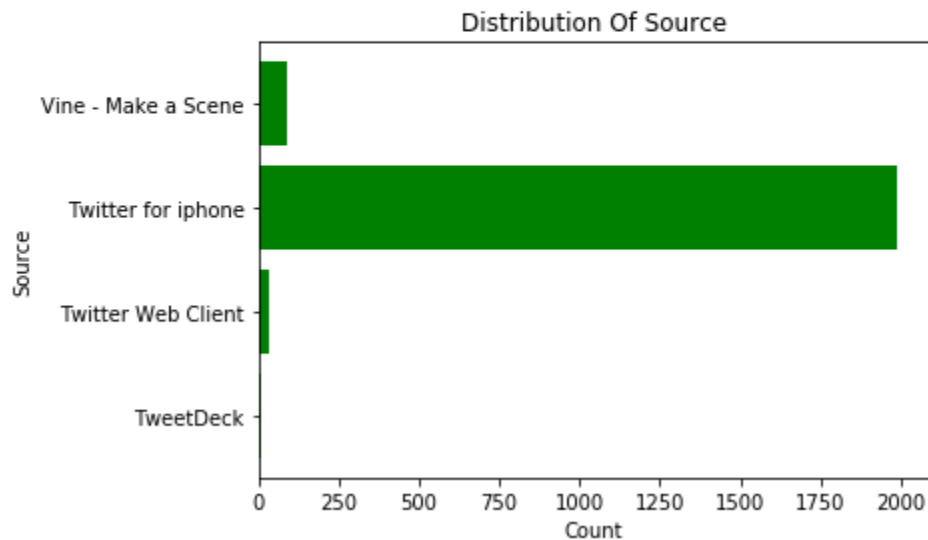


Data Analysis & Visualization

1- Distribution Of Source

- We see That the domionate source of tweets is from iphone twitter app. That means the twitter app is the main channel for people using to tweet ,retweet and others

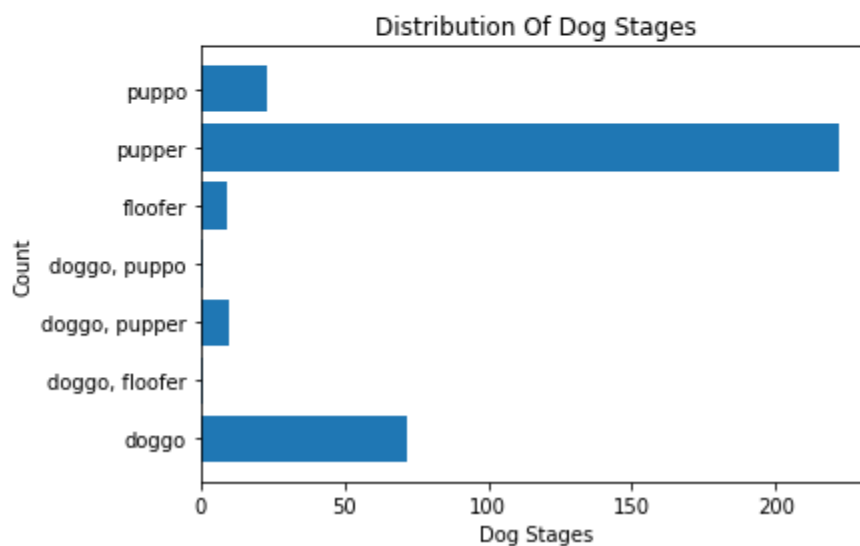
```
Twitter for iphone    1985
Vine - Make a Scene   91
Twitter Web Client    30
TweetDeck             11
Name: source, dtype: int64
```



2- Distribution Of Dog Stages

- This Shows that 'pupper' is the most popular dog stage, followed by 'doggo' , 'puppo' and 'floofer'. It should also be noticed that there's huge amount missing data in dog stages, thus the distribution may not reflect the truth.

```
pupper      222
doggo       72
puppo       23
doggo, pupper    10
floofer      9
doggo, puppo     1
doggo, floofer   1
Name: stage, dtype: int64
```



3- Relationship Between Retweet count & Favorite count

- I test the correlation between retweet_count and favorite_count and it is a high value showing a strong positive relationship between both.

	retweet_count	favorite_count
retweet_count	1.00000	0.91114
favorite_count	0.91114	1.00000

