Udacity Data Analyst Nano Degree Project 04: WeRateDogs
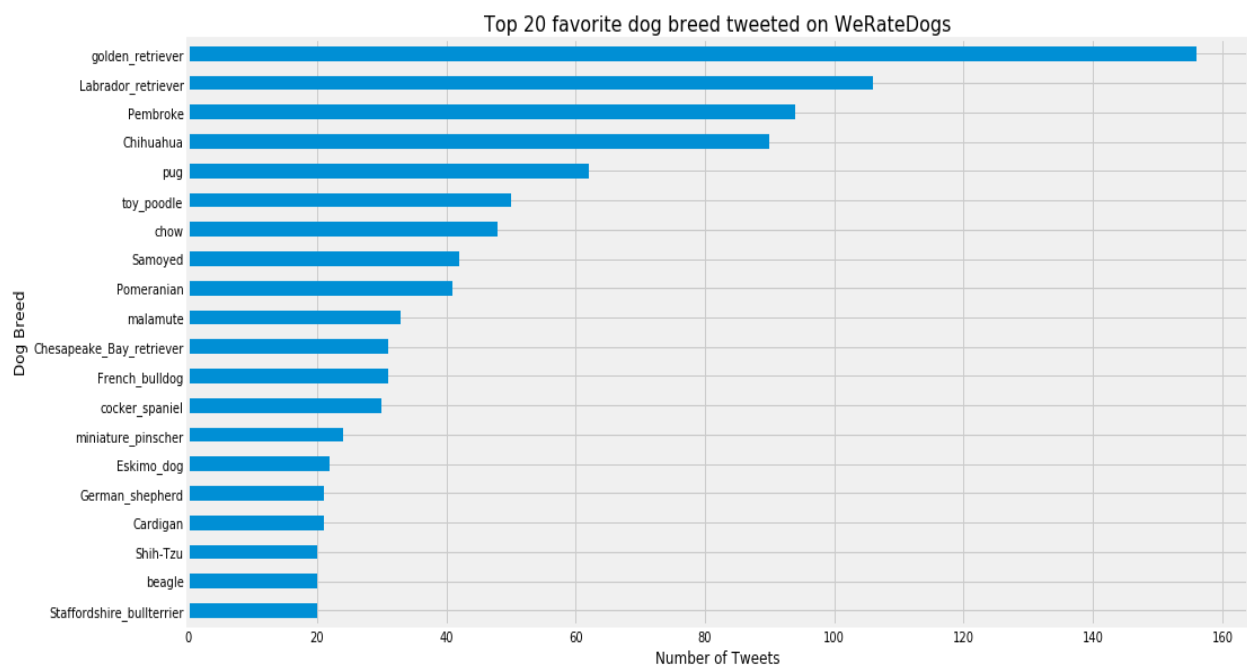
Swadeep Kumar Singh

Introduction

The intention of this report is to communicate the insights and displays the visualization(s) produced from our wrangled data. In this report we go one-by-one over the results of our visualizations which has been done in python (using its libraries) and Tableau. Tableau, which is data visualization tool for creating interactive plots and graphs.
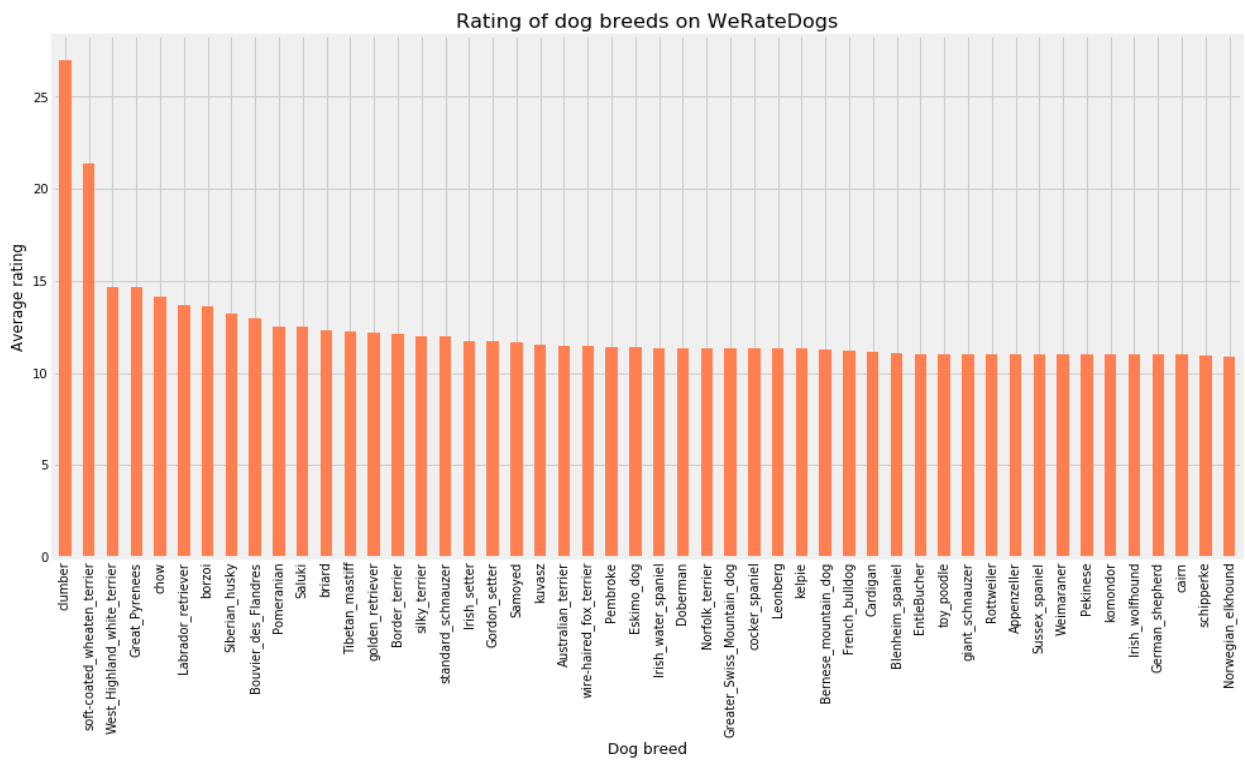
Visual Assessments:

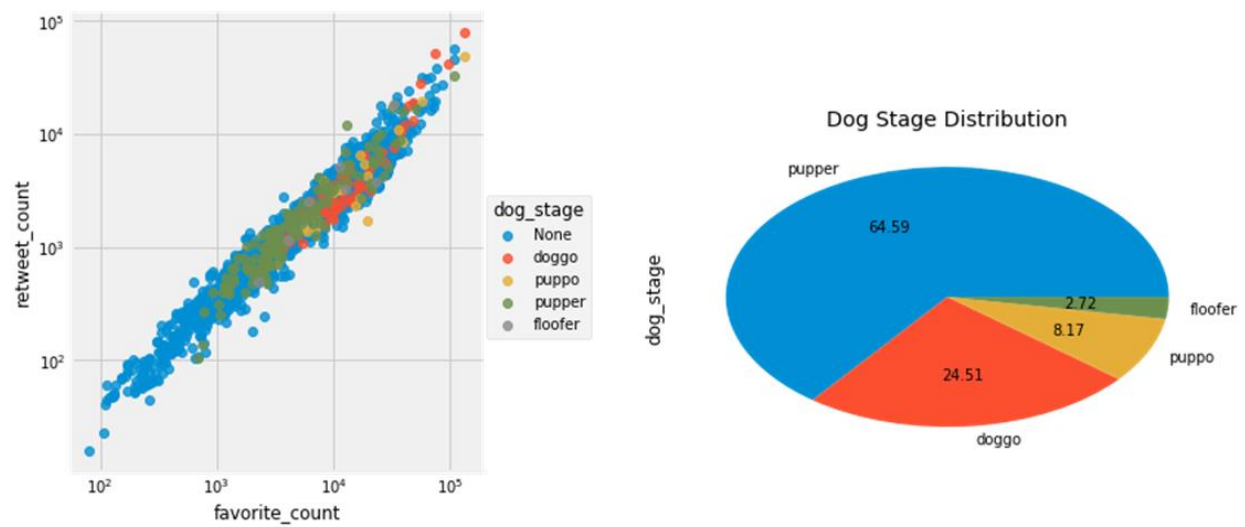### Visualization 1: Top favorite dog breed among twitter users

In the above graph we can see that ***golden retriever*** is the most tweeted breed of dog. The result is represented as vertical bar graph where *x-axis* represents the *number of records* and *y-axis* represents the *dog breeds*.

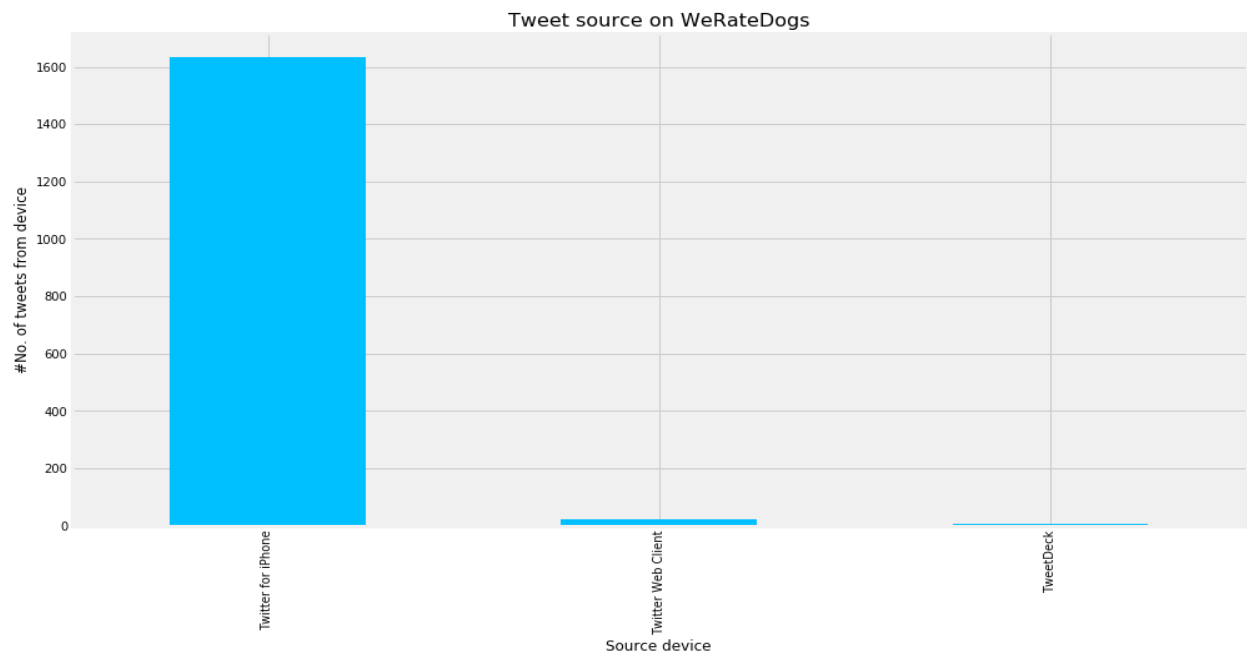## Visualization 2: Show average ratings of dog breeds

The result is represented as bar graph where *x-axis* represents the *dog breeds* and *y-axis* represents the *average ratings*. We can see that the average rating is approx. 12 and highly rated dogs appears most often.

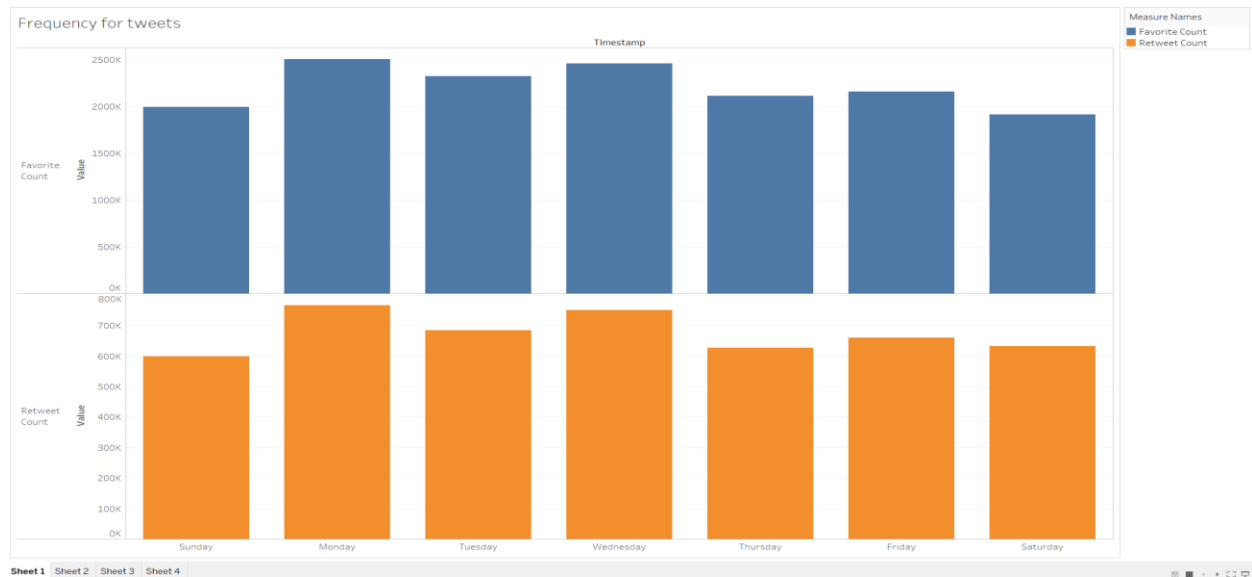## Visualization 3: Log distribution for popularity of dog stages

From log transformation we can see that people are more likely to favorite than to retweet. So when the post reaches a certain group of people it creates the viral effect. The *x and y-axis* have been log-transformed thus creating a linear correlation between retweets and favorites. The legends were shown for most favorite's dog stages. On right we can see the distribution for dog stages excluding *none*.

Visualization 4: Source for tweets

The above graph shows the twitter's users based on the device on which they rated the dogs. Thus, we can conclude that most of the tweets came from **iPhone**.
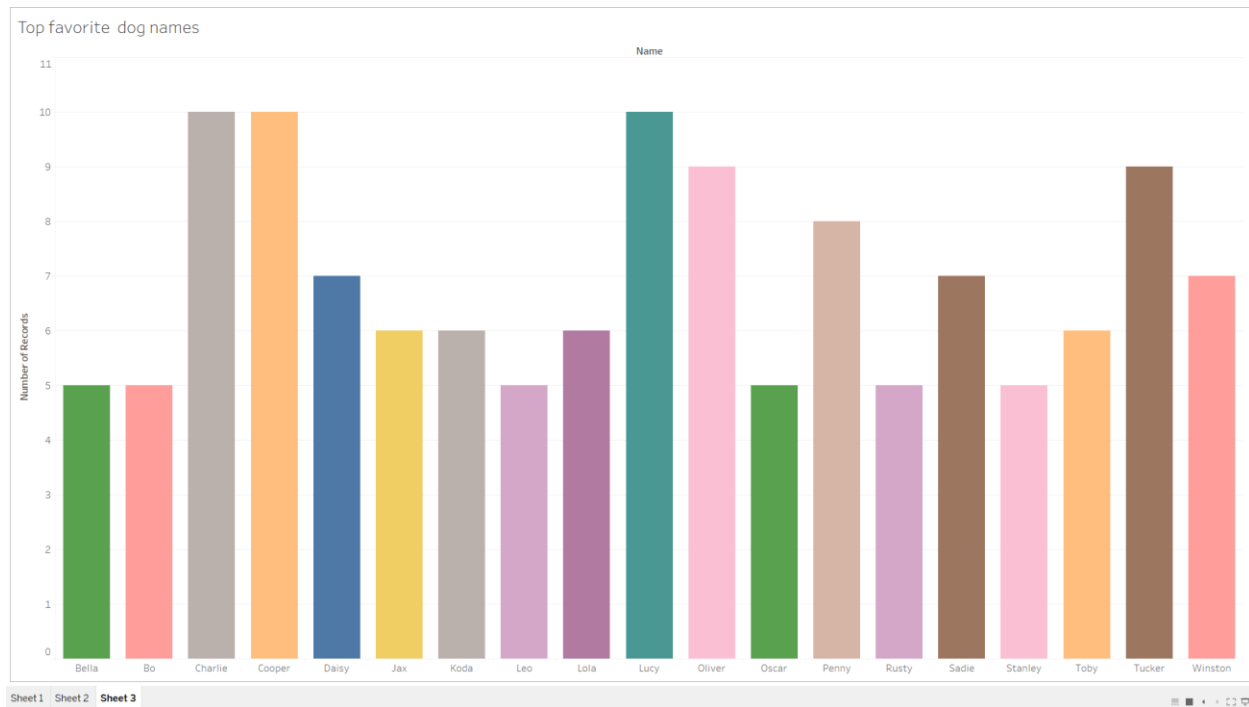
Visualization 5: Favorite count and Retweet count timeline

The above plot is visualized using Tableau, and it shows that most favorite count and retweet count comes on **Monday**.

Measure Names	Weekday of Timestamp	Measure Values
Favorite Count	Monday	2,505,900
Retweet Count	Monday	766,477

Visualization 6: Favorite dog names

Top favorite dog names

This is the interesting graph where we can see that which name is most favored among the dog owners. We can see that **Charlie, Cooper, Lucy, Oliver, and Tucker** are most famous names in this dataset.

Conclusions

After performing the above analysis we can be able to provide top insights as follows:

Top insights after analysis

1. Most favorite dog breed which is tweeted on WeRateDogs are:

- **Golden retriever: 156**

- **Labrador retriever: 106**

- **Pembroke: 94**

2. Most favorite count and retweet count comes on **Monday**.

3. Most favorite dog stage mentioned in tweet is **pupper**.

4. Famous dog names were **Charlie, Cooper, Lucy, Oliver, and Tucker**.

5. Most of the tweets came from **iPhone**.