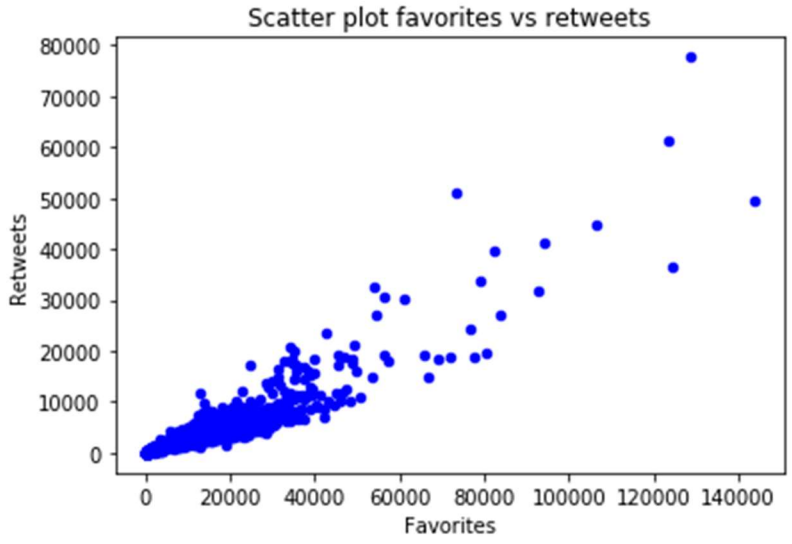
WeRateDogs - Data Analysis Report

Introduction :

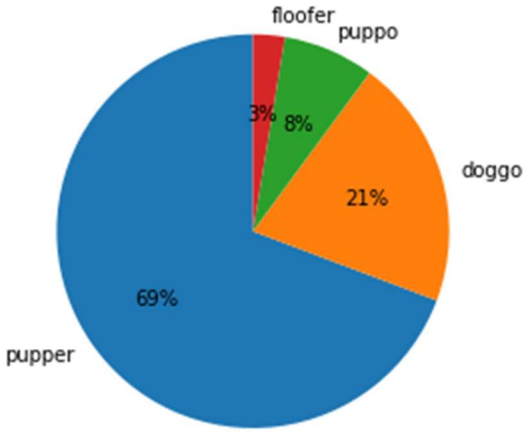
This report has been created by analysing the consolidated Dog Rates Twitter data obtained from three.different sources. Analysis was performed to understand several parameters like whether the tweets with high like counts are retweeted more, also which type of Dog Breed has been more popular and also ingeneral Dog_stage that has been be so common among the people.

Analysis 1 : Study of Relation between Favorite count and Retweet Count



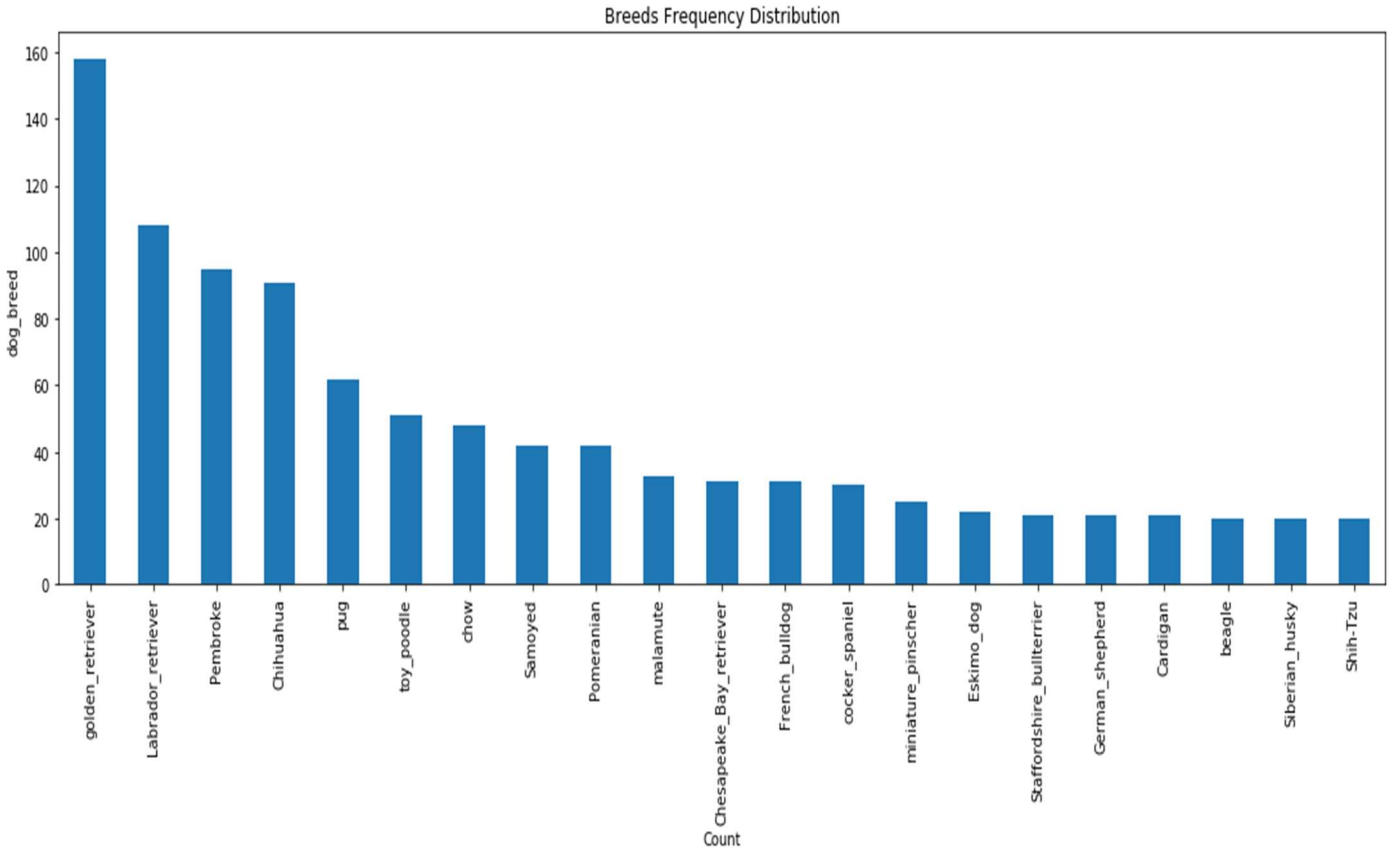
First we started to analyse if retweet count has correlation with the Favorite count. Based on the above plot, we could clearly infer that there is a positive correlation between them. This is something which is not surprising because, only if the given tweet is liked by a person, he/she would retweet the post. Mathematically it has very high correlation value of 0.91 which establishes a strong relationship between both the variables.

Analysis 2 : Dog_Stage_Type Distribution



Our next analysis was to understand under the distribution of Dog_stage from the given data. Surprisingly pupper dog_stage category had the highest share percentage of 69 %, followed by Doggo - 21%. **For this plot, we did not consider the dogs for whom stage was not known**.

Analysis 3 : Study of Dog Breed Histogram plot



As part of our final analysis, we tried to understand which dog_breed has been popular and over all we had around 110 types of Dog breeds. However for our plot we considered only the Dog_breeds which had minimum sum count of 20 records. That is for plotting, We ignored the Dog_breeds which had count of less than 20 . We could infer that dog breeds Golden Retriever and Labrador Retriever are more popular.

Conclusion:

To conclude , as part of our analysis we were able to clearly observe that Golden Retreiver Dog breed and Pupper dog stage have been popular - based on the given tweet data. However we can investigate other valuable information from this tweet dataset like how the tweets have varied over the time, distribution of Dog's rating based on the Dog's breed etc.