<u>**Analysing and Visualizing WeRateDogs - Data Analytics Project.**</u>

Introduction

WeRateDogs, which is also known as @dog_rates(twitter account), is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators however are almost always greater than 10 (11/10, 12/10, 13/10, etc.). With more than 4million followers, it has received international media coverage and has been recognised all over the world.

So, which dog type (breed) is most common dog in the tweet dataset? Which dog is the most rated and with the highest total retweet count? These questions and more are answered in the following insights:

Project Goals:

The goals for this project are listed as follows.

1. Using different data sources provided by Udacity, gather the relevant data from the provided sources.
2. Assess the data by outlining all the data qualities and tidiness issues in the data.
3. Cleaning the data by removing duplicates, unknown or incomplete records, removing outliers, fill missing data where possible etc.
4. Store the cleaned data.
5. Analysis and visualization from the cleaned data by creating at least 3 insights.
6. Report writing on the findings from the data.

Basic Insights:

1. Which dog is the most rated and with the highest total retweet count?
   Charlie is the most rated dog with 10 count of ratings and an average rating of 1.17 and total retweet count of 27,317.
2. Which stage of dog is the most present in the data?
   Pupper is the largest dog stage in the data with about 140.

Value Counts of Dog Class
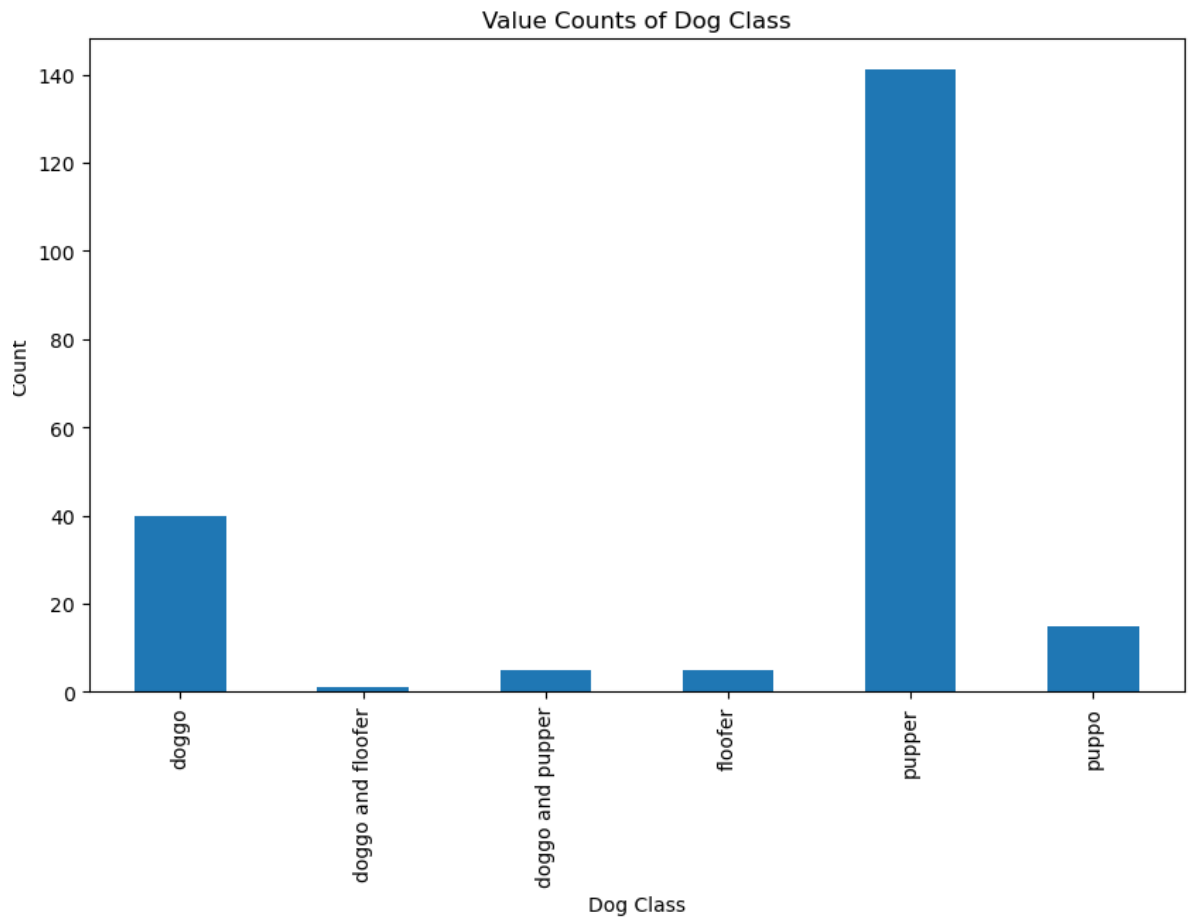
3. Can we conclude there is a relationship between retweet and most rated?

From the chart below, we can see that there is no strong relationship between retweet count and most rated dog. For example, dogs with 1.4 rating the highest in the group have lower favourite and retweet count.

Retweet and Favorite Count by Rating