Task: Votes Analysis

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

```python
In [1]: import pandas as pd
        dt = pd.read_csv(r"C:\Users\HP\OneDrive\Documents\Cognifyz Internship Program\Dataset.cs
        dt
```

`Out[1]:`

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longi
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...	121.02
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...	121.01
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...	121.05
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...	121.05
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...	121.05
...	...	...	...	...	...	...	...	
9546	5915730	Naml\ Gurme	208	��stanbul	Kemanke�� Karamustafa Pa��a Mahallesi, R\ht\m ...	Karak�_y	Karak�_y, ��stanbul	28.97
9547	5908749	Ceviz A��ac\	208	��stanbul	Ko��uyolu Mahallesi, Muhittin ��st�_nda�� Cadd...	Ko��uyolu	Ko��uyolu, ��stanbul	29.04
9548	5915807	Huqqa	208	��stanbul	Kuru�_e��me Mahallesi, Muallim Naci Caddesi, N...	Kuru�_e��me	Kuru�_e��me, ��stanbul	29.03
9549	5916112	A���k Kahve	208	��stanbul	Kuru�_e��me Mahallesi, Muallim Naci Caddesi, N...	Kuru�_e��me	Kuru�_e��me, ��stanbul	29.03
9550	5927402	Walter's Coffee Roastery	208	��stanbul	Cafea��a Mahallesi, Bademalt\ Sokak, No 21/B, ...	Moda	Moda, ��stanbul	29.02

9551 rows × 21 columns

```
In [2]:  # Calculate the number of votes for each restaurant
         votes = dt.groupby('Restaurant Name')['Aggregate rating'].count()
```

```
In [3]:  votes
```

Loading [MathJax]/extensions/Safe.js

```
Restaurant Name
#45                         1
#Dilliwaala6                1
#InstaFreeze                1
#OFF Campus                 1
#Urban Caf��                 1
                          ..
t Lounge by Dilmah          1
tashas                      1
wagamama                    1
{Niche} - Cafe & Bar        1
�ukura��a Sofras\            1
Name: Aggregate rating, Length: 7446, dtype: int64
```

In [4]:
```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 [5]:
```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()
```
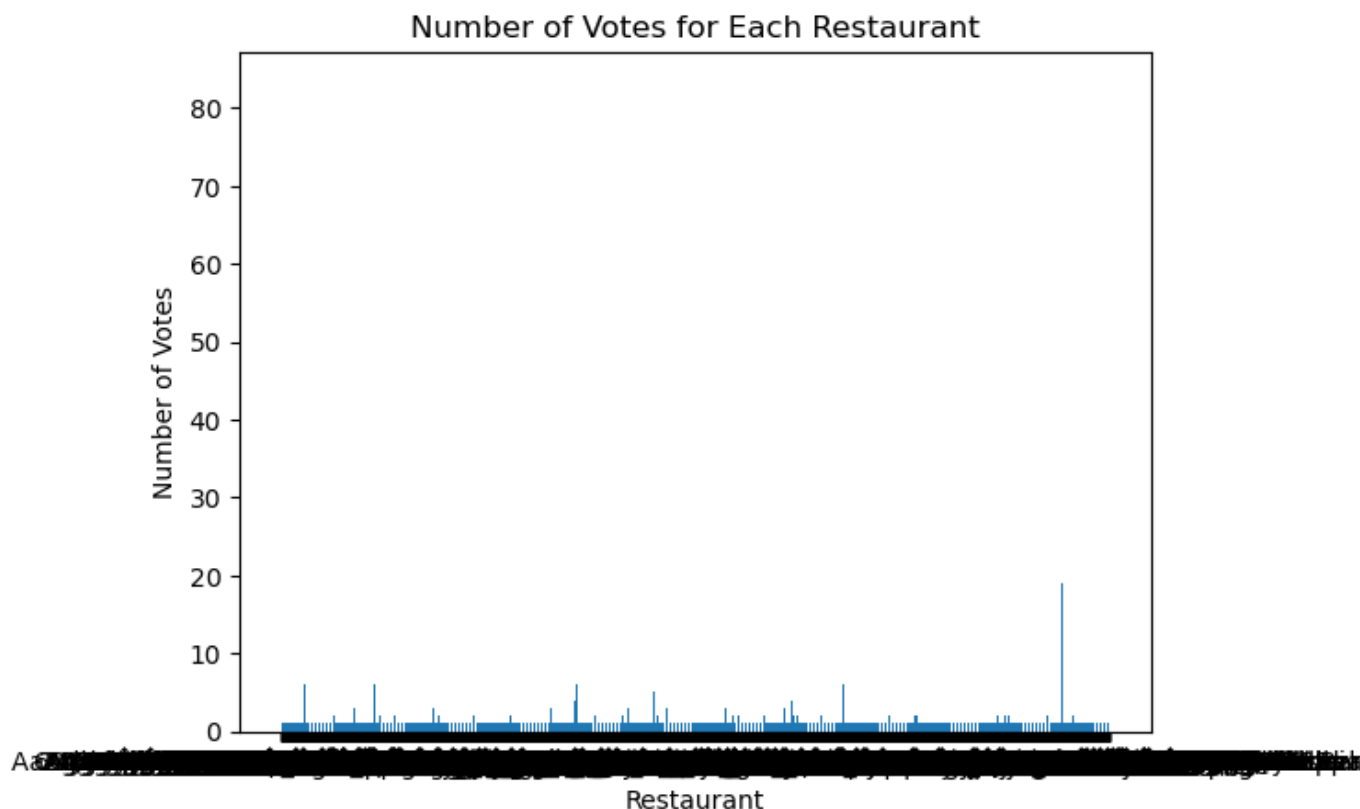

3.2.1 Analyze if there is a correlation between the number of votes and the rating of a restaurant

Loading [MathJax]/extensions/Safe.js

```
In [6]:   from scipy.stats import pearsonr
```

```
In [10]:  # Calculate Pearson correlation
          corr, p = pearsonr(dt['Votes'], dt['Aggregate rating'])
```
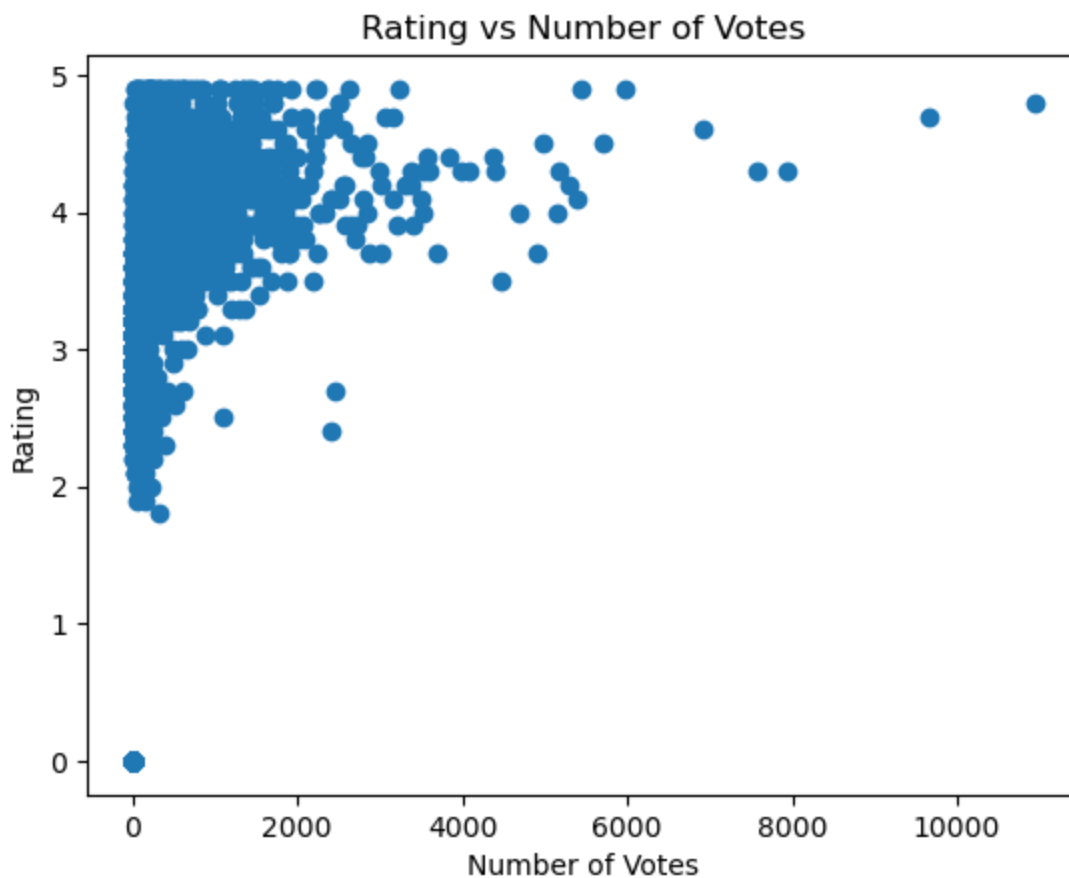
```
In [11]:  corr, p
```

```
Out[11]:  (0.3136905841954114, 4.215149194187134e-217)
```

```
In [12]:  print("Pearson Correlation:", corr)
```
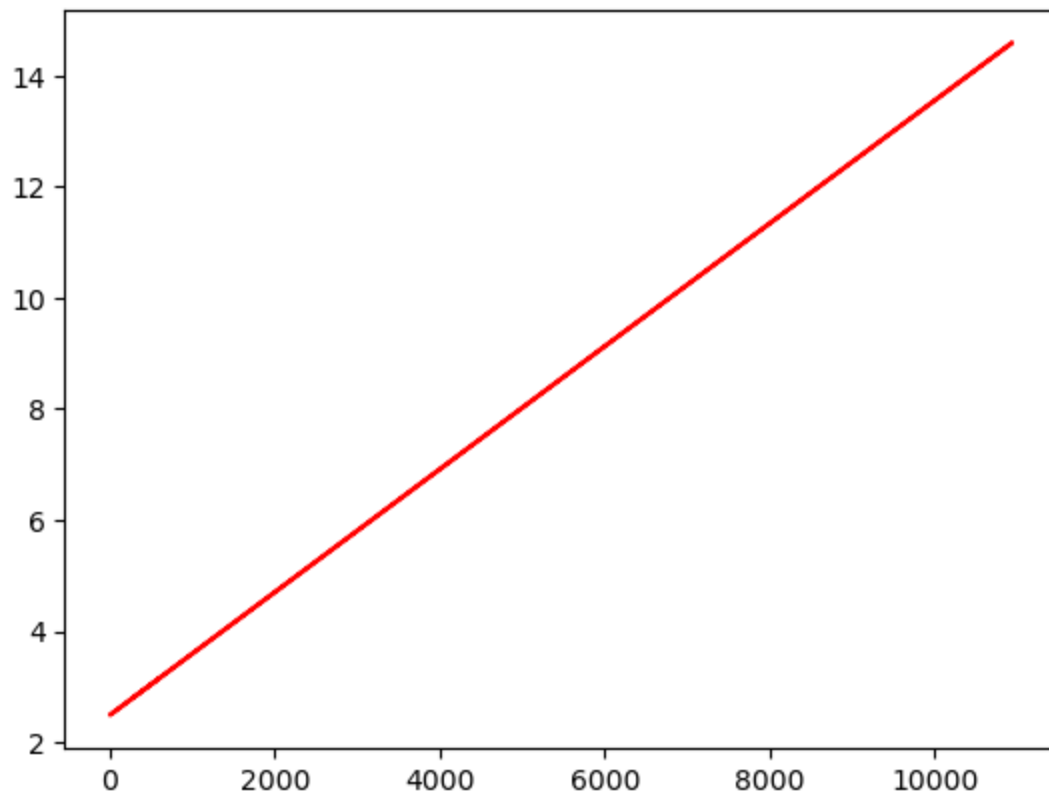
```
          Pearson Correlation: 0.3136905841954114
```

```
In [16]:  plt.scatter(dt['Votes'], dt['Aggregate rating'])
          plt.xlabel('Number of Votes')
          plt.ylabel('Rating')
          plt.title('Rating vs Number of Votes')
```

```
Out[16]:  Text(0.5, 1.0, 'Rating vs Number of Votes')
```



```
In [19]:  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 [ ]: