains present in the dataset.

In [45]:
```
df.groupby(df['Cuisines']=='Chinese').size()
```

Out[45]:
```
Cuisines
False    9551
dtype: int64
```

In [46]:
```
# Identify restaurant chains by finding restaurants with the same name in different locations<<<<<<<julia file>
restaurant_chains = df.groupby('Restaurant Name').filter(lambda x: len(x) > 1)
restaurant_chains
```

Out[46]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	
5	18189371	Din Tai Fung	162	Mandaluyong City	Ground Floor, Mega Fashion Hall, SM Megamall, ...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.056314	14.583764	
10	6309903	Silantro Fil-Mex	162	Pasig City	75 East Capitol Drive, Kapitolyo, Pasig City	Kapitolyo	Kapitolyo, Pasig City	121.057916	14.567689	
12	6318433	Silantro Fil-Mex	162	Quezon City	Second Floor, UP Town Center, Katipunan Avenue...	UP Town Center, Diliman, Quezon City	UP Town Center, Diliman, Quezon City, Quezon City	121.075419	14.649503	
35	6601589	Coco Bambu	30	Bras�_lia	Bras�_lia Shopping - Piso 2, SCN 5, Bloco A, A...	Bras�_lia Shopping, Asa Norte	Bras�_lia Shopping, Asa Norte, Bras�_lia	-47.889000	-15.786500	[I
39	6600427	Coco Bambu	30	Bras�_lia	SCES, Trecho 2, Conjunto 13/36, Setor de Clube...	Setor De Clubes Esportivos Sul	Setor De Clubes Esportivos Sul, Bras�_lia	-47.868500	-15.819000	[I
...	...	...	...	...	...	...	...	...	...	
9517	6001980	Timboo Cafe	208	Ankara	Armada AVM, Kat -1, Eski��ehir Yolu, No 6, Yen...	S�_���_t�_z�_, Yenimahalle	S�_���_t�_z�_, Yenimahalle, Ankara	32.809247	39.913206	
9528	6003668	Timboo Cafe	208	Ankara	Kentpark AVM, Kat -1, Mustafa Kemal Mahallesi,...	Kentpark AVM, ��niversiteler, �ankaya	Kentpark AVM, ��niversiteler, �ankaya, Ankara	32.776255	39.908957	
9534	6004089	D�_vero��lu	208	Ankara	Maltepe Mahallesi, Gen�_lik Caddesi, No 28, �a...	Maltepe	Maltepe, Ankara	32.842742	39.922536	
9535	6000921	D�_vero��lu	208	Ankara	��mitk�_y Mahallesi, 2432. Cadde (8. Cadde), N...	��mitk�_y	��mitk�_y, Ankara	32.701775	39.891564	
9538	5901782	Starbucks	208	��stanbul	Bebek Mahallesi, Cevdetpa��a Caddesi, No 30/A,...	Bebek	Bebek, ��stanbul	29.043734	41.077696	

2839 rows × 21 columns

In [47]:
```
# Display the restaurant chains
final_data=restaurant_chains[['Restaurant Name', 'City', 'Address']].sort_values(by='Restaurant Name').head(20)
final_data
```

Out[47]:

	Restaurant Name	City	Address
751	10 Downing Street	Bhopal	Third Floor, DB City Mall, Maharana Pratap Nag...
2333	10 Downing Street	Indore	Second Floor, Malhar Mega Mall, AB Road, Schem...
8848	221 B Baker Street	Noida	PG 30, TOT Mall, Sector 62, Noida
8498	221 B Baker Street	Noida	21, Jalvayu Vihar Market, Sector 25, Noida
8039	221 B Baker Street	Noida	10, Brahmaputra Shoping Complex, Sector 29, Noida
5547	34 Parkstreet Lane	New Delhi	Shop 7, Mukherjee Tower, Mukherjee Nagar, New ...
3903	34 Parkstreet Lane	New Delhi	DDA Market,Kala Sarai, Hauz Khas, New Delhi
7701	34, Chowringhee Lane	New Delhi	V 3 S Mall, Laxmi Nagar, New Delhi
5444	34, Chowringhee Lane	New Delhi	B-10, Opposite Metro Pillar 21, Model Town 2, ...
6166	34, Chowringhee Lane	New Delhi	Shop 9, Block A2, DDA Market, Paschim Vihar, N...
2691	34, Chowringhee Lane	New Delhi	115, Central Market, Ashok Vihar Phase 1, New ...
4462	34, Chowringhee Lane	New Delhi	UB-101, Kamla Nagar, New Delhi
4355	34, Chowringhee Lane	New Delhi	61-D, Ground Floor, Ber Sarai Market, Opposite...
6720	34, Chowringhee Lane	New Delhi	C8/354, Sector 8, Rohini, New Delhi
6499	34, Chowringhee Lane	New Delhi	Shop 68, Vasant Place Market, Sector 6, R K Pu...
7972	34, Chowringhee Lane	New Delhi	G-37, Ground Floor, Westend Mall, Janakpuri, N...
4093	34, Chowringhee Lane	New Delhi	23/1, Prem Nagar, Jail Road, New Delhi
7622	34, Chowringhee Lane	New Delhi	Shop 2, Plot 57, Under Dwarka Mor Metro Statio...
7000	34, Chowringhee Lane	New Delhi	93, Opposite Venkateswara College, Satyaniketa...
3431	4700BC Popcorn	New Delhi	Ground Floor, DLF Place Mall, Saket, New Delhi

Task

1. Analyze the ratings and popularity of different restaurant chains.

In [48]:
```
# Group by 'Restaurant Name' to analyze ratings and popularity
restaurant_analysis = df.groupby('Restaurant Name').agg({'Aggregate rating':'mean','Votes':'sum'}).reset_index(
restaurant_analysis
```

Out[48]:

	Restaurant Name	Aggregate rating	Votes
0	#45	3.6	209
1	#Dilliwaala6	3.7	124
2	#InstaFreeze	0.0	2
3	#OFF Campus	3.7	216
4	#Urban Caf��	3.3	49
...	...	...	...
7441	t Lounge by Dilmah	3.6	34
7442	tashas	4.1	374
7443	wagamama	3.7	131
7444	{Niche} - Cafe & Bar	4.1	492
7445	�ukura��a Sofras	4.4	296

7446 rows × 3 columns

In [49]:
```
# Rename columns for clarity
restaurant_analysis.columns = ['Restaurant Name', 'Average Rating', 'Total Votes']
```

In [53]:
```
# Sort by Total Votes to see the most popular restaurants
popular_restaurants = restaurant_analysis.sort_values(by='Total Votes', ascending=False).head(10)
# Sort by Average Rating to see the top-rated restaurants
top_rated_restaurants = restaurant_analysis.sort_values(by='Average Rating', ascending=False).head(10)
popular_restaurants
```

Out[53]:

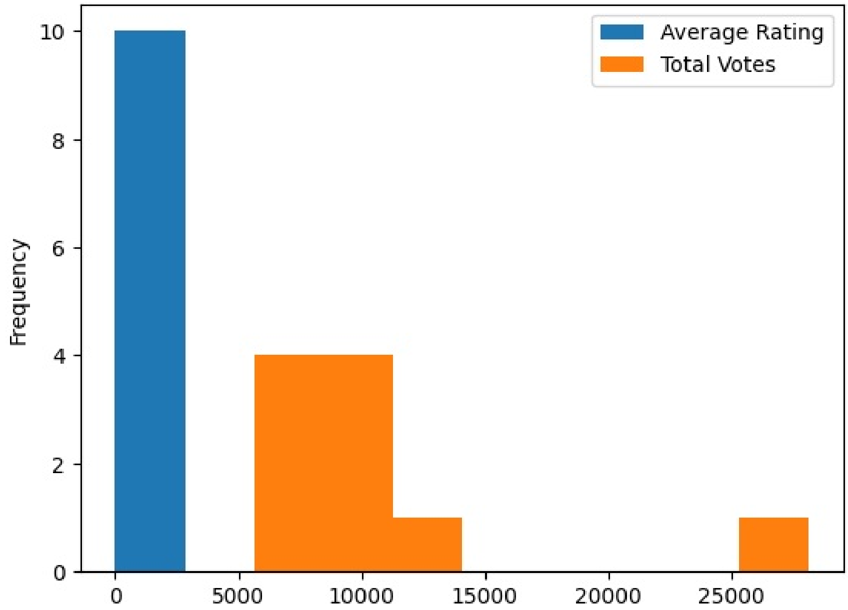
	Restaurant Name	Average Rating	Total Votes
663	Barbeque Nation	4.353846	28142
101	AB's - Absolute Barbecues	4.825000	13400
6943	Toit	4.800000	10934
785	Big Chill	4.475000	10853
2297	Farzi Cafe	4.366667	10098
6988	Truffles	3.950000	9682
1510	Chili's	4.580000	8156
2879	Hauz Khas Social	4.300000	7931
3261	Joey's Pizza	4.250000	7807
4902	Peter Cat	4.300000	7574

In [52]:

```
popular_restaurants.plot(kind='hist')
```

Out[52]:

```
<Axes: ylabel='Frequency'>
```



In [54]:

```
top_rated_restaurants
```

Out[54]:

	Restaurant Name	Average Rating	Total Votes
5946	Solita	4.9	162
6045	Spiral - Sofitel Philippine Plaza Manila	4.9	621
2354	Flat Iron	4.9	309
5477	Sagar Gaire Fast Food	4.9	427
651	Bao	4.9	161
7339	Yellow Dog Eats	4.9	1252
2559	Gaga Manjero	4.9	95
6171	Sushi Masa	4.9	605
5757	Shorts Burger and Shine	4.9	820
5726	Sheroes Hangout	4.9	77

In [ ]: