# Act_report

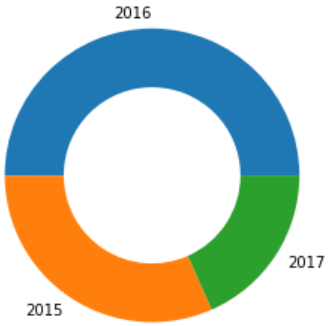
So, after gathering, assessing and cleaning the datasets final result is saved as ***twitter_archive_master.csv.***

Although, the dataset is quite small in size and also some data is not available for column like Dog name, still we can get some good insights about the tweet trends.

Let's discuss some of the observations:

1. **As per available Data, We have Maximum number of tweets posted in 2016**

We have ***timestamp*** column from which can extract year and the applied with *value_counts()* function to see the trends of tweets in year wise fashion. After, drawing a donut chart we can see maximum number of tweets is from 2016, and then followed by 2015 and 2017 respectively. The same is depicted below.

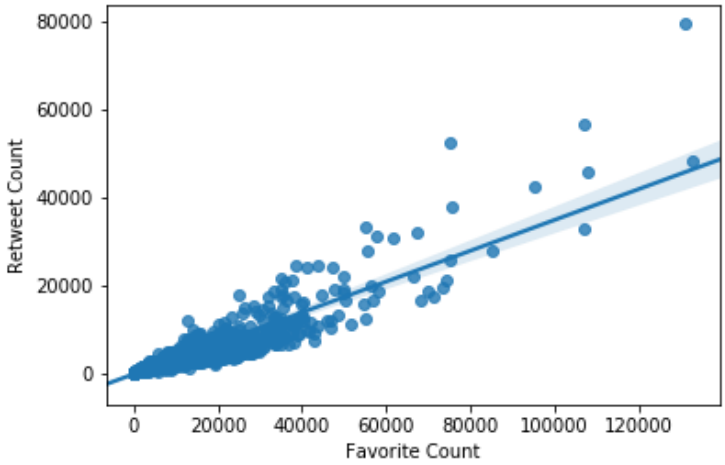



2. **Top 5 most rated Dogs.**

To see which are the most rated dogs by WeRateDogs in their tweets I select top 5 of them (*df_archive_clean['rating'].nlargest(5))* and a dog named as *Atticus* is highest rated one. Here comes the Winner!!!

### 3. Favorite Count V/s Retweet Count trend visualization

We have two good metrics in our data favorite count and retweet count. To see the relation between these two variables, a scatterplot would be best choice. A scatterplot of favourite count and re-tweet count shows a linear relationship between the two.




### 4. Source Distribution of the Tweets

The dataset contains information about the source of the tweet. Some of the tweets are posted as in form of a small video on Vine. A bar chart plot of the data shows most of the tweets are sent through Twitter's iphone app and next one in the list is video sent through Vine.

### 5. Top 10 most tweeted Dogs breeds

One more curious case is to check Dogs breeds amongst the several tweets. A bar chart plot of 10 most tweeted dogs shows the statistics clearly. ***Golden_retriever*** is most tweeted among them.