

Dogs on Twitter

Introduction

The data that we are interested in are the tweet archive from WeRateDogs. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. WeRateDogs has over 4 million followers and has received international media coverage. The archive in this study contains basic tweet data (tweet ID, timestamp, text, etc.) for thousands of tweets between November 2015 and August 2017.

Methodology

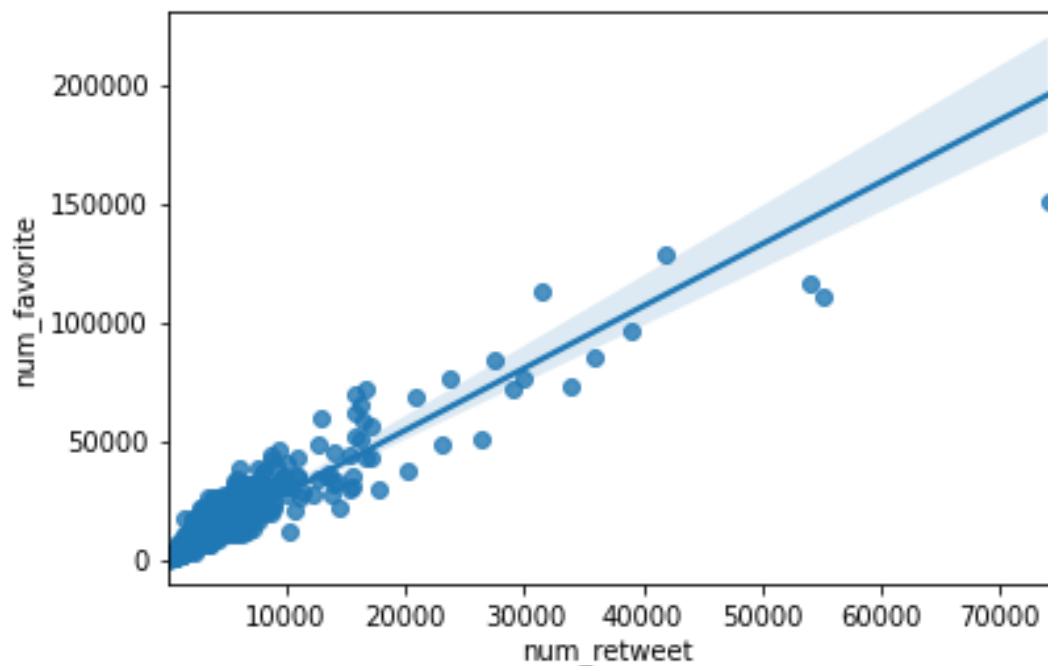
Real-world data rarely comes clean. Using Python and its libraries, we have gathered data from a variety of sources and in a variety of formats, assessed its quality and tidiness, and cleaned it.

The wrangling efforts are documented in a Jupyter Notebook, and these are showcased through analyses and visualisations.

Insights Gained

Some insights gained from finished wrangling of the data are:

- The most common Dogtionalary type is found to be pupper.
- The most common dog breed is found to be golden retriever, followed by Labrador retriever.
- Percentage of times a dog has been scored at or above 10, i.e., these dog owners must be beyond impressed with their dogs - this is found to be over 82%, or in other words, a large majority.
- Bedlington terrier is the highest retweeted & thumbed-up dog breed. Other dog breeds such as Afghan hound and French bulldog are also much liked on Twitter.
- There is a very strong positive correlation between number of retweet and favorite ratings, which is as expected. Also visualise this from the straight line below.



- A word cloud in a dog shaped contour is created below to show the most popular dog breeds on twitter between November 2015 and August 2017.

