

INSIGHTS AND VISUALIZATION OF DATA WRANGLING CARRIED OUT ON WeRateDog BETWEEN 2015 AND 2017

INTRODUCTION

About 5000 dataset of tweets were provided. Only 1994 tweets were analyzed after wrangling and cleaning. The analysis gave some insights about the dataset. The summary is provided as follows:

1. The highest source of tweets is iPhone.

Most tweets of the tweets analyzed were from **iPhone** (about 98%). A few are from **Web Client** (about 1%) while very few are from **TweetDeck** (less than 1%).

This can be seen in Figure 1.1

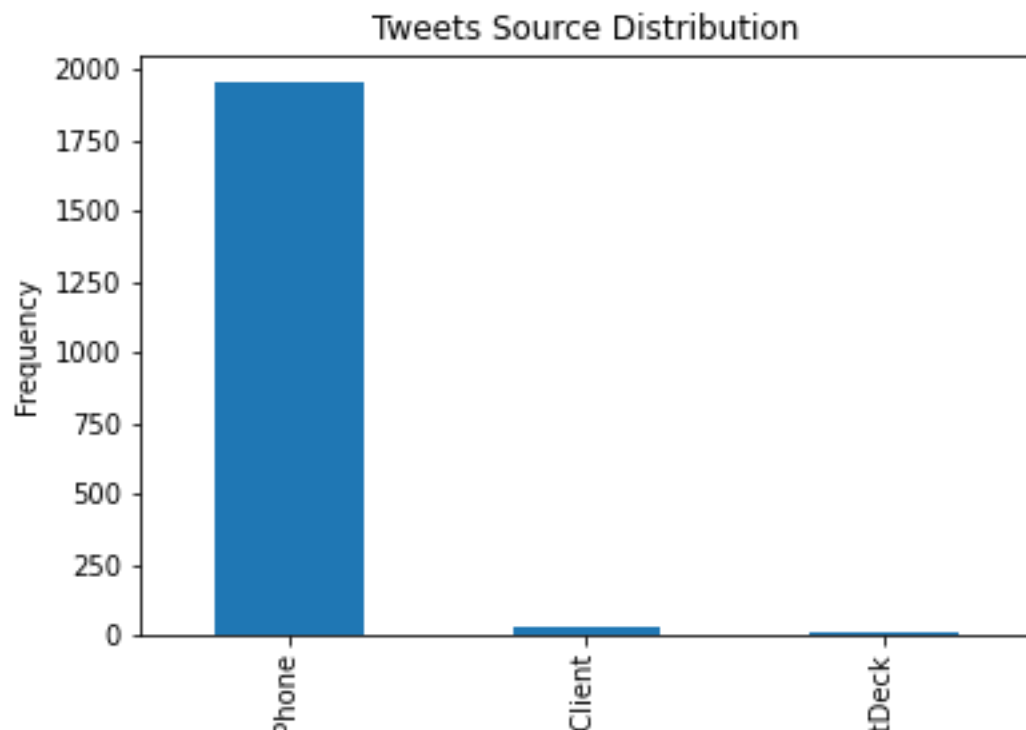


Figure 1.1: Tweets Source

2. Most dogs were rated between 10 and 13.

Dogs were rated on a scale of 10, although may be greater than 10. This depends on impressive people rating the dogs are. About 76% of the dogs were rated between 10 and 13. However, more than 50% of the total ratings were above 10 and few were above 100. Other ratings are displayed in Figure 1.2

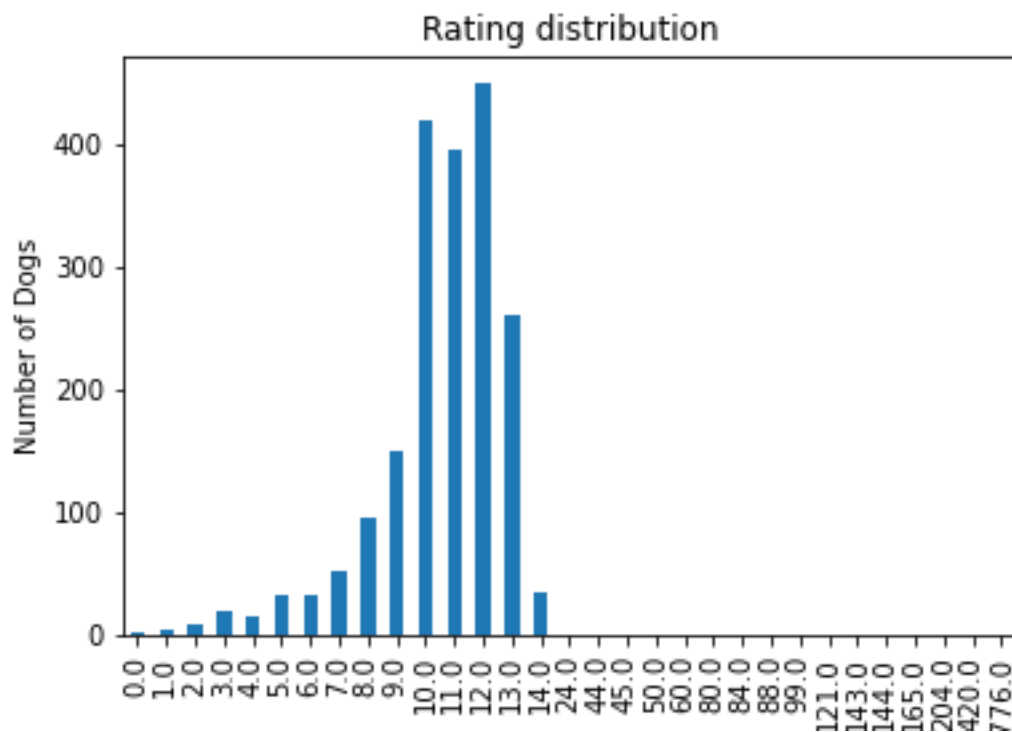


Figure 1.2: Dogs Ratings

3. The most common names of dogs are Cooper, Lucy and Oliver.

Out of 913 names of dogs in the dataset, Cooper, Lucy and Oliver are the top common names of dogs. Other name in the list of top 10 are: Penny, Tucker, Winston, Sadie, Toby, Lola and Daisy as displayed in Figure 1.3

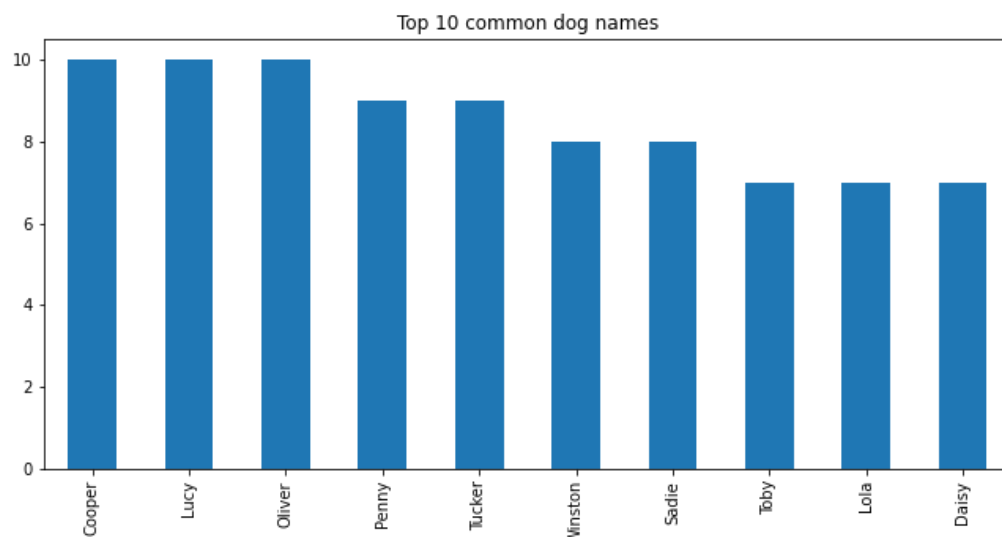


Figure: 1.3: Top 10 Dogs Common Names

4. Tweets trends over time.

The number of tweets in 2015 was high, about 30% of the total tweets analyzed. This trend, however, reduced drastically after 2015. The trend was almost uniform between 2016 and 2017 but later reduced further. Several factors might have contributed to this. But the available dataset may not give adequate insight of those factors.

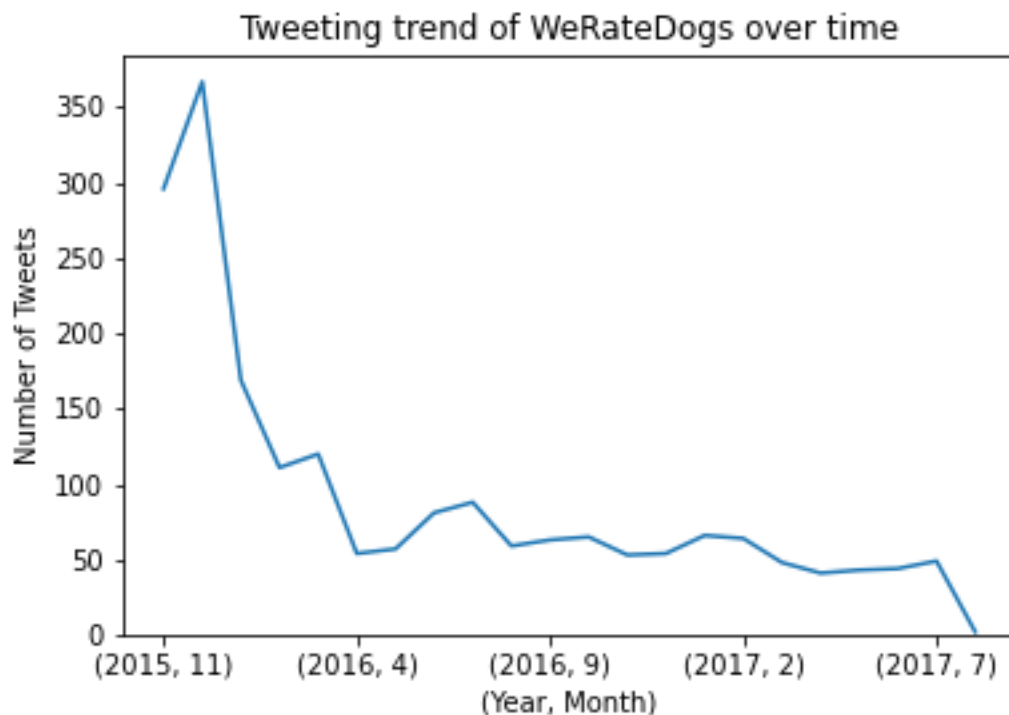


Figure 1.4: Tweets trends

5. Labrador Retriever, an American most popular breed, top the list.

111 breeds of dogs were predicted by the image prediction neural network. Top 10 breeds were analyzed and **Labrador Retriever** top the list.

* **Labrador Retriever** is America's most popular dog breed. It is friendly and companionable. It can bond with the whole family and socialize with dogs and humans alike.

Other breeds on the list are:

* **Pembroke** is a cattle herding dog breed that is strong, athletic, and lively.

* **Chihuahua** is a Mexican breed of toy dog and is among the smallest of all dog breeds.

* **Pug** is a breed of dog originally from China, with physically distinctive features of a wrinkly, short-muzzled face and curled tail.

* **Chow** is a spitz-type of dog breed originally from northern China.

* **Samoyed** is an ancient breed, selected by nomadic tribes in Siberia to herd their reindeer and double as sled pullers when needed.

* **Toy Poodle** is an extremely intelligent easily trained dog. It is an elegant-looking tiny dog with a slender muzzle and long neck.

* **Pomeranian** is a perky and bold toy dog breed with a long coat. It hails from what's now Germany and Poland.

* **Malamute** is the native Alaskan Arctic breed, cousin to the Samoyed of Russia. It possesses tremendous strength, energy, endurance, independence, and intelligence.

* **Cocker Spaniel** is breed of dog that was originally bred as hunting dogs in the UK. It is loving and easy-going.

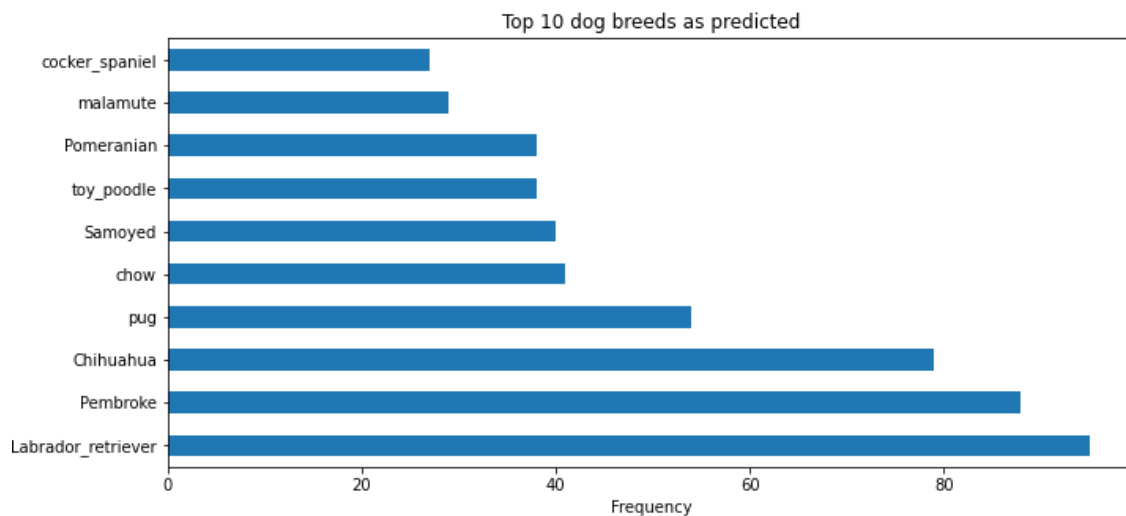


Figure 1.5: Top 10 dogs' breeds