

wrangle_report

June 25, 2022

0.1 Reporting: wrangle_report

- Create a **300-600 word written report** called "wrangle_report.pdf" or "wrangle_report.html" that briefly describes your wrangling efforts. This is to be framed as an internal document.

0.2 Title: Data-Wrangling Report On WeRateDogs Twitter Account Tweets From 2015-2017

0.2.1 Presented by: Ayobami Alaran

0.2.2 Date: 25-06-2022

0.3 Introduction

WeRateDogs is twitter account with username dog_rates where people post tweets about dogs and rate them. [click here](#) for more info on the twitter account

Data-Wrangling Analysis process which includes **Gathering, Assessing and Cleaning** was performed on the data obtained from the twitter account using three various method of gathering data for the gathering stage; >- data_1 = twitter_archive_enhanced.csv (which is already available for download in the classroom) >- data_2 = image_predictions.tsv (queried from the udacity server using request library) >- data_3 = twitter_data.json (scraped from twitter with twitter Api v1)

0.3.1 Project Objective

Perform data-wrangling analysis on the WeRateDog tweets from 2015 - 2017, which involves gathering, assessing and cleaning the datas

0.3.2 Assessing The Data

The Gathered data was assessed visually using Ms-Excel Spreadsheet and .head and .tail function in python to some quality issues(content issues) and tidiness(Structural issues). Moreso, the dataset was also assessed programmatically using some funtions in python to further investigate te dataset for hidden issues.

0.3.3 Cleaning

After assessing and documenting the detected issues, the cleaning stage was initialize to fix the issues highlighted to get the dataset ready for analysis by writing and automating some python codes to fix the issues.

0.3.4 Project Help Sources

Analysis and some syntax ideas was acquired from ALX DA Udacity classroom, data science community such as stackoverflow and geeks4geeks, datacamp and some code documentation websites such as tweepy docs and regular expression docs.