WeRateDogs Twitter Archive - ACT REPORT

The WeRateDogs Twitter archive consists of data gathered from the @dog_rates twitter account posted between November 2015 to August 2017. A total of 2356 tweets were extracted from 5000 tweets.

The data has been wrangled and eight quality issues as well as two tidiness issues were identified and sorted. After the four steps associated with the wrangling process, the data was analysed and visualised. Three insights and one visualisation were produced.

INSIGHT ONE: What dog stage has the highest favourite count?

Using the code **df.sort_values('favorite_count', ascending=False)**, the favourite_count column was sorted in a descending order.

expanded_urls	rating_numerator	rating_denominator	name	dog_stages	retweet_count	favorite_count
https://twitter.com/dog_rates/status/744234799360020481/video/1	13	10	None	doggo	70793.0	144978.0
https://twitter.com/dog_rates/status/744234799360020481/video/1	13	10	None	doggo	70784.0	144952.0
https://twitter.com/dog_rates/status/744234799360020481/video/1	13	10	None	doggo	70775.0	144938.0
https://twitter.com/dog_rates/status/822872901745569793/photo/1	13	10	None	puppo	39940.0	124162.0

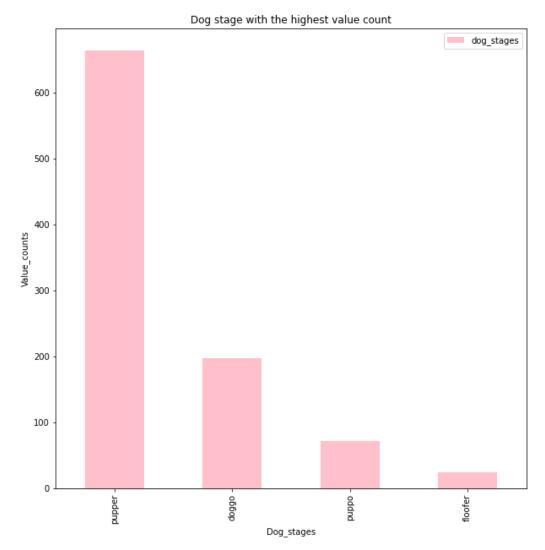
Finding: The highest favourite count can be found on the doggo dog stage with a value of 144978.0.

INSIGHT TWO: What dog stage has the highest value count

The value counts for the entries in the dog_stages column was sought using the code vis_1 = df.dog_stages.value_counts(). The resultant value counts were:

pupper 664 doggo 197 puppo 71 floofer 24

The result was visualised as seen below.



Finding: Pupper dog stage have the highest value_count

INSIGHT THREE: What dog name has the highest popularity?

Using the code **df.name.value_counts()**, the result was:

None 1889
Charlie 31
Cooper 30
Tucker 30
Oliver 30
...
Shaggy 2
Filup 2

Grizzie 2 Ron 2 Harry 1

Finding: Aside from the names that are stated as none, the dog name with the highest popularity is Charlie.