

ACT REPORT

AUTHOR: JAJA TAMUNO-OMI

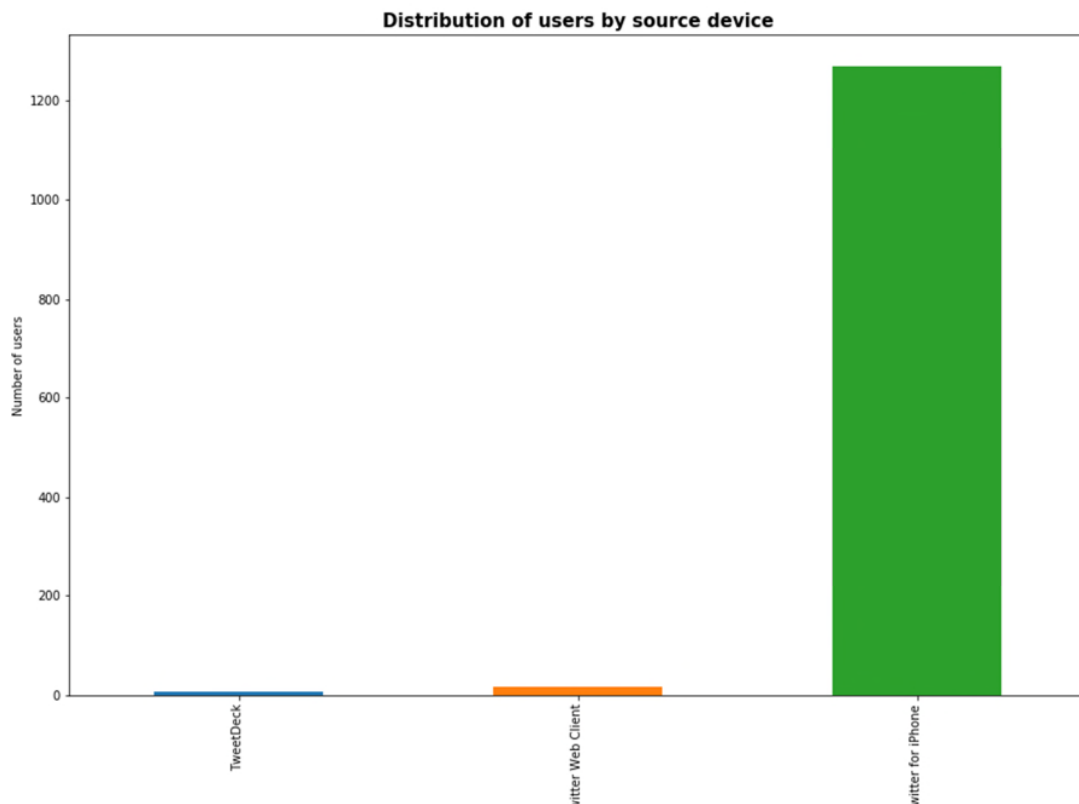
INTRODUCTION

Real-world data rarely comes clean. Using Python and its libraries, I gathered data from a variety of sources and in a variety of formats, assess its quality and tidiness, then clean it. The dataset used for this project is the tweet archive of Twitter user @dog_rates, also known as WeRateDogs which contain basic tweet information, additional gathering of retweet and favorite_count information was done and an image predictions file containing breed information generated by neural network model was provided by Udacity, after which assessing, and cleaning was done to enable analyses and visualizations

WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. The insights garnered are found below

1. What is the Distribution of user source devices?

I discovered that majority of the users access tweets via iPhone, with a very small number of users access their tweets through TweetDeck and Twitter web client.



2. Which is the most popular dog age group by likes?

The Dogtionary explains the various stages of dog: doggo, pupper, puppo, and floof(er) (via the #WeRateDogs book on Amazon),

Pupper

A small doggo usually younger. Can be equally, if not more mature than some doggos

Doggo

a big pupper, usually older. this label does not stop a doggo from behaving like a pupper

Puppo

A transitional phase between pupper and doggo. Easily understood as the dog equivalent of a teenager

Floof(er)

Any dog really. However, this label is commonly given to dogs with seemingly excess fur. Comical amounts of fur on a dog will certainly earn a dog this generic name

After analysing the dataset it was discovered that dogs described as **puppo** have the most likes

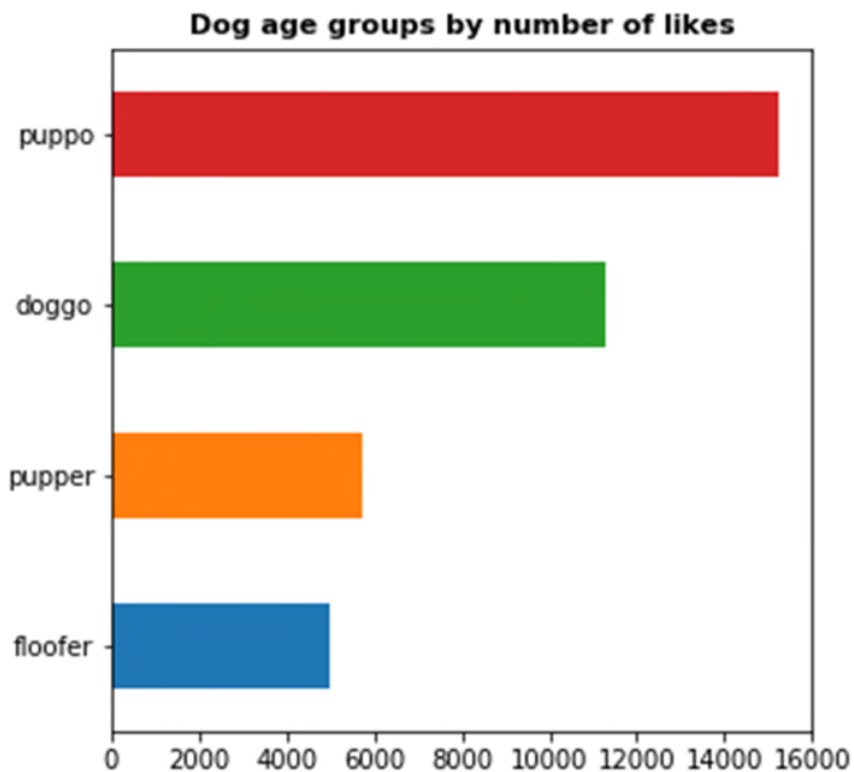




Fig: A puppo

3. Is there a correlation between likes and retweets?

There is a strong correlation between likes and retweets, where an increase in likes leads to an increase in retweets which indicates a positive relationship between both variables

