Report: act_report

Introduction

Brief description of the Dataset

This report displays and communicates the insights and displays the visualization(s) produced from the wrangled data. -Twiitter data, from the tweet data archive of the twitter page, dog_rates or WeRateDogs is is wrangled, analyzed and visaulized to provide important insights.

-This twitter account rates dogs using humorous comments abou the dogs, it containd ratings that are characterized, by a denominator which almost always with a value of 10. The denominator however is usually more than 10, this is because "they're good Dogs Brent"

-This account have over 4 million followers and has received huge international coverage. -The project was completed outside udacity workspace. reports were generated using jupyter notebook and converted to a PDF report.

Storing Data

-Cleaned data was stored in a master pandas dataframe in a csv format and named; twitter_archive_master.csv, for future data analysis and visualizations.

## Data analysis and visualization

- Analysis and visualization was performed on the cleaned data to provide useful insights about the data.
- Two key points of anlysis were used to provide insights. They include;

(a) Proportion of dog stages, in terms of percentages

(b) The realtionship between retweet count and favorite count

(a) Dog proportion in percentages

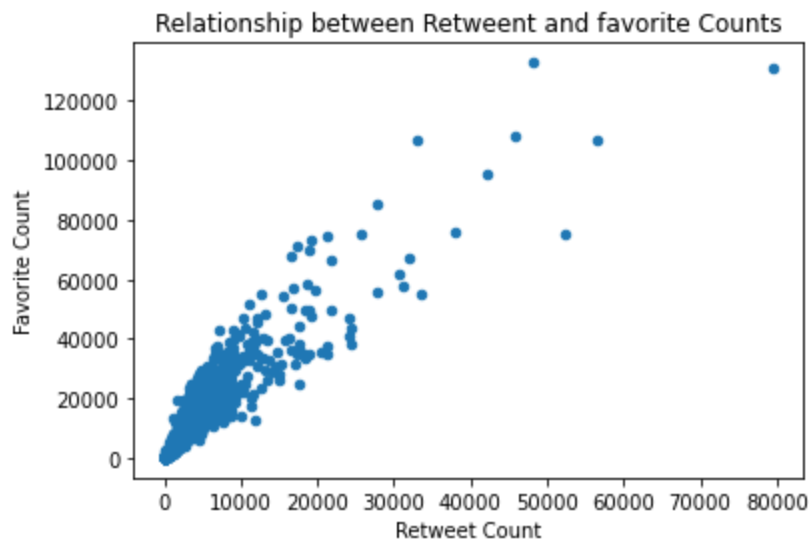- To report key findings, a pie chart was generated

Dog stage proportions
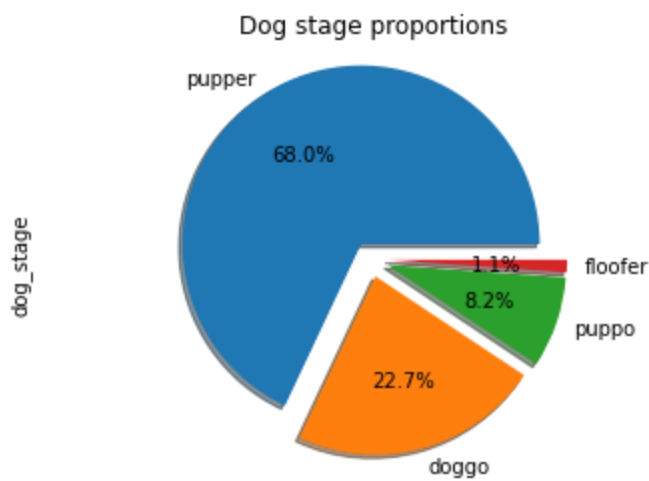
Insights:

1.Pupper takes the highest proportion, with 68%

2.doggo comes second with 22.7%, followed by puppo

3.floofer has the lowest proportion

(b). Relationship between the retweet and favorite count

To explain the relationship between the two variables, a scatter plot was generated.

Insights.

- From the visualization, it is clear that; there is a positive linear relationship between the two variable.