

Report: act_report

- This report communicates the insights and displays the visualization(s) produced from the wrangled data. This is to be framed as an external document.

The following insights were derived from the data.

rating numerator

- 1.71 percent of dogs are above the average rating numerator value. The average rating numerator value is 19.19/10, The highest rating numerator value is 2016 while the lowest rating numerator value is 0. The modal rating numerator value is 12, the median value is 11, while the mean rating_numerator value is 19.91. The rating numerator distribution is right skewed.

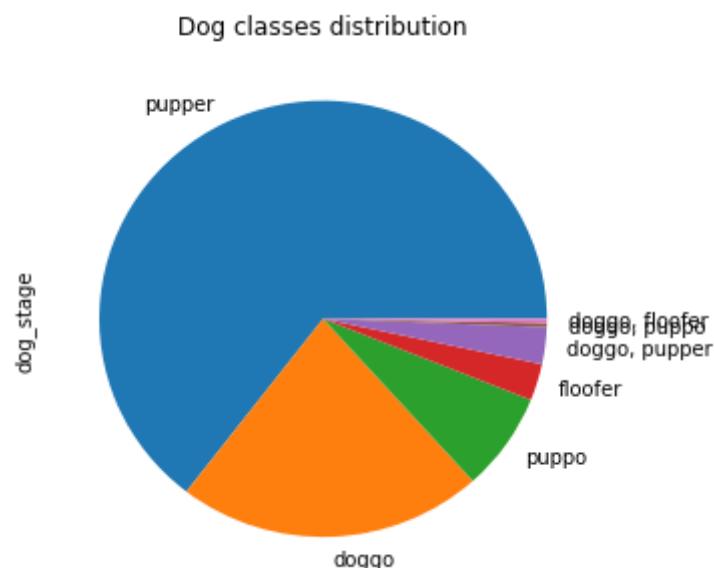
favourite count

- 29.55 percent of dogs are above the average favourite_count. The average favourite_count value is 7054.12. The favourite_count is right skewed with 75% entries between 0 and 8604.

favourite count grouped by rating numerator

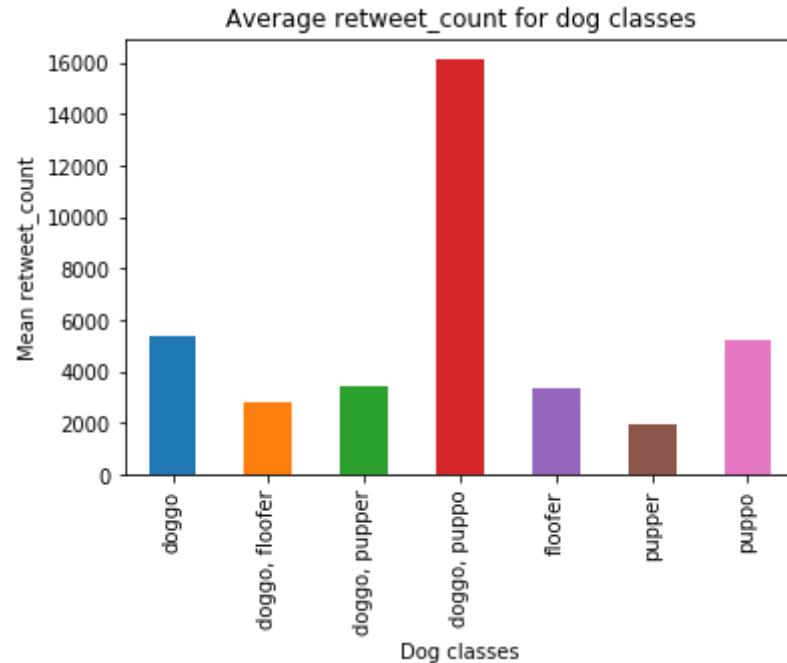
- The [boxplot.png](https://viewf6b31853.udacity-student-workspaces.com/view/Images/boxplot.png) (<https://viewf6b31853.udacity-student-workspaces.com/view/Images/boxplot.png>) shows that the favourite_count grouped by the rating_numerator values. While considering only the integer values ranging from 10 to 14, the rating_numerator of 14 has the maximum favourite_count and the highest median favourite_count followed by 13, 12, 11 and then 10. Also, from the boxplots, the higher the rating, the greater the favourite_counts if the numerator values considered are between 10 and 14. This shows a kind of positive correlation between the rating numerator values between 10 and 14 and the favourite_count values

The following chart shows the dog classes distribution.



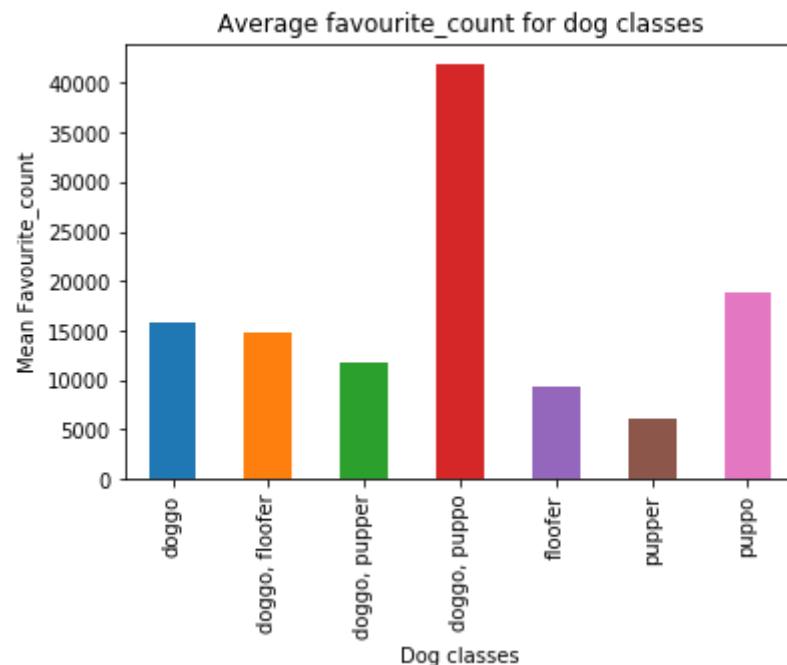
- This pie chart shows that the pupper class has the highest number of dogs, followed by the doggo class.

Average retweet count by dog classes



- The average retweet_count plot shows that the doggo, pupper class has the highest mean retweet count, followed by the doggo and pupper class. This shows that if we had the same number of dogs in all the classes, the doggo, pupper class would have the highest retweet_count, followed by the duggo and pupper class.

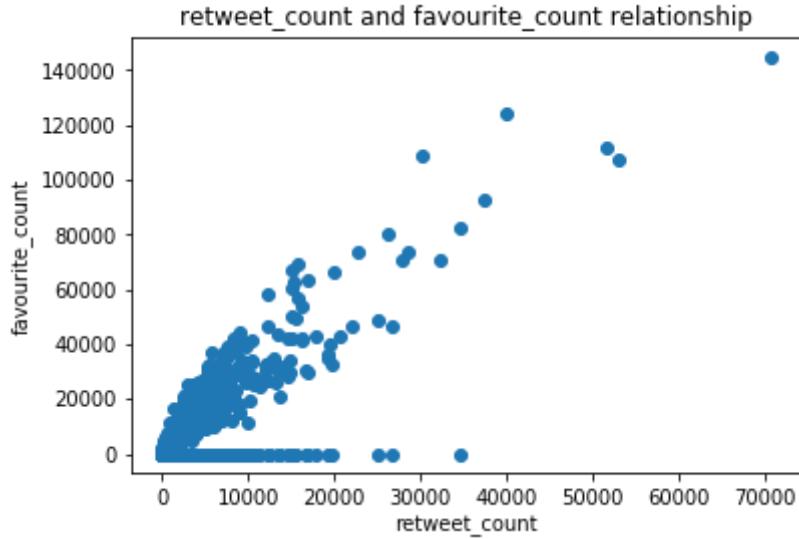
Average favourite count by dog classes



- Considering the mean favourite_counts for the dog classes, doggo, pupper class has the highest average number of favourite_count of above 42,000. This is followed by pupper, and the doggo class.

It seems that the doggo, puppo class has the highest retweet and favourite counts, and the charts are somewhat in agreement, hence the need to find the correlation between the favourite count and retweet.

Correlation between favourite_count and retweet_count



- The correlation coefficient is **0.80**. There is a **high positive correlation** between retweet_count and favourite_count. The retweet_count and favourite_count are positively correlated.

Top Favourite dogs

```
In [87]: favourite_dogs.head(5)
```

```
Out[87]:
```

	favourite_count	jpg_url	rating_numerator	retweet_count	first_prediction	name
842	144883.0	https://pbs.twimg.com/ext_tw_video_thumb/74423...	13.0	70732.0	Labrador_retriever	None
342	124115.0	https://pbs.twimg.com/media/C2tugXLXgAArJO4.jpg	13.0	39923.0	Lakeland_terrier	None
439	111697.0	https://pbs.twimg.com/ext_tw_video_thumb/80710...	13.0	51681.0	Chihuahua	Stephan
117	108906.0	https://pbs.twimg.com/media/DAZAUFBXcAAG_Nn.jpg	13.0	30210.0	French_bulldog	Jamesy
877	107250.0	https://pbs.twimg.com/ext_tw_video_thumb/73923...	13.0	52899.0	Eskimo_dog	None

The table above shows the top five most favourite dogs.

Most favourite dog 1



This is the Number 1 favourite dog. It is a Labrador_retriever. It was rated as 13.0 /10, favourite count is 144883.0, retweet count is 70732.0

Most favourite dog 2



This is the Number 2 favourite dog. It is a Lakeland_terrier. It was rated as 13.0/10, favourite count is 124115.0, retweet count is 39923.0

Highest rated dog

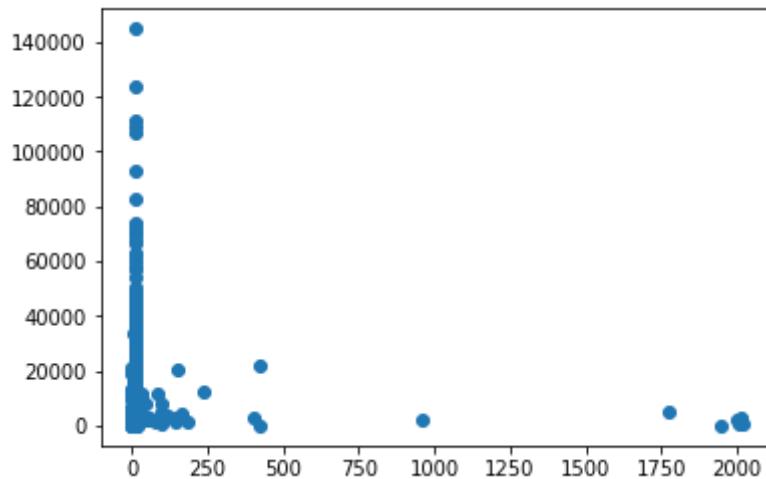


This is a dog with a highest rating. It is a standard_poodle. It was rated as 14.0/10, favourite count is 82708.0, retweet count is 34546.0

- The dog with the highest rating is **not** the dog that has the most favourite count.

To see if there is a correlation between rating_numerator and favourite_count

Correlation between rating_numerator and favourite_count



- The correlation coefficient is **-0.03**. There is **no correlation relationship** between the rating_numerator and the favourite_count.