WERATEDOGS

WRANGLE & DATA ANALYSIS

UDACITY DATA WRANGLING & MODELING twitter.com/dog_rates

OFF TO THE RACES!

WeRateDogs, an account on the popular Twitter website, will make your adorable pups famous! This Twitter account will showcase your beloved family's pets and will allow people to vote and make humorous comments about them. Matt Nelson, the creator of the website, launched the Twitter account on November 15, 2015. It is currently active and has been up for nearly seven years.

DATA ANALYSIS AND VISUALIZATION

Insights

After gathering, cleaning, and storing data on the popular Twitter account called WeRateDogs (@dog_rates) it is now time to present the data for everyone's entertainment and curiosity. The project raised several questions for analysis:

1. Which of the dog breed prediction algorithms was the most accurate?
2. What were the top 20 dog names in the tweet dataset?
3. Which year had the most active tweet favorite count interaction?

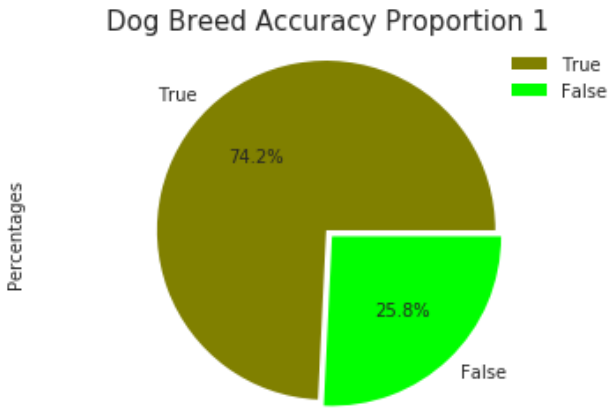
Dog Breed Predictions Algorithms

The images from WeRateDogs were ran through a neural network that can determine the breeds of dogs. The neural network used three different dog breed prediction algorithms to determine a dog's breed by the pictures uploaded. There was p1 algorithm, p2 algorithm, and p3 algorithm. The data for each algorithm, the pictures, and whether the algorithms were correct or not (True or False) are included in the image predictions tsv file.

Figure 1 Puppy Einstein

Prediction #1 Accuracy Rate

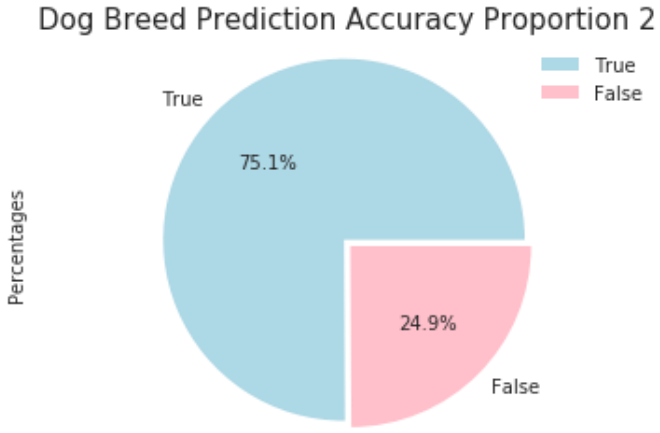
True 1463
False 508
Name: p1_dog, dtype: int64



The dog breed prediction algorithm #1 has an accuracy rate of 74.2% and only 25.8% of being inaccurate.

Prediction #2 Accuracy Rate

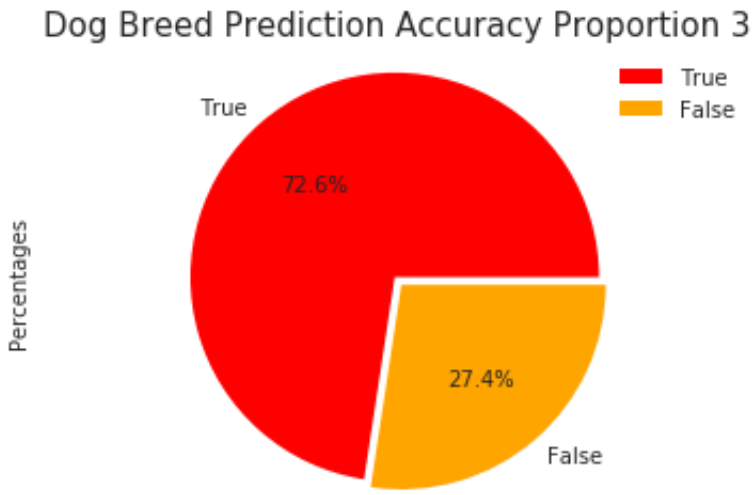
True 1408
False 491
Name: p2_dog, dtype: int64



The dog breed prediction algorithm #2 has an accuracy rate of 75.1% and only 24.9% of being inaccurate.

Prediction #3 Accuracy Rate

True 1431
False 540
Name: p3_dog, dtype: int64



The dog breed prediction algorithm #3 has an accuracy rate of 72.6% and only 27.4% of being inaccurate.

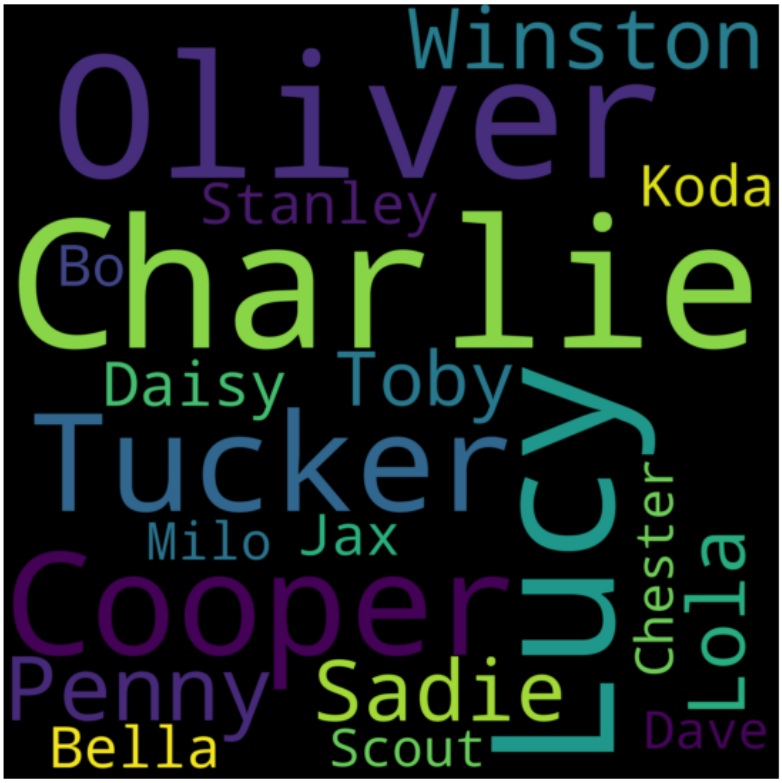
Dog Breed Prediction Algorithms Conclusion

All three of the prediction algorithms were relatively close, with the prediction #2 algorithm having a slightly higher accuracy rate (75.1%). Overall, all three prediction algorithms performed well.

Top 20 Dog Names

Charlie	11
Lucy	10
Oliver	10
Cooper	10
Tucker	9
Penny	9
Winston	8
Sadie	8
Lola	7
Toby	7
Daisy	7
Stanley	6
Koda	6
Jax	6
Bella	6
Bo	6
Scout	5
Dave	5
Milo	5
Chester	5

Name: name, dtype: int64

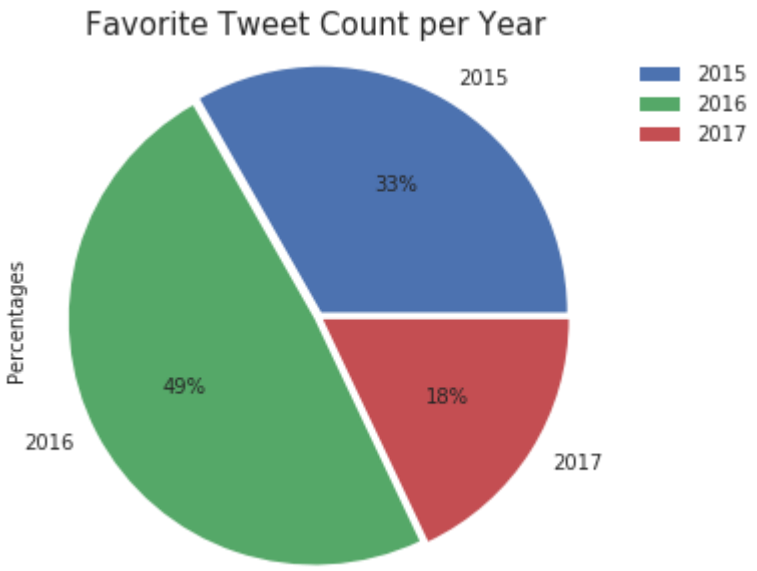


Top 20 Dog Names Conclusion

The top dog's name is Charlie with 11 occurrences in the dataset, followed by Lucy, Oliver, and Cooper with 10 occurrences. Tucker and Penny tie with 9. Winston, and Sadie tie with 8. Lola, Toby, and Daisy have 7 occurrences. Stanley, Koda, Jax, Bella, Bo all have 6 occurrences, followed by Scout, Dave, Milo, and Chester finishing with 5 occurrences.

Favorite Tweet Count Per Year

timestamp
2015 655
2016 962
2017 354
Name: favorite_count, dtype: int64



Favorite Tweet Count Per Year Conclusion

The year 2015 had 655 favorite tweet counts, while 2016 had the most favorite tweet counts with 962 favorite tweets. The year 2017 had the fewest favorite tweets with 354 favorite tweets.