Act Report

The aim of this report is to communicate the insights and display the visualizations produced from the data wrangling of "WeRateDogs" dataset.

First impressions of the data:

After taking a look at df_master_clean dataframe using several pandas functions like shape, info and describe, there are some observations that can be done:

- There are 1941 rows and 10 columns.
- There are some incomplete tweets:
    - 564 tweets without name,
    - 302 tweets without dog_breed
    - 1,628 tweets without dog_stage.
- The rating average is 10.53, which is very high. 25% of the tweets have a rating lower than 10. The minimum rating is 0 and the maximum 14.
- The favorite_count average is 7,943.82. 75% of the tweets have a favorite_count higher than 9,907. The minimum favorite_count is 67 and the maximum 148,346.
- The retweet_count average is 2,317.64. 75% of the tweets have a retweet_count higher than 2,643. The minimum retweet_count is 11 and the maximum 72,874.
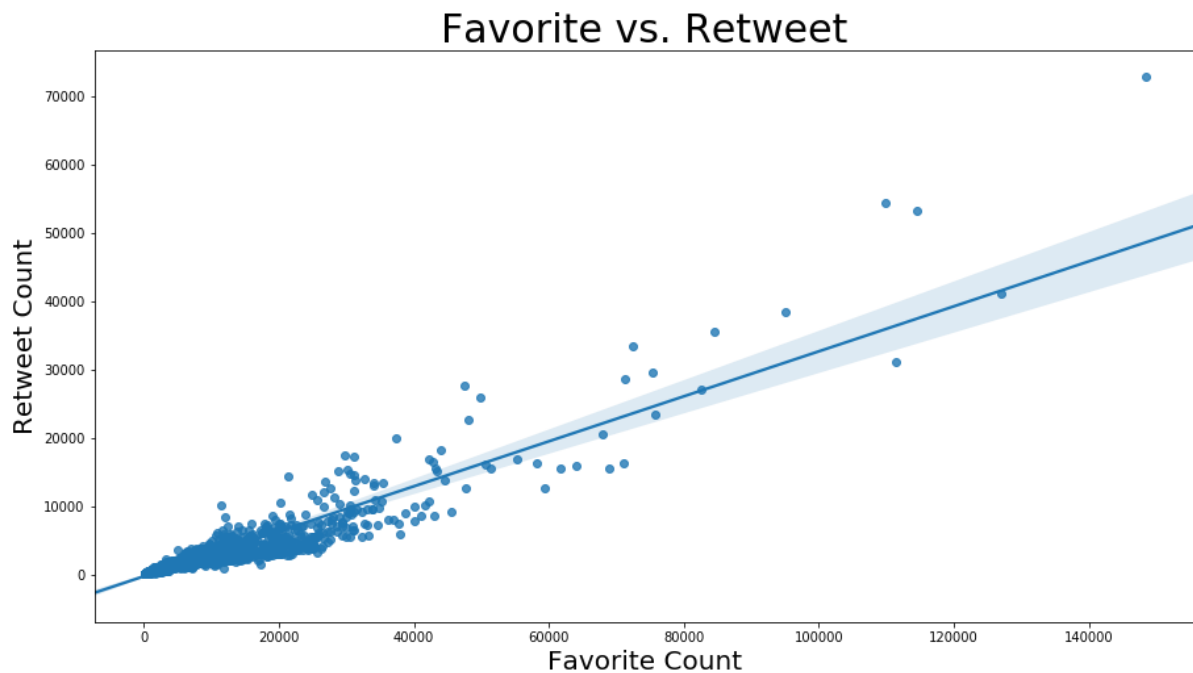
Questions:

Let's define the following questions that will be answered analyzing the data:

1. Is there any correlation between any variable?
2. How popular is "WeRatedDogs" over time?
3. Which is the most common dog stage? and the most rated one?
4. Which are the top 10 breeds with highest ratings?
5. Which are the top 10 most liked dog breeds?
6. Which are the top 10 most popular dog breeds?
7. Which are the top 5 most common dog names?
8. How is the distribution of Tweet sources?

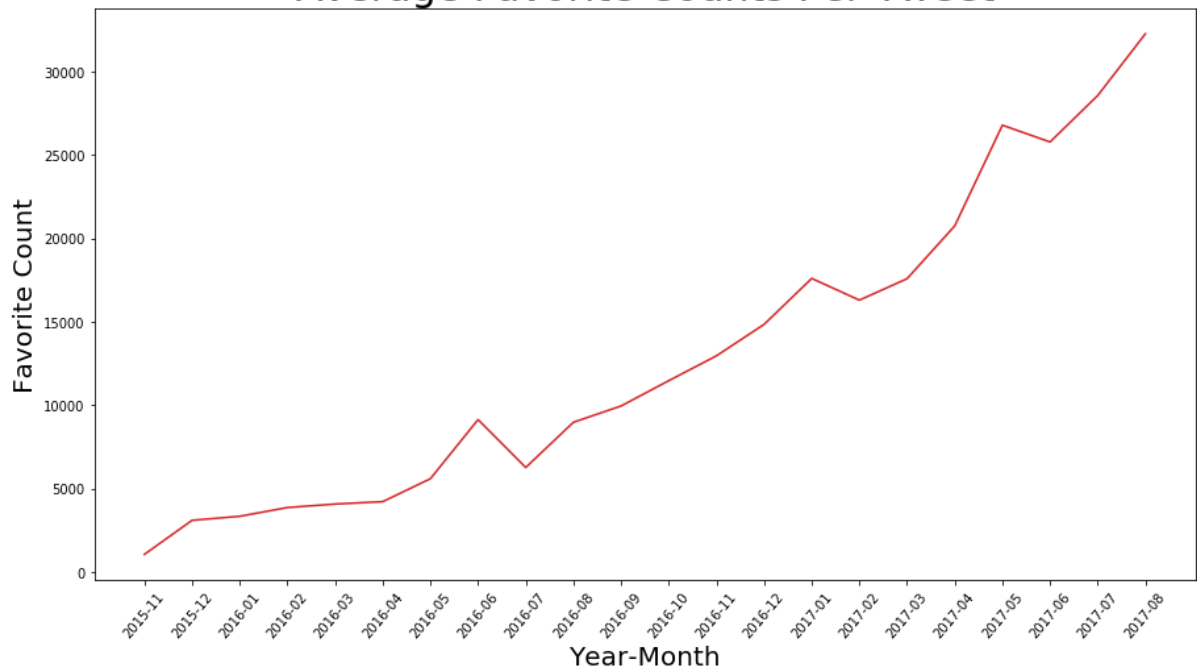# Question 1. Is there any correlation between any variable?

There is a strong positive correlation between favorite_count and retweet_count (0.9279) which means that when a tweet increases its number of retweets then the number of likes increases as well.
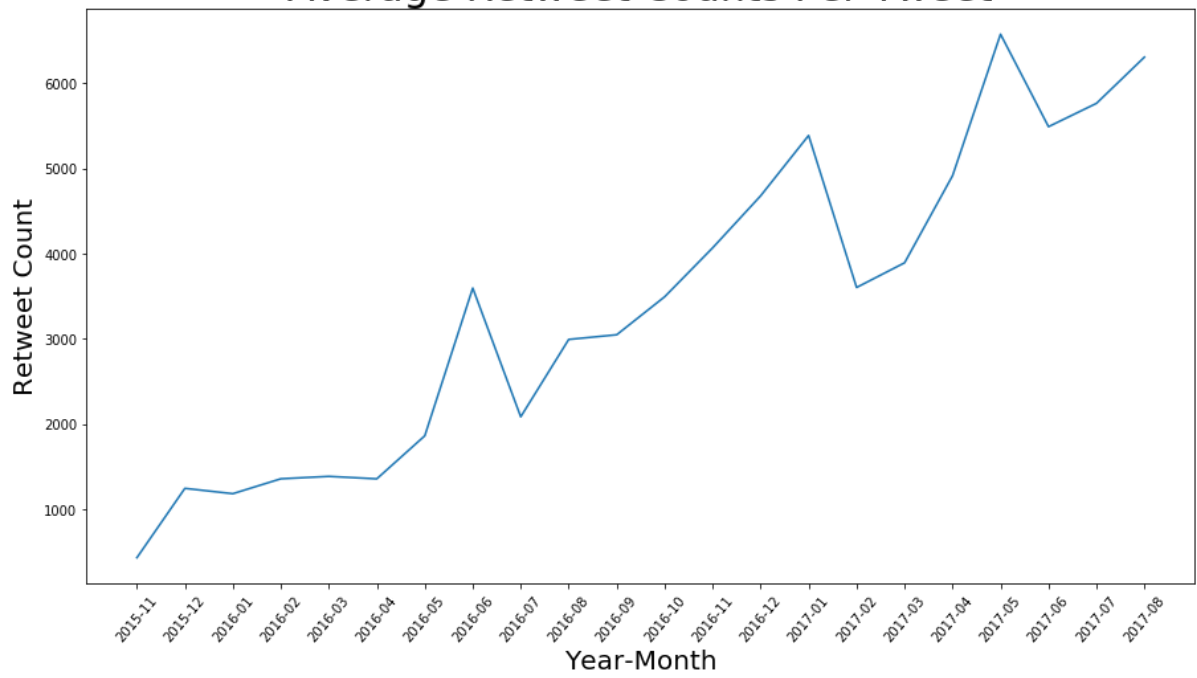

Favorite vs. Retweet

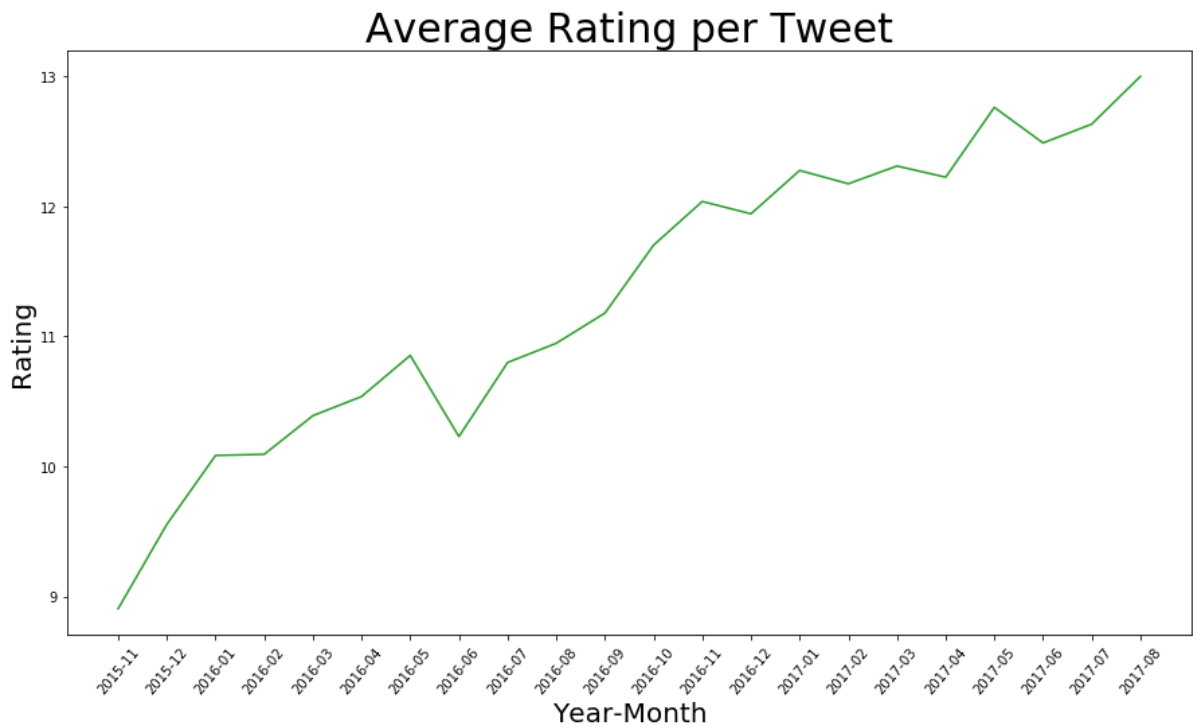# Question 2. How popular is "WeRateDogs"?

Taking a look at the average of favorite count and retweet count per tweet, we can say that WeRatesDog has increased in popularity over the time. The trend of favorite count, retweet count and rating was decidedly upward. The highest peak of likes was in August, 2017, reaching around 30,000 likes per tweet.
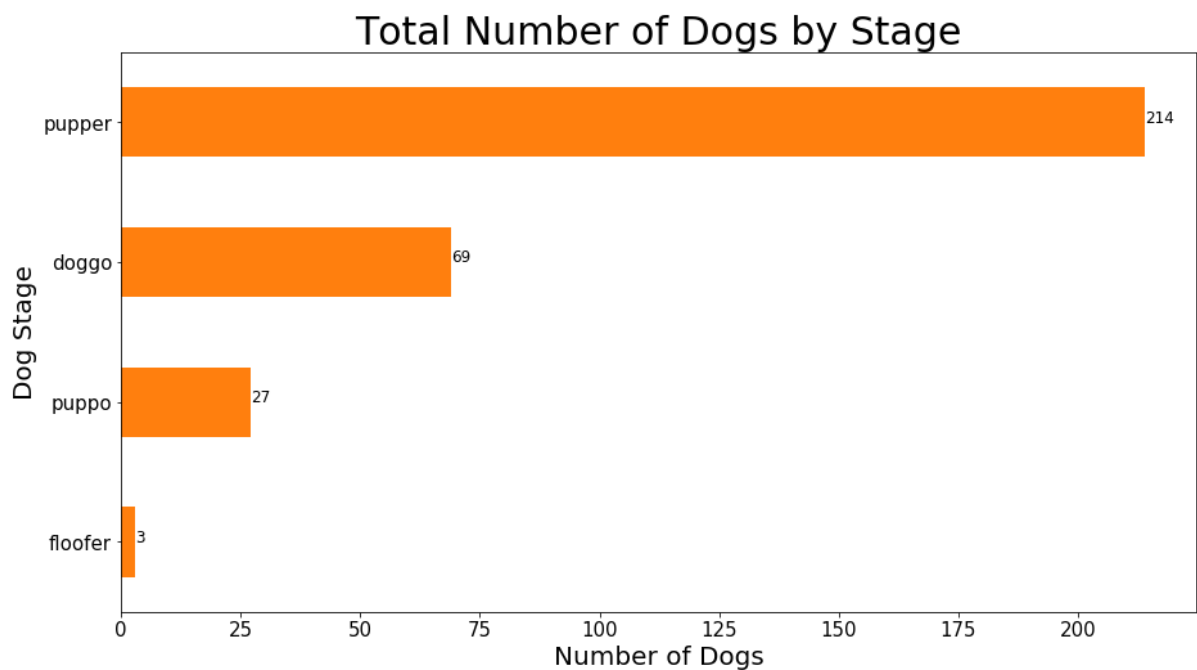
Average Favorite Counts Per Tweet

Average Retweet Counts Per Tweet
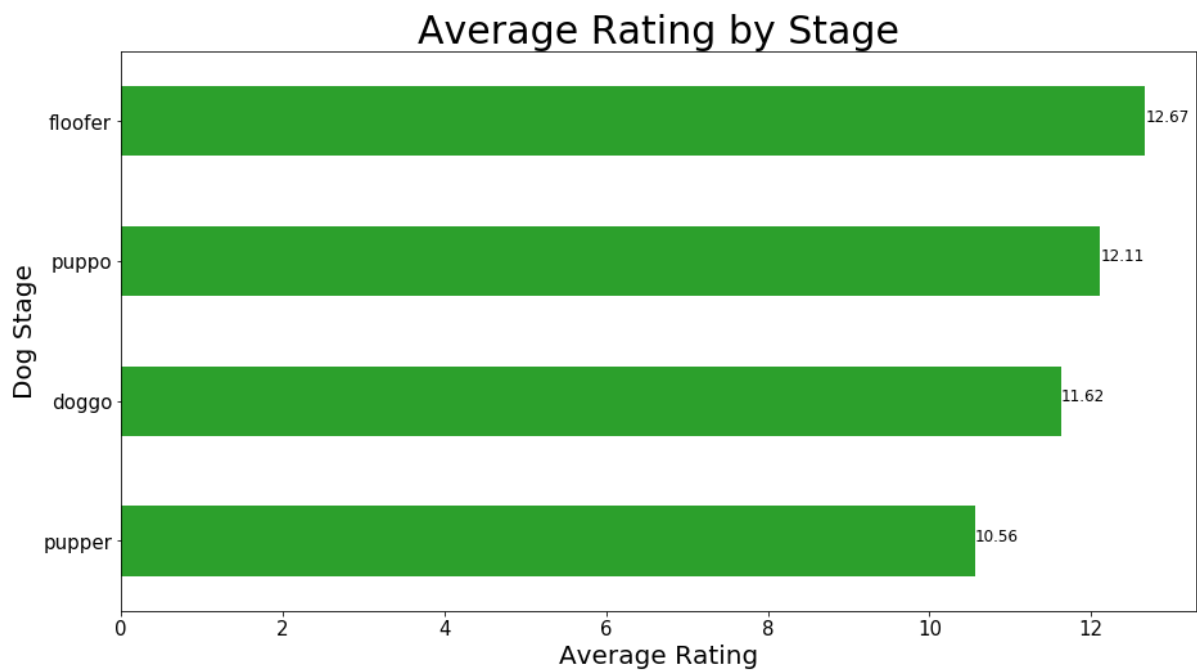
Average Rating per Tweet

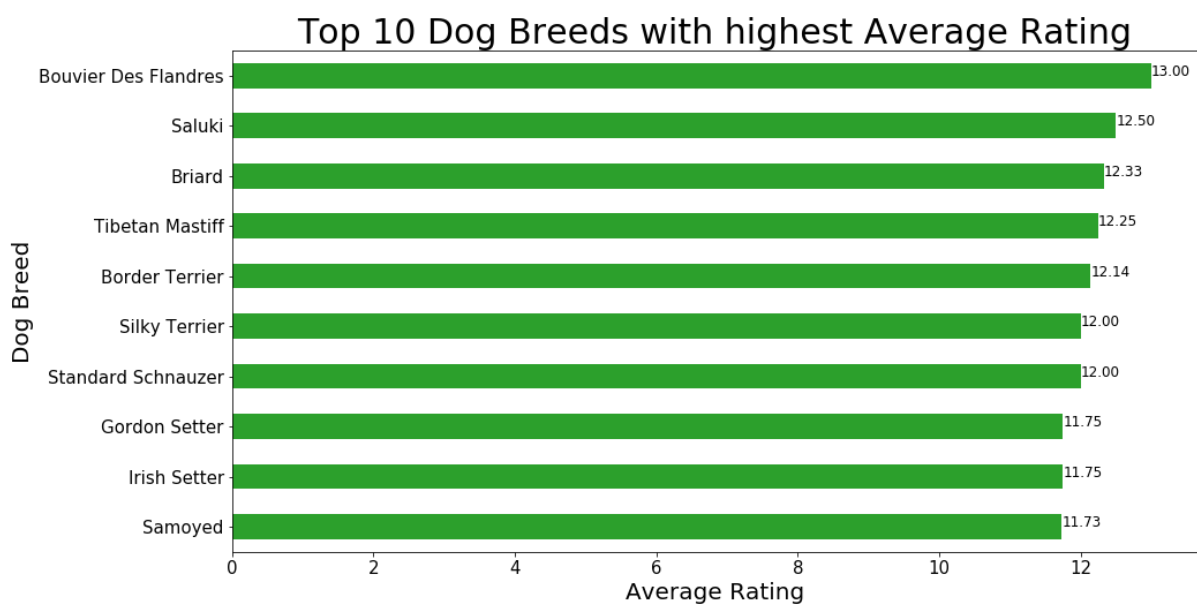## Question 3. Which is the most common dog stage? and the most rated one?

The most common dog stage is pupper (214 dogs). However, pupper is the dog stage with the lowest average rating (10.56). The dog stage with the highest average rating is floofer, but it's also the less common dog stage (only 3 dogs).

Total Number of Dogs by Stage
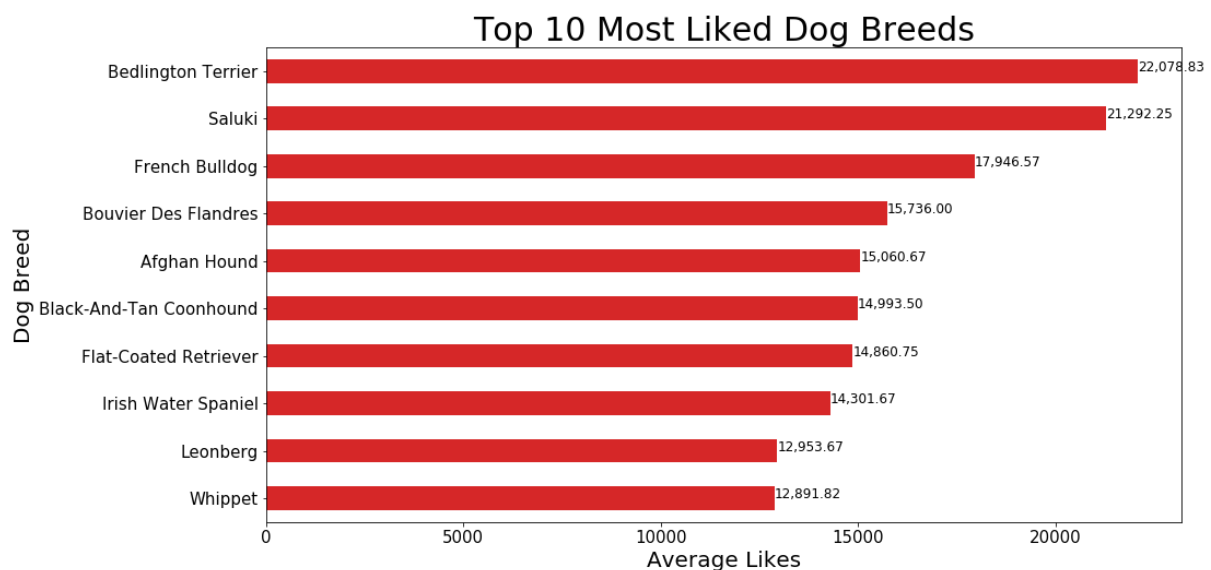
Average Rating by Stage

Question 4. Which are the top 10 breeds with highest ratings?

Bouvier Des Flandres is the dog breeds with the highest average rating (13.00). Followed very closely by Saluki with an average rating of 12.50. In the third position is Briard (12.33). Fourth position is for Tibetan Mastiff (12.25), fifth position is for Border Terrier (12.25), sixth position is for Silky Terrier (12.00). There is a tie in the seventh and eighth position, Gordon Setter and Irish Setter have both an average rating of 11.75. And the top ten list lasts with Samoyed in tenth position with an average rating of 11.73.

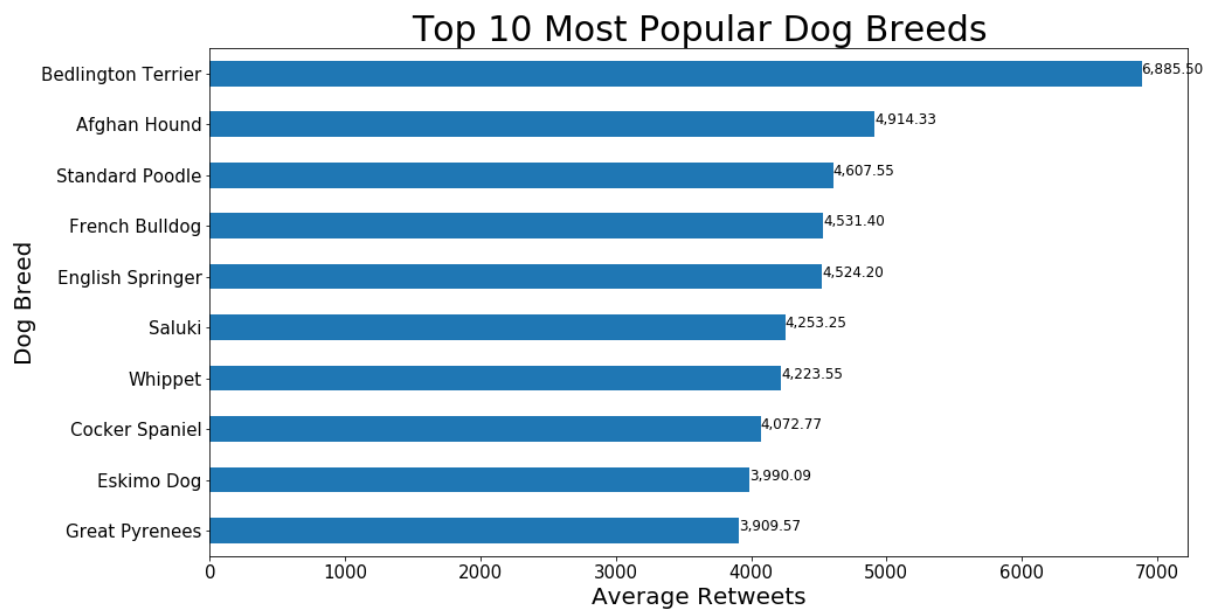Top 10 Dog Breeds with highest Average Rating

Question 5. Which are the top 10 most liked dog breeds?

Bedlington Terrier is the most liked dog breed. The second position is for Saluki. Followed very closely by Saluki with an average rating of 12.50. In third position is French Bulldog. Bouvier Des Flandres is the dog breed in fourth position. Afghan Hound is in fifth position, followed very closely by Black-And-Tan Coonhound in sixth position and Flat-Coated Retriever in seventh position. Eighth position is for Irish Water Spaniel. In the last positions of top 10 most liked dog breeds are found Leonberg and Whippet respectively.
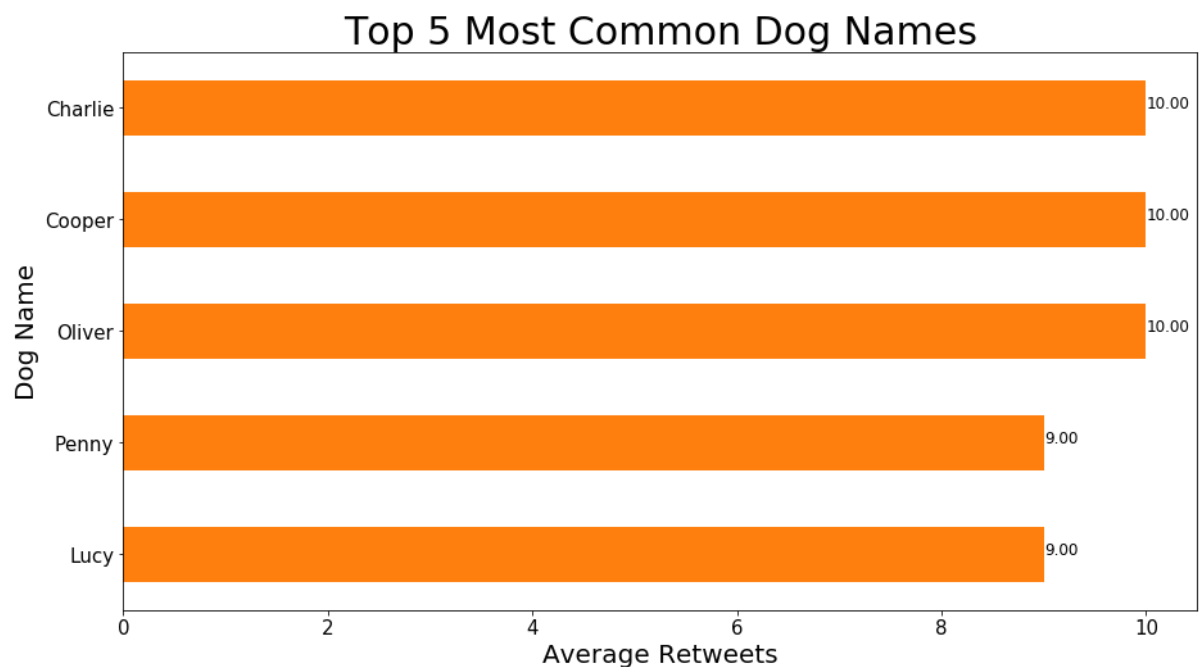

Top 10 Most Liked Dog Breeds

Question 6. Which are the top 10 most popular dog breeds?

The dog breed with the highest average of retweets is Bedlington Terrier with a wide difference from the second position occupied by Afghan Hound. The third position is for Standard Poodle. Fourth position is for French Bulldog, followed very closely by English Springer. Saluki is the sixth most popular dog breed, followed by Whippet in seventh position, Cocker Spaniel in eighth place, Eskimo Dog in ninth place and Great Pyrenees is in tenth position.

Top 10 Most Popular Dog Breeds

Question 7. Which are the top 5 most common dog names?

There is a triple tie in the first three positions, Charlie, Cooper and Oliver are the most common dog names. Again, there is a tie in fourth and fifth positions and they are occupied by Penny and Lucy respectively.

Top 5 Most Common Dog Names

Question 8. How is the distribution of Tweet sources?

The majority of the tweets are from 'Twitter for iPhone' (98.1%). The two other sources are very small, Twitter Web Client (1.4%) and TweetDeck (0.5%).