Analysing and visualising data

The insights and visualisations which were made were from twitter_archive_master data set. A scatter plot of retweet count and favorite count was made. There was a positive correlation between favorite_count and retweet_count, the more a tweet gets retweeted the more it reaches a karge audience and gets more favorites. For tweets to go viral it would have been entertaining, or thought provoking.

Most of the tweets were coming from Twitter for iPhone, followed by tweets from Twitter Web Client, then by TweetDeck. The number of tweets from Twitter Web Client and TweetDeck were small, and they were less than tweets from Twitter for iPhone even when combined.

Although most of the dogs did not have the stage specified, for those specified most of the dogs were at pupper stage, followed by doggo, then puppo, and lastly by floofer.

The most favorited tweet was liked 132810 times, it was the only tweet which was liked more than 100000 times. The tweet which has less favorite was liked 81 times. The median of favorite tweets was 4110, 50% of the favorite tweets was lower than the average of the favorite tweets of 8907.66, this is the value which cuts data in half. The total number of original tweets was 1961. Standard deviation (from the mean) was 12238. The 25th percentile is 1971, this is the number at which 25% of the tweets lie below. The 75th percentile is 11363, 75% of the tweets lie above this value.

There were no sensitive tweets, tweets were safe for consumption, hopefully without triggering negative emotions from the audience.