act_report

September 6, 2022

0.1 Report: act_report

- Create a **250-word-minimum written report** called "act_report.pdf" or "act_report.html" that communicates the insights and displays the visualization(s) produced from your wrangled data. This is to be framed as an external document, like a blog post or magazine article, for example.

0.1.1 Introduction

The dataset that to be analyzed and visualized is the tweet archive of Twitter user [@dog_rates](https://twitter.com/dog_rates), also known as WeRateDogs. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. Why? Because "they're good dogs Brent." WeRateDogs has over 4 million followers and has received international media coverage.

0.1.2 Gather

The project gather data from the following sources.

```
1. The WeRateDogs twitter archived data
2. The tweet image prediction
3. Twitter API and phyton's tweepy library
```

0.1.3 Assessing data

After the data was gathered I began to assess data on both quality and tidness issue.

There are four main issue in quality dimensions

```
1. Completenece
2. Validity
3. Accuracy
4. Consistency
```

There are three main requirement for tidness

```
1. Each variable forms a column
2. Each observation form row
3. Each type of observation unit forms table
```

0.1.4 Clean

The cleaning process involves three steps

1. Define
2. Code
3. Test

0.1.5 Analysis and Visualizations

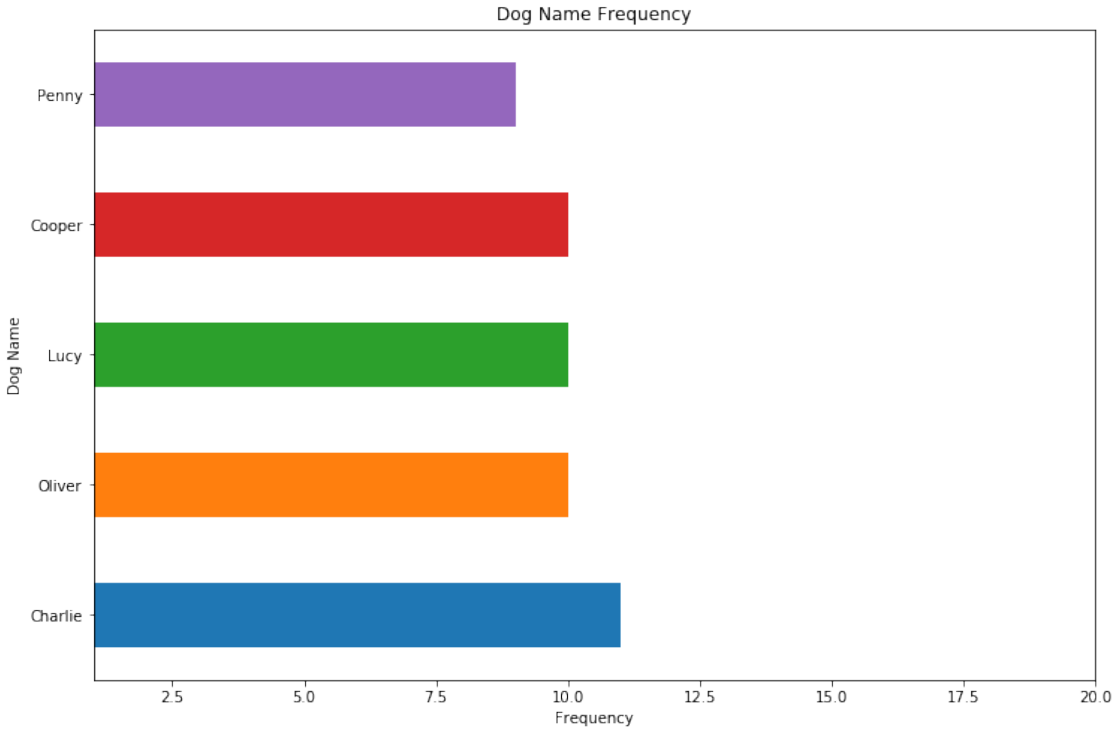
The following inferences were drawn by analyzing and visualizing the data:

0.1.6 Popular Dog Name

The first five dog names are :

1. Charlie
2. Oliver
3. Lucy
4. Cooper
5. Penny

In [4]:

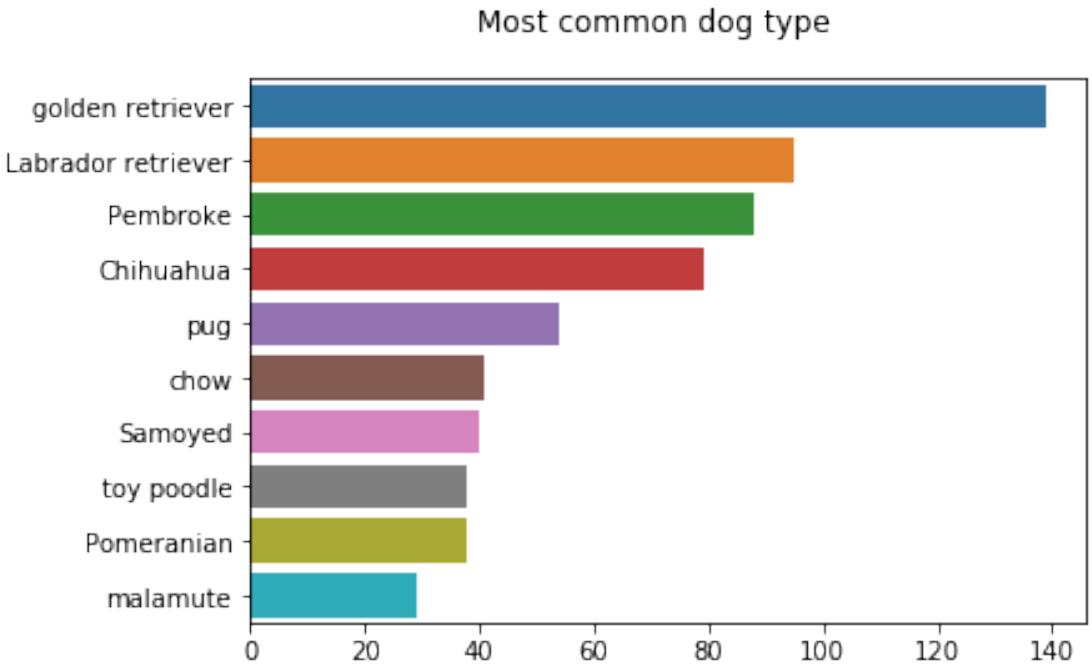


0.1.7 Common dog type

The most common dog type is golden retriever

In [8]:

Out[8]: Text(0.5,0.98,'Most common dog type')

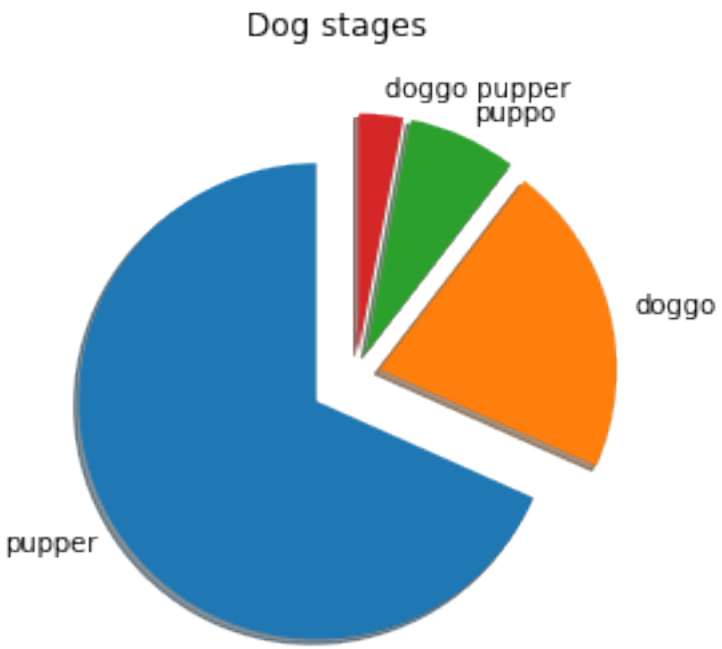


0.1.8 Common Dog stage

The most common dog stage is pupper

In [9]:

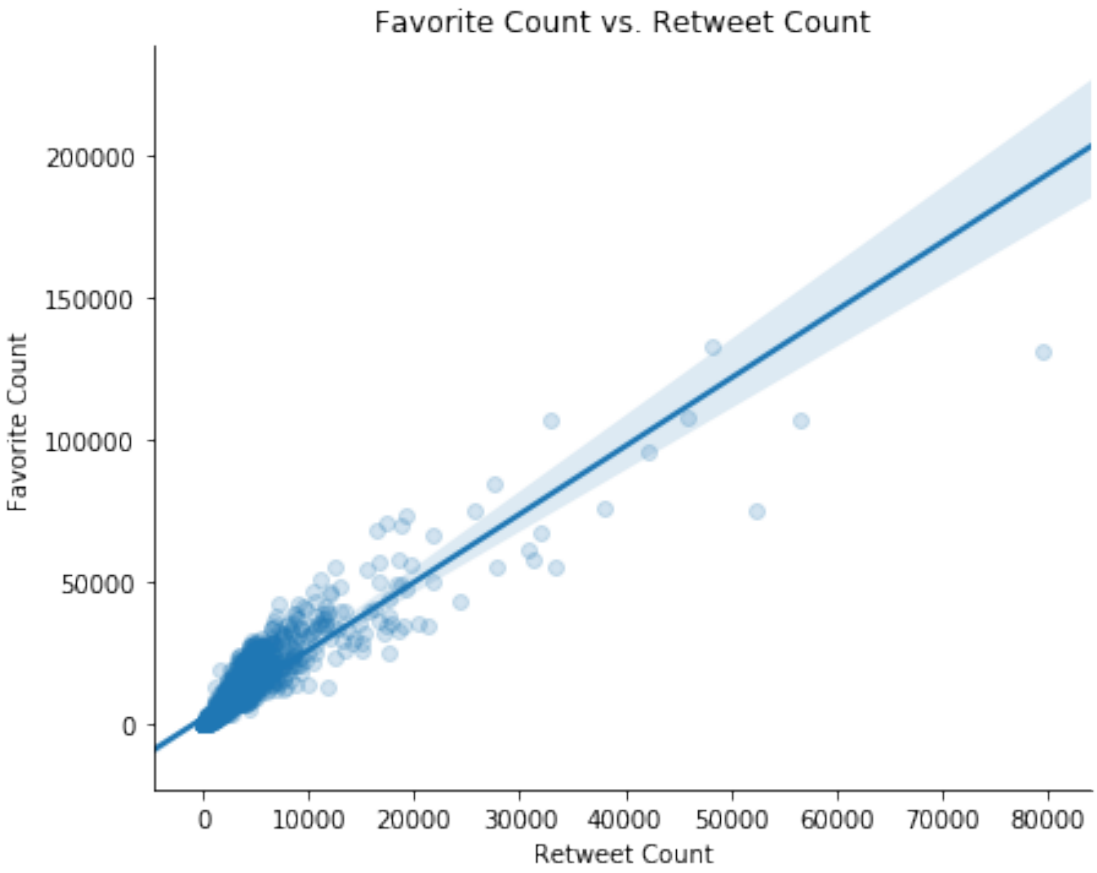
Out[9]: Text(0.5,0.98,'Dog stages')



0.1.9 Favorite Vs Retweet count

There is a positive correlations between favorites and retweet counts.

In [10]:



0.1.10 Conclusion

The write up offers a straight look at the data wrangling process. There is so much more that can be done with the dataset provided.

In [ ]: