Qasem M. Dhamad

Udacity - Data Analyst - MCIT

Act Report

This report communicated the insights and showed the visualization that conducted from the wrangled data.

The dataset of twitter archive that was analyzed and visualized is from the Twitter account named "@dog_rates or called WeRateDogs. To illustrates, WeRateDogs is an account in Twitter which is rates people's dogs with some comments about their dogs. However, some ratings are greater than 10. For instances, 13/10, 11/10, 12/10, and so on. This is because some people think that their dogs' ratings deserve more than 10. The twitter account of @dog_rates has more than 9 million followers.

This project was completed on the Udacity Classroom on the Project Workspace. As I choose to generate this report as a PDFs using Microsoft Word.

Sorting Data

After gathering, assessing, and cleaning the datasets, we stored the data as a CSV file names twitter_archive_master.csv to analyze and visualize data.

Analyze and Visualize Data

The master data "cleaned data" was analyzed, visualized, and insights based on the following results.

Insights:

1. Pupper dog has the highest total number while floofer has the lowest number:

2. Pupper dog has the highest total number compared to other:
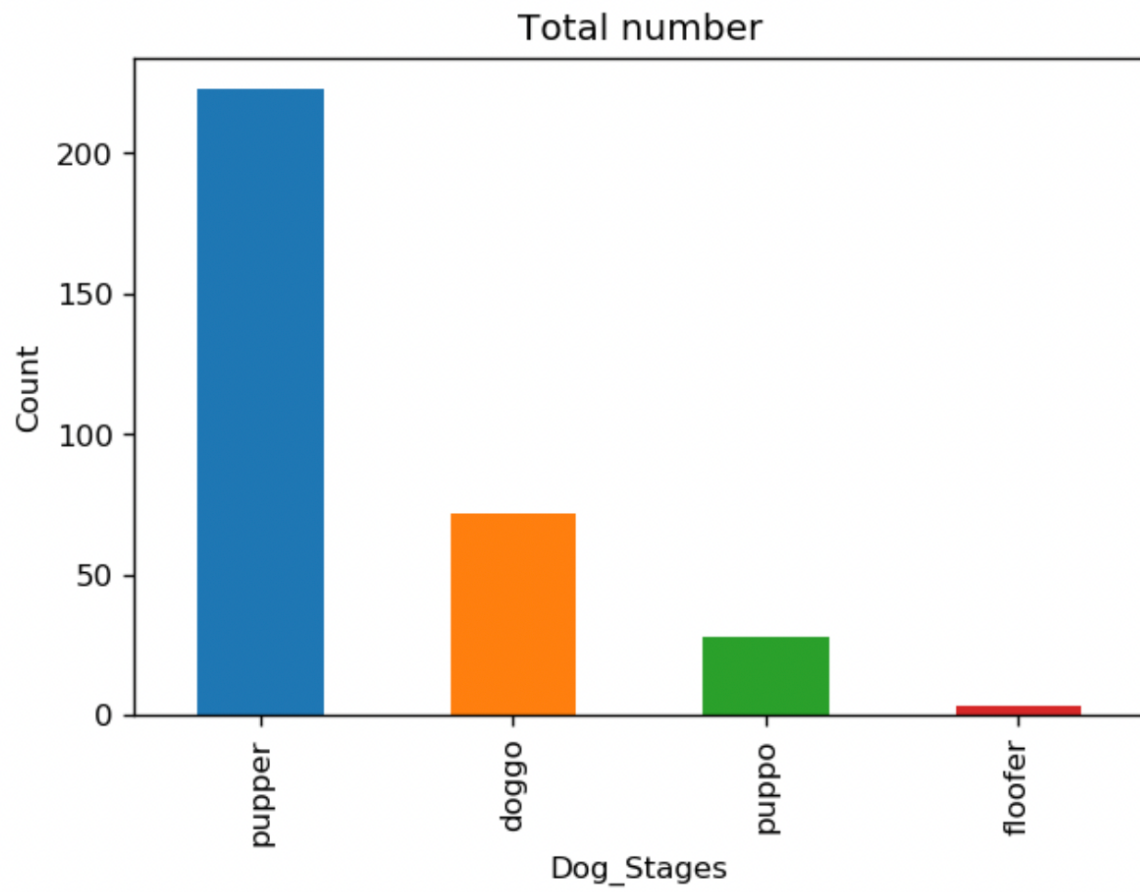
3. Rating_numerator of "12" has the highest total number:

Visualization

Displaying total number of each dog stages:

```
In [179]:  plt.figure(figsize=(6,4),dpi=120)
           twitter_archive_master['dog_stages'].value_counts().plot(kind='bar')
           plt.title ('Total number')
           plt.xlabel ('Dog_Stages')
           plt.ylabel ('Count');
```
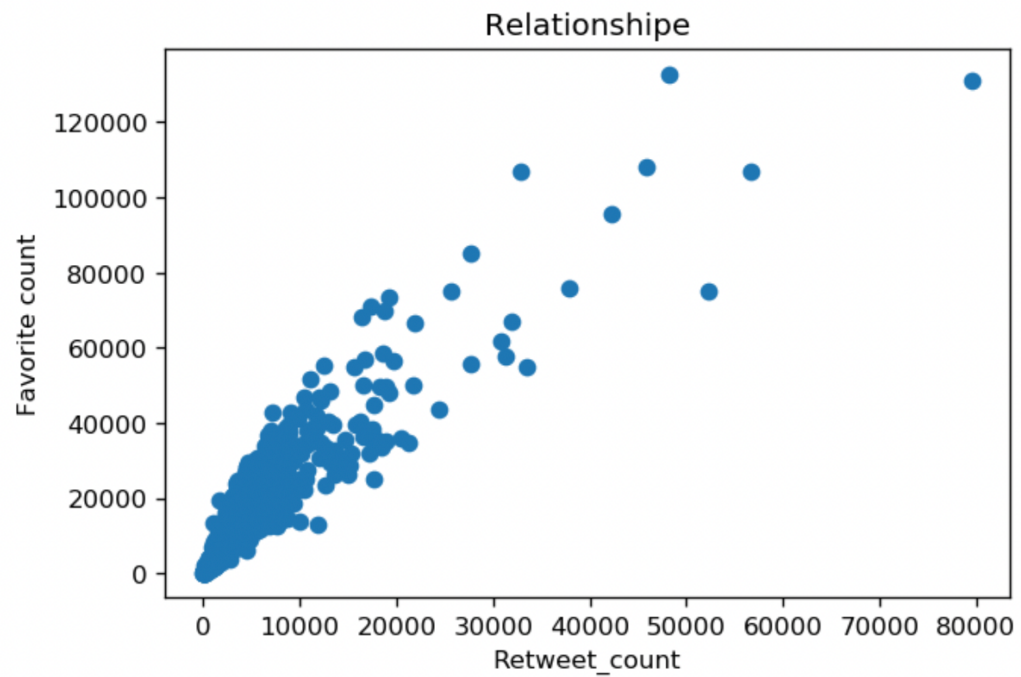
Qasem M. Dhamad

Udacity - Data Analyst - MCIT

Qasem M. Dhamad

Udacity - Data Analyst - MCIT

Dsiplaying the relationshipe between retweet_count and favorite count:

As shown below there is linear relationship between between retweet_count and favorite count

```
In [180]: plt.figure(figsize=(6,4),dpi=120)
          plt.scatter(twitter_archive_master['retweet_count'], twitter_archive_mas
          plt.title ('Relationshipe')
          plt.xlabel ('Retweet_count')
          plt.ylabel ('Favorite count');
```

Retweet_count

Dsiplaying the total number of rating_numerator:

Rating_numerator of "12" has the highest total number:

```
In [181]: plt.figure(figsize=(6,4),dpi=120)
          twitter_archive_master['rating_numerator'].value_counts().plot(kind='bar
```