

# Act\_Report

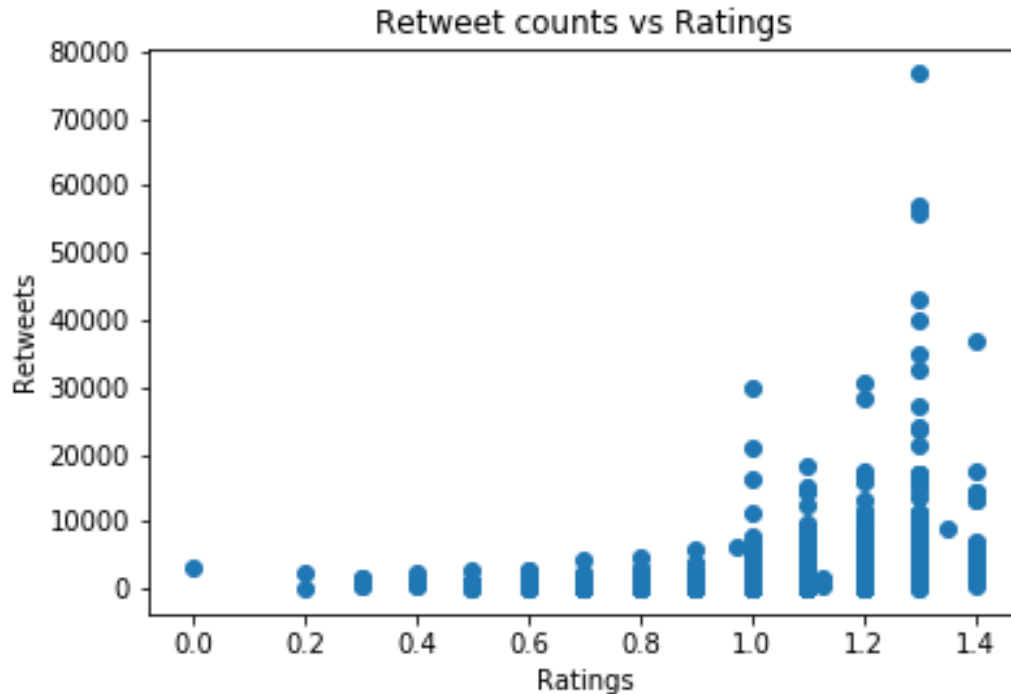
August 17, 2020

## 1 Analysing and visualising Data: By Hiten Naran

Now having spent a considerable amount of time cleaning data, I have been able to obtain a few observations and insights based on a series of questions against the cleaned data. **Question 1:** Do we see a relationship between ratings and retweet counts?

```
In [2]: from PIL import Image
        pil_im = Image.open('retweet_ratings.png')
        pil_im
```

Out[2]:

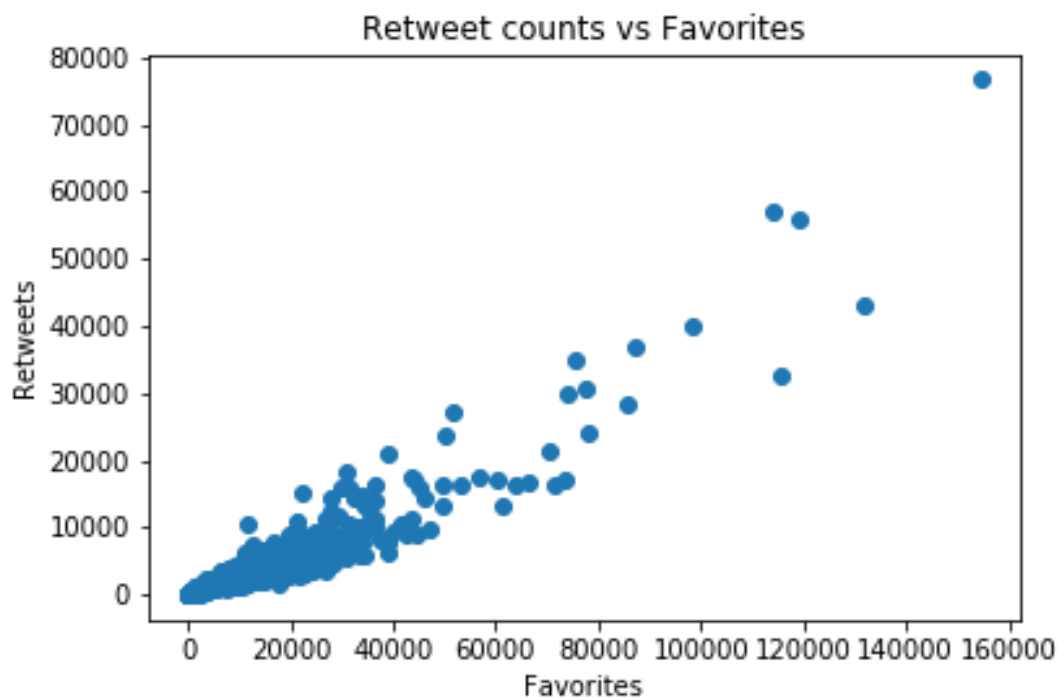


**Conclusion:** The relationship between between retweets and ratings looks weak. Though when ratings exceed 1.0, it does seem like we see a higher volume of retweeted tweets. I wouldn't confidently declare a strong relationship between retweets and ratings

**Question 2:** Do we see a relationship between favorites and retweet counts?

```
In [4]: pil_im_2 = Image.open('retweet_favorites.png')
        pil_im_2
```

Out[4]:

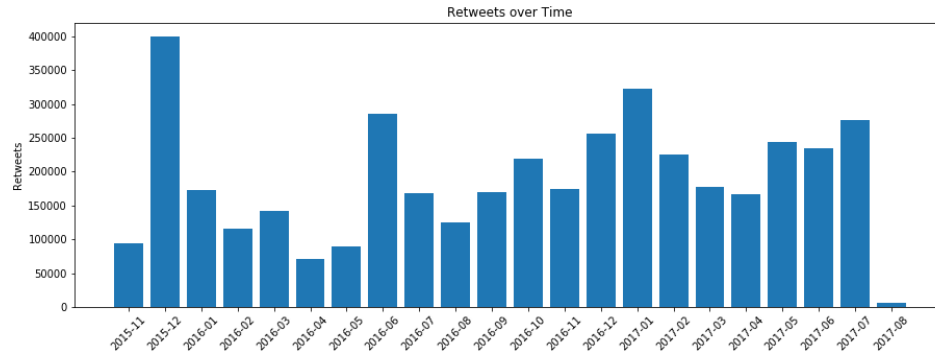


**Conclusion:** There looks to be a fairly strong relationship between retweets and favorites. From looking at this visual I would be confident in saying that the larger the number of retweets the more likely it is to generate a favorite. This is probably due to the fact that more people are seeing the tweet hence the greater likelihood in a tweet being favorited.

**Question 3:** Is there a trend to be seen for retweet counts over time?

```
In [5]: pil_im_3 = Image.open('retweet_time.png')
        pil_im_3
```

Out[5]:



**Conclusion:** There doesn't seem to be an obvious trend observed over time. The only stand-out stat is that 'December 2015' was the month which generated the most cumulative number of retweets.

**Question 4:** What was the most rated Dog Type?

```
In [11]: import pandas as pd
         pd.read_csv('dog_type_ratingcount.csv').head(10)
```

```
Out[11]:
```

	type_dog	rating_count	percentage_rating_count
0	golden_retriever	157	0.093564
1	Labrador_retriever	106	0.063170
2	Pembroke	95	0.056615
3	Chihuahua	91	0.054231
4	pug	62	0.036949
5	toy_poodle	50	0.029797
6	chow	48	0.028605
7	Pomeranian	42	0.025030
8	Samoyed	41	0.024434
9	malamute	33	0.019666

**Conclusion:** It looks like 'golden\_retriever' and 'Labrador\_retriever' were the most rated dog types, featuring across ~ 15.6% of tweets.

```
In [ ]:
```