

Act Report

The goal of data visualization is to come up with insights that could have been hidden into the dataset. It also helps communicate the information in the data using visual representations like scatter plots and bar charts.

In this visualization, scatter plots and bar plots were used to visualize the data. In addition to visualizations, I calculated the proportions of different columns to find out percentage of the values.

Visualizations

The figure 1 below shows a bar plot representation of dog stages count.

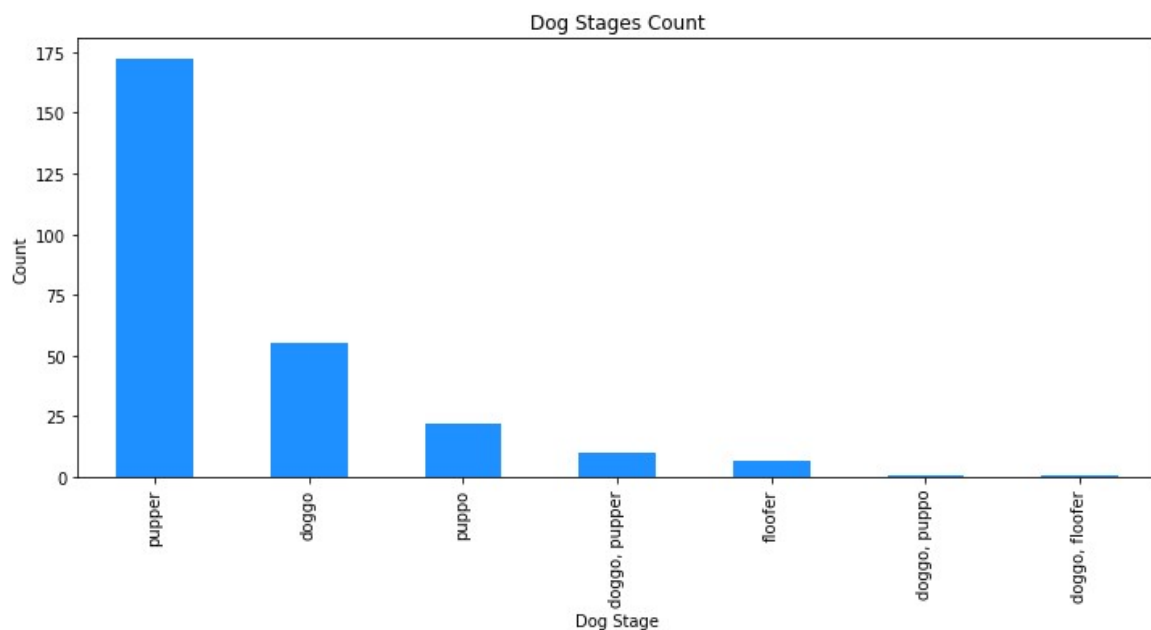


Figure 1: Dog stages count.

From the figure 1 above, it is clear that pupper is the most popular dog stage followed by doggo, then puppo, doggo pupper, floofer, doggo puppo and finally doggo floofer.

Figure 2 is a scatter plot showing the relationship between favorite count and retweet count. The plot reveals a strong positive correlation between favorites and retweets which means that people are more likely to retweet favorite tweets.

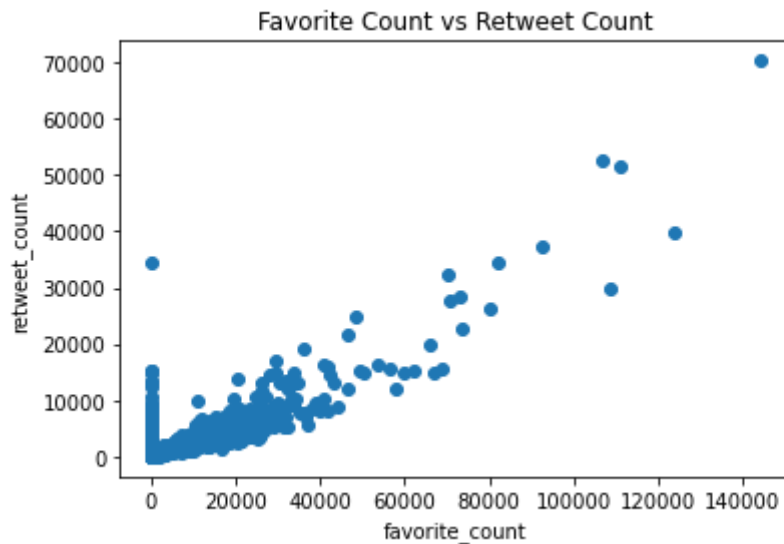


Figure 2: Relationship between favorite count and retweet count.

Insights from the Dataset

After performing a successful data visualization, the following insights were made:

- High positive correlation between favorite count and retweet count indicates that people are more likely to retweet tweets that they find favorite
- Most tweets are from twitter for iPhone (about 98%)
- The most popular dog stage is pupper at about 64%