

# act\_report

June 28, 2022

## 0.1 Report: act\_report

- Create a **250-word-minimum written report** called "act\_report.pdf" or "act\_report.html" that communicates the insights and displays the visualization(s) produced from your wrangled data. This is to be framed as an external document, like a blog post or magazine article, for example.

## 0.2 Summary:

This gives a summary of the insights and visualisations found during the process.

```
In [2]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

```
In [3]: df = pd.read_csv('twitter_archive_master')
df.describe()
```

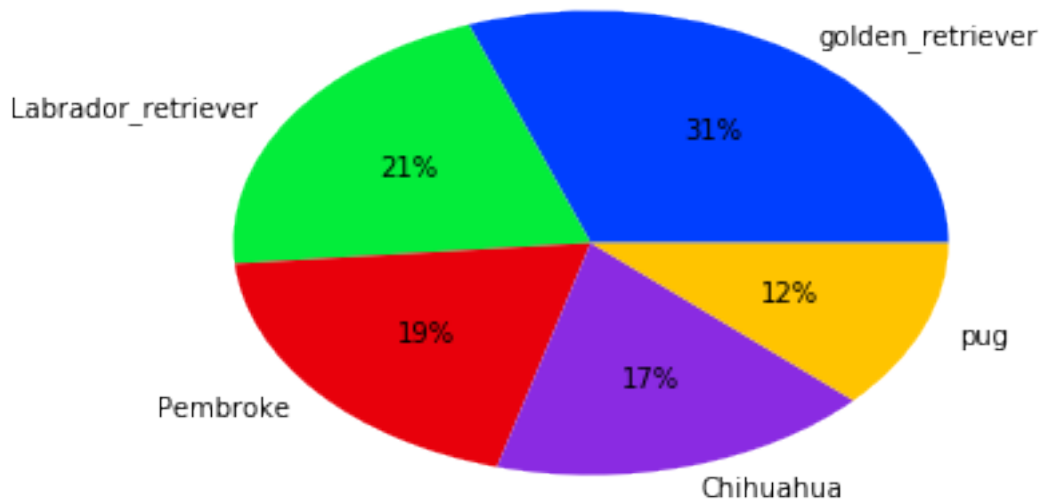
```
Out[3]:
```

	tweet_id	rating_numerator	rating_denominator	retweet_count	\
count	1.994000e+03	1994.000000	1994.000000	1994.000000	
mean	7.358508e+17	12.280843	10.532096	2766.753260	
std	6.747816e+16	41.497718	7.320710	4674.698447	
min	6.660209e+17	0.000000	2.000000	16.000000	
25%	6.758475e+17	10.000000	10.000000	624.750000	
50%	7.084748e+17	11.000000	10.000000	1359.500000	
75%	7.877873e+17	12.000000	10.000000	3220.000000	
max	8.924206e+17	1776.000000	170.000000	79515.000000	

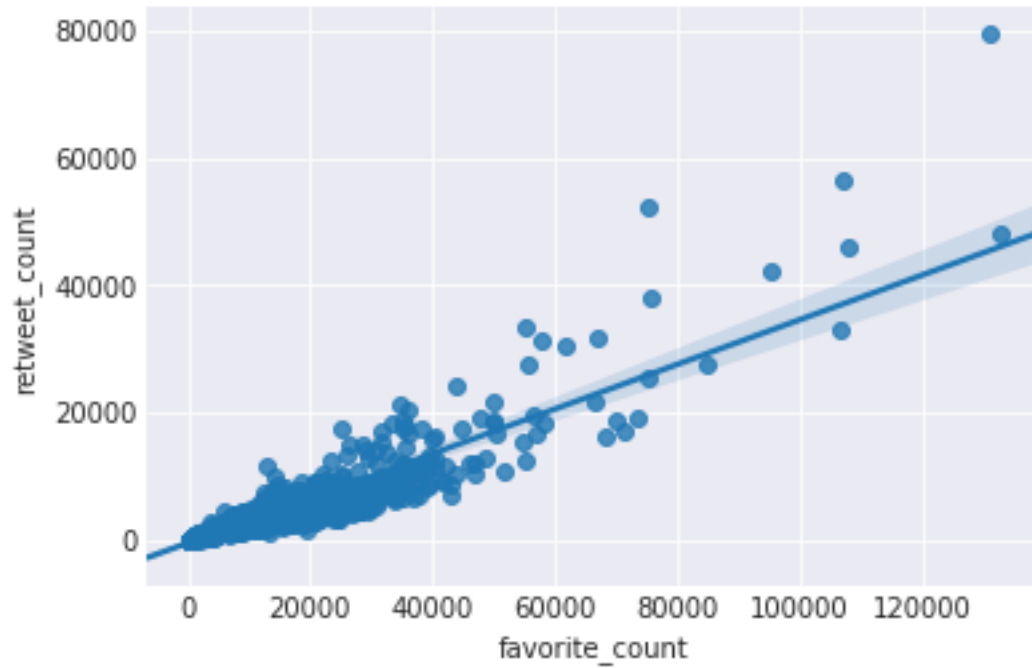
  

	favorite_count	img_num	p1_conf	p2_conf	p3_conf
count	1994.000000	1994.000000	1994.000000	1.994000e+03	1.994000e+03
mean	8895.725677	1.203109	0.593941	1.344195e-01	6.024848e-02
std	12213.193181	0.560777	0.271954	1.006807e-01	5.089067e-02
min	81.000000	1.000000	0.044333	1.011300e-08	1.740170e-10
25%	1982.000000	1.000000	0.362857	5.393988e-02	1.619283e-02
50%	4136.000000	1.000000	0.587635	1.174550e-01	4.950530e-02
75%	11308.000000	1.000000	0.846285	1.951377e-01	9.159438e-02
max	132810.000000	4.000000	1.000000	4.880140e-01	2.734190e-01

```
In [6]: # Showing the top 5 most common dogs
dog_types=['golden_retriever', 'Labrador_retriever', 'Pembroke', 'Chihuahua', 'pug']
color_ = sns.color_palette('bright')
plt.pie(df['p1'].value_counts().head(), labels=dog_types, colors=color_, autopct='%0f%%')
plt.show()
```



```
In [8]: # Showing the relation between favorite_count and retweet_counts
sns.set_style('darkgrid')
ax= sns.regplot(x='favorite_count', y='retweet_count', data=df)
```



### 0.2.1 Insights:

- The average numerator rating is at 12.28
- Dogs with no names contribute to over 27%
- The top 3 dog types are the Golden retriever, Lambrador retriwer and Pembroke
- The minimum amount of likes(favorite) is at 81 likes while the most liked is at over 132k likes

In [ ]: