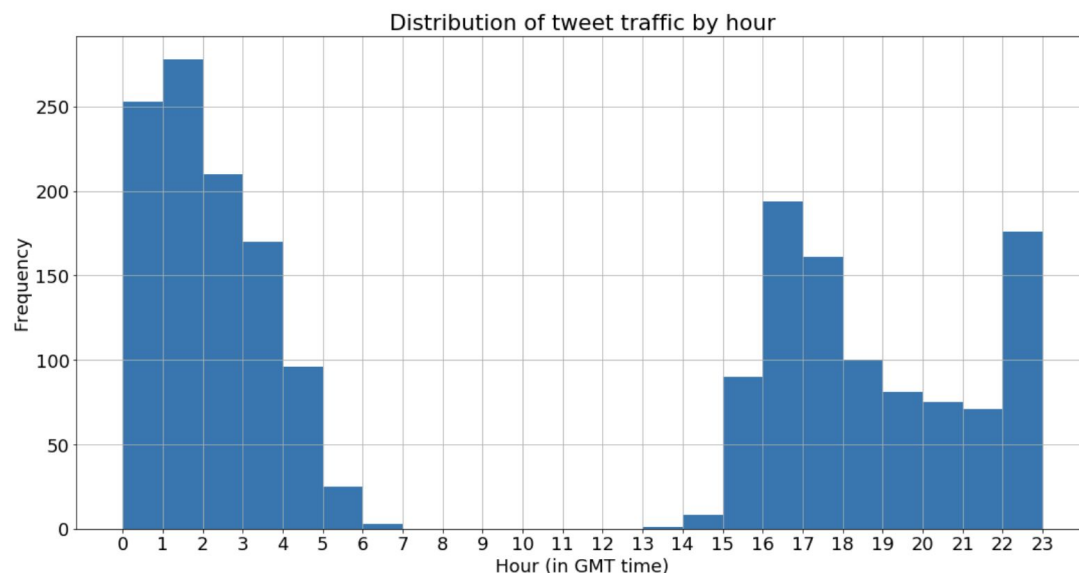
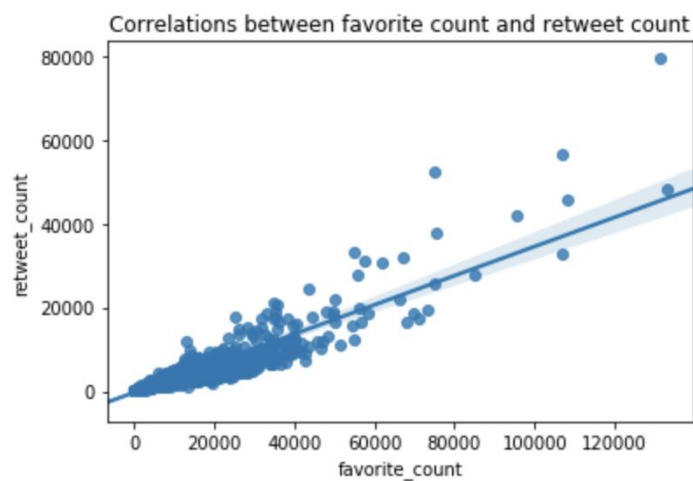


Insight #1: I would like to see the tweeting pattern of WeRateDogs account in terms of timing. What time did this user usually write tweets? Was it around the clock, everyday? The best way to see it to plot a histogram showing how many tweets were sent out at each time interval. Below, you can see that tweets were sent out mostly between 0000-0500 GMT (8PM-1AM in US eastern time) and 1500-2300 GMT (11AM-7PM US eastern time). There are almost no tweets between 0600-1500 GMT (2AM-11AM US eastern time). As Weratedogs founder Matt Nelson stated, he operates out of his parents' home in West Virginia, which probably explains the lull between 0600 and 1500 GMT (He needs sleep, you know). (Source: <http://money.com/money/5225272/weratedogs-matt-nelson-interview/>)



Insight #2: I would also like to see any correlation between retweet_count and favorite count. Intuitively, they seem correlated, but I'd like to confirm this with specific number and visualization. Python's calculation shows a correlation coefficient of 0.913 on the scale of -1 to 1, which again confirms high correlations.



Insight #3: At this point, I would like to find out if retweet count and stages have any bearing on dog's rating. I built a linear regression model to see if this holds. Regression results below show that retweet_count has a low p-values, but its coefficient is way too low to be of any significance to the eventual rating. Hence, retweet count or dog stage information appear not to have any practical impact on dog rating. Low R-squared value also suggests that this model probably does not do very good job of explaining the variability of these variables.

Dep. Variable:	rating	R-squared:	0.006
Model:	OLS	Adj. R-squared:	0.004
Method:	Least Squares	F-statistic:	2.557
Date:	Mon, 17 Jun 2019	Prob (F-statistic):	0.0258
Time:	06:50:45	Log-Likelihood:	-2732.5
No. Observations:	1993	AIC:	5477.
Df Residuals:	1987	BIC:	5511.
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
intercept	1.0341	0.026	39.037	0.000	0.982	1.086
retweet_count	1.586e-05	4.64e-06	3.416	0.001	6.75e-06	2.5e-05
blep	0.1318	0.478	0.276	0.783	-0.806	1.069
doggo	0.0336	0.110	0.304	0.761	-0.183	0.250
floof	0.0748	0.159	0.472	0.637	-0.236	0.386
pupper	0.0014	0.066	0.021	0.984	-0.128	0.131

Omnibus:	5857.374	Durbin-Watson:	1.978
Prob(Omnibus):	0.000	Jarque-Bera (JB):	239964249.172
Skew:	39.764	Prob(JB):	0.00
Kurtosis:	1701.046	Cond. No.	1.21e+05