

# WERATEDOGS

ANALYSIS AND INSIGHTS



# ANALYSIS QUESTIONS

- When I am done with wrangling the 3 files and merged them in the finale DataFrame I asked myself what do I want to know from this data and what came to my mind as follows:
  1. What is the most used name for dogs in this period?
  2. Is there a relation between favorite count and retweet count?
  3. What is the most favorited breed per year?
  4. What is the most used source over the period to direct any future campaigns “if-any”?
  5. How many tweets per day get a high rating.



## WHAT IS THE MOST USED NAME

I am using Tableau within this project and to answer this question I created a word cloud and it turns out that the top 3 names are “Charlie, Oliver and cooper” those are male names and for female name the top 2 are Sadie and Lucy.



# IS THERE A RELATION BETWEEN FAVORITE COUNT AND RETWEET COUNT

- For this part I did create a viz on tableau that show a good relation between both also I have added some elements and the description of this viz is as follows:

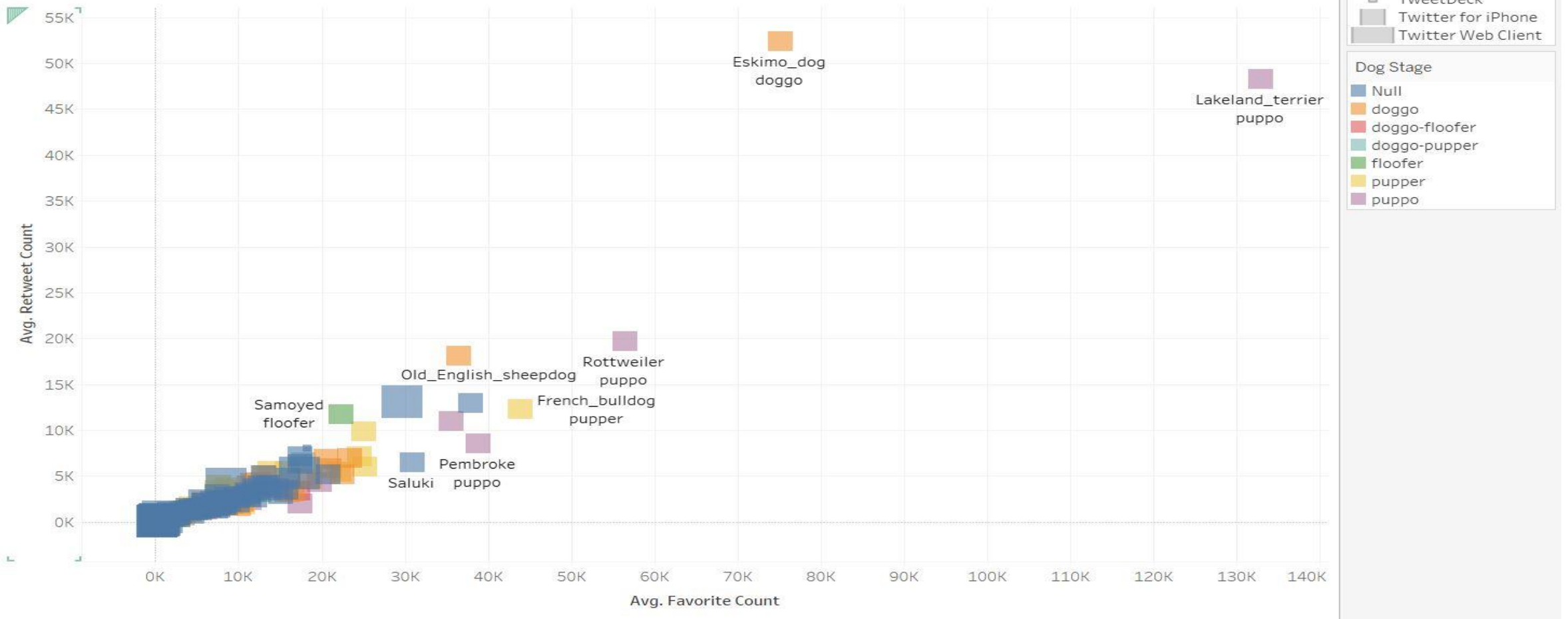
1. Average of Favorite Count vs. average of Retweet Count. Color shows details about Dog Stage. Size shows details about Source. The marks are labeled by Breed and Dog Stage. The view is filtered on Breed and Exclusions (Breed,Dog Stage,Source). The Breed filter keeps 81 of 78 members. The Exclusions (Breed,Dog Stage,Source) filter keeps 188 members.

2. Also:

<b>Rows:</b>	Avg. Retweet Count
<b>Columns:</b>	Avg. Favorite Count
<b>Filters:</b>	Breed, Exclusions (Breed,Dog Stage,Source)
<b>Text:</b>	Breed and Dog Stage
<b>Color:</b>	Dog Stage
<b>Size:</b>	Source

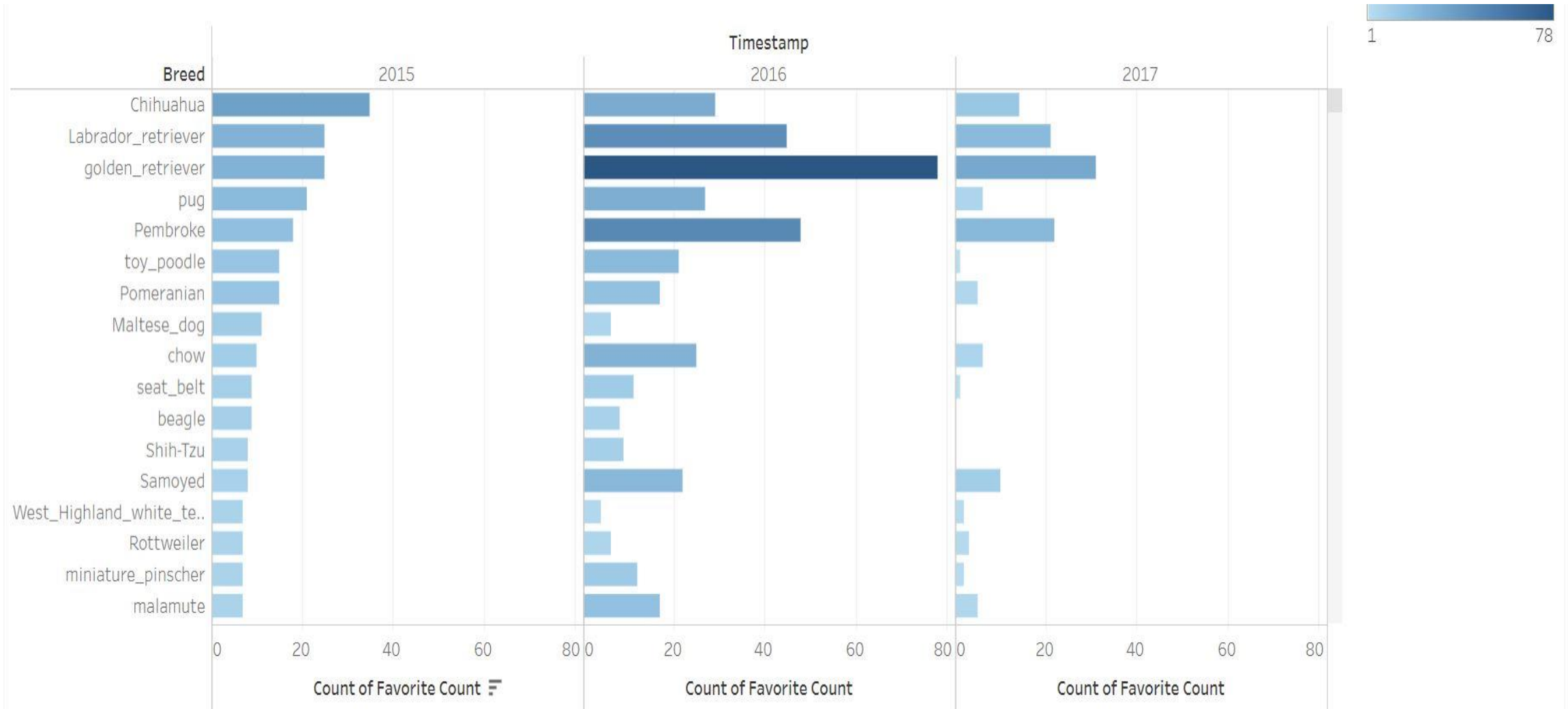
# IS THERE A RELATION BETWEEN FAVORITE COUNT AND RETWEET COUNT

Average retweet vs average favorite





# WHAT IS THE MOST FAVORITED BREED PER YEAR?



# WHAT IS THE MOST FAVORITED BREED PER YEAR?

- Count of Favorite Count for each Breed broken down by Timestamp Year. Color shows count of Favorite Count. The data is filtered on Timestamp Year, which ranges from 2015 to 2017
  1. For 2015 it is chihuahua with 35 records
  2. For 2016 it is golden retriever with 78 records
  3. For 2017 it is golden retriever with 31 records

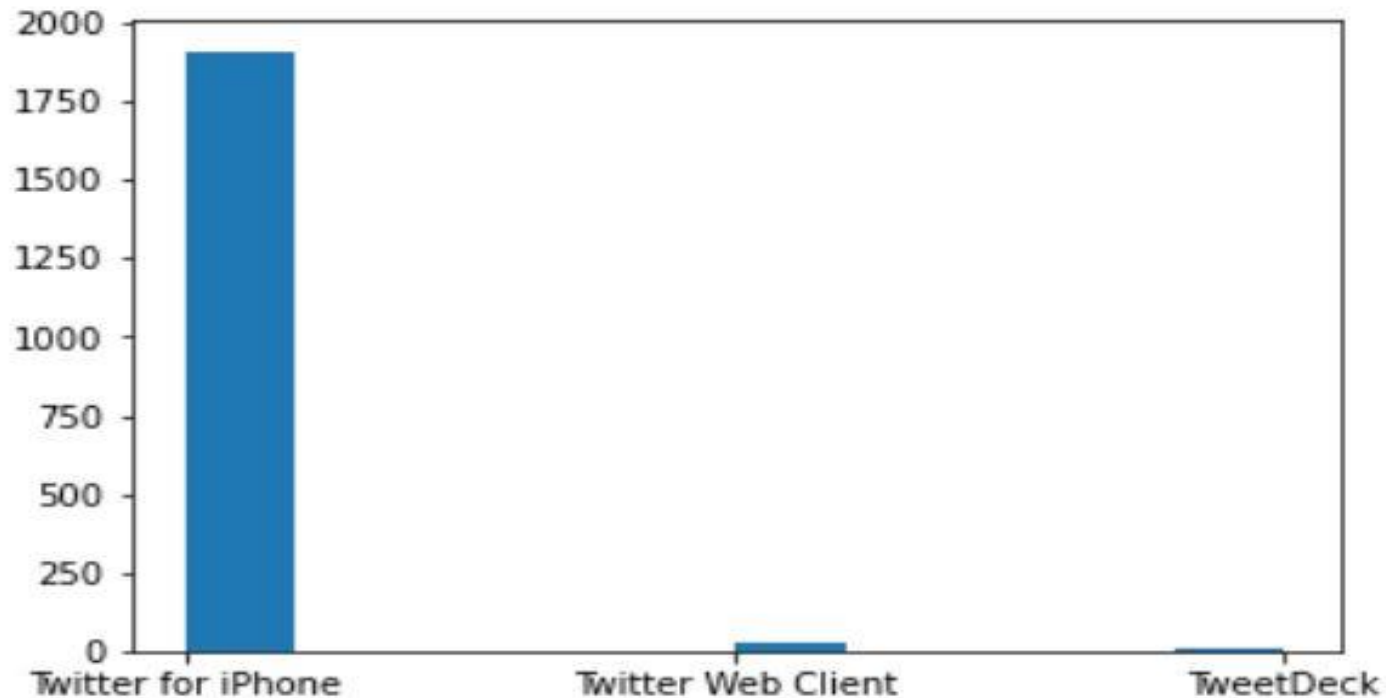
Its seems that retrievers are the most liked breed “I like them myself” 😊

<b>Rows:</b>	Breed Sorted descending by count of Favorite Count within Breed, broken down by 2015.
<b>Columns:</b>	Year of Timestamp, Count of Favorite Count
<b>Filters:</b>	Year of Timestamp
<b>Color:</b>	Count of Favorite Count



## WHAT IS THE MOST USED SOURCE OVER THE PERIOD TO DIRECT ANY FUTURE CAMPAIGNS “IF-ANY”?

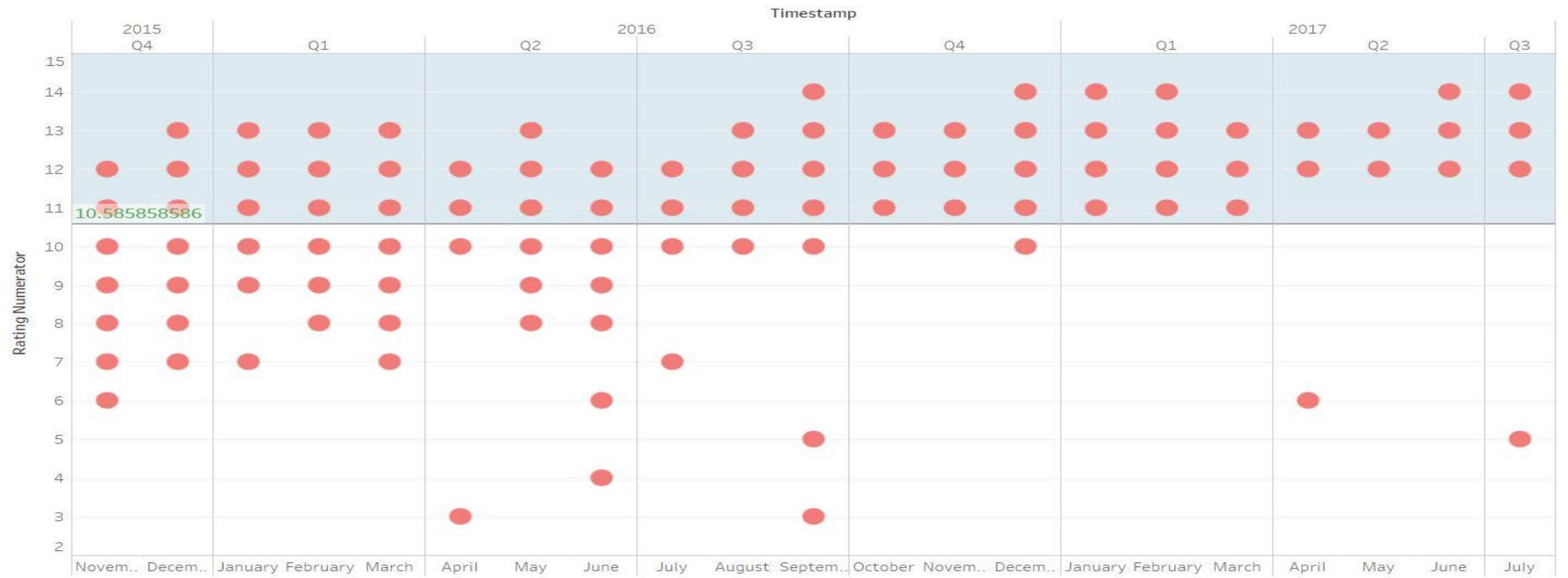
- For this one I used a simple pyplot function to get the frequency using a histogram:
- And it seems twitter for iphone got it all also I noticed that there is no twitter deck for 2017 but not showing in this viz.



# HOW MANY TWEETS PER DAY GET A HIGH RATING.

- It will be useful to know how many tweets getting a high rating per day and check those urls to know why they have a higher rating and other not but for clarity I did the viz by quarter/month. And the average about 10.58.

count of tweets with high rating monthly





## HOW MANY TWEETS PER DAY GET A HIGH RATING.

- Rating Numerator for each Timestamp Month broken down by Timestamp Year and Timestamp Quarter. The data is filtered on maximum of Breed, which excludes None.

**Rows:** Rating Numerator

**Columns:** Year of Timestamp, Quarter of Timestamp, Month of Timestamp

**Filters:** Max. Breed

**Level of detail:** Count of Tweet Id