

# Analysis and Insights Into Final Data

By Mariam Alabkari

## Introduction

The aim of this project is to collect dataset from Twitter by using Twitter API. The Twitter API will help to collect the dogs images form 'We Rate Dogs' Twitter account. This account provides many dogs images with people rates.

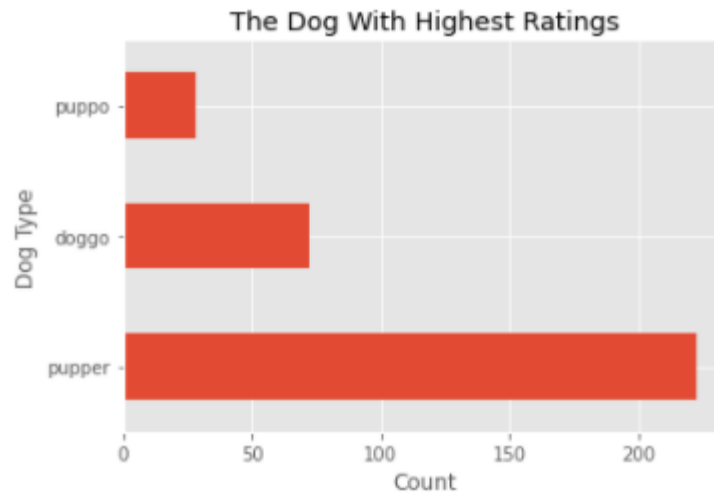
The dataset collected contains some inaccurate data, such as the url images since it contains some non-dogs images. Also, The rating must be out of ten, but people rated the dogs with higher numbers, or rated more than one dog.

After the cleaning process, I tried to answer some of the most common answered questions. Which include:

- Which type of dog has a higher number of ratings?
- Which dog stage is the favorite?
- What are the top 10 retweets count?
- What are the top 10 frequent dog names?
- What is the relationship between the retweets and counts?

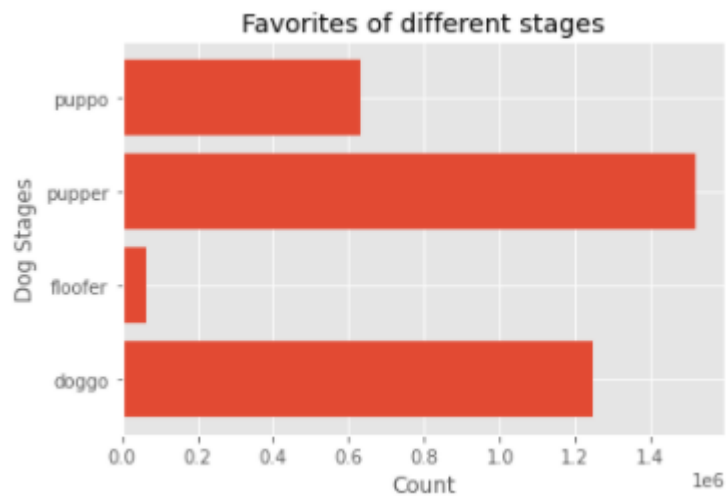
## Analysis and Insights

- ✦ Which type of dog has a higher number of ratings?



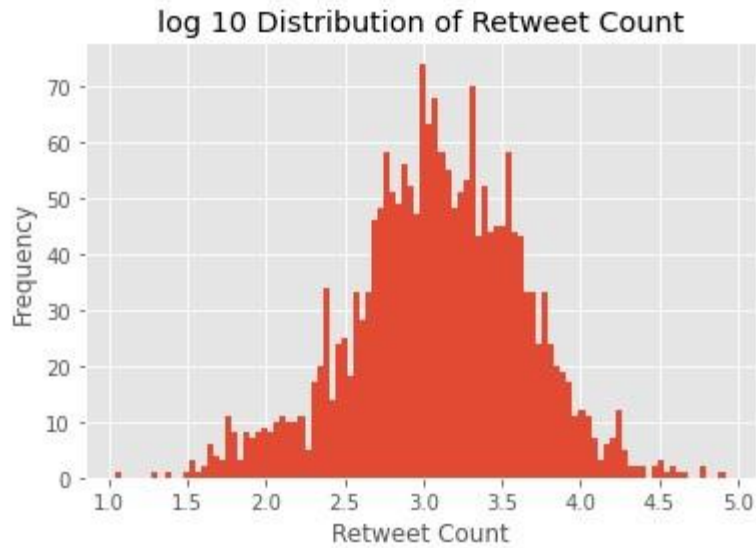
The above figure shows the numbers of each dog. It is clear that pupper dogs got a higher amount of ratings than doggo dogs.

✦ Which dog stage is the favorite?



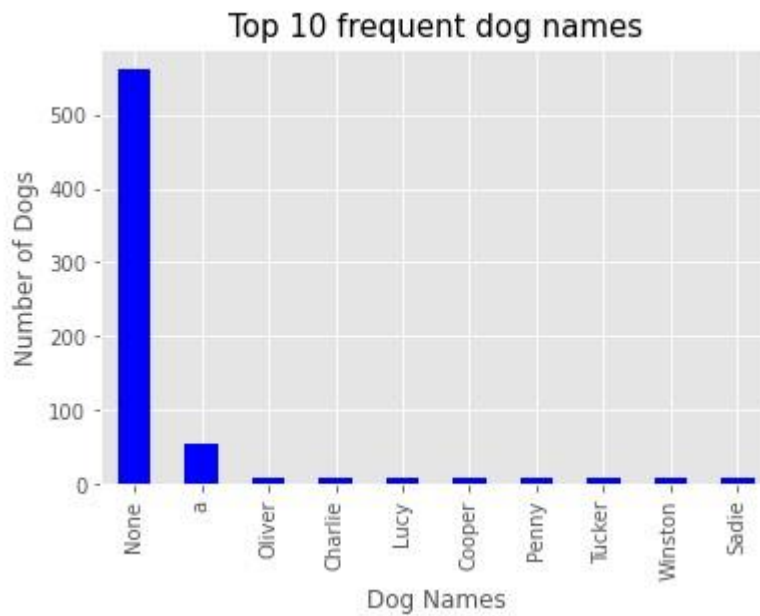
From the figure it is clear that pupper dogs are more popular among the other types.

- What are the top 10 retweets count?



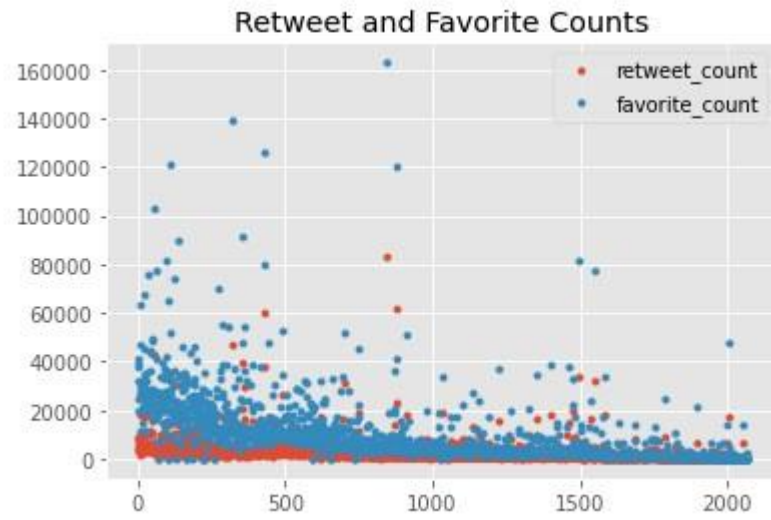
From the histogram it is clear that the retweets count has a normal distribution.

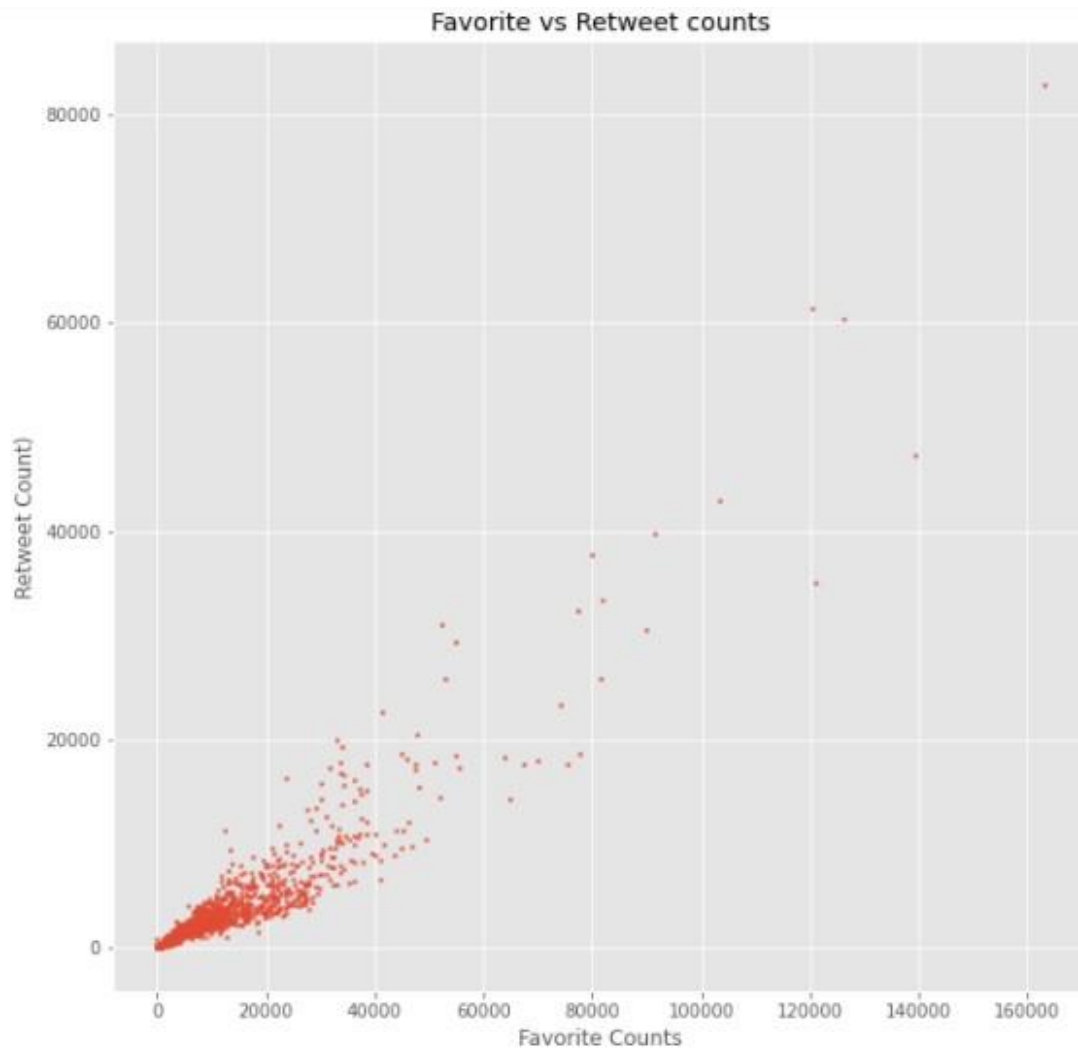
✦ What are the top 10 frequent dog names?



From the bar chart it is clear that there are many none names. And the total number of the names are almost colse.

✦ What is the relationship between the retweets and counts?





From the scatter plots it is clear that the tweets got more favorites than retweets.