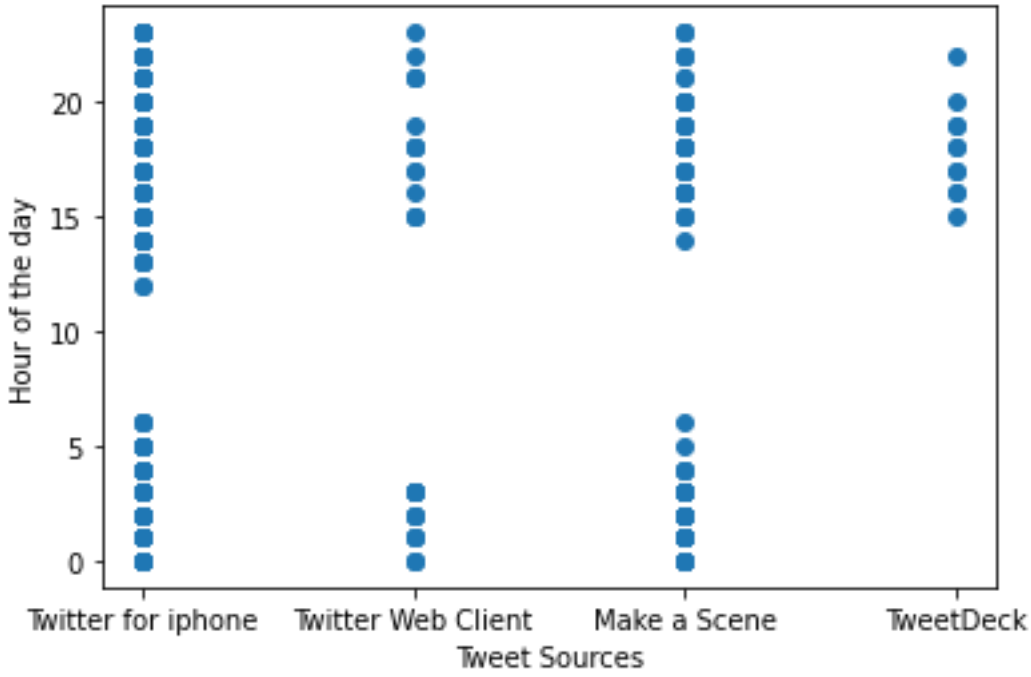
In analysis and visualization phase I was able to gather 3 analytics and one done with the aid of a visualization plot which is a scatter plot.

For my first analysis, I wanted to know how tweets were sent throughout the day(to know which time tweeters are active), so I had to manipulate my timestamp data and extract the hours and create a separate column for it. I also manipulated the source string values where I converted them from links to easily understandable texts such as 'Twitter for iphone'

The following was the plot I made:



As per the visualization I was able to analyze the following:

From the scatter plot above we can draw some straightforward conclusions:

1. Many of the sources do not tweet between 0600 hours and 1200 hours.
2. Most users use iphones to tweet.
3. Tweetdeck users mostly tweet later in the day before midnight and during afternoon hours.

For my second analysis I wanted to know which dog image had the most likes(Highest number of favorites) to do this I pulled the tweet id corresponding to highest favorite count and extracted the image link. The image link that was still functional was the one from the image-predictions dataset. Here it is

For my third analysis I wanted to know which image was best predicted by the neural network by extracting the one with the highest value of confidence, (p1_conf) where the highest was 0.999956 and the dog breed was **Komondor** and its id was 668636665813057536

The following was the best predicted image in the first prediction: