Analyzing and Visualizing Data

Maha Alossiami
22 June 2020

Introduction:

This report contains the insights and visualizations produced from your wrangled data project.

Finding :

Tweet sources are :

After analyzing the sources of tweets we find out these percentage

- Twitter for iPhone   98.1%
- Twitter Web Client   1.2%
- TweetDeck   0.7%

We can see that Twitter for iPhone is the dominator.

Top three popular dogs :

Using the numerator as the significant factor the ranking;

1. pupper 122
2. doggo 38
3. Puppo 15

Puppers are the most popular dogs which received the highest numbers.

Top five names:

The top five names are:

1. Charlie   11
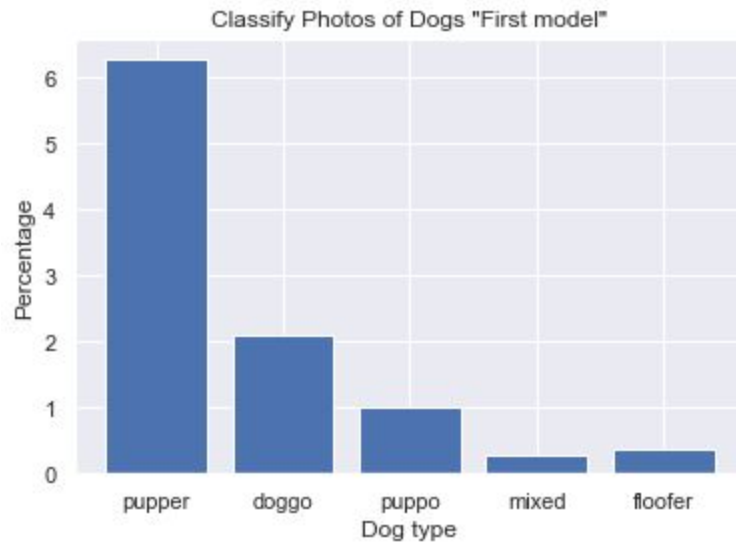2. Oliver   10
3. Lucy   10
4. Cooper   10
5. Tucker   9

Accuracy prediction for each model:

We used a simple method to calculate times where the prediction was correct.

The three models have slightly different rates of correct prediction with 77%, 78.4%, and 75.4% respectively.

Classify Photos of Dogs "First model" :

To find which dog stag was easily predicted, the percentage of true prediction was calculated for each stage. The Pupper dogs were the easiest to be predicted using the first model.

Classify Photos of Dogs "First model"

Density Plot of Retweets by the type of dog:

A density plot is a representation of the distribution of a numeric variable. In this plot, we can see the distribution of retweets count for each dog's type.

Density Plot of Retweets by the type of dog