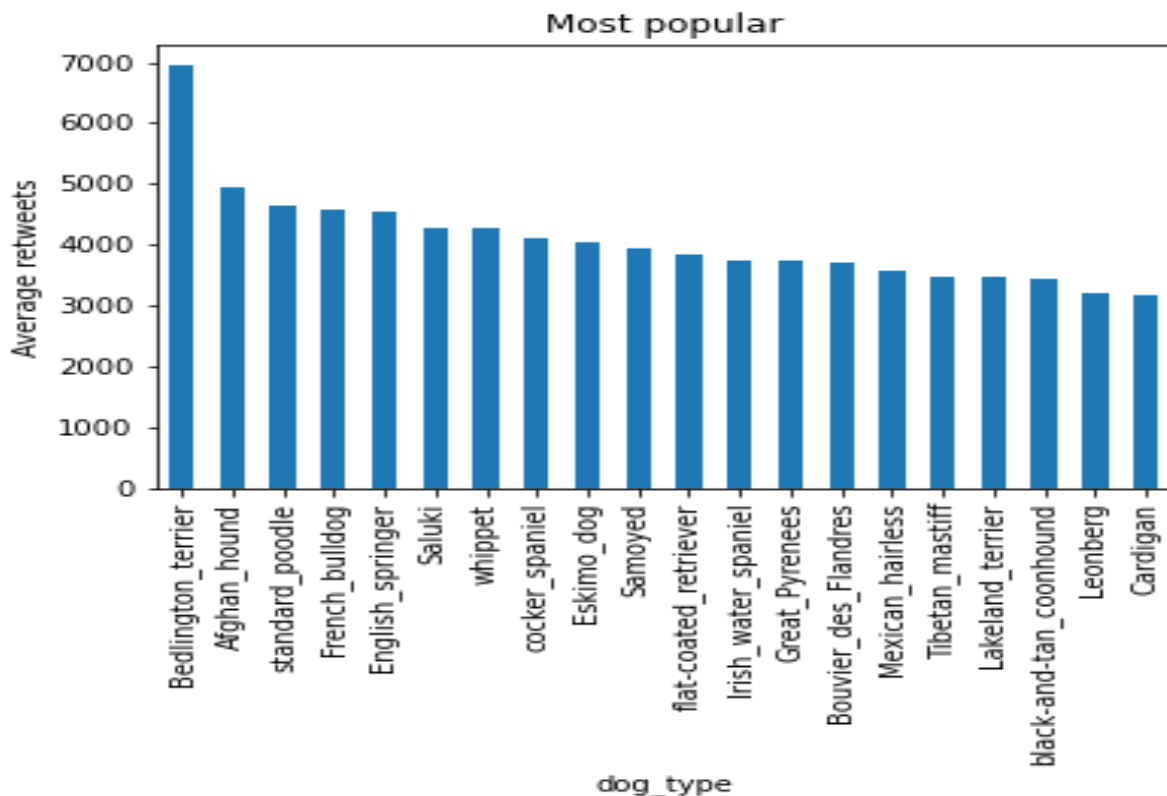


A plain vanilla query on dog type having maximum tweets with say that **golden\_retriever** is the most popular /favorite/ liked/ loved etc. etc. dog types as it has got maximum total number of tweets.

However, the story is a bit when we try to dig further with criteria's like dog type getting maximum retweets for each tweet shared. For this if we do the mean of retweet counts for each dog type, **Bedlington terrier** came up as a winner in most popular/sort after breed among all dogs.



**Fig1: Bar chart depicting average number of retweets for each dog type**

Another way to investigate the data is using rating numerator and rating denominator provided. There are various values of denominator present instead of just 10. Thus, to normalize the rating attribute, we have divided numerator by denominator so that all rating is now between 0 and 1. Now to find highest rated dog type, we have applied a criteria of number of tweets for that breed type of at least 3. This has covered 99 out of total 113 dog types. This criterion has been applied to avoid the bias that might occur if very few numbers of tweets are present for a breed, and it has got very good rating. After performing this analysis, the result was **Saluki and Pomeranian** at number 1 and 2 highly rated dog types.

Next for the purpose of analysis, we have created a shareability index. If we take marking as favorite as a proxy for viewing the tweet, the shareability index tells us what percentage of users who viewed the tweet, shared the tweet as well. This index captures the real impact for a tweet

on the audience. According to this index, top 3 dogs in terms of shareability are **Gordon\_setter**, **Maltese\_dog**, **Ibizan\_hound**

**Above analysis on from various angles gave a better picture on dog types that popular, highly rated and loved by many.**