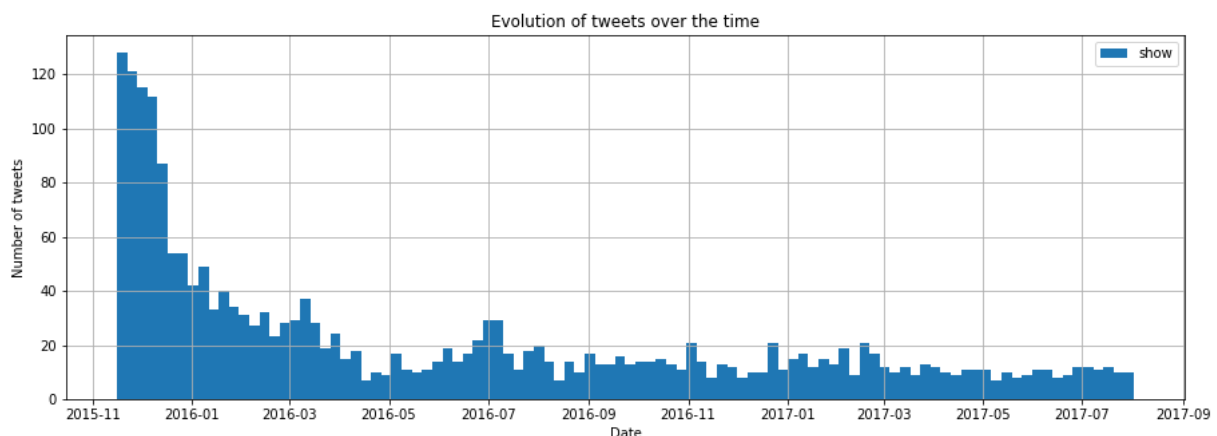# Analyzing and visualizing DATA

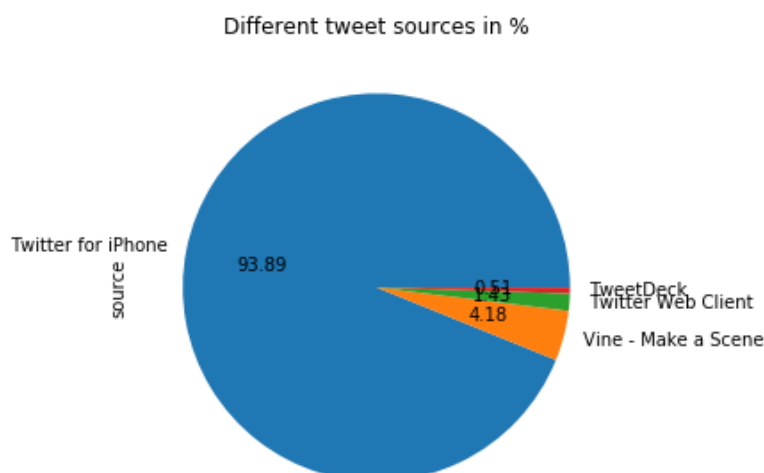In the Exploratory Data Analysis we will try to reply to these questions:

- How does the activity evolve over the time for Twitter user @dog_rates?
- Which source is the most common?
- What is the most frequent rating?
- Which names are the most common?
- Which category is the most common and best rated?
- According to image predictions, what are the 10 most posted breeds?
- According to image predictions, what is the percentage of non dog breed in the images?

1) How does the activity evolve over the time for Twitter user @dog_rates?

Twitter user @dog_rates starts tweeting in November 2015. During the first month this account posted massively compared to next period and the number of tweets keeps dropping till August 2017.

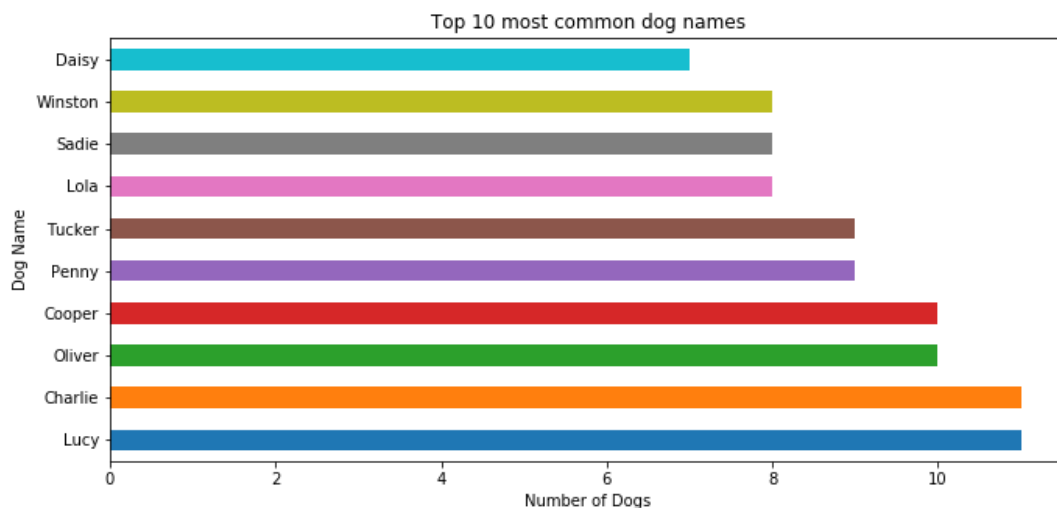## 2) Which source is the most common?

From the four sources of tweets, twitter for iphone is far ahead with 93.89% compared to Vine with 4.18% wich is the second source of tweets.

3) What is the most frequent rating?

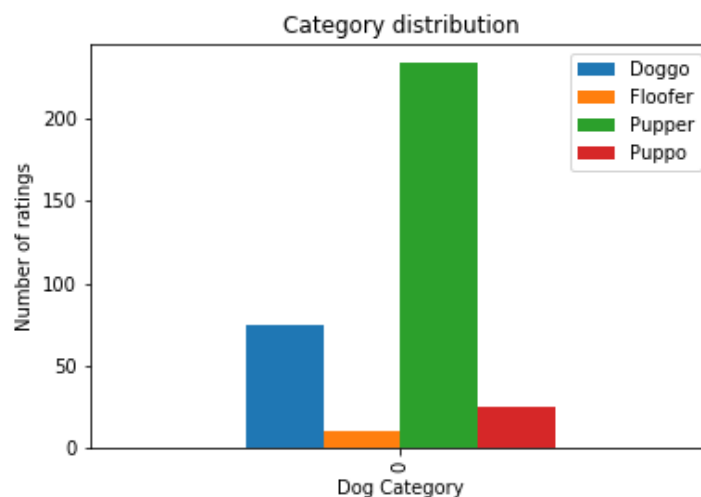| Rating Numerator	Frequency
11.0	1299
12.0	500
13.0	307
14.0	43

The most frequent dog rating for @dog_rates user was 11/10 with 1299 ratings followed by 12/10 with 500 ratings, then 13/10 and 14/10 with respectively 307 and 43 ratings. However, the frequency of 11/10 could be biased by the ratings under 11/10 that we have normalized to 11/10.
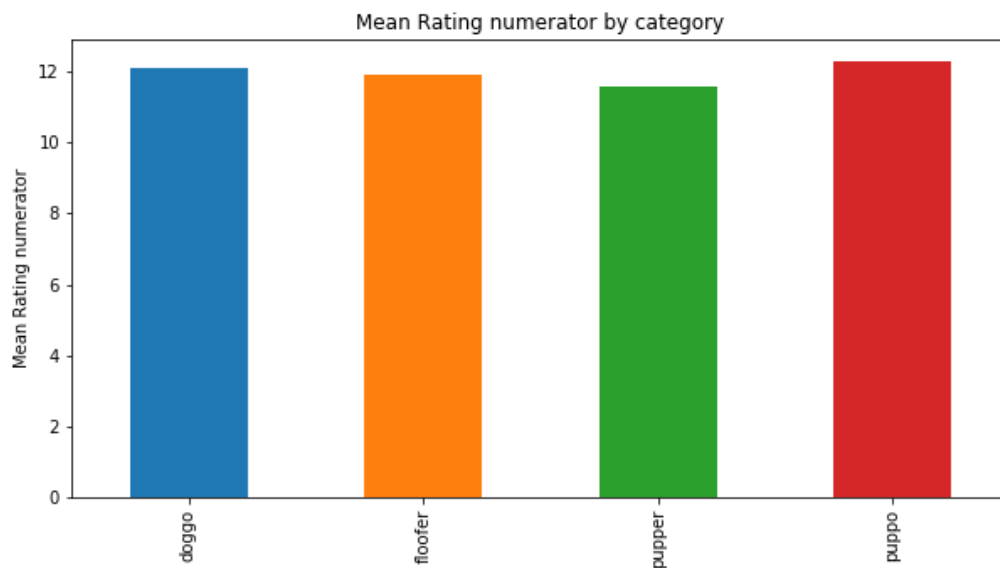
4) Which names are the most common?

Lucy and Charlie are the most common dog names posted with 11 posts followed by Oliver and Cooper with 10 posts and Penny and Tucker with 9 posts.

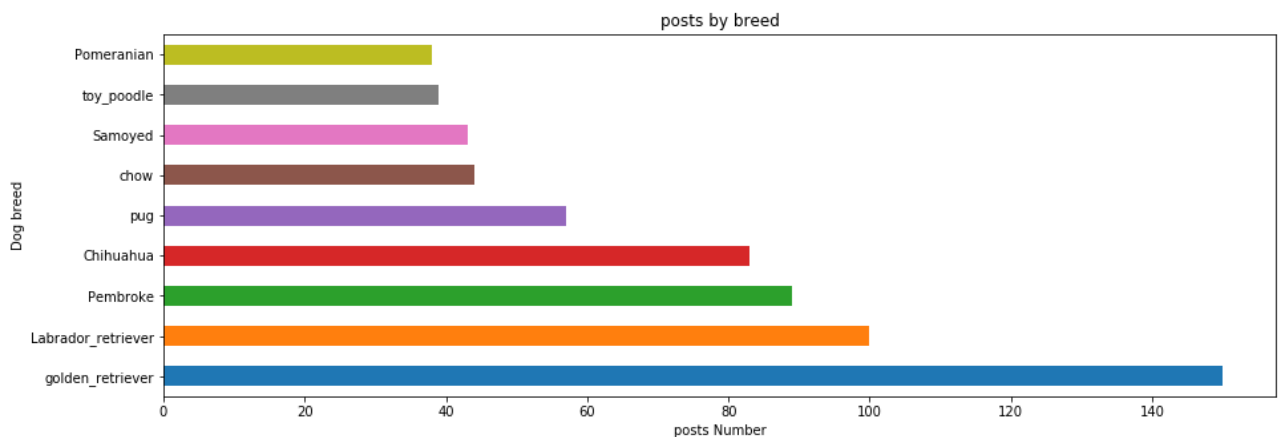### 5) Which category is the most common and best rated?

Pupper is most common category with over 220 posts followed by Doggo with over 70 posts.

Mean Rating numerator by category

Comparing the mean ratings for the four categories, we can say that Puppo have the best rating in average followed by Doggo category, but the difference was insignificant. In addition, large ratings (outliers) could distort this result.

6) According to image predictions, what are the 10 most posted breeds?

posts by breed

According to image predictions inferred by the first algorithm, Golder retriever is the first posted breed with about 150 image posts followed Labrador retriever with approximately 100 image posts.

7) According to image predictions, what is the percentage of non dog breed in the images?

According to the first prediction algorithm, 26.16% of the images posted are not dogs and referring to other species.