
WeRateDogs Data Analysis Report

documentation of analysis and insights into final data

Analyze and visualize your wrangled data in your wrangle_act.ipynb Jupyter Notebook. At least three (3) insights and one (1) visualization must be produced.

- A. load DataFrame and fix category and Datetime format
- B. explore DataFrame by info() describe()
- C. ask question:
 - 1) which dog breed get most interactions and ratings
 - 2) which dog stage gets most interactions and ratings
 - 3) how is images with no dogs affect rating.
 - 4) How does rating affect interaction.
 - 5) type of relation between follower count and retweet and favorite count.
- D. Plot scatter matrix
- E. Calculate rating ratio
- F. Using plotting, rating ratio outliers were spotted
- G. Identify rating ratio outliers and delete index rows.
- H. 1 tweet with rating = 0 , remove row
- I. create new columns to with Boolean True and False to study relation between dog name and rating and interactions.
- J. create new interaction column with sum favorite and retweet.
- K. create new column relative_interaction to study the relation between user count and interaction
- L. record insights
- M. save modifies DataFrame.