

# Data analysis and visualization report

July 6, 2019

## 1 Introduction

This report summarises the key finding of the data analysis of the 'Wrangle and Analyze Data' project.

The goal of this project is to perform data wrangling on the WeRateDogs Twitter data to create interesting and trustworthy analysis and visualizations.

WeRateDogs 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. and numerator almost always greater than 10: 11/10, 12/10, 13/10.

## 2 Data wrangling

The data wrangling for this project consisted of the following steps:

- gathering WeRateDogs twitter data from three different sources
- assessing the data to find possible quality and tidiness issues
- cleaning and merging data into a single dataset to perform data analysis and create visualizations.

## 3 Data analysis

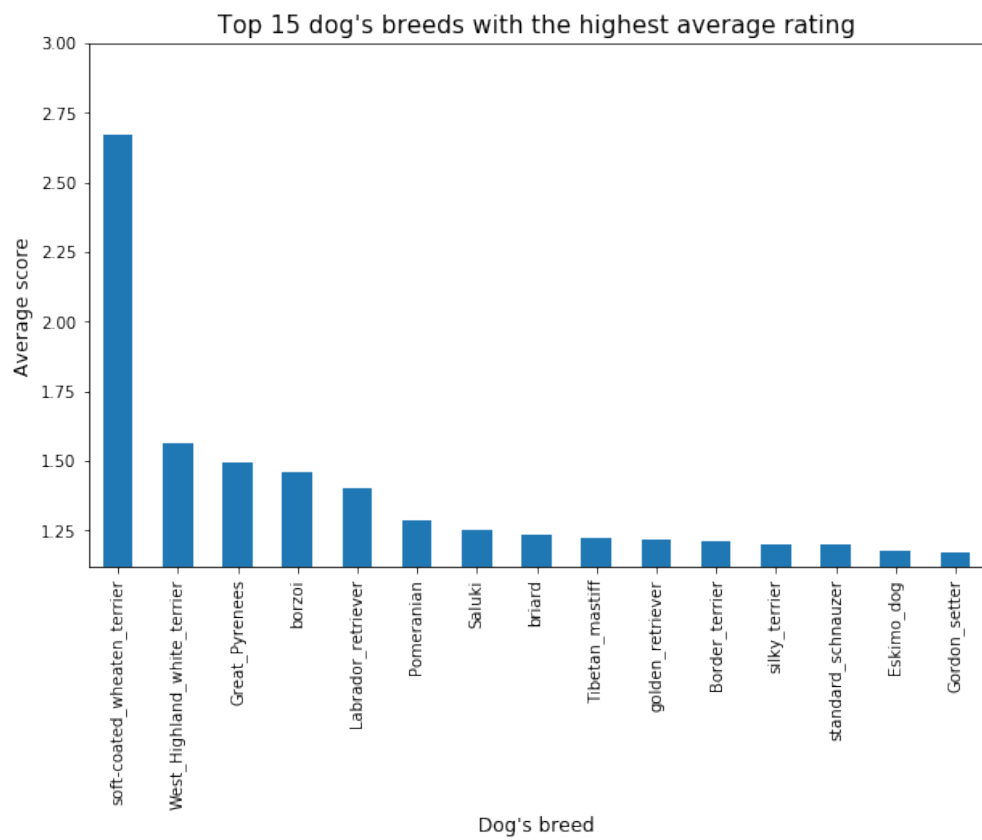
Before performing data analysis the following questions have been posed:

- What are the most common dog breeds?
- Which dog breeds have the highest rating?
- Which dog breeds have the highest favorite counts?
- What dogs have the highest retweet count?
- What is the relationship between favorite and retweet count?
- Which dog stages are associated with the highest favorite and retweet counts?

### 3.1 Most common dog breeds

This pie chart shows the 10 most common dog breeds. We can see that golden retriever is the most common dog's breed, following by Labrador retriever, pembroke and chihuahua.

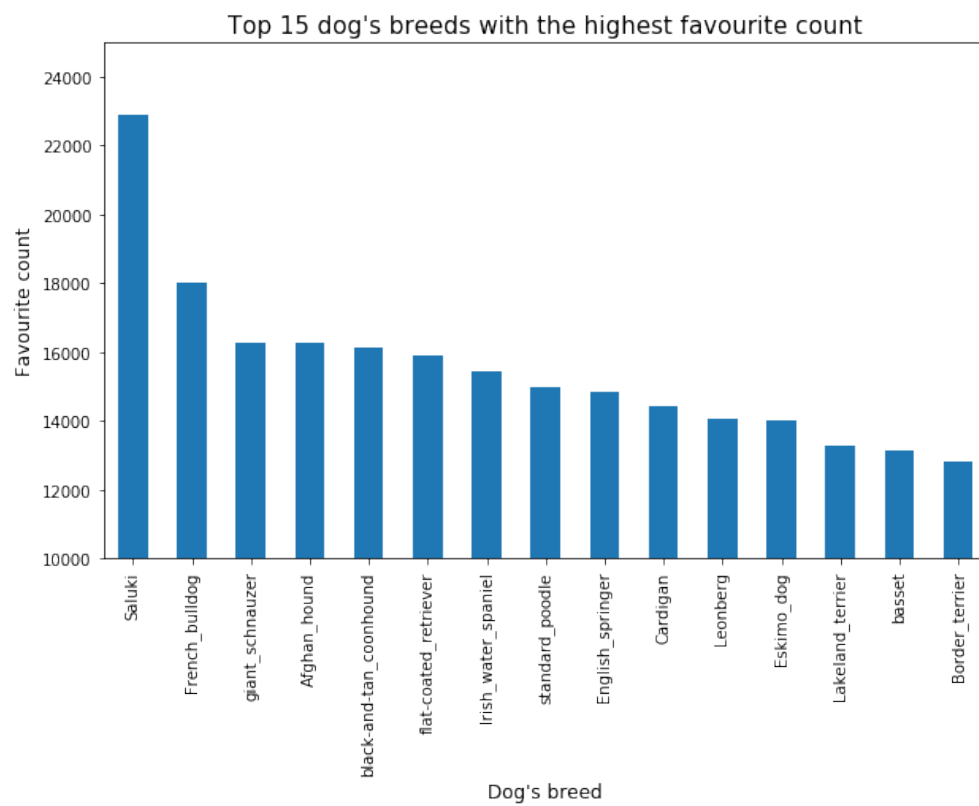




alt text



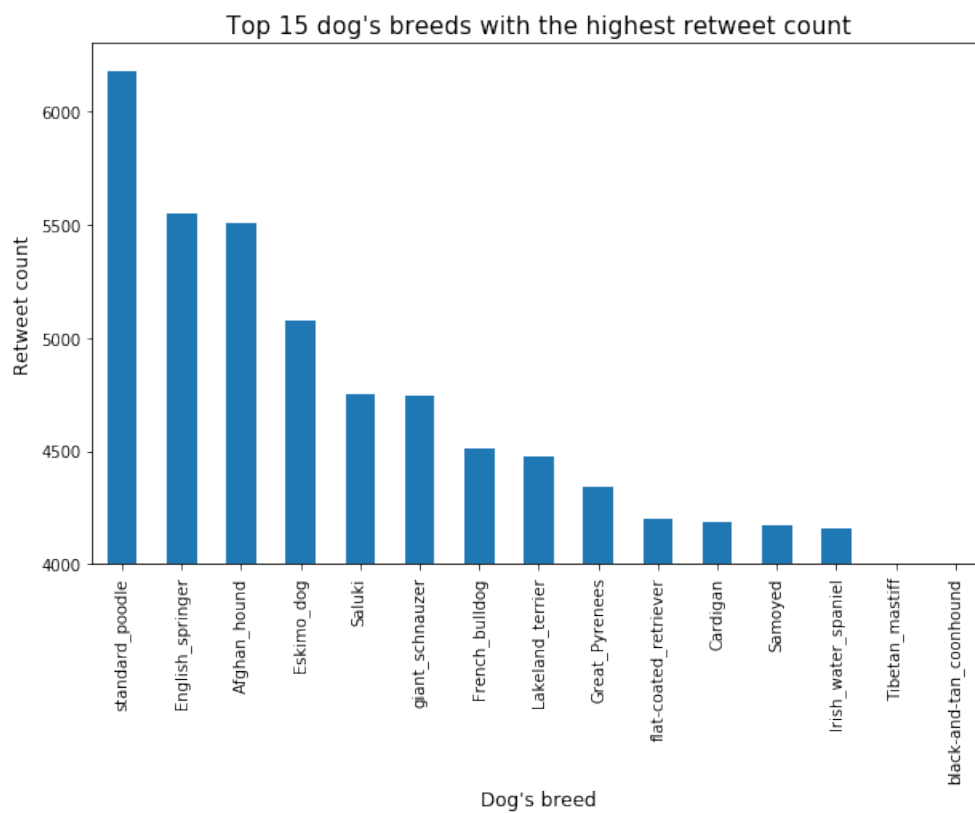
alt text



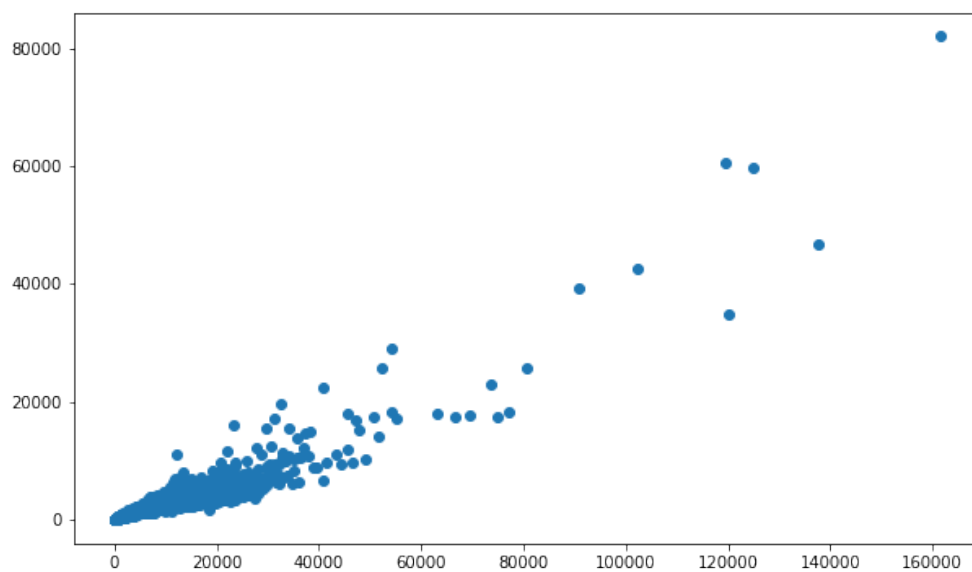
alt text



alt text

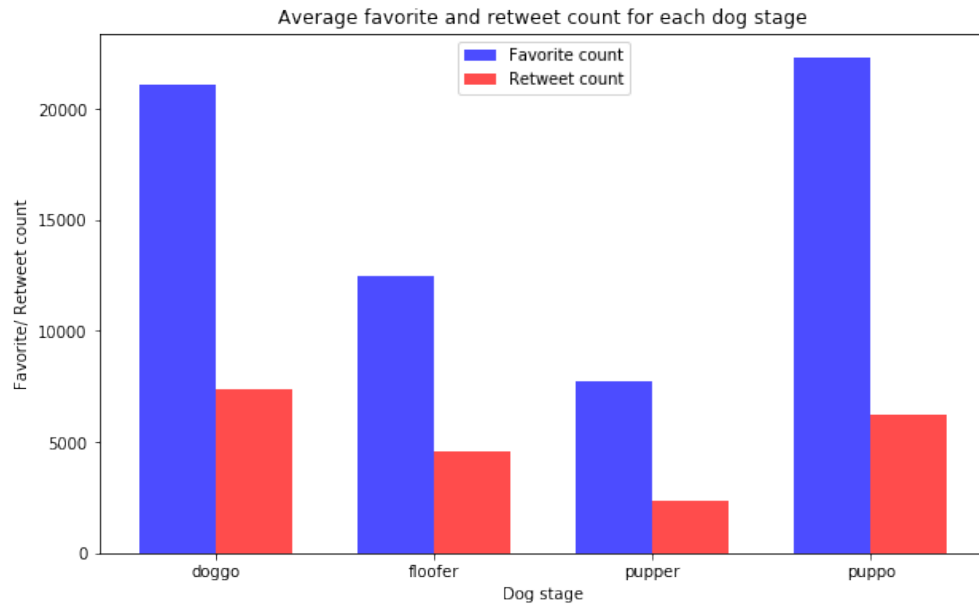


alt text



alt text





alt text

### 3.6 Favotite vs Retweet count for each dog stage

The bar chart below shows the average favorite and retweet count for each dog breed. Overall, the favorite count is higher that retweet count for all dog stages. It can be also observed that doggo and puppo have the highest favorite and retweet counts, while pupper have the lowest both favorite and retweet counts.