

Udacity Data Analysis Second Project

[We Rate Dogs]

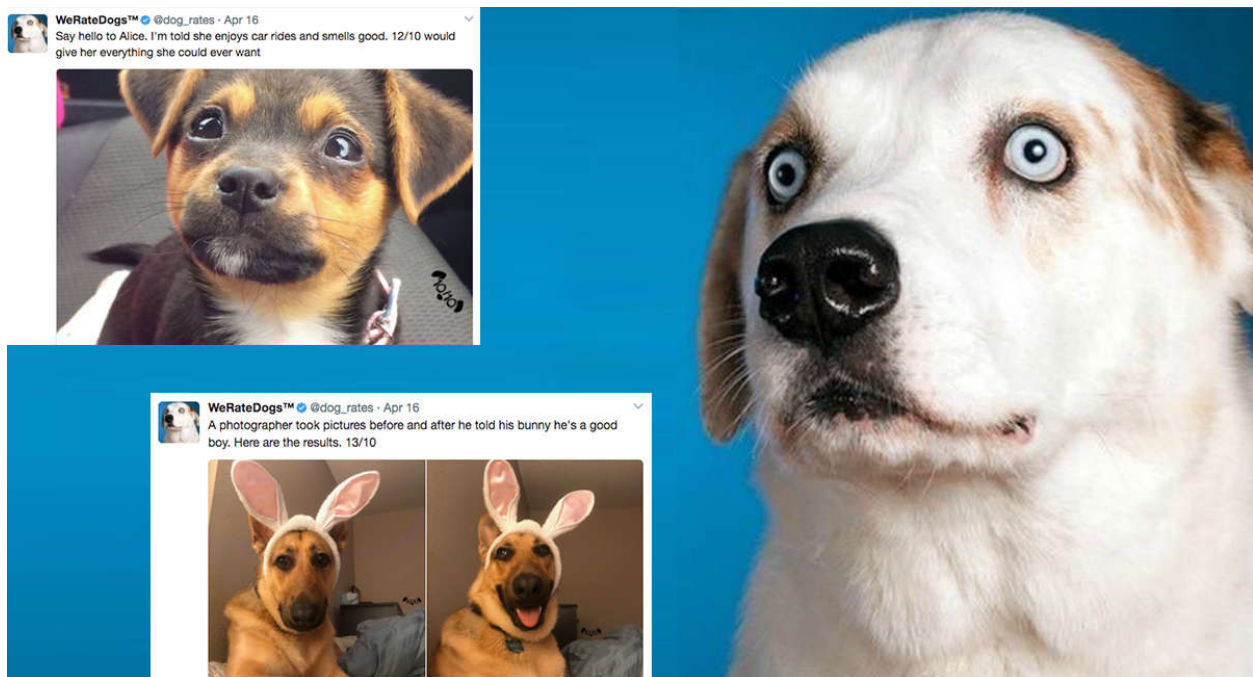
Data Wrangling Project



By: Yousef Ezzeldeen

Introduction

Real-world data rarely comes clean. Using Python and its libraries, I will gather data from a variety of sources and in a variety of formats, assess its quality and tidiness, then clean it. This is called data wrangling. I document my wrangling efforts in a Jupyter Notebook attached in the project folder, plus showcase them through analyses and visualizations using Python (and its libraries) and/or SQL.

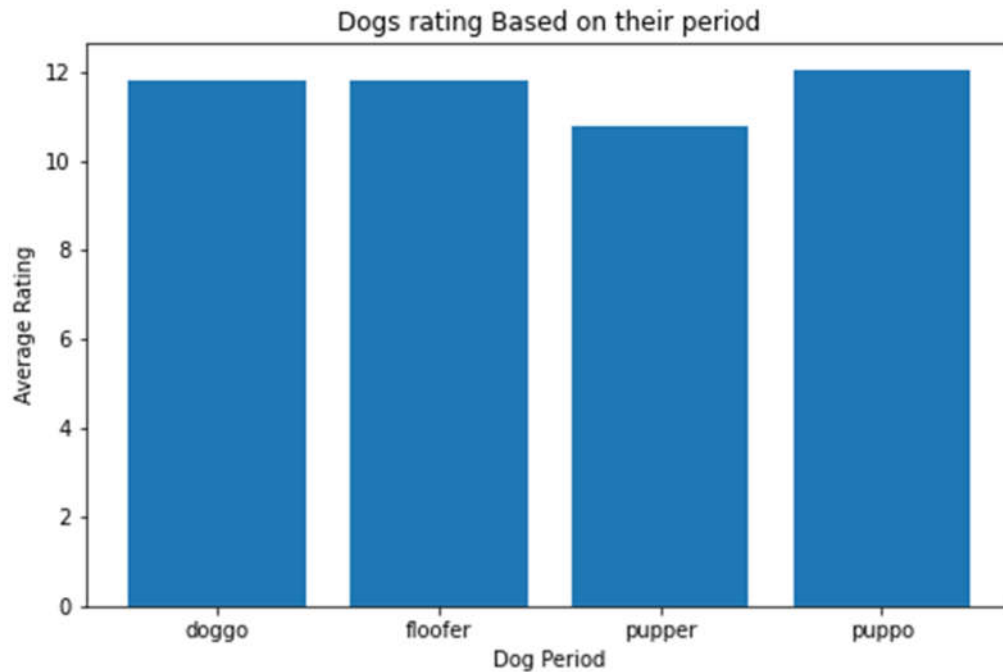


The dataset that I will be wrangling (and analyzing and visualizing) is the tweet archive of Twitter user [@dog_rates](#), also known as [WeRateDogs](#). is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. Why? Because "they're good dogs Brent." WeRateDogs has over 4 million followers and has received international media coverage. Let's take a look.

Data Visualization

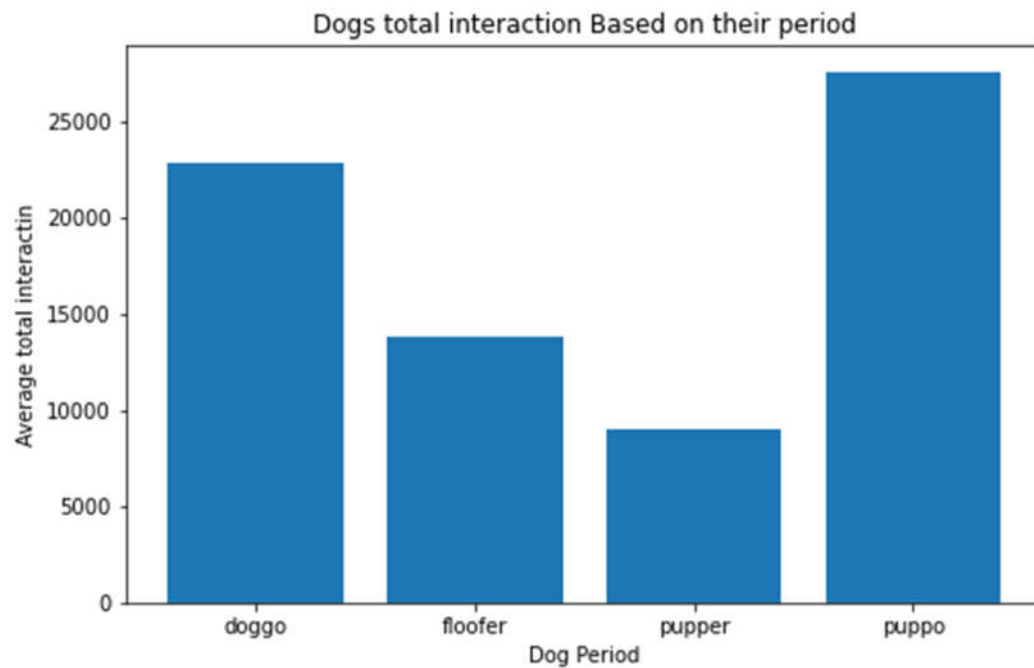
1- Dogs Rating Insights

1-1- Dogs rating Based on their period.



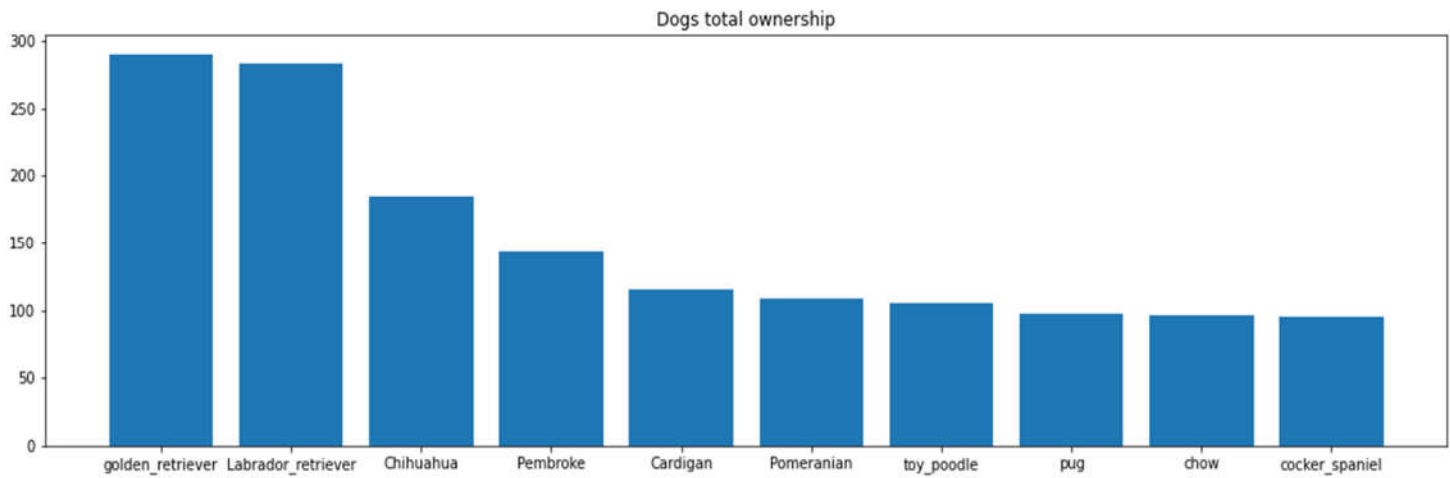
As we see, there is some slight difference between each period of a dog and their rating, but it seems like these people really love dogs rather than rating them.

1-2- Dogs total interaction Based on their period.



When we come to the total interaction of tweets, we notice that people tend to interact with tweets that have a puppo the most. Interesting!

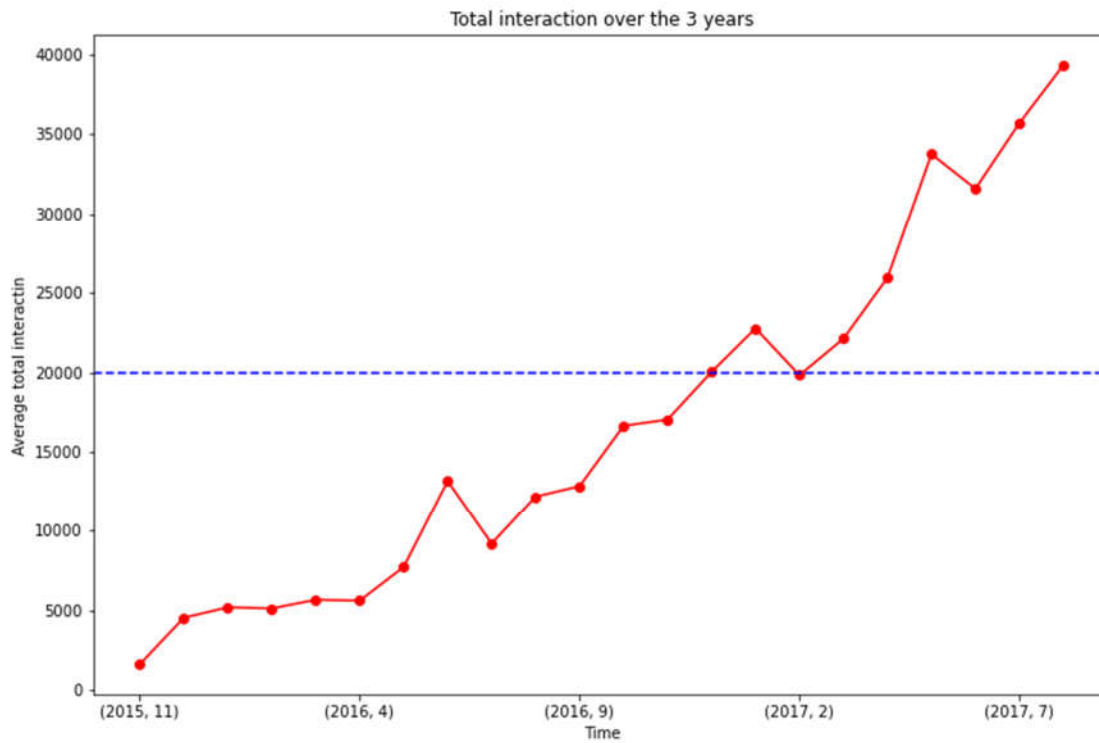
1-3- Dogs total ownership.



Dog ownership can give us a strong insight about their price, most of people have a Golden Retriever because of its low price compared to a cocker spaniel for example.

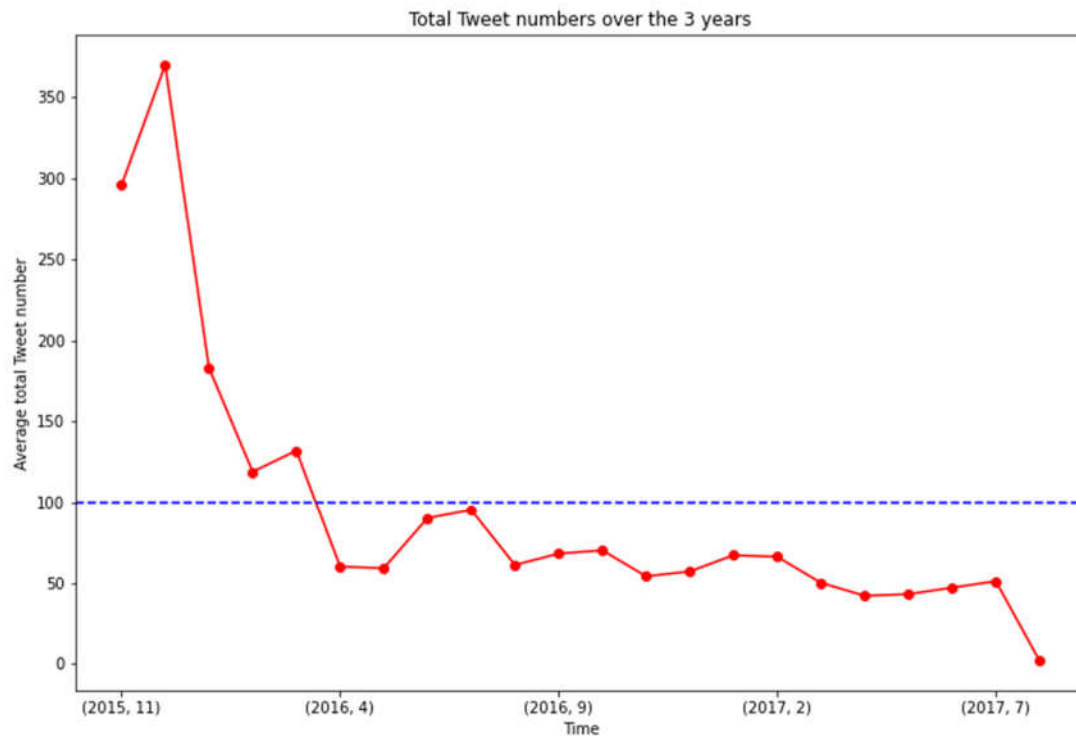
2- Account Insights

2-1- Total interaction over the 3 years



The success of an account depends on interaction. And here we can see that over the 3 years, the account is a success. Great!

2-2- Total Tweet numbers over the 3 years.



It doesn't matter if tweets are many or few, over the 3 years tweets have been reduced, yet still the interaction increases. Awesome!

