

Project: Wrangle and Analyze Data

Analysis and Insights

Question: Which are the most common scores on the tweets?

Insight: The most common scores are over 10/10 and in the top ten (see visualization) you see that 12/10 has the top1 with 500 tweets. Other popular scores are 11/10 and 13/10 with scores of 425 and 307 respectively.

This means that most of the owners love their pets in order to put a classification that is over 100% (over 10/10). The exact score of 10/10 has a count of 442, being popular too.

We see also that there are some outliers with extreme as 1776/10 or other with a denominator that is not the base 10: 44/40, 121/110, 165/150, etc

We see also that there are some bad notes: 1/10 with a count of 7, 2/10 with a count 9, 3/10 with a count 19 and 4/10 with a count of 15, for a total of 50!

Code:

```
tweet_data_clean['score'] = tweet_data_clean['rating_numerator'].astype(str) + "/" + \
```

```
tweet_data_clean['rating_denominator'].astype(str)
```

```
tweet_data_clean.score.value_counts()
```

Result:

12/10	500
10/10	442
11/10	425
13/10	307
9/10	155
8/10	98
7/10	53
14/10	43
5/10	36
6/10	32
3/10	19
4/10	15
2/10	9
1/10	7
420/10	2
4/20	2
0/10	2
44/40	1
27/10	1
121/110	1
9/11	1
11/15	1
7/11	1
165/150	1

26/10	1
666/10	1
75/10	1
99/90	1
50/50	1
1776/10	1
20/16	1
143/130	1
80/80	1
84/70	1
15/10	1
17/10	1
144/120	1
45/50	1
204/170	1
88/80	1
24/7	1
960/0	1
60/50	1
182/10	1
1/2	1

Visualization:

