WeRateDog Report

Introduction

WeRateDog is a Twitter handle that rate dog, upload dog image with a short description and give its followers a chance to like and re-tweet them. They are unique for using improper fractions as rating scores. We asked WeRateDog for their twitter-archive-enhanced and they mailed it to us named twitter-archive-enhanced. I programmatically downloaded, opened it, and observed it contained basic information but not everything. Twitter-archive-enhanced lacked essential information like favorite and re-tweet counts. Also, it doesn't contain information about the breed but thanks to Udacity for classifying the dog image uploaded using neural network procedure and saving it as image-prediction.

To create meaningful insight from this archive data, I needed more information such as re-tweet and favorite counts. So, I applied for a Twitter developer account which will allow me to query Twitter API with the help of the tweepy library. I fetched re-tweet and favorite counts for each tweet-id in the twitter-archive-enhanced and saved it as tweet-df.

Thereafter, I formed a complete table, wrangle it, and generated the following insight.

1. A high favorite count does not correspond to a high retweet count.

The analysis of the cleaned data infers that the fact that a dog breed has a good favorite count doesn't guarantee a high re-tweet count. Though **Labrador_retriever breed has the highest favorite count of 144,334 and the highest retweet count of 70,378.** This does not apply to others. For example, **Lakeland_terrier breed came second on the favorite count ranking but was in the fourth position on the retweet table.**

Labrador retriever

Lakeland terrier

Similarly, Eskimo_dog has the second-highest retweet count but was the fifth most favorite breed. Also, Malamute which is eighth, and Golden_retriever which is ninth with the highest favorite count are not among the top ten with retweet count

Eskimo

2. The age of the dog doesn't correspond to the retweet count.

It was observed that the age of the drug does not mean the dog will receive a good re-tweet count. Floolfer and puppo have the highest average rating of 1.2 but the dog with the highest re-tweet counts is from the age class doggo which has the third highest average rating of 1.18.

3. 98.08% of the tweet came from an iPhone device but the dog breed tweeted from TweetDeck has the highest average rating. This does not infer that people favor one tweet

source over the other. TweetDeck only represents 0.44% of the tweet source