Insights from we rate dogs

This report gives insights drawn form the we rate dogs tweet, this tweet data is spread across three data frames; the tweet archive, the image predictions and scraped data from we rate dog twitter API.

Data Analysis and Visualization

1. The popular names of dogs in we rate dogs tweets

The image below shows the most used names for dogs to be Oliver, Cooper and Charlie as dogs bearing these names appear 10 times each in the we rate dogs tweets

See the popular names of dog

```
In [92]: #Check to see the popular dog names
         tweet_df[['name']].value_counts()

Out[92]: name
         Unknown          587
         Oliver            10
         Cooper            10
         Charlie           10
         Penny              9
         Tucker             9
         Lucy               9
         Sadie              8
         Winston            8
         the                7
         Toby               7
         Daisy              7
         Lola               7
         Bella              6
         Bo                 6
         Jax                6
```

2. The source of the we rate dog tweets

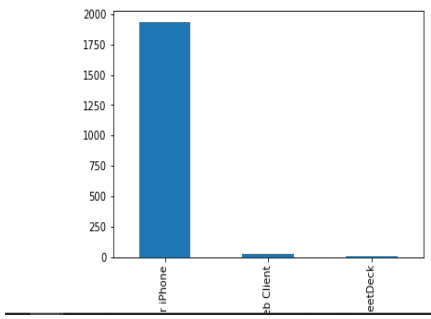
2 The sources of we rate dog tweets

```
In [93]: tweet_df['source'].value_counts()

Out[93]: Twitter for iPhone    1931
         Twitter Web Client      28
         TweetDeck               9
         Name: source, dtype: int64

In [94]: tweet_df['source'].value_counts().plot(kind='bar')

Out[94]: <AxesSubplot:>
```



From the image above we can see that most tweets in the we rate dogs tweets were made from an Iphone, with tweetdeck being the least possible source of tweet with just 9 tweets from it.

3. Engagements

I check for the dog and tweet with most engagements, I did this by adding up the retweets and favorite count, the dog with the highest engagement had no name, it is a doggo and had a rating of 13/10.the tweet was made from an Iphone.

```
[95]: #create a total engagement column so we can get which we rate dogs tweet got more reactions from fans
      tweet_df['total_engagement'] = tweet_df['retweet_count'] + tweet_df['favorite_count']
```

```
[96]: # Using sort values to see which dog has the most engagements
      tweet_df.sort_values(['total_engagement'], ascending=False)
```

t[96]:

	tweet_id	timestamp	source	text	
762	744234799360020481	2016-06-18 18:26:18+00:00	Twitter for iPhone	Here's a doggo realizing you can stand in a pool. 13/10 enlightened af (vid by Tina Conrad) https://t.co/7wE9LTEXC4	https://twitter.com/dog_
305	822872901745569793	2017-01-21 18:26:02+00:00	Twitter for iPhone	Here's a super supportive puppo participating in the Toronto #WomensMarch today. 13/10 https://t.co/nTz3FtorBc	https://twitter.com/dog_
394	807106840509214720	2016-12-09 06:17:20+00:00	Twitter for iPhone	This is Stephan. He just wants to help. 13/10 such a good boy https://t.co/DkBYaCAg2d	https://twitter.com/dog_
797	739238157791694849	2016-06-04 23:31:25+00:00	Twitter for iPhone	Here's a doggo blowing bubbles. It's downright legendary. 13/10 would watch on repeat forever (vid by Kent Duryee) https://t.co/YcXgHfp1EC	https://twitter.com/dog_
107	866450705531457537	2017-05-22 00:28:40+00:00	Twitter for iPhone	This is Jamesy. He gives a kiss to every other pupper he sees on his walk. 13/10 such passion, much tender https://t.co/wk7TfysWHr	https://twitter.com/dog_
58	879415818425184262	2017-06-26 19:07:24+00:00	Twitter for iPhone	This is Duddles. He did an attempt. 13/10 someone help him (vid by Georgia Felici) https://t.co/UDT7ZkcTgY	https://twitter.com/dog_
328	819004803107983360	2017-01-11 02:15:36+00:00	Twitter for iPhone	This is Bo. He was a very good First Doggo. 14/10 would be an absolute honor to pet https://t.co/AdPKrI8BZ1	https://twitter.com/dog_

4. The most featured dog stage in the dataset

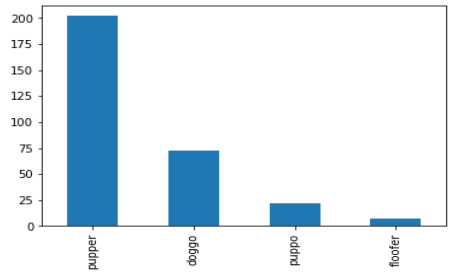
From the distribution of dogstages in the dataset, pupper(a doggo that is inexperienced or younger) is the most frequent dogstage, there are limitations to this as we have a lot of missing data in the dogstage so we cant really draw conclusions from this distributions

```
]1]: tweet_df['dog_stages'].value_counts()
```

```
]1]: pupper     202
     doggo      73
     puppo      22
     floofer    7
     Name: dog_stages, dtype: int64
```

```
]2]: tweet_df['dog_stages'].value_counts().plot(kind='bar')
```

```
]2]: <AxesSubplot:>
```



The floofer is the least represented dog stage in the dataset.

5. The relationship between the dog stages and the engagements

I checked to see the relationship between the dog stages and its effect on reactions to the tweet, from the insight drawn, the more a dog is a doggo or pupper the more engagements it got from we rate dogs fans

```
107]: df.plot(kind='bar')

107]: <AxesSubplot:xlabel='dog_stages'>
```