

Data analysis of WeRateDogs

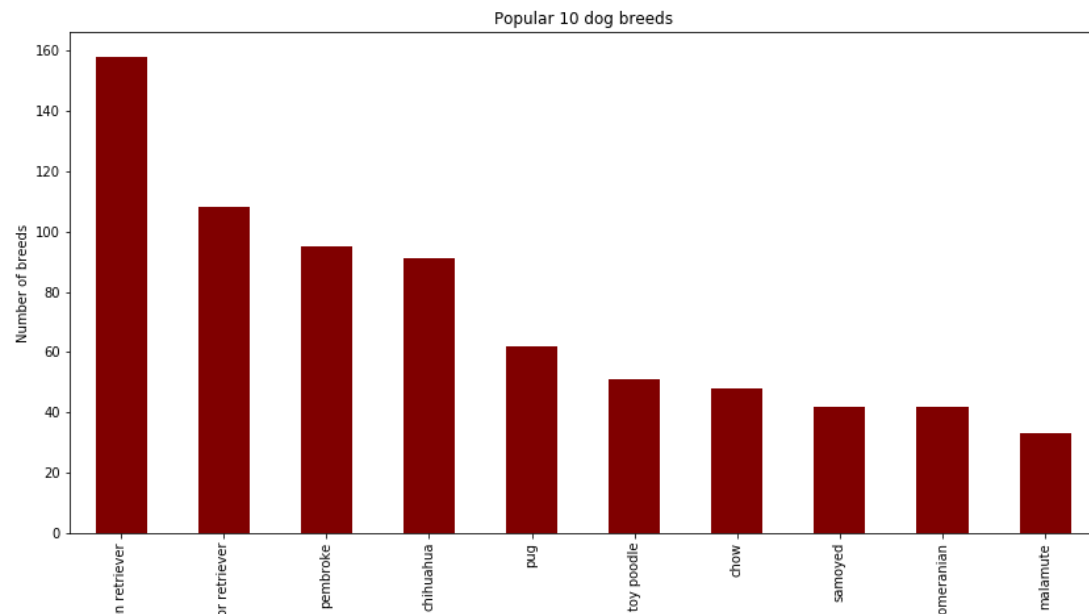
Introduction

Real-world data rarely comes clean. Using Python and its libraries, I will gather data from a variety of sources and in a variety of formats, assess its quality and tidiness, then clean it. This is called data wrangling. I will document my wrangling efforts in a Jupyter Notebook, plus showcase them through analyses and visualizations using Python (and its libraries) and/or SQL.

The dataset that I will be wrangling (and analyzing and visualizing) is the tweet archive of Twitter user [@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.

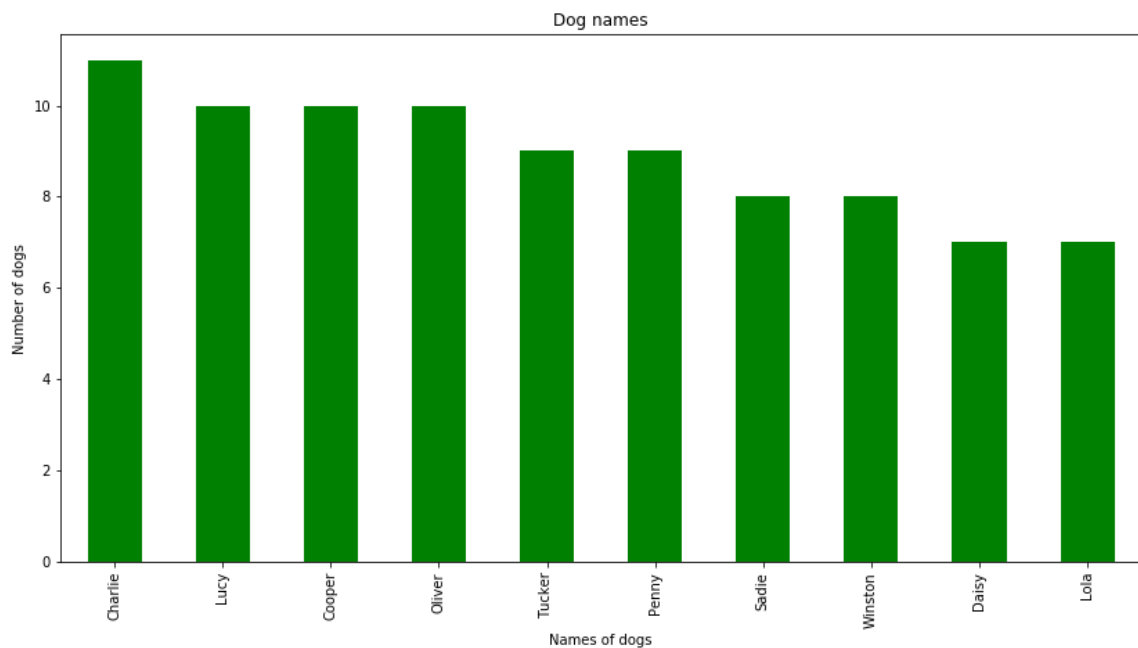
So Which dog breed is most common in the tweet dataset? What dog names are the most popular? What is the most used source? What is the number tweets in every month? What is distribution of rating numerator? What is the the most frequent dog stage? These questions and more are answered in the following insights.

Which dog breed is most common in the tweet dataset?



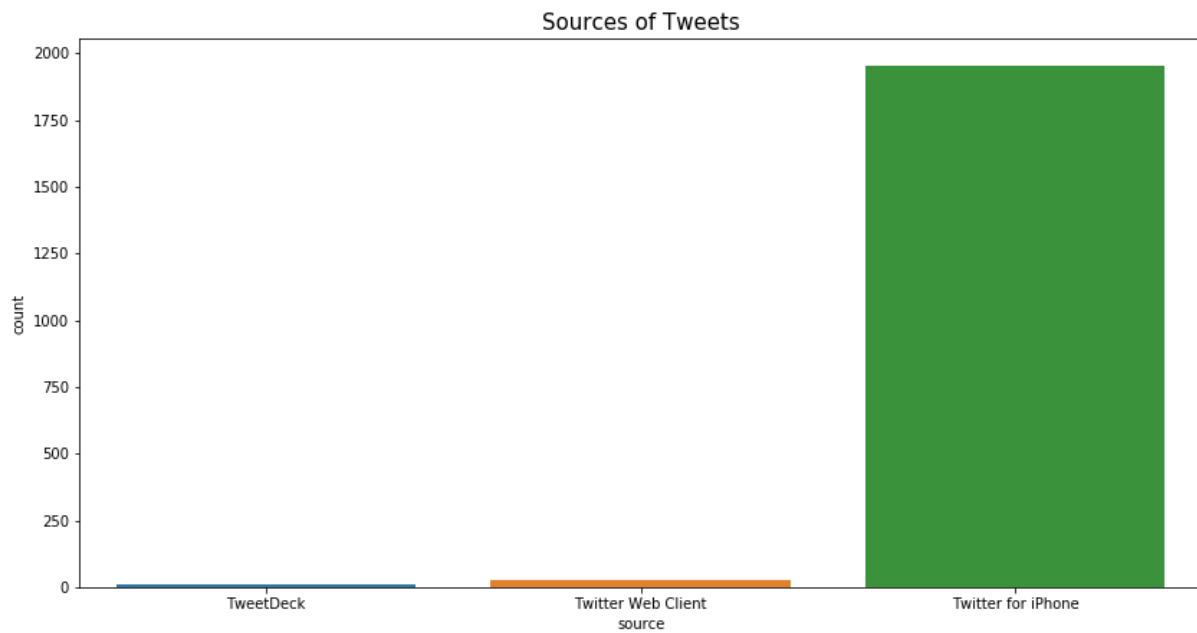
The most common dog breed is Golden retriever in the dataset with 158 predictions.

What dog names are the most popular?



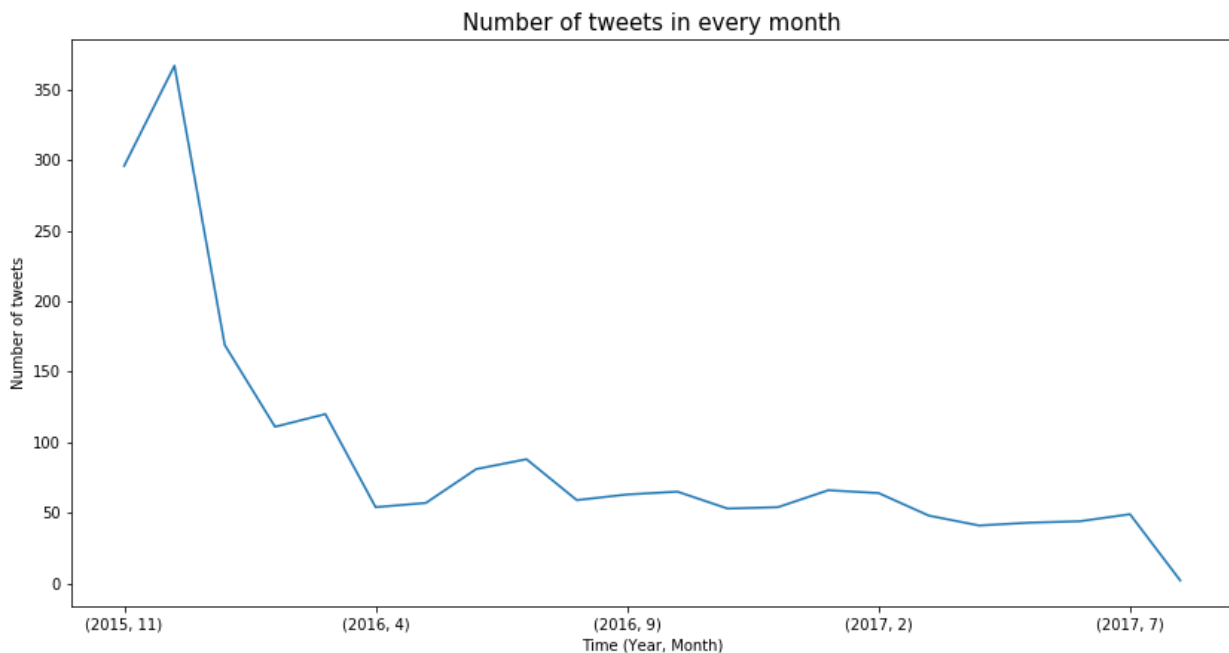
Charlie is the most popular dog name. Meanwhile Lucy, Cooper and Oliver are also famous.

What is the most used source?



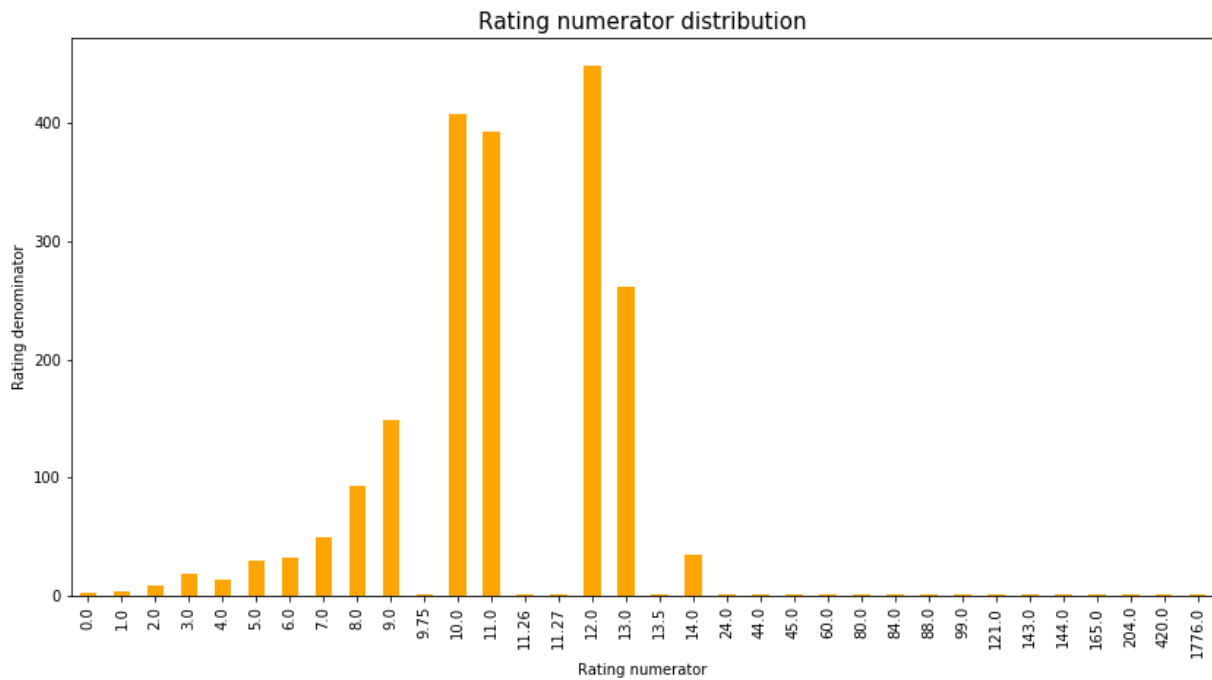
The most popular source is Twitter for iPhone. The rest has small amounts of counts.

What is the number tweets in every month?



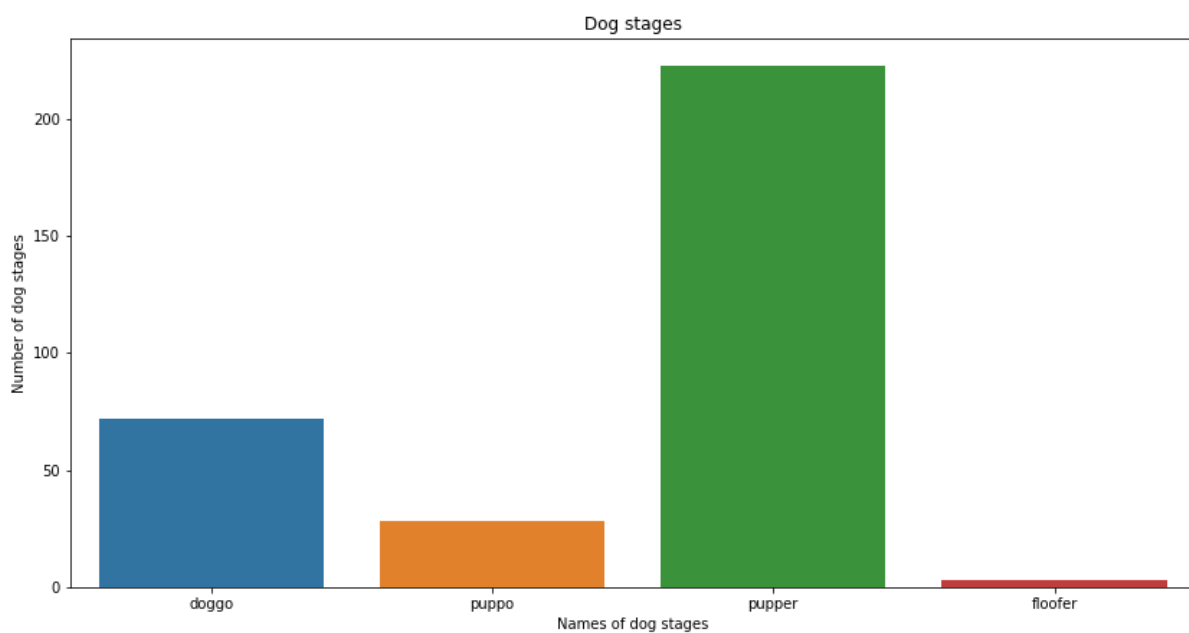
As you can see, December 2015 has the most of tweets. After that period, number tweets decreased drastically. It carried on decreasing till July 2017.

What is distribution of rating numerator?



1156 dogs have ratings which are over than 10.

What is the the most frequent dog stage?



Pupper is the most frequent one with 223. Whereas doggo, puppo and floofer are 72, 28 and 3 respectively.