Act report

Author: Iloezumma Ifeanyi

Date: September 5, 2022

The report will give information on the insights gained from the project as well as visualizations needed to complete the wrangle and analyze project in the Udacity Data Analysis Nanodegree.
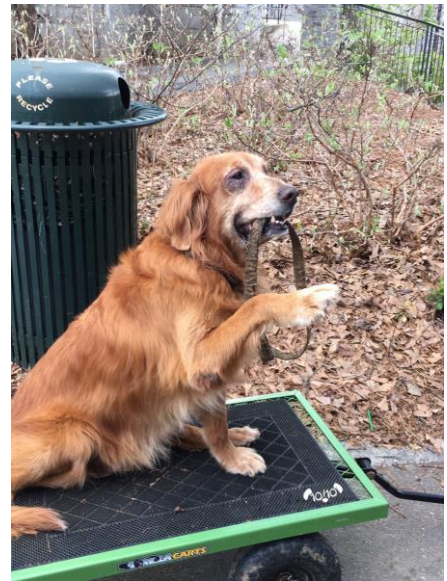
1.0 Insights

Three insights was drawn from the analyzing the cleaned data and they are given below;

1. There is a strong positive correlation between retweet count and favourite count (0.92). This can also be seen in the line chart below. This denotes that retweet count and favourite count varies linearly. The data used for to calculate the correlation is retweet count and favourite count in the twitter_archive_master cleaned data.

2. The Golden Retriver tops the most popular dog breed with a value count of 96 from the prediction_1 column. Also the names of the top five Dog from this breed are Zoey, Barney, Bella, Alfy, Ruby. The images of theses dogs are displayed below. The data used to gain this insight is the prediction_1 in the twitter_archive_master cleaned data.
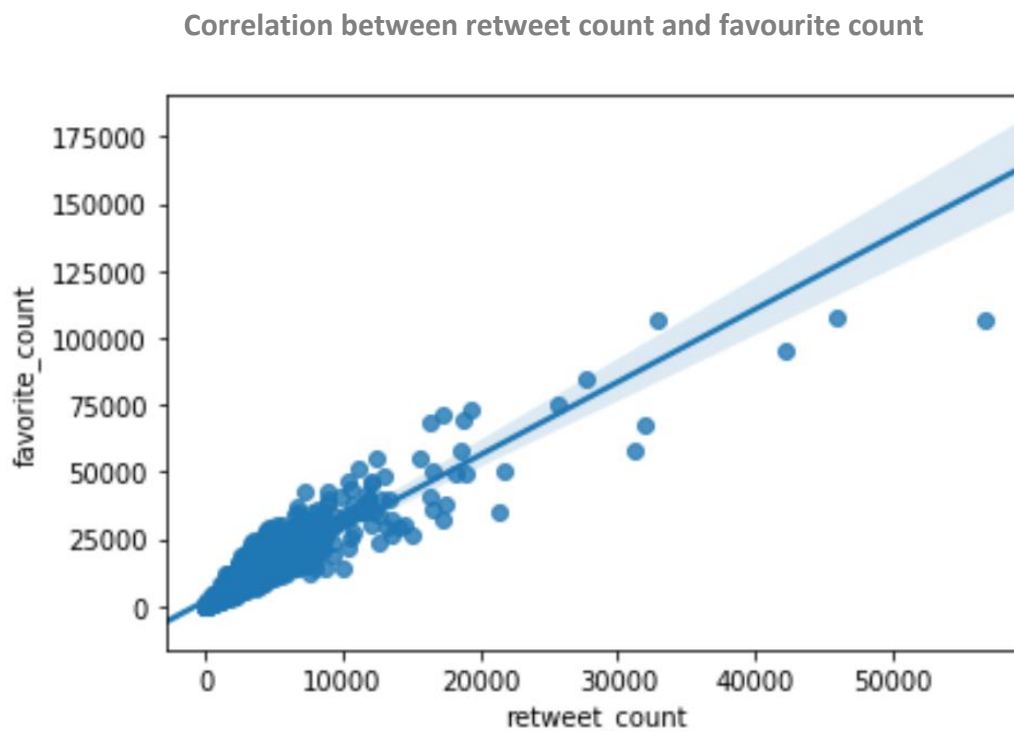

Name: Zoey
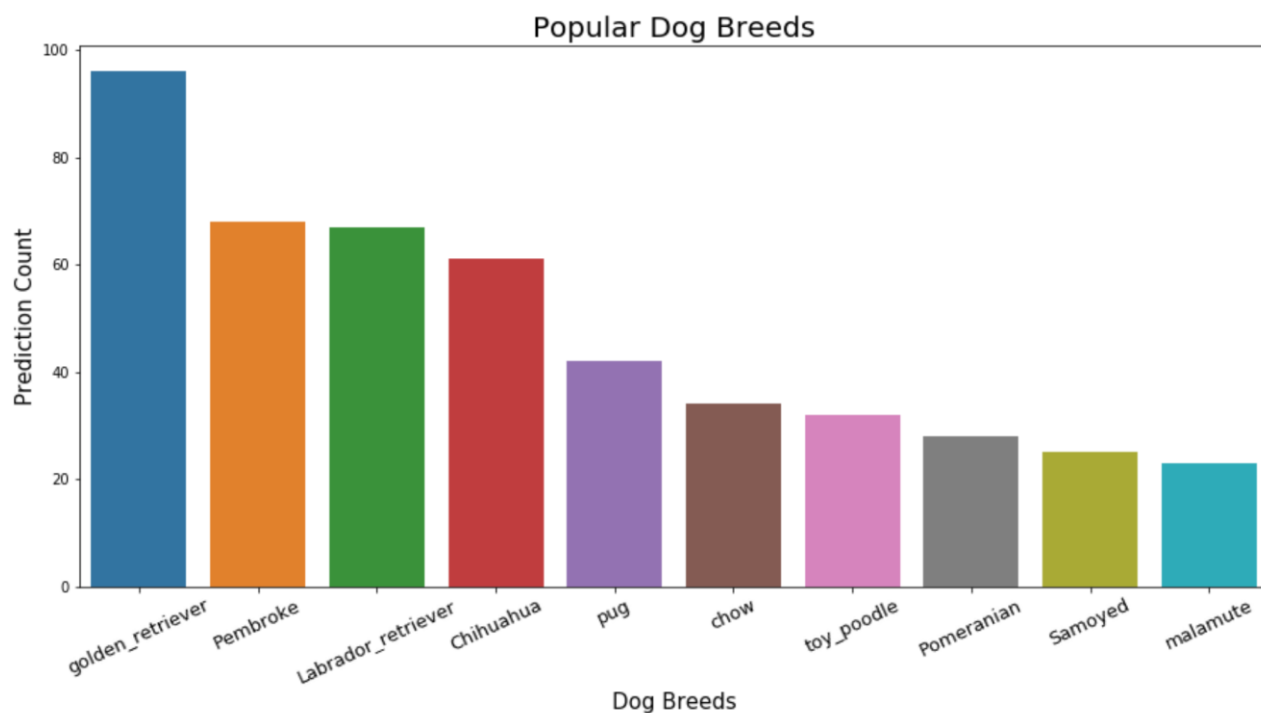

Name: Barney

Name: Bella

Name: Alfy

Name: Ruby

3. The are 145 dogs with the lowest rating of 1. The top five on the list are fountain, grille, groenendael, guenon, hand_blower dog breed. The data used for this insight is the prediction_1 column and the rating_numerator column in the twitter_archive_master cleaned data.

2.0 Visualizations

Visualization 1: There is a strong positive correlation between retweet count and favourite count (0.92). There is also included in the chart a regression line. This can also be seen in the line chart below.

Correlation between retweet count and favourite count

Visualization 2: Using the data in prediction_1 column in the twitter_archive_master cleaned data, we got the count of the prediction for the different dog breeds and then display it on a bar chart. From the bar chart, we can deduce that that the most popular dog breed is the Golden Retriver while the Malamute dog breed is least popular.

Note: See the file wrangle_act.ipynb in the project workspace to see the relevant codes to achieve the visualizations and insights.