

We Rate Dogs | Data Analysis & Visualization

Sarah Aranda

After a structured process of wrangling data from the WeRateDogs Twitter account, I moved onto analysis and visualization.

For the analysis, I investigated which dog names were the most popular; which dog stages were given the highest ratings, retweet and favorite count; then I moved on to looking at the most popular dogs according to retweet count, favorite count, and rating.

The visualizations section of this project was also part of the analysis. I decided to focus on experimenting with heatmaps since I am less familiar with them than with histograms, scatterplots, and boxplots; I was interested in trying my hand at something new.

I created two heatmaps. One heatmap compares the mean favorite counts for the three most common dog breeds (Golden Retrievers, Labrador Retrievers, and Pembrokes) at the four stages. The other heatmap compares the month and day of the week to the mean retweet counts.

Here are some of the insights I gathered from the analyses and visualizations:

- 1. The top five dog names are Cooper, Oliver, Charlie, Tucker, and Lucy.
- 2. When comparing the mean retweet count, mean favorite count, and ratings, only the Saluki breed appears in the top five for all three categories.
- 3. In comparing Golden Retrievers, Labrador Retrievers, Pembrokes across different life stages, young Pembroke puppies (puppos) are the most favorited followed by mature (doggo) Labrador Retrievers.
- 4. When looking at the means of the retweet counts, the highest number of retweets was found in Fridays in June followed by Thursdays in August.
- 5. At the end of the year (throughout November and December) retweet means seem to stay consistently low no matter what day of the week it is.

Example Visualization from Project: Heatmap comparing month and day of the week with retweet counts

