INSIGHTS AND VISUALIZATIONS REPORT

The following are the insights I gathered from the master dataset:

1. Most Used Tweet Sources
2. Top Dog Breeds Featured by WeRateDogs
3. Distribution of the calculated rating
4. Distribution of favorite count
5. Distribution of retweet count
6. Favorite count over time
7. Retweet count over time
8. Relationship between retweet count and favorite count

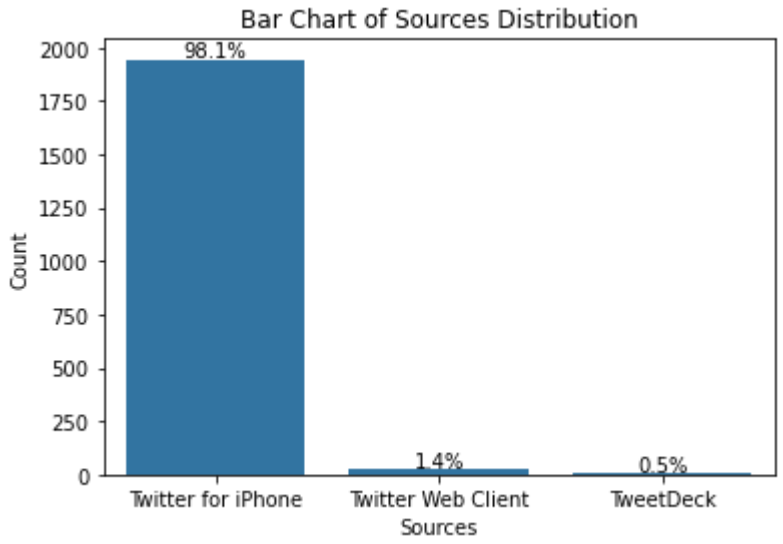
INSIGHT #1: Most Used Tweet Sources

This insight will show which tweet sources are most used. Here's a table displaying the insight:

Sources	Value Counts
Twitter for iPhone	1944
Twitter Web Client	28
TweetDeck	10

VISUALIZATION

Here's the visualization of this insight:



CONCLUSION

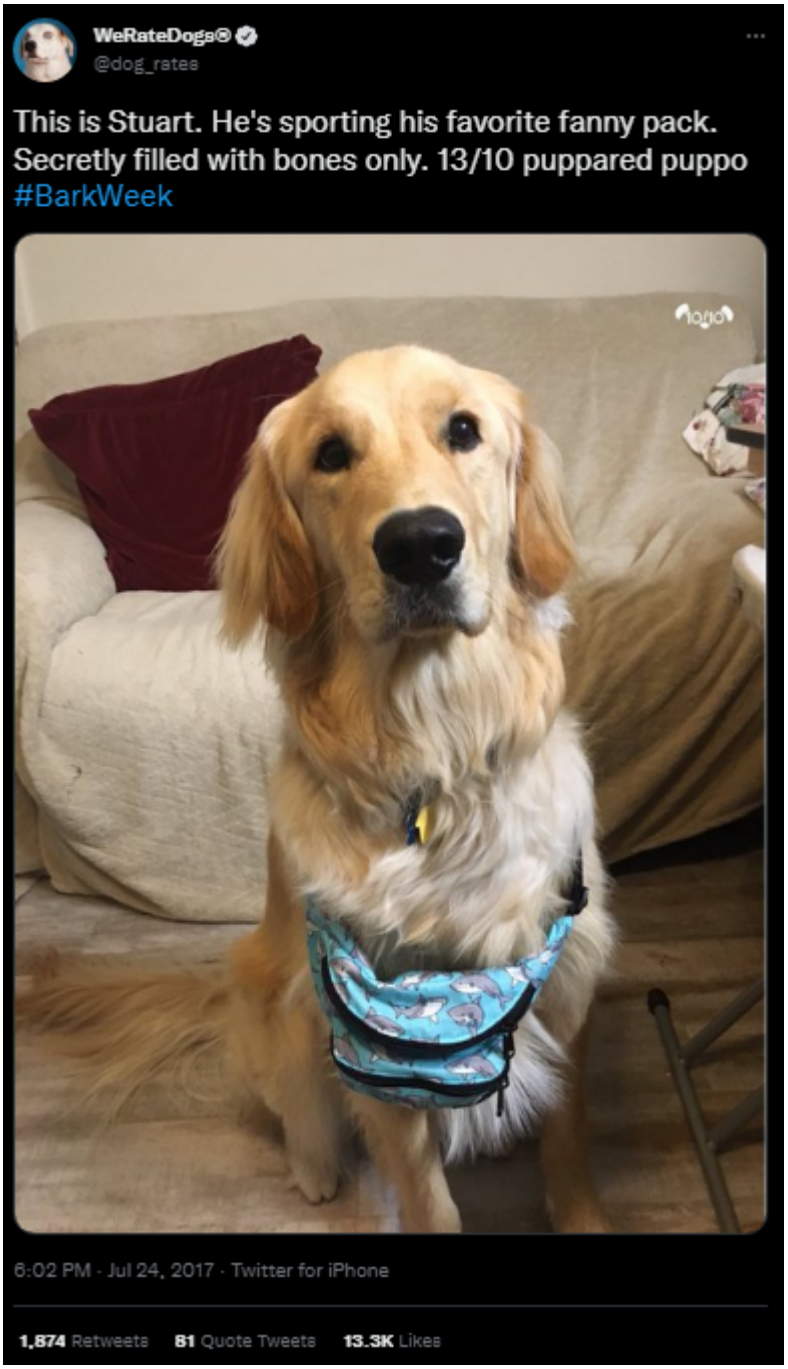
From the insight and the visualization, it's evident that the most used tweet source is **Twitter for iPhone**.

INSIGHT #2: Top Dog Breeds Featured by WeRateDogs

This insight will tell us the top dog breeds the image prediction neural network determines from the tweets made by WeRateDogs. Here's a table displaying the top breeds featured by WeRateDogs:

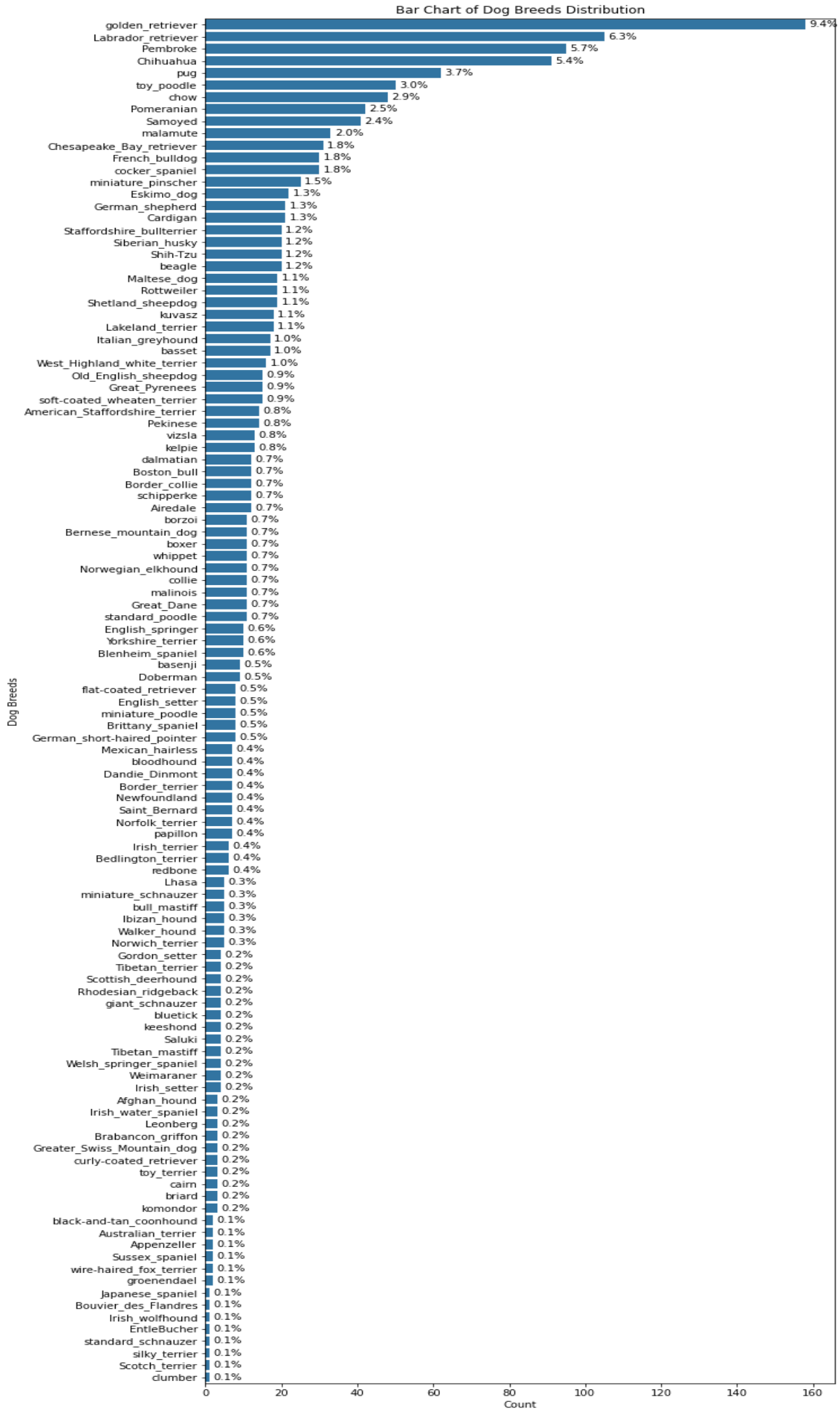
Dog Breed	Value Count
Golden Retriever	158
Labrador Retriever	105
Pembroke	95
Chihuahua	91
Pug	62
Toy Puddle	50
Chow	48
Pomeranian	42

Samoyed	41
Malamute	33



Stuart is a golden retriever.

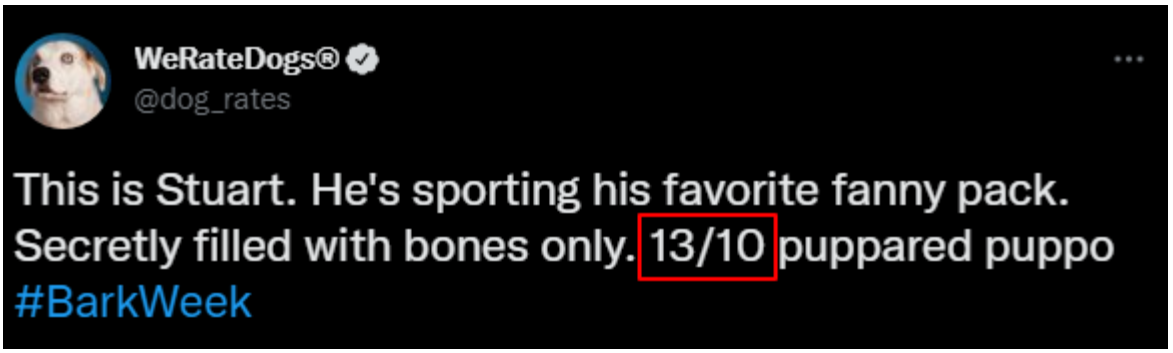
VISUALIZATION


Bar Chart of Dog Breeds Distribution
golden_retriever 9.4%
Labrador_retriever 6.3%
Pembroke 5.7%
Chihuahua 5.4%
pug 3.7%
toy_poodle 3.0%
chow 2.9%
Pomeranian 2.5%
Samoyed 2.4%
malamute 2.0%
Chesapeake_Bay_retriever 1.8%
French_bulldog 1.8%
cocker_spaniel 1.8%
miniature_pinscher 1.5%
Eskimo_dog 1.3%
German_shepherd 1.3%
Cardigan 1.3%
Staffordshire_bullterrier 1.2%
Siberian_husky 1.2%
Shih-Tzu 1.2%
beagle 1.2%
Maltese_dog 1.1%
Rottweiler 1.1%
Shetland_sheepdog 1.1%
kuvasz 1.1%
Lakeland_terrier 1.1%
Italian_greyhound 1.0%
basset 1.0%
West_Highland_white_terrier 1.0%
Old_English_sheepdog 0.9%
Great_Pyrenees 0.9%
soft-coated_wheaten_terrier 0.9%
American_Staffordshire_terrier 0.8%
Pekinese 0.8%
vizsla 0.8%
kelpie 0.8%
dalmatian 0.7%
Boston_bull 0.7%
Border_collie 0.7%
schipperke 0.7%
Airedale 0.7%
borzoi 0.7%
Bernese_mountain_dog 0.7%
boxer 0.7%
whippet 0.7%
Norwegian_elkhound 0.7%
collie 0.7%
malinois 0.7%
Great_Dane 0.7%
standard_poodle 0.7%
English_springer 0.6%
Yorkshire_terrier 0.6%
Blenheim_spaniel 0.6%
basenji 0.5%
Doberman 0.5%
flat-coated_retriever 0.5%
English_setter 0.5%
miniature_poodle 0.5%
Brittany_spaniel 0.5%
German_short-haired_pointer 0.5%
Mexican_hairless 0.4%
bloodhound 0.4%
Dandie_Dinmont 0.4%
Border_terrier 0.4%
Newfoundland 0.4%
Saint_Bernard 0.4%
Norfolk_terrier 0.4%
papillon 0.4%
Irish_terrier 0.4%
Bedlington_terrier 0.4%
redbone 0.4%
Lhasa 0.3%
miniature_schnauzer 0.3%
bull_mastiff 0.3%
Ibizan_hound 0.3%
Walker_hound 0.3%
Norwich_terrier 0.3%
Gordon_setter 0.2%
Tibetan_terrier 0.2%
Scottish_deerhound 0.2%
Rhodesian_ridgeback 0.2%
giant_schnauzer 0.2%
bluetick 0.2%
keeshond 0.2%
Saluki 0.2%
Tibetan_mastiff 0.2%
Welsh_springer_spaniel 0.2%
Weimaraner 0.2%
Irish_setter 0.2%
Afghan_hound 0.2%
Irish_water_spaniel 0.2%
Leonberg 0.2%
Brabancon_griffon 0.2%
Greater_Swiss_Mountain_dog 0.2%
curly-coated_retriever 0.2%
toy_terrier 0.2%
cairn 0.2%
briard 0.2%
komondor 0.2%
black-and-tan_coonhound 0.1%
Australian_terrier 0.1%
Appenzeller 0.1%
Sussex_spaniel 0.1%
wire-haired_fox_terrier 0.1%
groenendael 0.1%
Japanese_spaniel 0.1%
Bouvier_des_Flandres 0.1%
Irish_wolfhound 0.1%
EntleBucher 0.1%
standard_schnauzer 0.1%
silky_terrier 0.1%
Scotch_terrier 0.1%
clumber 0.1%
Dog Breeds
0
20
40
60
80
100
120
140
160
Count

CONCLUSION

The top dog breeds featured by WeRate are golden retriever, Labrador retriever, Pembroke, Chihuahua, pug, toy poodle, Pomeranian, chow, Samoyed, and malamute.

INSIGHT #3: Distribution of calculated rating

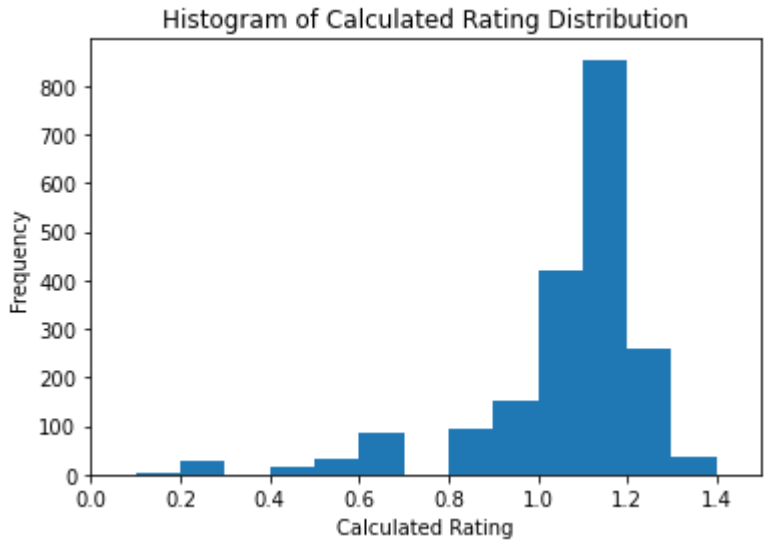


The insight will answer the question “how is the calculated rating distributed?" Here’s a table showing the relevant statistics for the ***calculated_rating*** column:

Statistics	Value
mean	1.170577
std	4.079183
min	0.100000
25%	1.000000
50%	1.100000
75%	1.200000
max	177.600000

VISUALIZATION

Here’s the visualization for this insight:



CONCLUSION

From the histogram above, it can be concluded that most dogs have a rating between 1.0 and 1.3.

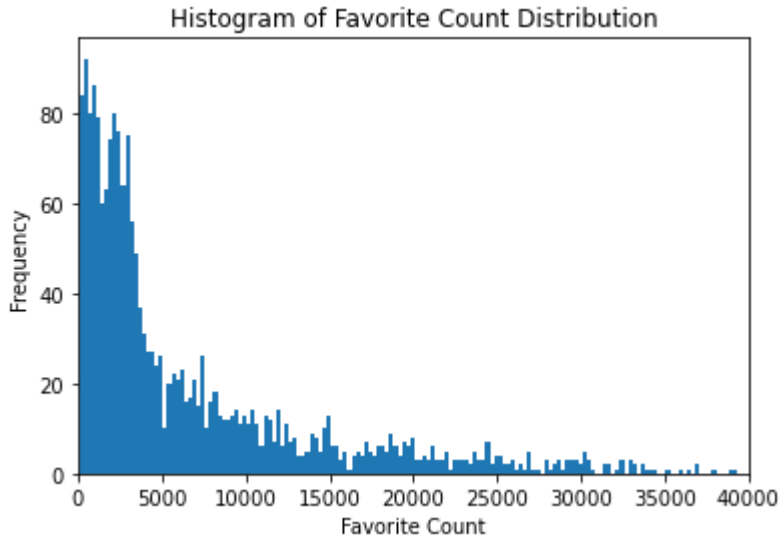
INSIGHT #4: Distribution of favorite count

This insight will answer the question “how is the favorite count distributed?”. Below is a table of the summary statistics for the ***favorite_count*** column:

Statistics	Value
mean	7686.266902
std	11347.240584
min	66.000000
25%	1627.250000
50%	3449.500000
75%	9526.250000
max	144401.000000

VISUALIZATION

Here's the visualization for the insight:



CONCLUSION

The histogram above shows the frequency count reduces with increasing favorite count values.

INSIGHT #5: Distribution of retweet count

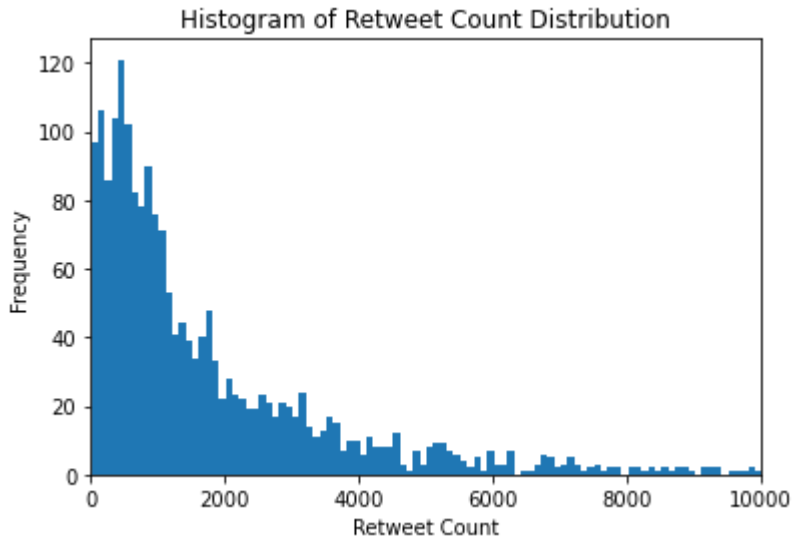
This insight will answer the question, "How is the retweet count distributed?". Here's a table showing some relevant statistics for the ***retweet_count*** column:

Statistics	Value
mean	2236.331988
std	4006.935669
min	11.000000
25%	493.250000
50%	1077.000000
75%	2547.250000

max	70427.000000

VISUALIZATION

Below is the visualization of the insight:



CONCLUSION

This shows a similar trend to the favorite retweet count distribution. As the retweet count value increases, the frequency decreases.

INSIGHT #6: Favorite count over time

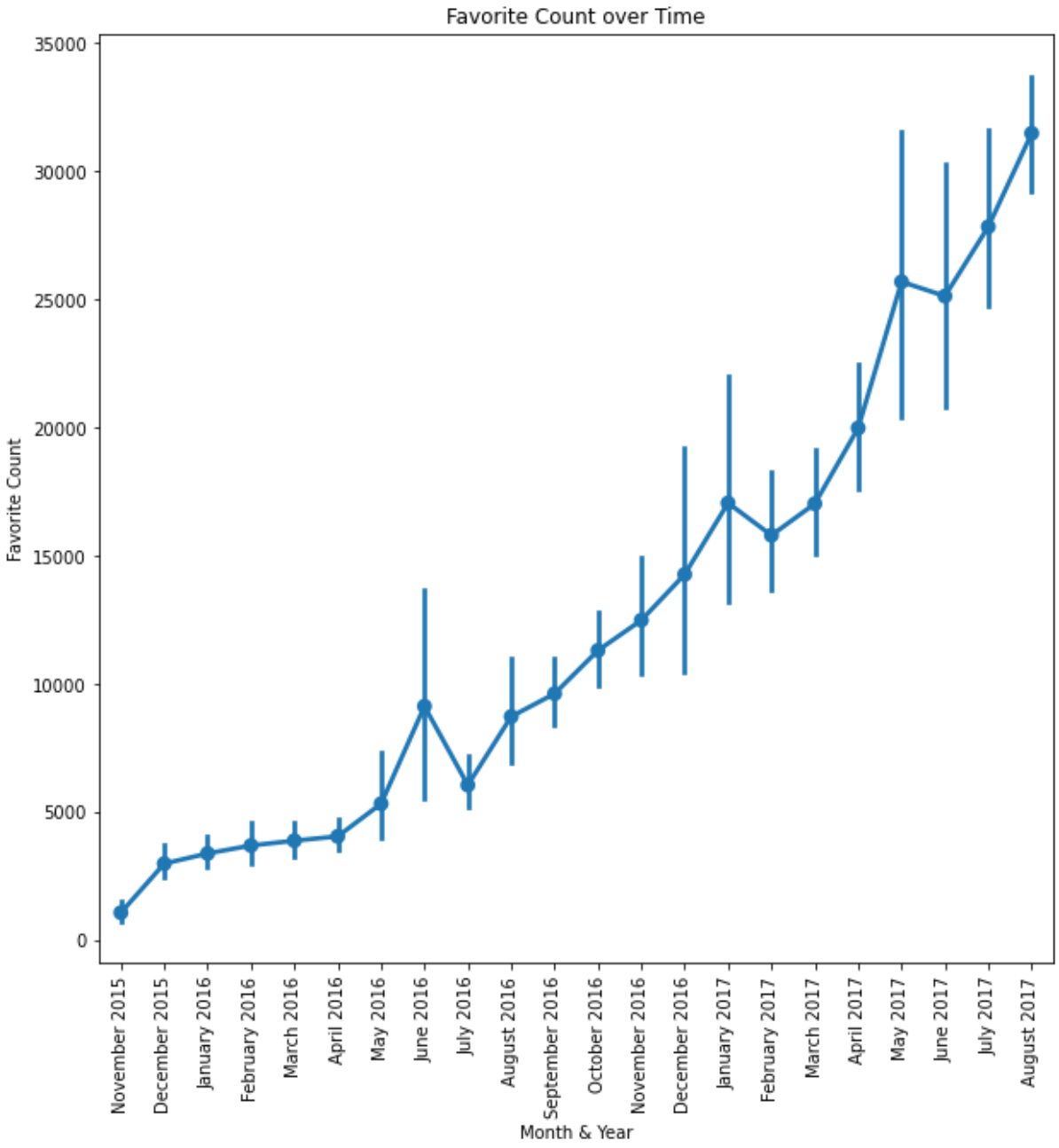
This insight will answer the question, “What is the average favorite count each month?” Here’s a table showing the insight:

month_year	Average Favorite Count
November 2015	1084.908784
December 2015	2998.122951
January 2016	3395.928994
February 2016	3705.063636
March 2016	3893.883333
April 2016	4060.962963

May 2016	5335.385965
June 2016	9138.112500
July 2016	6062.103448
August 2016	8735.051724
September 2016	9629.451613
October 2016	11315.234375
November 2016	12495.660377
December 2016	14282.685185
January 2017	17067.969697
February 2017	15808.500000
March 2017	17040.434783
April 2017	19994.268293
May 2017	25699.325581
June 2017	25139.534884
July 2017	27846.489796
August 2017	31491.500000

VISUALIZATION

Here's a line chart showing the average favorite count each month from November 2015 to August 2017:



CONCLUSION

The line series chart shows that the average favorite count rises steadily from November 2015 to August 2017, but then falls slightly in July 2016, February 2017, and June 2017.

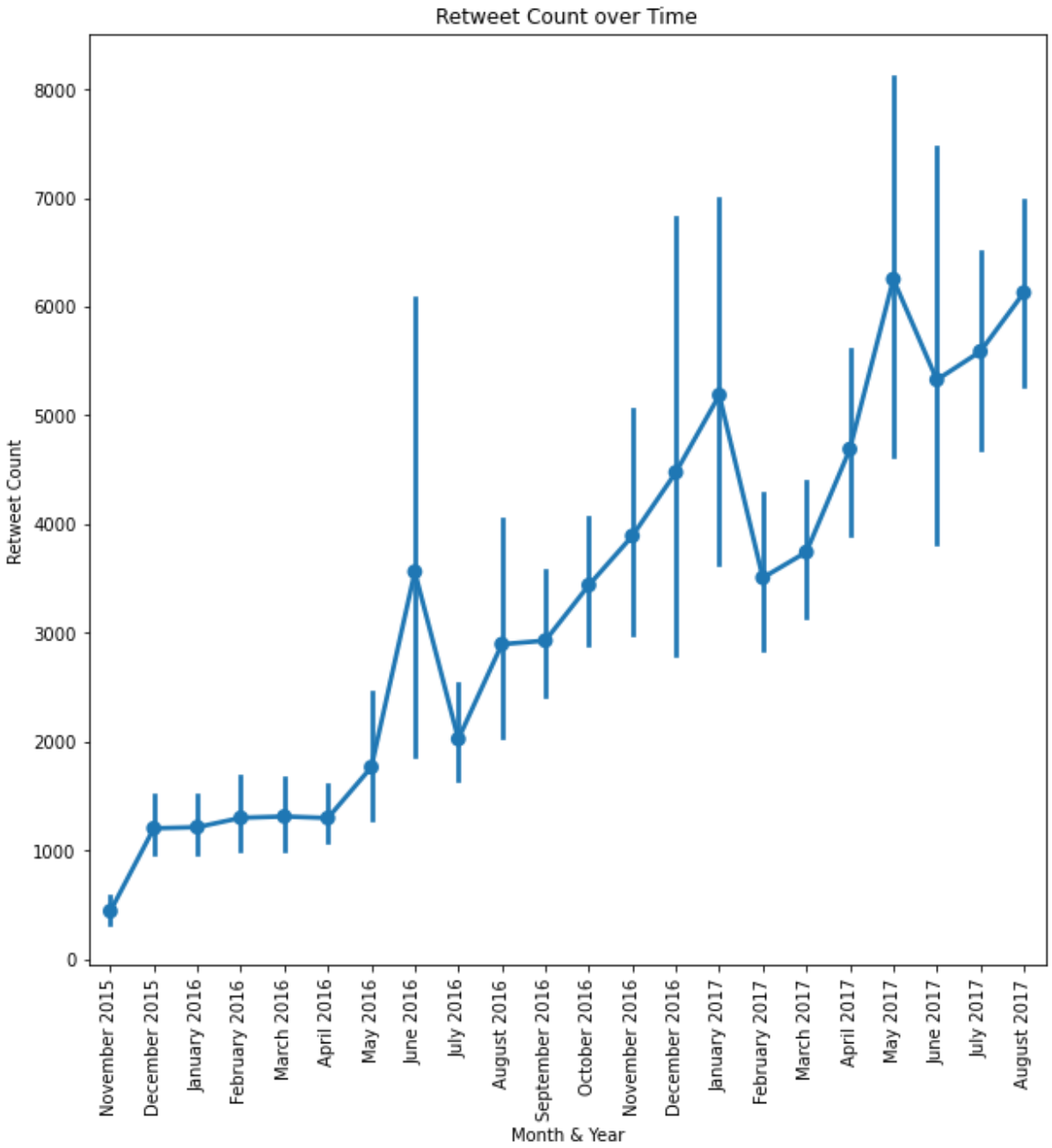
INSIGHT #7: Retweet count over time

This insight will answer the question, “What is the average retweet count each month?”. Here’s a table showing the insight:

month_year	Average Retweet Count
November 2015	439.577703
December 2015	1201.833333
January 2016	1212.106509
February 2016	1298.336364
March 2016	1310.983333
April 2016	1296.277778
May 2016	1764.210526
June 2016	3562.425000
July 2016	2023.402299
August 2016	2895.379310
September 2016	2928.483871
October 2016	3438.890625
November 2016	3891.415094
December 2016	4476.500000
January 2017	5187.090909
February 2017	3507.161290
March 2017	3743.086957
April 2017	4689.487805
May 2017	6253.162791
June 2017	5327.674419
July 2017	5588.020408
August 2017	6129.500000

VISUALIZATION

Here's a line chart showing the average retweet count each month from November 2015 to August 2017:



CONCLUSION

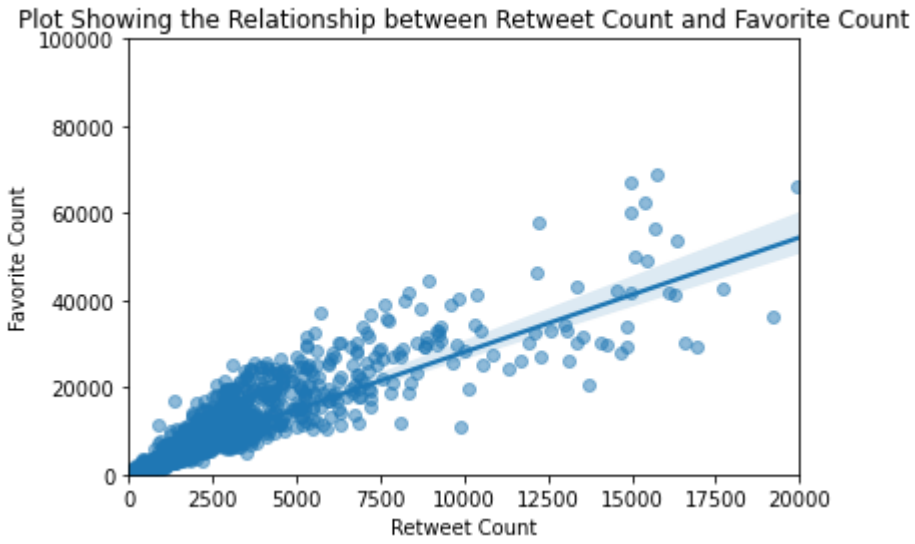
The line series chart shows that the average retweet count rises steadily from November 2015 to August 2017 but falls slightly in July 2016, February 2017, and June 2017.

INSIGHT #8: Relationship between retweet count and favorite count

This insight will show the relationship between retweet count and favorite count using a scatter plot.

VISUALIZATION

Here's the visualization of the insight using a scatter plot:



CONCLUSION

From the visualizations above; we can conclude that the favorite count will generally increase with an increasing retweet count and vice versa.