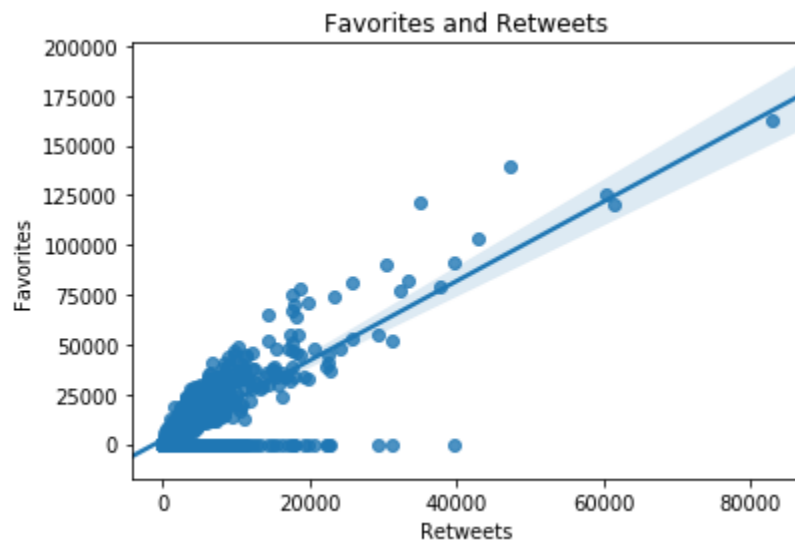WeRateDogs Data Wrangling Project

Reem Tarek

After gathering data from different sources, assessing and cleaning it, I've attempted visualizing some of the parameters in the dataset based on some questions. And it follows as shown below:

Question 1:

Is there a correlation between retweets and favorites counts?

Answer:

It looks like both are highly correlated with a value of almost 0.8
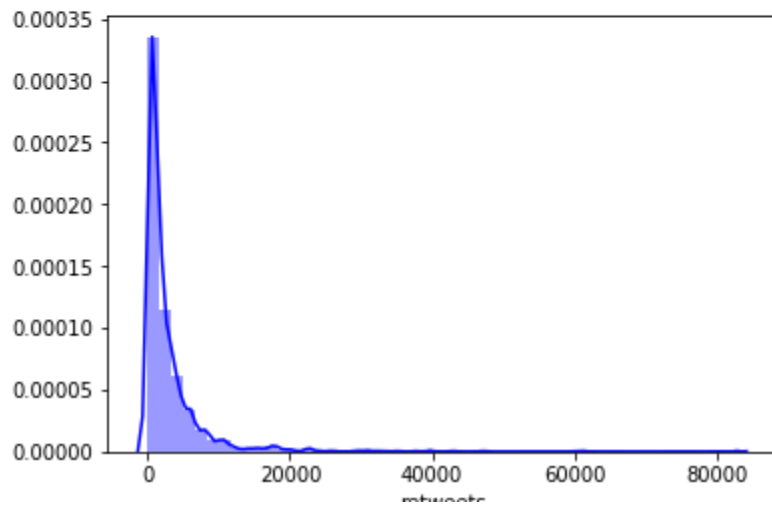
Question 2:

How are both of retweets and favorites distributed?

Answer:
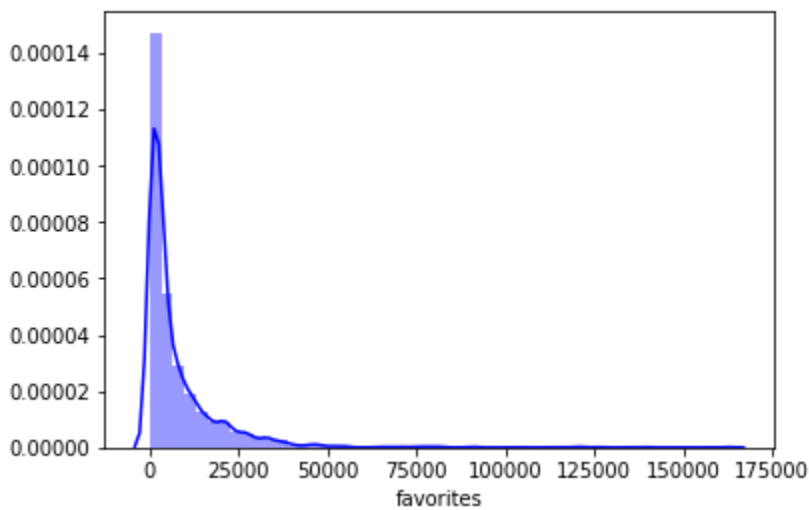
Retweets:

```
count      2339.000000
mean       2899.461736
std        4889.049948
min           0.000000
25%         583.000000
50%        1353.000000
75%        3379.500000
max       82868.000000
Name: retweets, dtype: float64
```

<u>Favorites:</u>

```
count        2339.000000
mean         7888.054724
std         12224.649732
min             0.000000
25%          1366.500000
50%          3431.000000
75%          9658.000000
max        162835.000000
Name: favorites, dtype: float64
```


Question 3:

What is the most and least tweeted dog stage?

Answer:

pupper has been tweeted 138 times, while puppo has been tweeted 20 times.