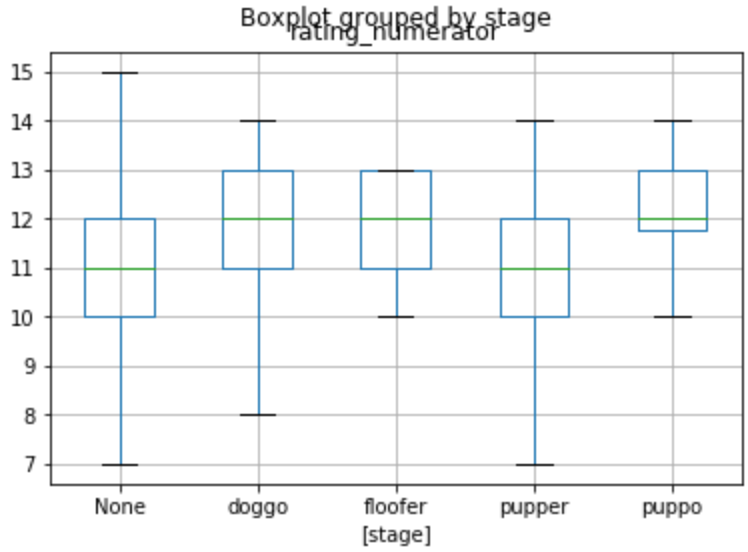
**Act Report**

## Introduction

This report contains analyses of data from the WeRateDogs Twitter account.

## Analysis 1: Relationship between dog stage and rating

To analyse if there is a relationship between which rating a dog gets and the stage it is in, I made a box plot of the dog photo ratings at each of the dog stages:




The plot indicates that dogs on the doggo, floofer and puppo stages get the highest ratings.

One should note that both stage and rating is decided by the owner of the WeRateDogs Twitter account, so the correlation mainly says something about that person :)

## Analysis 2: Correlation between retweets and dog rating

We also analysed if there is a correlation between the dog rating (as decided by the owner of the abovementioned Twitter account) and how many times the tweet is retweeted.

To do this, I used data from the Twitter archive plus the data from the Twitter API.

I found a correlation of 1.5%, which suggests that there is no significant correlation between the two variables.

We can therefore conclude that a high rating does not predict a lot of retweets.

## Analysis 3: Regression to predict how many times a tweet has been favorited

Finally, I made a linear regression model to predict how many times a tweet has been favorited, based on other available information.

I used the following as predictive variables:

- Number of retweets
- Rating of the dog
- Whether the most likely content of the tweet image is a dog

I estimated the following model using OLS:

| | | | |
|---|---|---|---|
| **Dep. Variable:** | favorite_count | **R-squared:** | 0.865 |
| **Model:** | OLS | **Adj. R-squared:** | 0.865 |
| **Method:** | Least Squares | **F-statistic:** | 4187. |
| **Date:** | Mon, 29 Jul 2019 | **Prob (F-statistic):** | 0.00 |
| **Time:** | 17:51:18 | **Log-Likelihood:** | -19404. |
| **No. Observations:** | 1969 | **AIC:** | 3.882e+04 |
| **Df Residuals:** | 1965 | **BIC:** | 3.884e+04 |
| **Df Model:** | 3 | | |
| **Covariance Type:** | nonrobust | | |

| | **coef** | **std err** | **t** | **P>\|t\|** | **[0.025** | **0.975]** |
|---|---|---|---|---|---|---|
| **const** | 1152.6277 | 214.047 | 5.385 | 0.000 | 732.845 | 1572.411 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **retweet_count** | 2.5585 | 0.023 | 111.902 | 0.000 | 2.514 | 2.603 |
| **rating_numerator** | 0.5406 | 2.544 | 0.213 | 0.832 | -4.448 | 5.529 |
| **p1_dog** | 1165.0324 | 236.878 | 4.918 | 0.000 | 700.475 | 1629.590 |

| | | | |
|---|---|---|---|
| **Omnibus:** | 521.128 | **Durbin-Watson:** | 0.788 |
| **Prob(Omnibus):** | 0.000 | **Jarque-Bera (JB):** | 17973.551 |
| **Skew:** | 0.549 | **Prob(JB):** | 0.00 |
| **Kurtosis:** | 17.761 | **Cond. No.** | 1.53e+04 |

The results suggest that:

- The number of times that a tweet has been retweeted is predictive of how many times it is favorited. This is not a suprise, as many people will both favorite and retweet at the same time.
- A picture of a dog is more likely to be favorited. Also not a surprise, as the Twitter account is about dogs, so that's what its followers expect and want.