

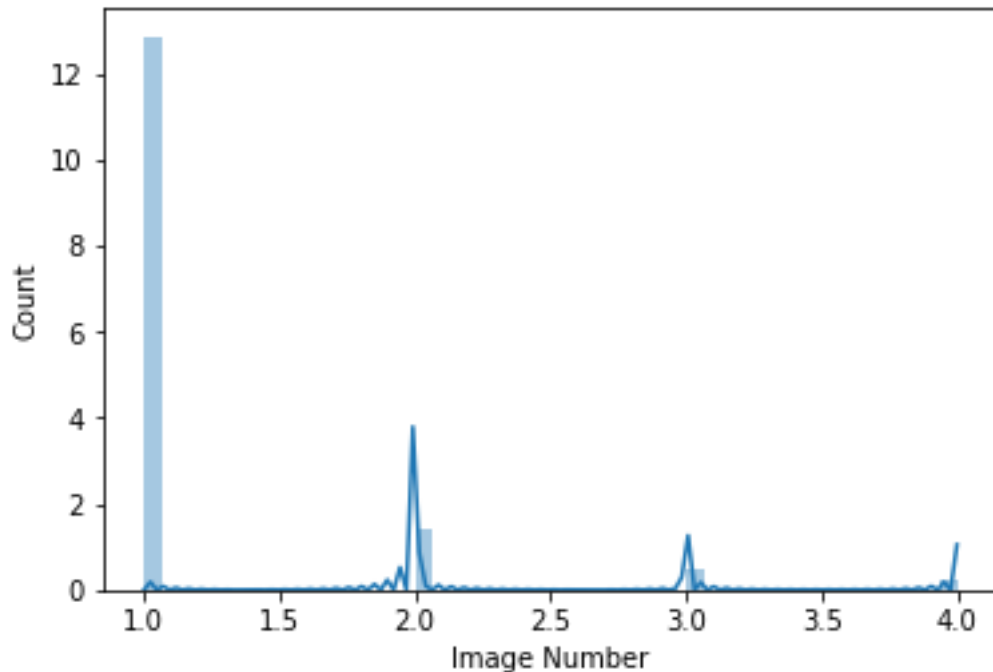
# **ACT REPORT**

## **WE RATE DOGS**

*-Vishnu Kumar V H*

## 1

Since our analysis also involves predicting the images, I felt it is right only if I found out the stats for the image for which we are most confident with.



As we can see the first image had the highest number as confident. There is a gradual decrease in the number of images in each Image number.

## 2

I was interested in going in detail, the number of predictions that were true for each of the predictions p1, p2 and p3. I used the `value_counts()` function for p1\_dog, p2\_dog and p3\_dog to find the number of values which were True and the predictions that were False.

I found out that there were 1532 predictions which were True and 543 predictions which were False for p1\_dog. 1553 predictions were True and 522 predictions were False for p2\_dog. And for p3\_dog, 1499 predictions were True and 576 predictions were found to be False.

The conclusion that I obtained from this insight is most of that more than 66% predictions were True and the rest were False.

```
In [102]: df_predictions_clean.p1_dog.value_counts()
```

```
Out[102]: True      1532  
False      543  
Name: p1_dog, dtype: int64
```

```
In [103]: df_predictions_clean.p2_dog.value_counts()
```

```
Out[103]: True      1553  
False      522  
Name: p2_dog, dtype: int64
```

```
In [104]: df_predictions_clean.p3_dog.value_counts()
```

```
Out[104]: True      1499  
False      576  
Name: p3_dog, dtype: int64
```

```
In [121]: len(remove_dog)
```

```
Out[121]: 324
```

Another important insight is the number of values present in `remove_dog` list. It is the number of rows in which all the predictions are False. I.e. They are not dog tweets.

### 3

Till this point we have seen the insights about predictions. I felt interested in going in detail about the users. Precisely, the way they have downloaded the content from WeRateDogs. I used `value_counts()` function to count the different values and the count of these values.

```
In [122]: df_archive_clean.source.value_counts()
```

```
Out[122]: <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a>      1964  
<a href="http://vine.co" rel="nofollow">Vine - Make a Scene</a>      91  
<a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>      31  
<a href="https://about.twitter.com/products/tweetdeck" rel="nofollow">TweetDeck</a>      11  
Name: source, dtype: int64
```

We can see that most of the users use twitter on their iPhones. The other numbers are very low. Vine App stands second with 91. And TweetDeck stands last with 11 users using it.