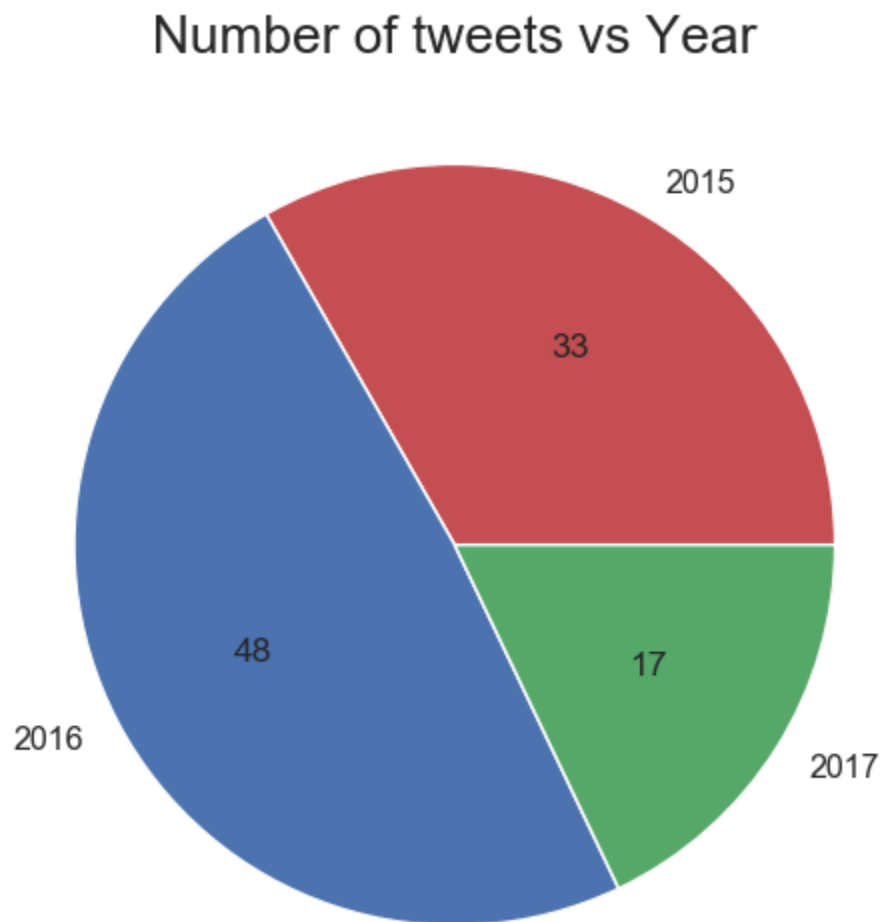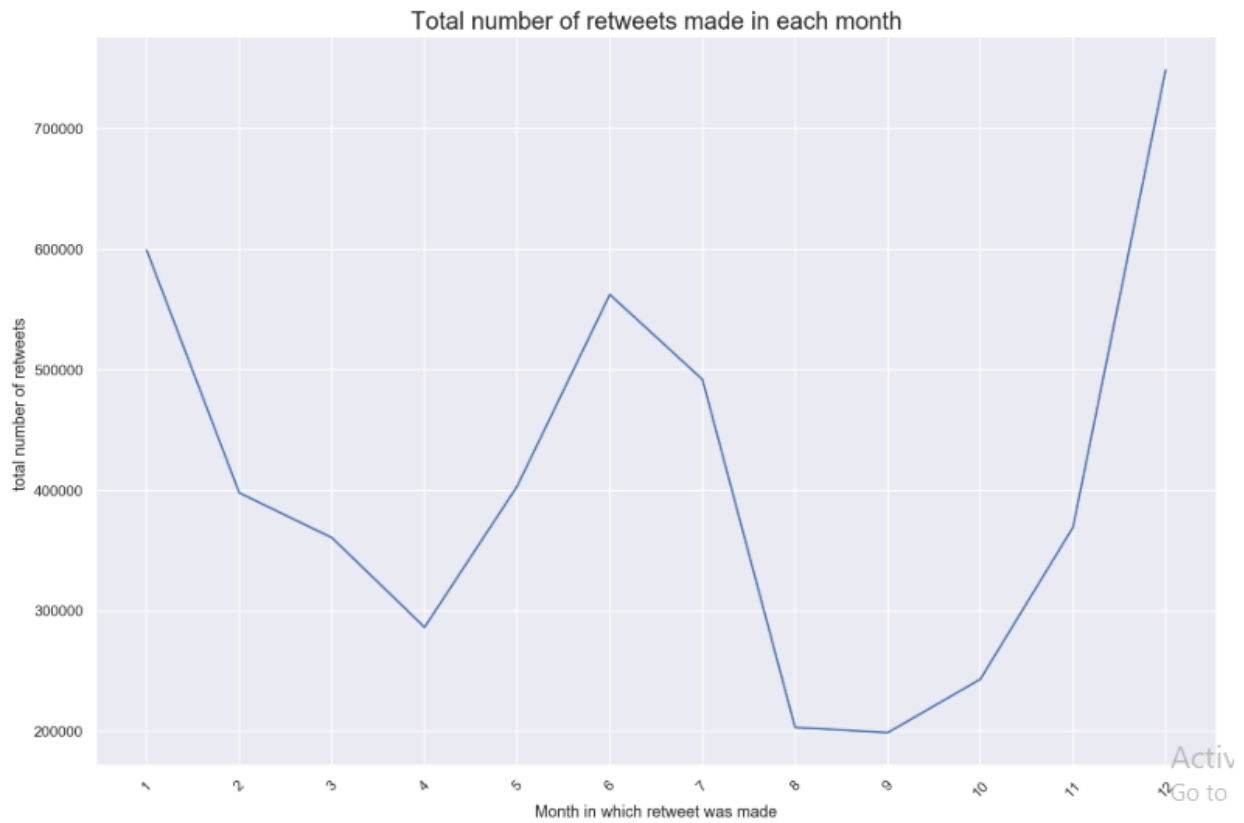# ANALYSIS AND INSIGHTS OF WERATEDOGS TWITTER ARCHIVE

The purpose of this project is to wrangle and analyse the dataset which is gathered from csv file , tsv file which is hosted on udacity server and must be downloaded programmatically and the entire set of json data from WeRateDogs Twitter archive from twitter API using Python's Tweepy library which needs to be assessed ,cleaned and analyzed.

NUMBER OF TWEETS MADE BY USERS IN EACH YEAR
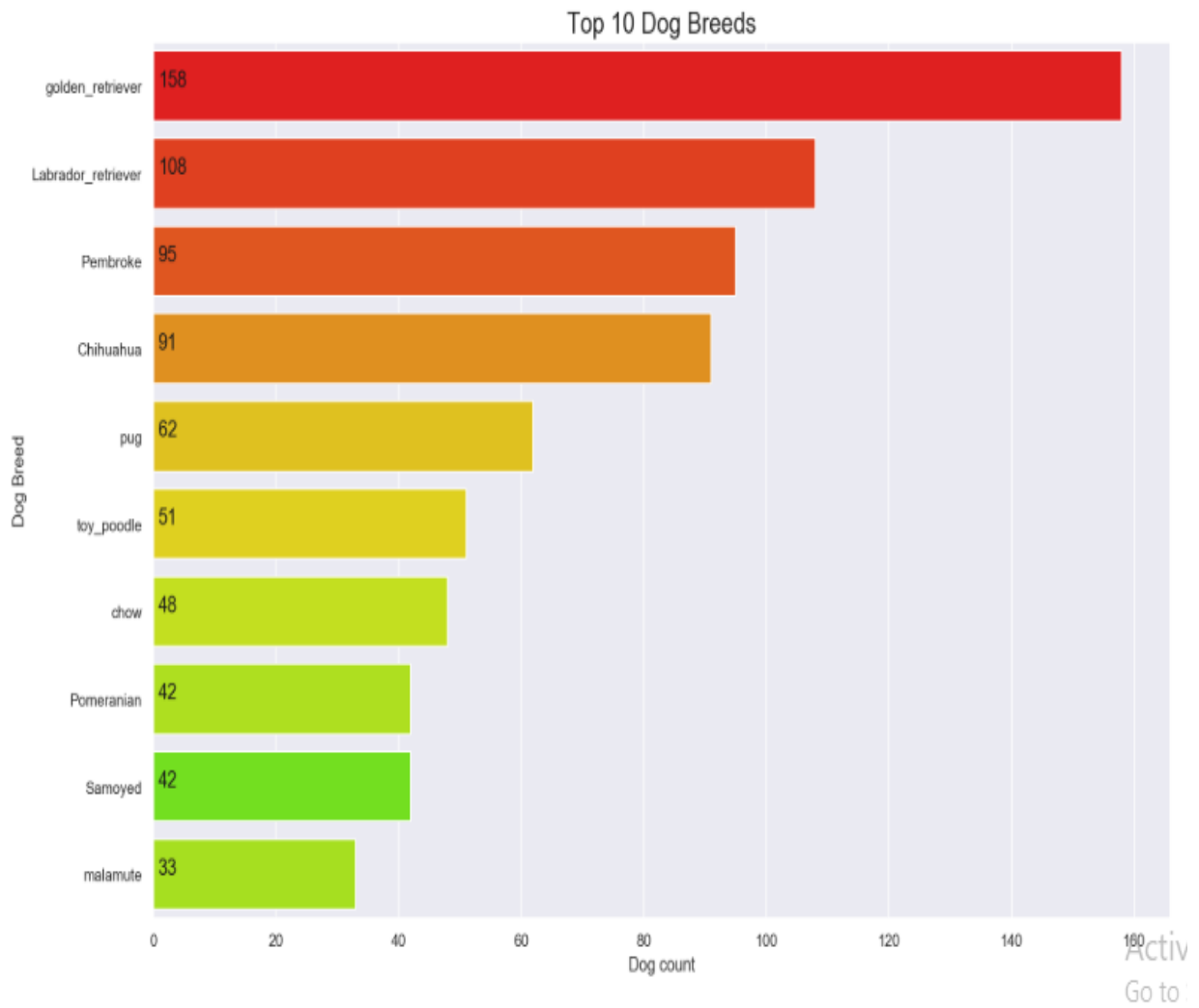
Number of tweets vs Year

We see that 48% of the tweets are made in the year 2016,while 33% and 17% of the tweets in the year 2015 and 2017.

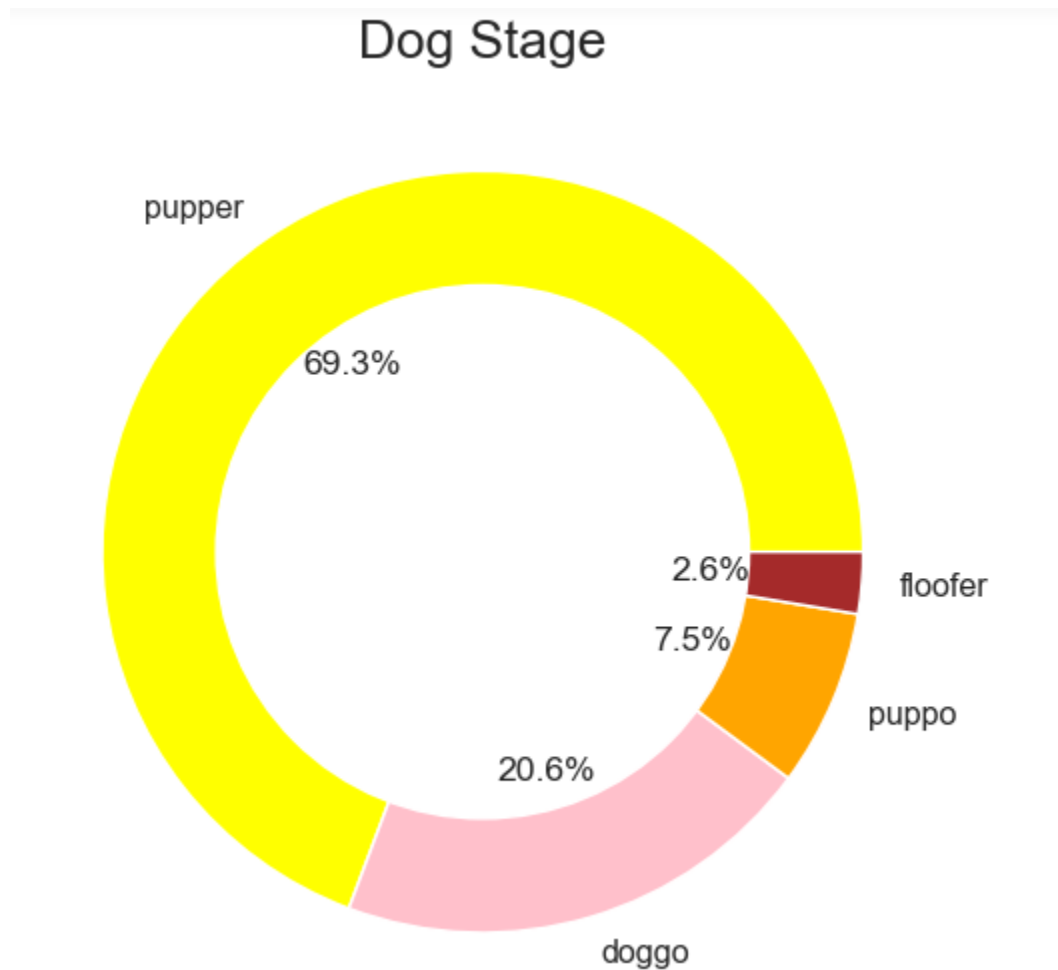# NUMBER OF RETWEETS MADE BY USERS IN EACH MONTH

We see that most number of retweets were made in the month of December while least number of retweets were made in the month of September.

TOP 10 DOG BREEDS

Top 10 Dog Breeds

Dog Breed	Dog count
golden_retriever	158
Labrador_retriever	108
Pembroke	95
Chihuahua	91
pug	62
toy_poodle	51
chow	48
Pomeranian	42
Samoyed	42
malamute	33

We see that golden retriever is the most common dog breed in this dataset.

VARIOUS STAGES OF DOGS

Dog Stage

pupper
69.3%

2.6%   floofer

7.5%

puppo

20.6%

doggo

We see that most of the dogs are puppers i.e. 69.3% while floofers are least i.e. 2.6% in the dataset.

HIGHEST FAVORITE COUNT DOG

We see that the highest favorite count dog is Labrador retriever which is same as our prediction with 82% confidence.

HIGHEST CONFIDENCE DOG

We see that the highest confidence dog is komondor which is same as our prediction with 99.99% confidence.