act_report

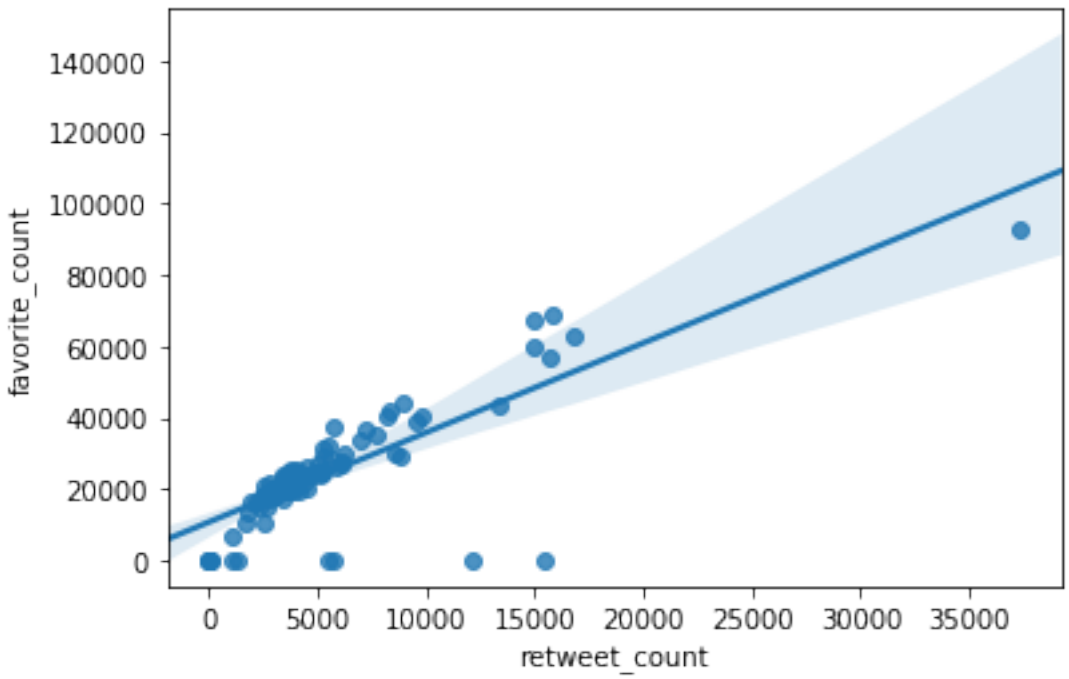
August 2, 2022

1 Project: Wrangling and Analyze Data

1.1 Data Gathering

Analyzing and Visualizing Data

```
In [45]: sb.regplot(data=df, x='retweet_count', y='favorite_count')
         plt.title=('retweet and favorite')
```



1.1.1 Insight:

1.The graph is skewed to the left and this shows there is a positive correlation between favorite count and retweet count because as one variable increases the other one increases too.

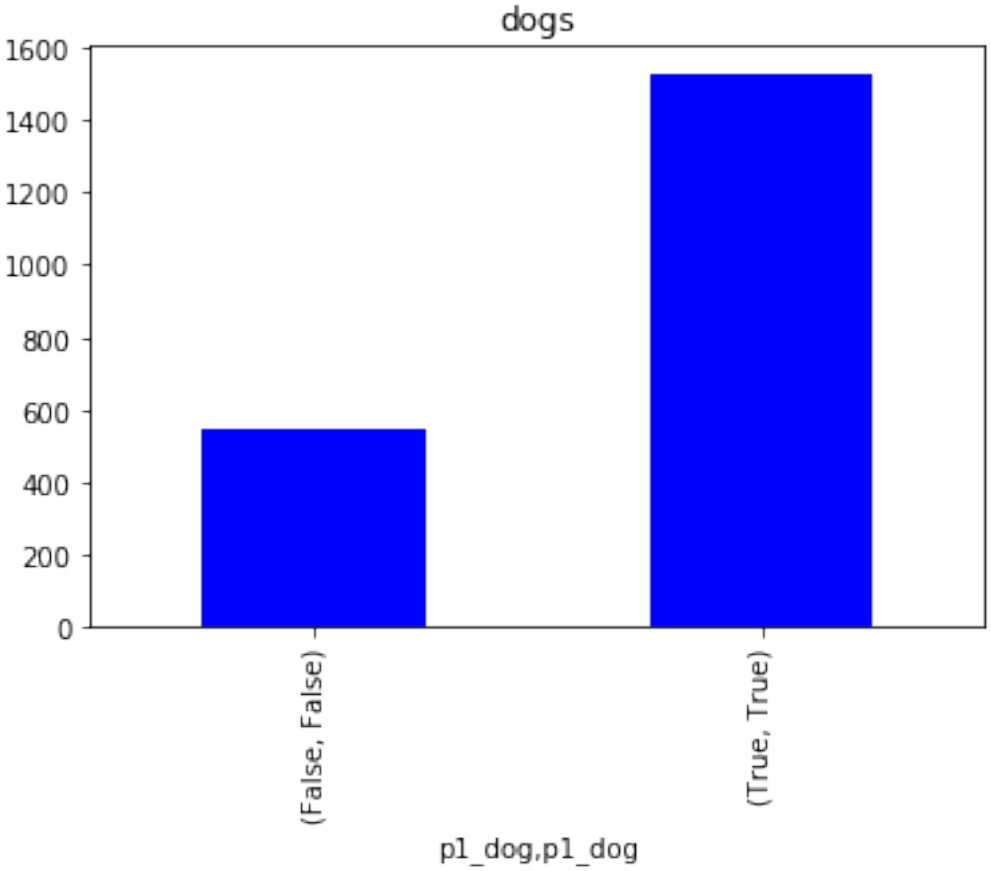
```
In [46]: p1_dog=df_img.p1_dog==True
         p1_dog=df_img.p1_dog==False

In [47]: p1_dog.value_counts()

Out[47]: False    1532
         True      543
         Name: p1_dog, dtype: int64

In [48]: df_img.groupby('p1_dog').p1_dog.value_counts().plot(kind='bar',color='blue',title='dogs

Out[48]: <matplotlib.axes._subplots.AxesSubplot at 0x7fdd7577bc88>
```



This graph showS that in p1_dog have values which are for dogs(true) and have the highest value and others(false) which means these are not dogs.

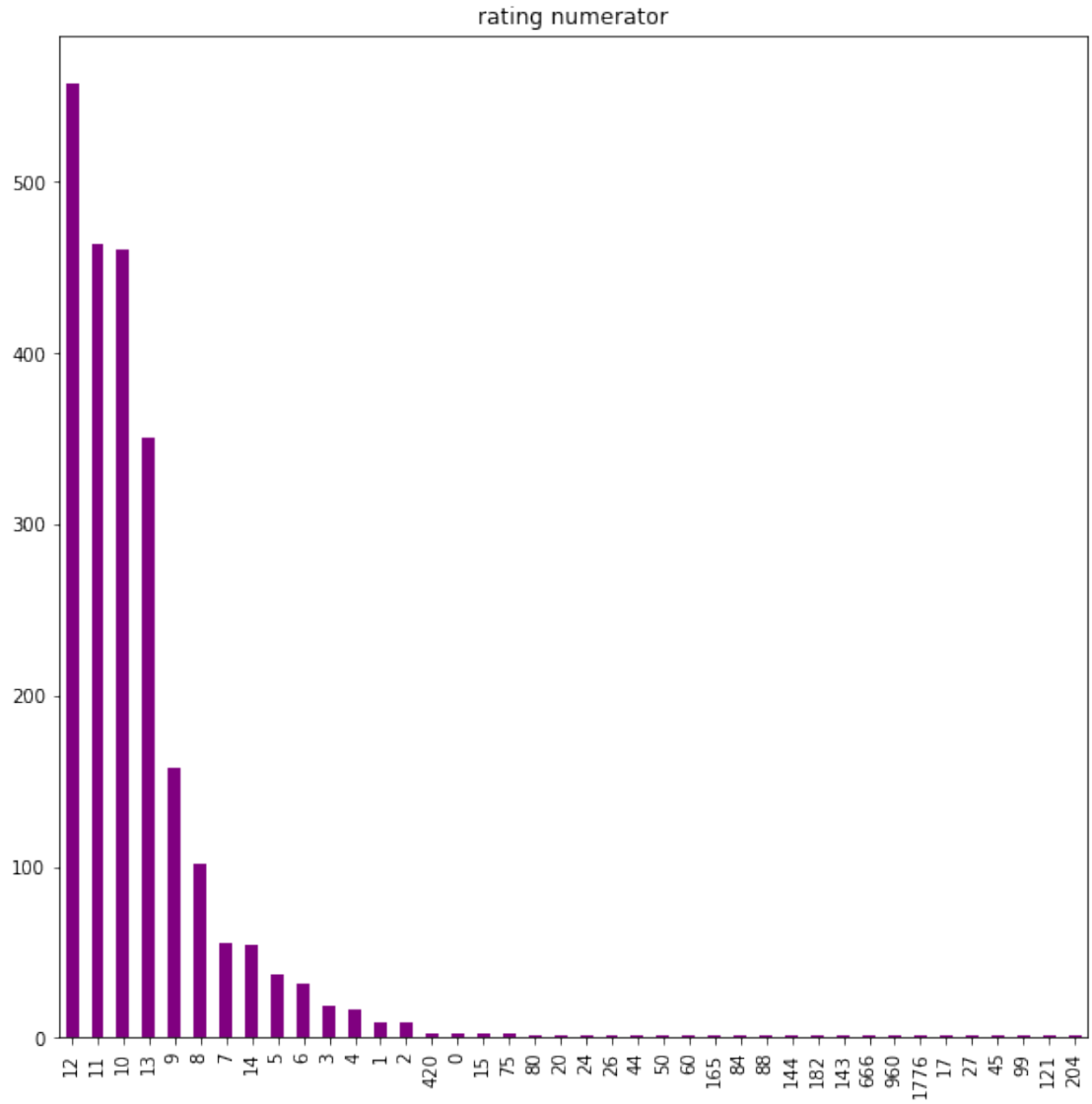
```
In [49]:  df_arc.rating_numerator.value_counts()

Out[49]: 12      558
         11      464
         10      461
```

```
13      351
9       158
8       102
7        55
14       54
5        37
6        32
3        19
4        17
1         9
2         9
420       2
0         2
15        2
75        2
80        1
20        1
24        1
26        1
44        1
50        1
60        1
165       1
84        1
88        1
144       1
182       1
143       1
666       1
960       1
1776      1
17        1
27        1
45        1
99        1
121       1
204       1
Name: rating_numerator, dtype: int64
```

In [50]:
```
plt.figure(figsize=[10,10])
df_arc['rating_numerator'].value_counts().plot(kind='bar', color='purple',title='rating
```

Out[50]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7fdda01dea90>
```



This graph is skewed to the left and this shows that most ratings fall between 10 and 14