Task: Votes Analysis

3.2.1 Identify the restaurants with the highest and lowest number of votes.

In [7]:
```python
import pandas as pd
```

In [2]:
```python
dt = pd.read_csv("C:\\Users\\MAHESH\Desktop\cognifyz\Dataset.csv")
dt
```

Out[2]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	...	Currency	b
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	...	Botswana Pula(P)	
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	...	Botswana Pula(P)	
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	...	Botswana Pula(P)	
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	...	Botswana Pula(P)	
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	...	Botswana Pula(P)	
...	...	...	...	...	...	...	...	...	...	...	...	...	
9546	5915730	Namlı Gurme	208	İstanbul	Kemankeş Karamustafa Paşa Mahallesi, Rıhtım ...	Karaköy	Karaköy, İstanbul	28.977392	41.022793	Turkish	...	Turkish Lira(TL)	
9547	5908749	Ceviz Ağacı	208	İstanbul	Koşuyolu Mahallesi, Muhittin Üstündağ Cadd...	Koşuyolu	Koşuyolu, İstanbul	29.041297	41.009847	World Cuisine, Patisserie, Cafe	...	Turkish Lira(TL)	
9548	5915807	Huqqa	208	İstanbul	Kuruçeşme Mahallesi, Muallim Naci Caddesi, N...	Kuruçeşme	Kuruçeşme, İstanbul	29.034640	41.055817	Italian, World Cuisine	...	Turkish Lira(TL)	
9549	5916112	Aşk Kahve	208	İstanbul	Kuruçeşme Mahallesi, Muallim Naci Caddesi, N...	Kuruçeşme	Kuruçeşme, İstanbul	29.036019	41.057979	Restaurant Cafe	...	Turkish Lira(TL)	
9550	5927402	Walter's Coffee Roastery	208	İstanbul	Cafeağa Mahallesi, Bademaltı Sokak, No 21/B, ...	Moda	Moda, İstanbul	29.026016	40.984776	Cafe	...	Turkish Lira(TL)	

9551 rows × 21 columns

In [3]:
```python
votes = dt.groupby('Restaurant Name')['Aggregate rating'].count()
```

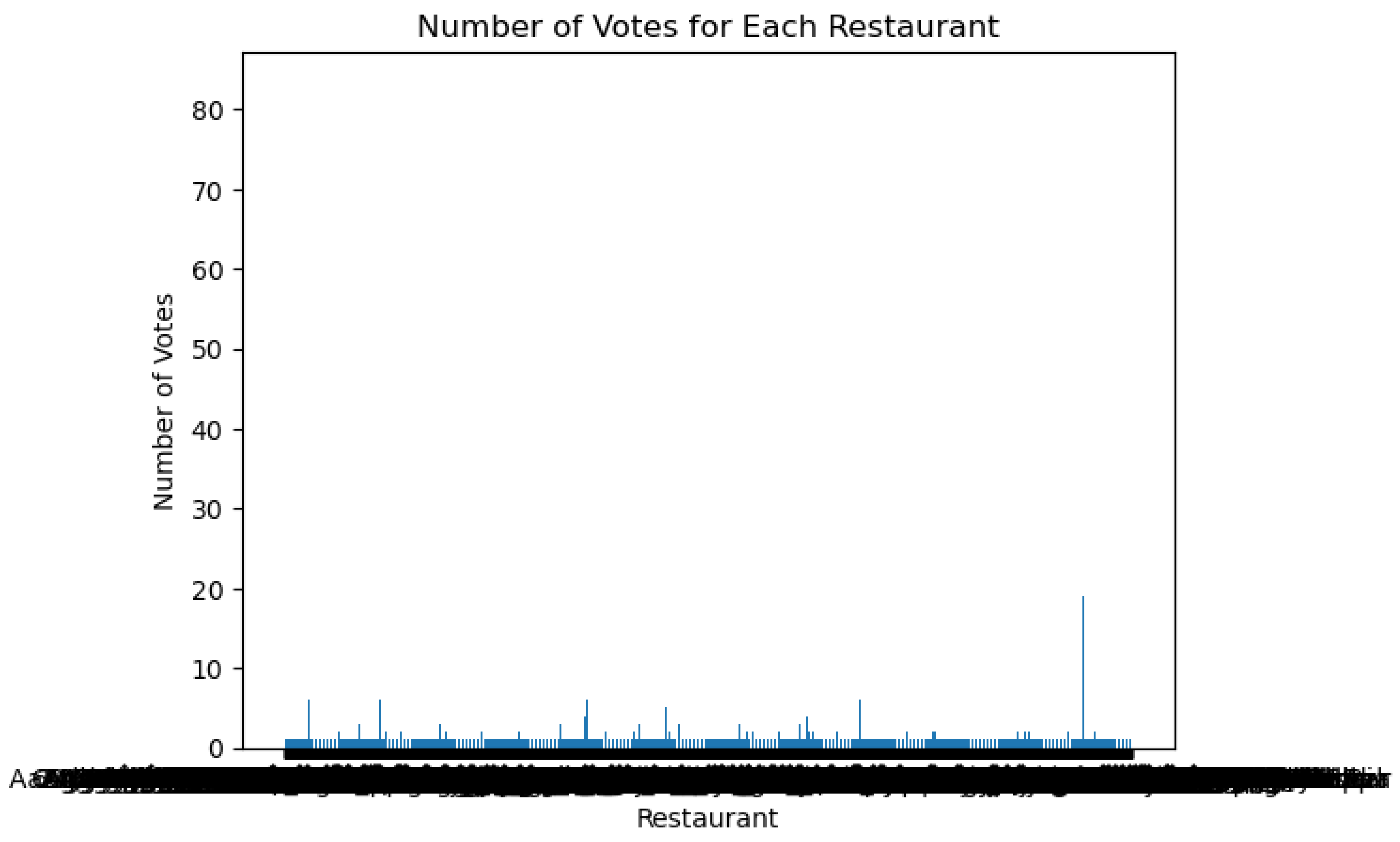
In [4]:
```python
votes
```

Out[4]:
```
Restaurant Name
#45                        1
#Dilliwaala6               1
#InstaFreeze               1
#OFF Campus                1
#Urban Café              1
                          ..
t Lounge by Dilmah         1
tashas                     1
wagamama                   1
{Niche} - Cafe & Bar       1
Ükuraå Sofrası          1
Name: Aggregate rating, Length: 7446, dtype: int64
```

In [5]:
```python
# Identify the restaurants with the highest and lowest number of votes
highest_votes = votes.sort_values(ascending=False).index[0]
lowest_votes = votes.sort_values(ascending=True).index[0]
# Print the restaurants with the highest and lowest number of votes
print(f'Restaurant with the highest number of votes: {highest_votes}')
print(f'Restaurant with the lowest number of votes: {lowest_votes}')
```

```
Restaurant with the highest number of votes: Cafe Coffee Day
Restaurant with the lowest number of votes: #45
```

In [6]:
```python
import matplotlib.pyplot as plt
# Create a bar chart of the number of votes for each restaurant
plt.bar(votes.index, votes.values)
plt.xlabel('Restaurant')
plt.ylabel('Number of Votes')
plt.title('Number of Votes for Each Restaurant')
plt.show()
```



3.2.1 Analyze if there is a correlation between the number of votes and the rating of a restaurant

In [8]:
```python
from scipy.stats import pearsonr
```

In [9]:
```python
corr, p = pearsonr(dt['Votes'], dt['Aggregate rating'])
```

In [10]:
```python
corr, p
```

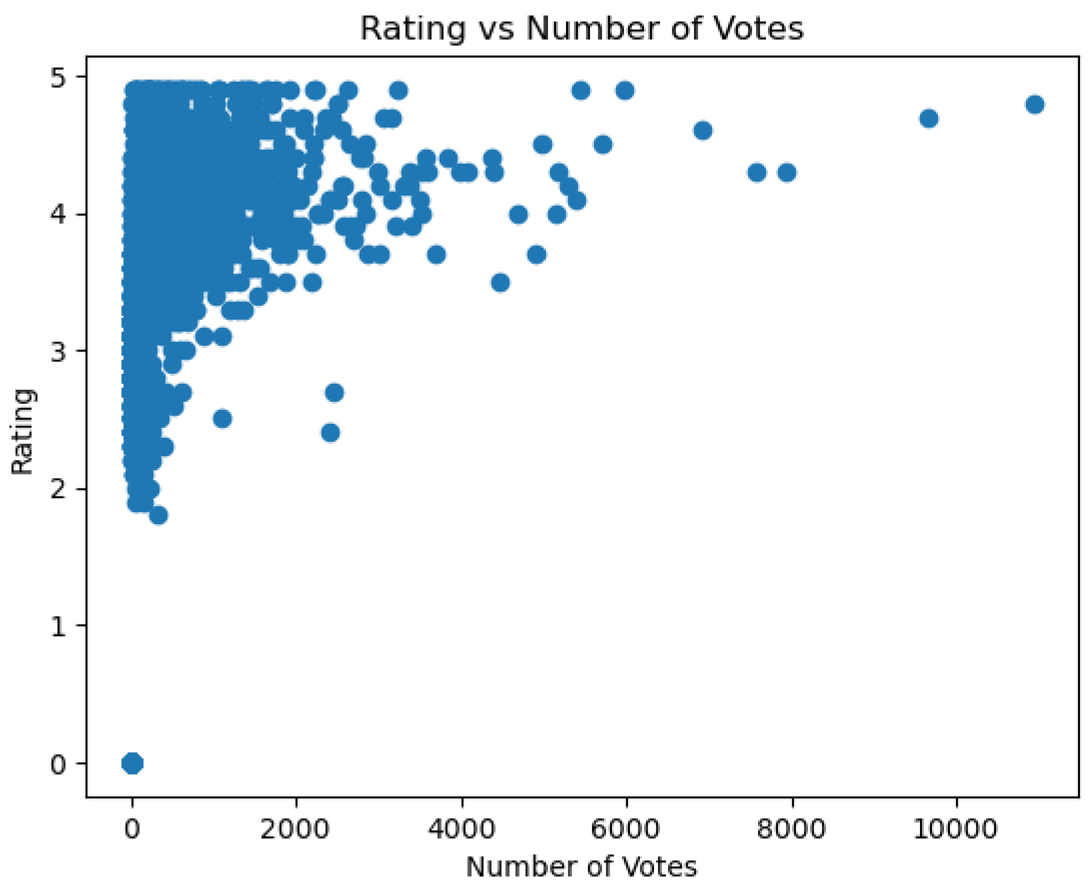
Out[10]:
```
(0.3136905841954111, 4.215149194293825e-217)
```

In [11]:
```python
print("Pearson Correlation:", corr)
```

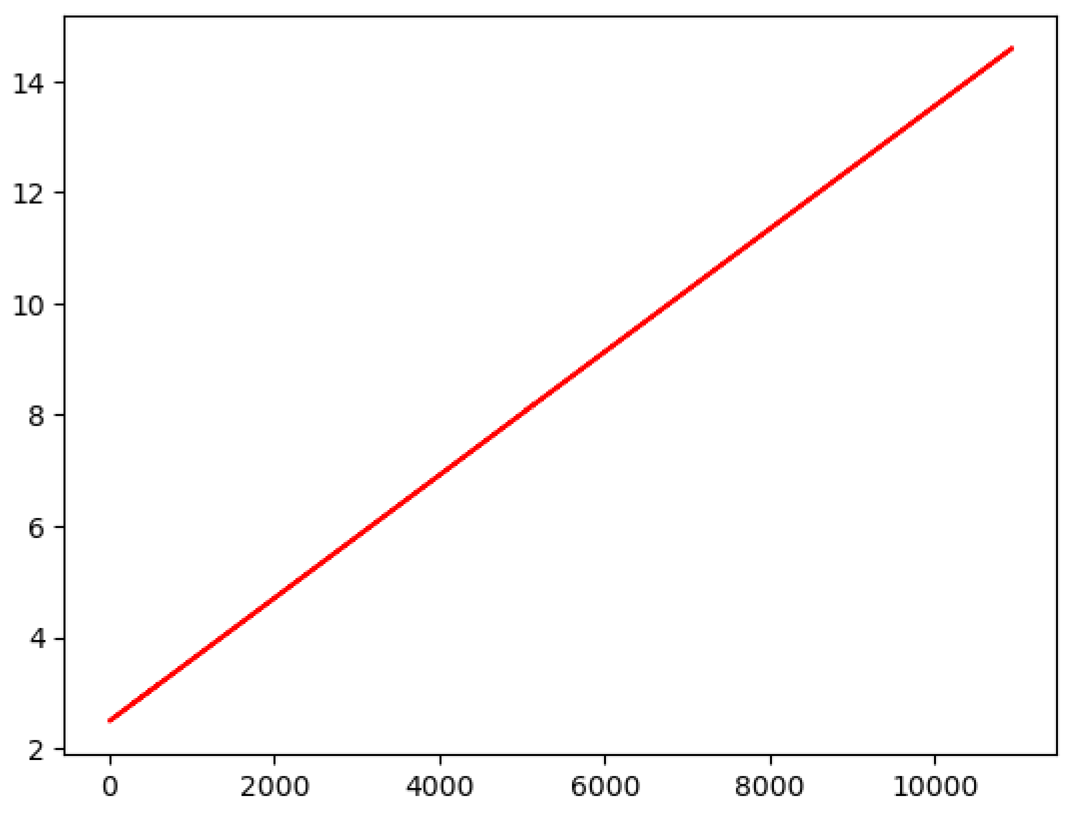
```
Pearson Correlation: 0.3136905841954111
```

In [12]:
```python
plt.scatter(dt['Votes'], dt['Aggregate rating'])
plt.xlabel('Number of Votes')
plt.ylabel('Rating')
plt.title('Rating vs Number of Votes')
```

Out[12]:
```
Text(0.5, 1.0, 'Rating vs Number of Votes')
```



In [13]:
```python
from scipy.stats import linregress
m, b = linregress(dt['Votes'], dt['Aggregate rating'])[:2]
plt.plot(dt['Votes'], m*dt['Votes'] + b, 'r')
plt.show()
```



In [ ]: