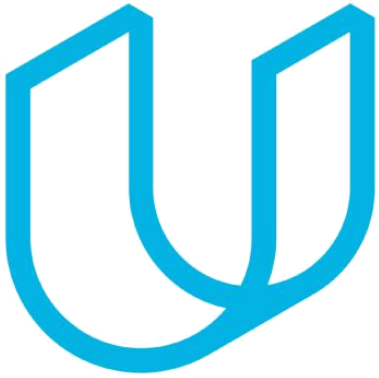


UDACITY

WRANGLE AND ANALYZE DATA

ACT REPORT

PRESENTED BY: Sunny Paul

CONTENTS

WRANGLE AND ANALYZE DATA

PURPOSE

The purpose of this document is to communicate the insights and visualizations produced as a result of my wrangling efforts against the WeRateDogs Twitter archive.

INSIGHTS:

INSIGHT #1 - HIGHEST AND LOWEST RATED DOG

The highest rated dog is named Atticus and he was rated with 1776/10

However, the lowest rated tweet was not related to dogs at all

This emphasizes the concept to re-iterate the process of Wrangling Data. Picking the second lowest rated dog, the lowest rated dog is named Crystal and she was rated with 2/10 for being irresponsible

INSIGHT #2 – TOP FAVORITE COUNT OF A TWEET

The top favorited tweet has been favorited 144175 times which was rated with 13/10

INSIGHT #3 – TWEET SOURCES

I was interested to find out the top tweet sources and it turned out there are only 3 sources:

Source	Count
Twitter for iPhone	2000
Twitter Web Client	29
TweetDeck	11

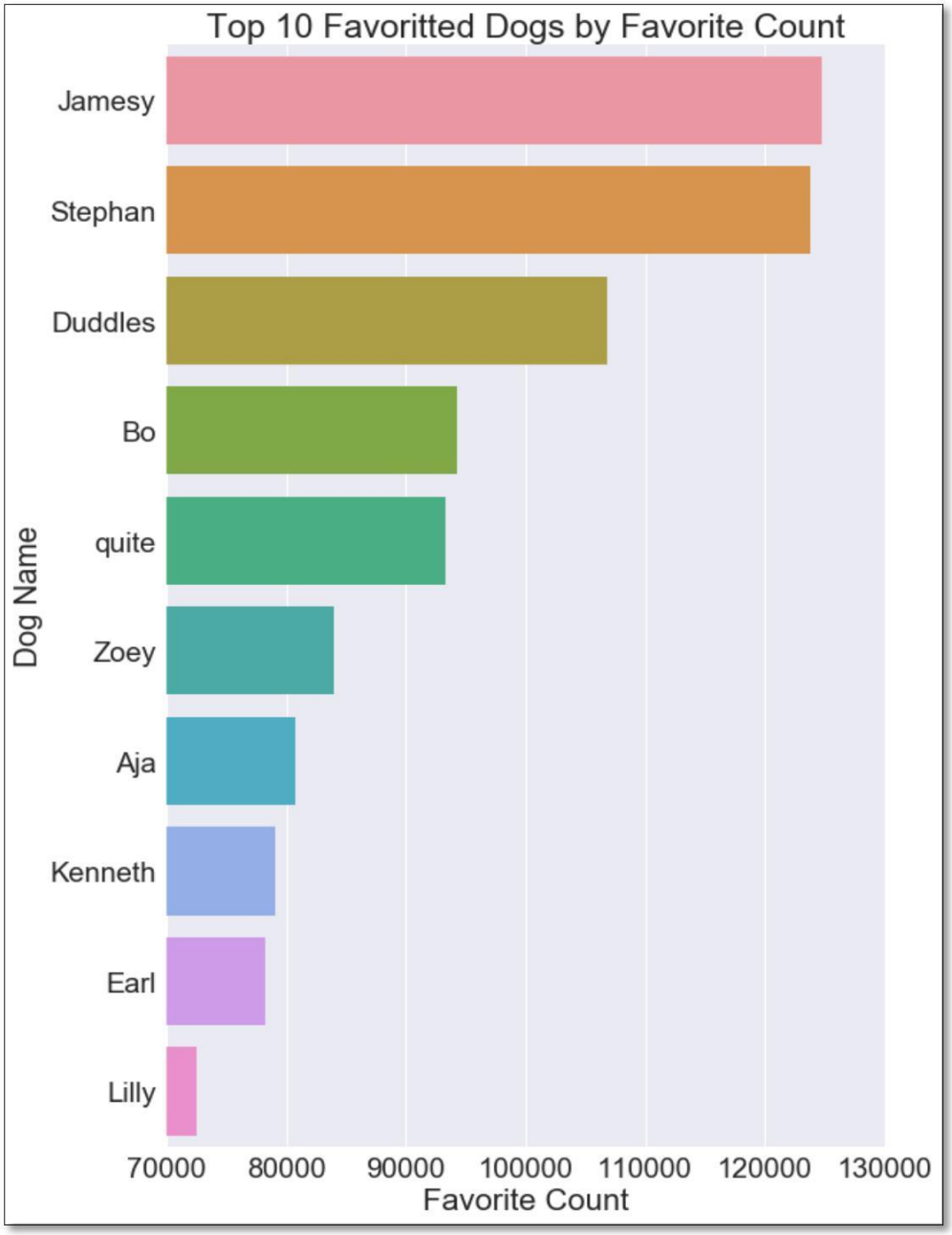
VISUALIZATION:

I have created (2) visualizations of my own interest against the wrangled tweets archive

VISUALIZATION #1 – TOP 10 FAVORITTED DOGS BY FAVORITE COUNT

This visualization shows the top favorited dog by favorite count for named dogs. For this reason, the result found in **Insight #2 – Top favorite count of a tweet** is not listed part of this visualization

The top favorited dog is named Jamesy and he earned 124742 favorite counts



VISUALIZATION #2 – SUCCESSFUL BREED PREDICTION FOR ALGORITHM P1

This visualization shows the successful breed prediction when using the algorithm P1. The algorithm highest successful breed prediction is Golden Retriever which was successfully predicted 148 times. However, the lowest successful breed predictions are **standard_schnauzer**, **groenendael**, **EntleBucher**, **silky_terrier**, **Scotch_terrier**, **clumber**, and **Japanese_spaniel** which were successfully predicted 1 time only