

# wrangle\_report

May 20, 2019

## 1 done

- making doggo, floofer, pupper and puppo tidy as column called stage
- merging twitter\_archive\_clean and tweet\_likes\_retw
- removing tweets that are retweets or replies from twitter\_archive
- making new column called  $\text{rating\_out\_of\_10} = \text{rating\_numerator} / \text{rating\_denominator} * 10$  to get one column to compare and get rid of some outliers
- renaming image\_predictions\_clean columns
- making prediction1, confidence1, dog1, prediction2, confidence2, dog2, prediction3, confidence3 and dog3 tidy

## 2 more to be done

- collecting data about the actual dog type to analyze models sensitivity and specificity
- completing stage data