**WeRateDogs – Twitter Data**

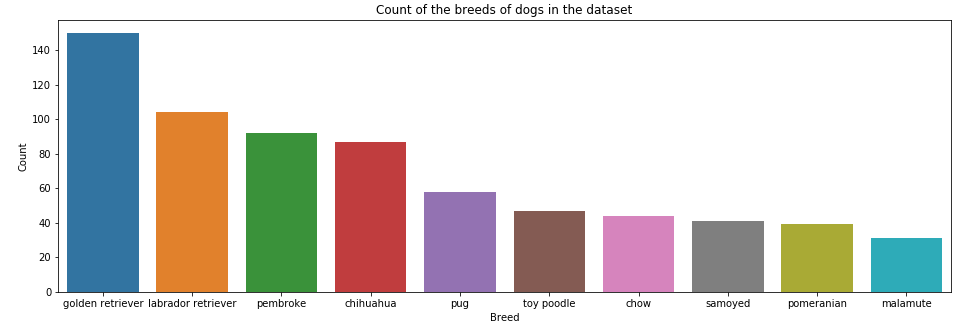
**Data Analysis**

**Questions:**

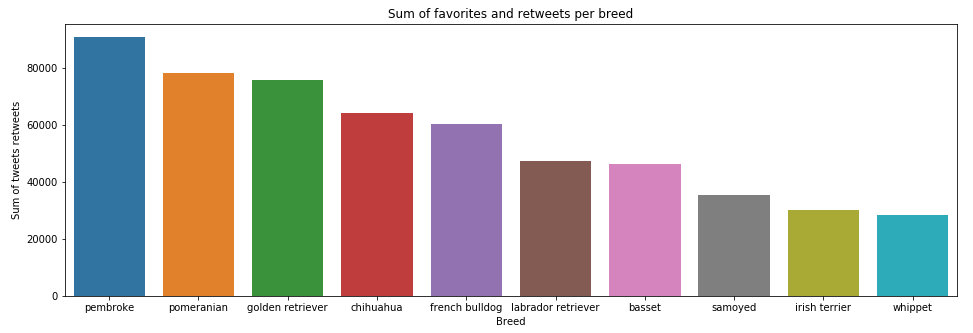
**1.**Based on the predicted, most likely uploaded dog breed: Which breed gets retweeted and favorited the most overall?  
**2.**What is the frequency of tweets being done by the users?  
**3.**What is the most predicted dog breed according to each of the algorithms?

**Based on the predicted, most likely uploaded dog breed: Which breed gets retweeted and favorited the most overall?**

For this we will look at the breed column in our cleaned dataframe and find the most likely dog breed



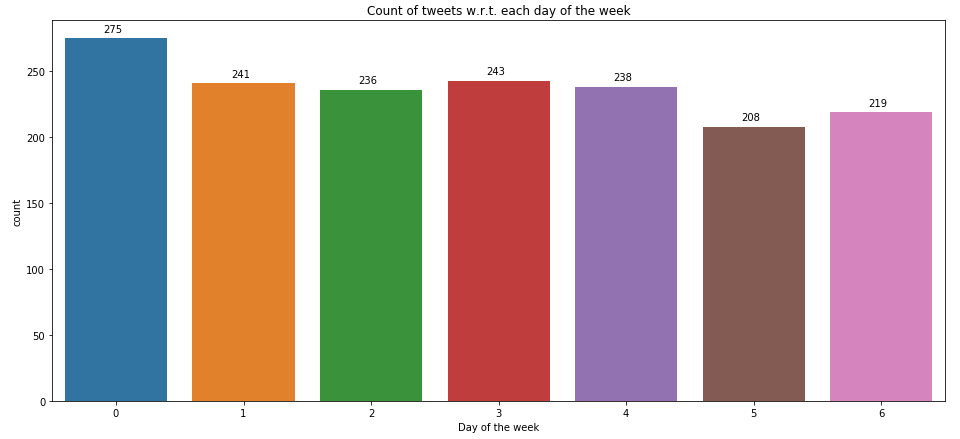
We find here that the most common breed among the ones tweeted is Golden Retriever



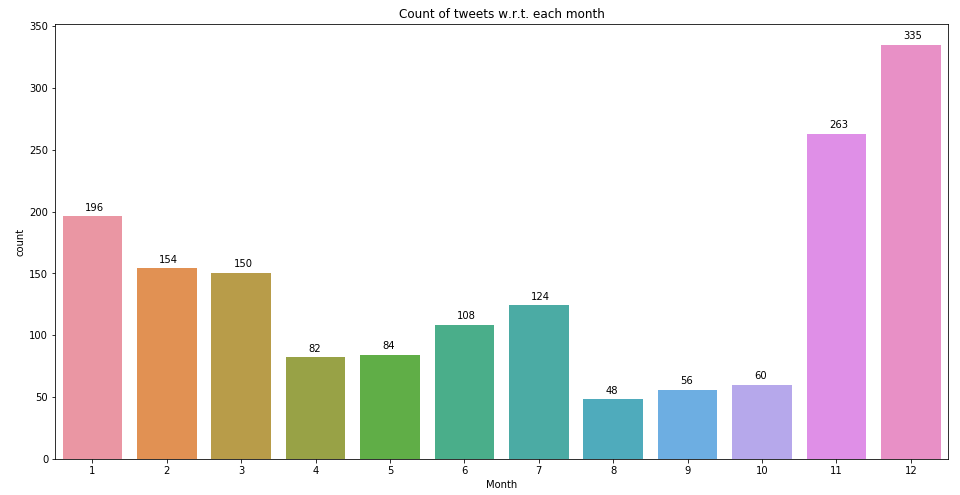
We find with this graph that the most tweeted dog breed is the Pembroke

**What is the frequency of tweets being done by the users?**

**For this we extracted the day of the week for each tweet being done as well as the month of each tweet using the lambda function**

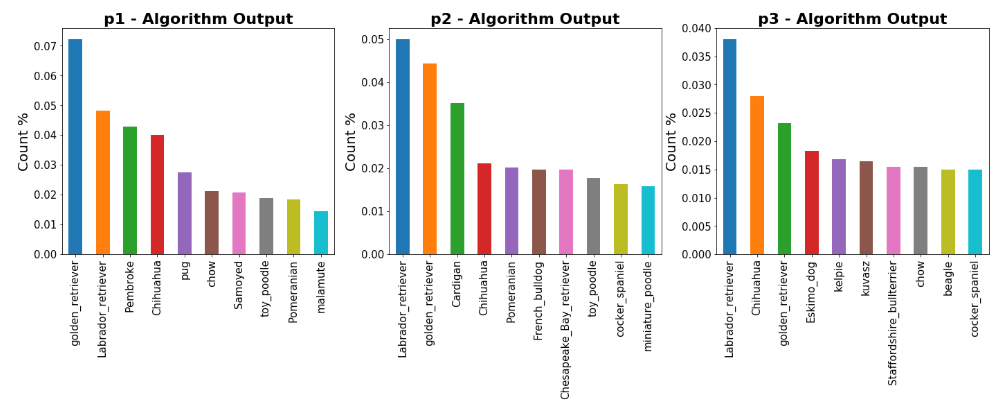
****

Monday was the day when most tweets were done on the page.



The maximum tweet frequency was in the month of December.

**What is the most predicted dog breed according to each of the algorithms?**



We can see here that the most predicted dog breed is mainly the Golder Retriever or the Labrador Retriever.

From all the analysis done above, we can say that most people that tweet own Retrievers and most of the tweet data is about them. Also many times the users do not give the correct ratings as per the 10 point format and it takes a lot of time in cleaning such type of data.