Udacity
Data Analyst Nanodegree Program
Data Wrangling
Wrangle and Analyze Data

Act Report

Wrangling WeRateDogs insights & visualizations

Done by:

Zyad alatar

CONTENT

❖ Introduction:

In this project I will put the fundamentals learned in videos in order to complete the project. In this project I will be wrangling & analyzing WeRateDogs dataset to make in best form to be analyzed correctly. Typically, it consists of 4 question each one have given better understanding of the data.

❖ Q1: Who's the favorite dog:

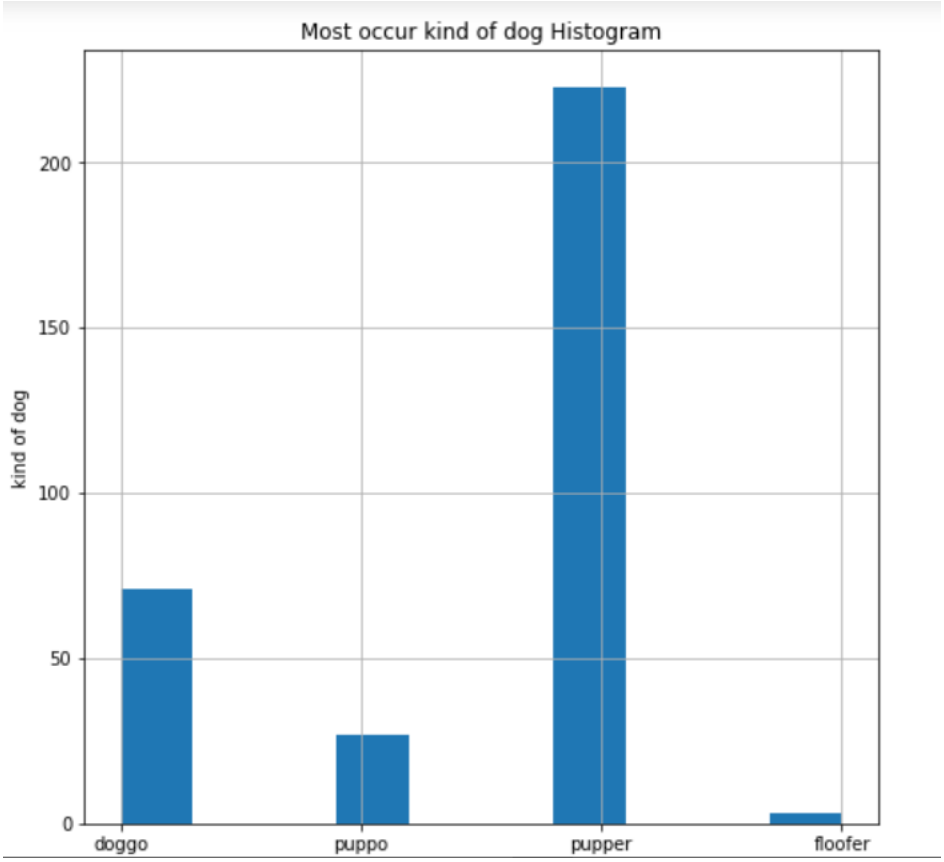
The answer:

After exploring the data, I found that the tweet with tweet_id 744234799360020481 is the most favorite dog. This is his/her picture:

❖ Q2: Which kind of dog occurs most?

The answer:

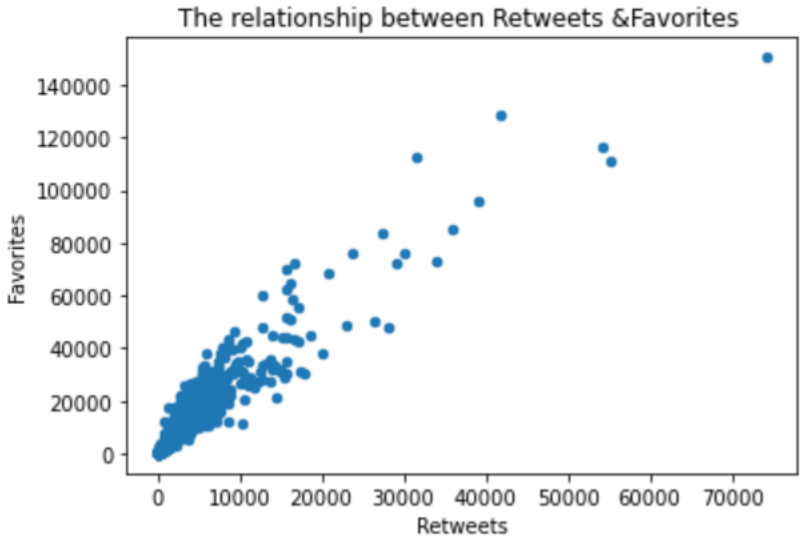
After exploring the data, this is a histogram that show which kind of dog occurs most. It's clear that pupper has occurred the most with significant difference where floofer is the Lowest. Please note that there is xlable, but cutting the graph was not smooth process



❖Q3: Describe the relationship between retweets and favorites:

The answer:

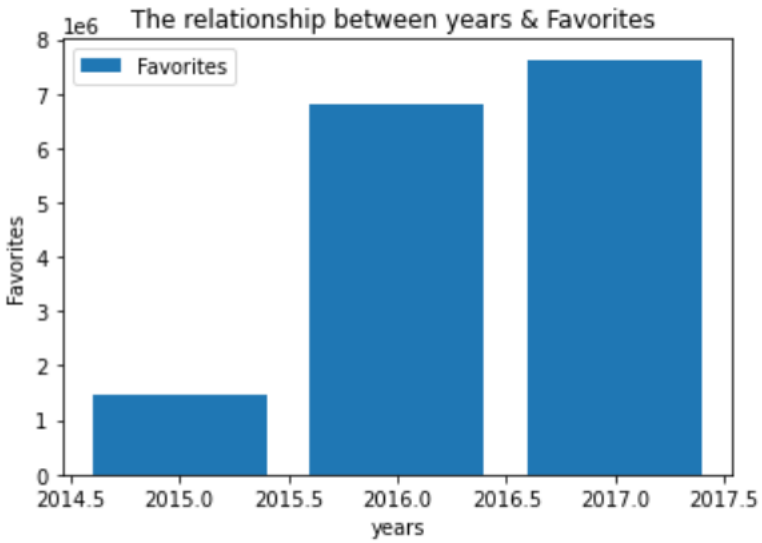
After exploring the data, this is a scatter plot that shows the relationship between retweets and favorites and it clear the more retweets you have that gives an indicator that you got more favorites. The relationship between them is positively strong.



❖ Q4: Does favorite changes significantly over the years

The answer:

After exploring the data, this is a bar chart shows relationship between years and favorites. It's clear that there is a relationship between them and there is a significant change from 2015 to 2016. 2017 got the highest favorites which allow me to predict that we may get more favorites in 2018 since, number of favorites is going in increase m anner.



❖ conclusion:

In conclusion, this project was amazing the I have learned a lot and I will be happy to practice what I learned in my career. The insight provided and visualization was built using the pawedness of python libraries