

Data Visualization and Analysis for Processed Twitter Data of WeRateDogs Account

Introduction

Tweet's information of WeRateDogss account were gathered through three different sources: given file, downloading file programmatically, and using twitter API. Three data frames out of those sources. The data frames were assessed for quality and tidiness issues. After that there were cleaned, and stored in csv file. Finally, data visualization was applied over the final product, in addition to analyzing this data and come out with some insights and observations about it.

In this report only the data visualization and analysis process are explained. The data wrangling process were explained in the "wrangle report". Hence, only the insights and observations made over the processed data are going to be shown in this report.

Brief description of the WeRateDogs account is needed for the reader to understand the information what is being analyzed. This account basically rates dogs out of 10. The rating could be above or below 10. So, the tweets and the rating of the dogs is being gathered, wrangled, and analyzed to find, for example, what are the dogs that have the most likes, retweets, or the highest ratings.

Data visualization

Starting with data visualization instead on the data analysis is the method used here. First the cleaned data were visualized using word cloud to find what are the most popular dog breed mentioned in the data frame. The following result was received:



Figure 1: Word cloud showing the most popular dog breed

As shown in the figure above, the most three popular dog breeds are: Chihuahua, Labrador retriever, and golden retriever. As a result, the analysis was conducted over those three types of dogs to find which one has the highest retweets, likes, and average rating.

Data analysis

Even though Chihuahuas are mentioned most, the analysis shows that they are not liked the most neither have the higher ratings. The results are shown in the below table, which summarize the whole analysis.

Dog breed	Number of retweets received for the tweets related to this breed	Number of likes received	Average rating by WeRateDogs account
Chihuahua	422,798	1,437,109	10.7
Labrador retriever	925,977	3,234,187	11.3
golden retriever	878,525	2,986,730	11.5

From the above table we can see that Chihuahuas have the lowest number of retweets, likes, and ratings. However, Labrador and golden retriever are having almost similar retweets and likes (even though Labrador are slightly higher). Moreover, their average ratings are almost identical with 0.2 higher in the favor of golden retriever dogs.