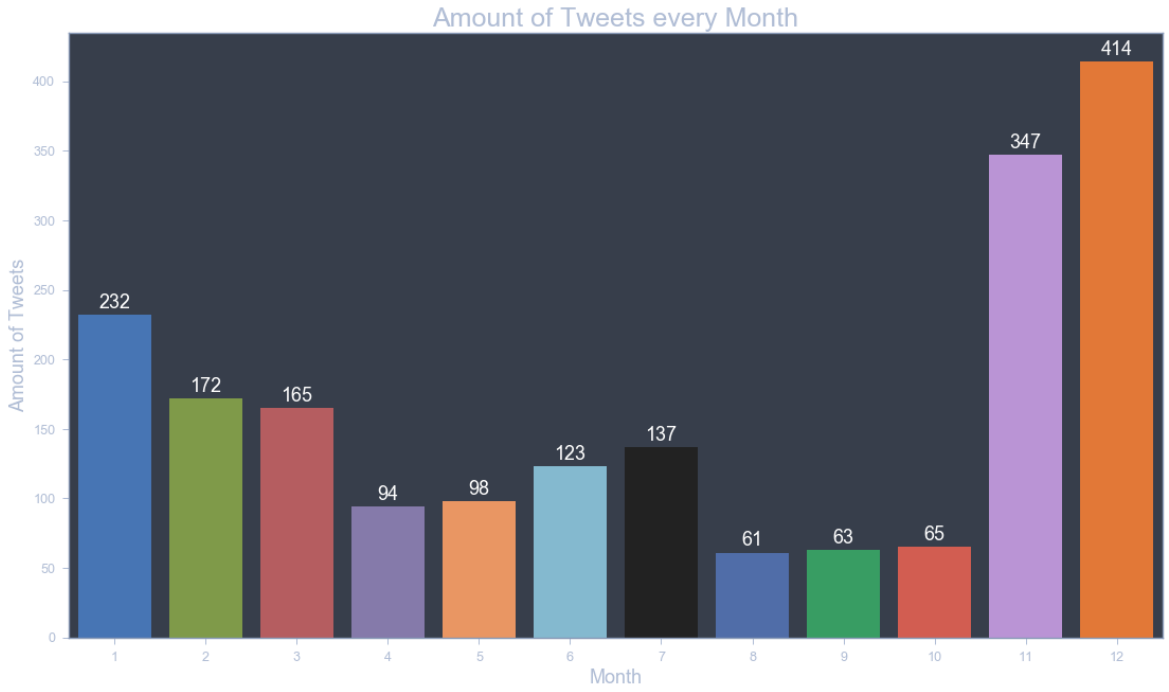
Act Report

This is an analysis report using the wrangled data from the other report. There are 3 kind of insights that I found from the data.

Amount of Tweets by Month

Here I found the amount of analysis for every month ranging from 2015 to 2018.



The amount of tweets in the end year was usually high, compared to beginning or middle of the year. The admin of WeRateDogs looks like doesn't really active in the summer and fall season. Mostly they're active in the winter season. Now we know that If we want to look at so much WeRateDogs tweets in the future, just wait until winter.

Which dog name is the most and least mentioned on WeRateDogs tweets?

Let's see who is the most mentioned by WeRateDogs

```
ChNonerlie    11
Lucy          10
Cooper        10
Oliver        10
Tucker         9
              ..
Smiley         1
Jett           1
Steve          1
CNoneryl       1
Mike           1
```

“ChNonerlie” seems the most mentioned by the admins. Do you think that’s a real name of a dog? Well, I don’t think so. Let’s find out who the real name is

```
192     This is Charlie. He wants to know if you have ...
197     This is Charlie. He's wishing you a very fun a...
241     Meet Charlie. She asked u to change the channe...
249     This is Charlie. He fell asleep on a heating v...
284     This is Charlie. He wins every game of chess h...
588     This is Charlie. He works for @TODAYshow. Supe...
684     This is Charlie. He pouts until he gets to go ...
1051    This is Charlie. He's a West Side Niddlewog. M...
1123    Meet Charlie. He likes to kiss all the big mil...
1331    Say hello to Charlie. He's scholarly af. Quite...
1671    This is Charlie. He was just informed that dog...
Name: text, dtype: object
```

11 of tweets here says that they’re Charlie. So we can assume the “None” here is actually “a”. Well I messed up by not assessing this column. We can do this fix in the future Data Wrangling. What about the least mentioned dogs?

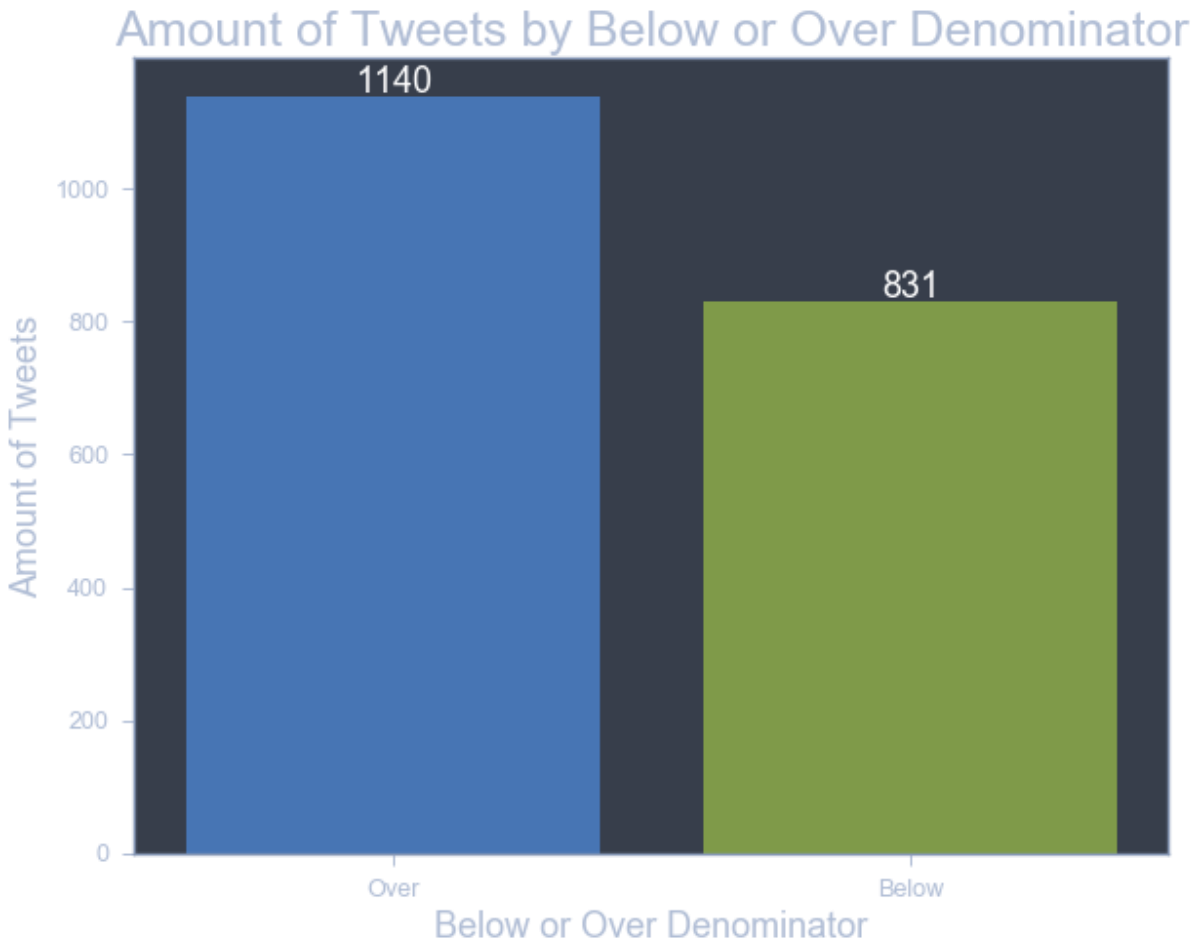
```
1             Tilly
3         DNonerlNone
11            Bruno
14          StuNonert
16               Jim
             ...
1918        CNonerll
1924         Jockson
1927           Josep
1928        LugNonen
1930      Christoper
Name: name, Length: 705, dtype: object
```

Well I found a lot of name that was mentioned only once, and a lot of dirty data. At least now we know that the word "None" should be replaced as "a" (e.g ChNonerlie => Charlie, CNonerll => Carll).

But the most important insight here, Charlie is the most mentioned one and there are 705 other dogs only mentioned once

How many dogs rated below and over its denominator?

the uniqueness of WeRateDogs is on their rating numerator is over than its denominator. But I'd like whether most of the ratings are higher or lower than its denominator. Here is the comparison between them



Here is the proportion

```
Below Denominator Proportion :  0.4216133942161339
Over Denominator Proportion :  0.578386605783866
```

So, most of the ratings are over its denominator. It's interesting that the "below" one isn't that much of difference. Now we know that their uniqueness is actually by giving ratings over their denominators

Conclusion

WeRateDogs admin mostly rate the dogs over its denominator. And it's interesting that they usually make tweets in the winter season. We can see a lot of dog tweets in winter.

Wrangling the data is very important and could reflect to the final analysis. I didn't clean the **name** column which reflects to the most mentioned and least mentioned name. Always remember to do Data Wrangling before doing analysis!