WeRateDogs Twitter Archive Act-Report

Introduction

This data wrangle report shows the insights and visualization for the tweet archive of Twitter user @dog_rates, also known as WeRateDogs. WeRateDogs is a Twitter account that rates people's dogs with humorous comments about the dog. These ratings almost always have a denominator of 10.

Using data wrangling techniques this large volume of data was gathered, analyzed and assessed and combined to provide insights.

Insights and Visualization

This last requirement of the project was to create visualizations from the insights gotten. The following questions were asked with the hope of finding answers with the visualization.

Insights One

What year did we have the highest tweets by the users?

The rate of tweets by users over the years from the year 2015 to 2017

The Year 2015 we had 393 tweets which is 28.2% ,then there was a high increase in 2016 of

712 tweets which is 51.1% and in the year 2017 a decline of 288 tweets which is 20.7%

It clearly indicates that most tweets occured in 2016 which indicates a high traffic of tweet users and in 2017 the tweets decreased drastically maybe due to some factors.

Insights Two

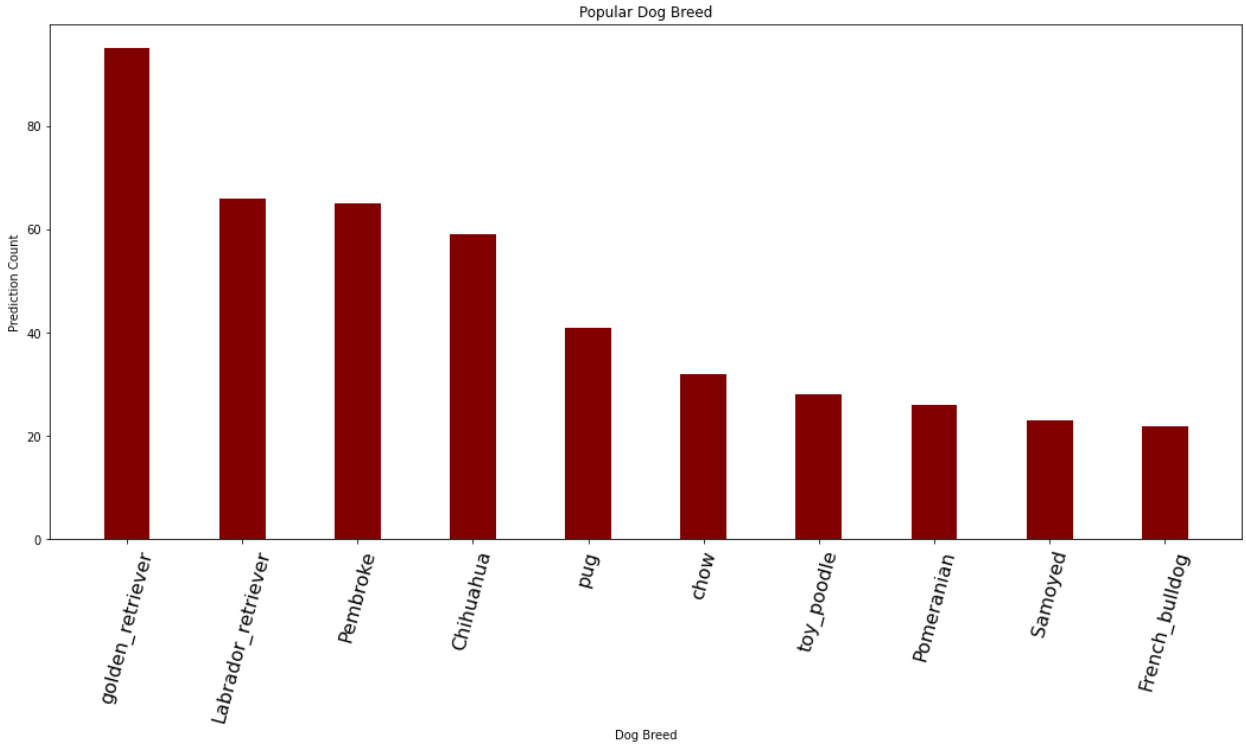
What is the most frequent predicted dog breed?

The dataset has different types of dog breeds that the users favor.

The golden retriever is the most frequently predicted dog breed in the Prediction 1 and has a prediction confidence rate of 99%, and a Prediction dog which predicts whether the dog is really a dog or not is certainly True

This indicates it is the most popular choice for pet dogs lovers followed by the Labrador and .Pembroke.

The visualization below


Popular Dog Breed
Prediction Count
0
20
40
60
80
golden_retriever
Labrador_retriever
Pembroke
Chihuahua
pug
chow
toy_poodle
Pomeranian
Samoyed
French_bulldog
Dog Breed


@dog_rates

Insights Three

What is the most common dog name?

Names are important to dogs

.This dataset contained a lot of different names for the dogs both in lowercase structure and uppercase structure, and characters such as “a’.and others

It also had missing values represented as ‘None’

.Cleaning was done and the ‘a’ and ‘None’ were changed to NaN and dropped ..

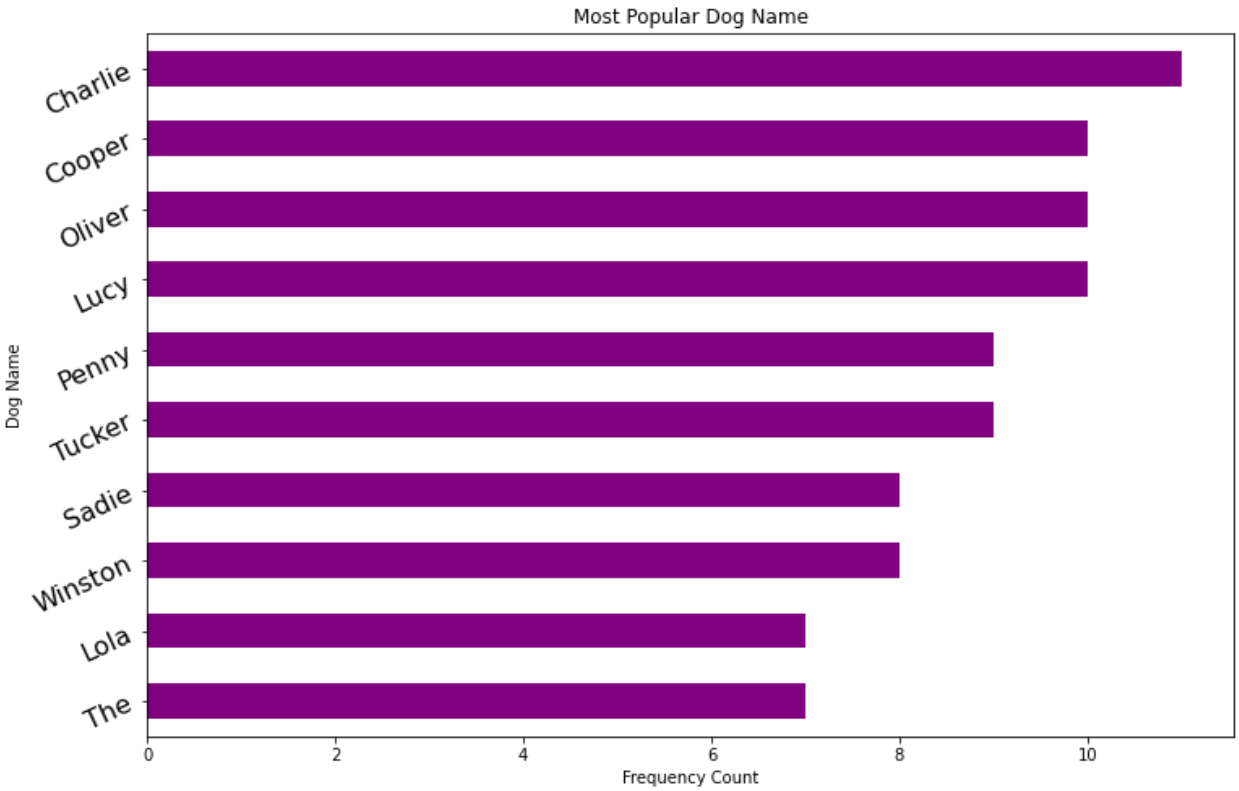
The other name value in the dataset that was in lowercase was changed to uppercase in order

to have a consistent uppercase structure..

During the analysis and visualization it shows the most favored name.

The most popular dog name is Charlie as shown below.

.
Visualization


Most Popular Dog Name
Charlie
Cooper
Oliver
Lucy
Penny
Tucker
Sadie
Winston
Lola
The
Dog Name
0
2
4
6
8
10
Frequency Count