# wrangle_report

March 15, 2021

## 0.1 Wrangle Report

## 0.2 Projects details

About this project

Gathering Data

Assesing Data

Cleaning Data

Storing, analyzing and visualizing your wrangled Data

## 0.3 Gathering Data

This project involved data from different sources such as we gonna see below

Enhanced Twitter Archive

The archive called "twitter-archive-enhanced-2" was provided by Udacity Course and i was used pd.read_csv('twitter-archive-enhanced-2.csv') to import them into a DataFrame

Imagen prediction

The Imagen prediction was a File (image_predictions.tvs)hosted on Udacity servers and was download using a python library and URL information about it

Twitter API

I use a Python library called Tweepy API. The Tweepy API returned the data in JSON format that contains retweet count and favorite count for each tweet to obtain through Twitter API

Storing, analyzing and visualizing your wrangled Data

## 0.4 Assesing Data

First off the data have been gathered from different resourses such as enhanced twitter archive, downloaded archive through Udacity and tweepy API through Twitter. Then I separated the issues detected in quality and tiddiness issues and prepare the data to cleaning it.

## 0.5 Cleaning Data

When I had to clean the data, the first thing that should do was assessed for issues, the second thing was to merge all of them into the main data frame. Although I created a copy of this data

frame cleaned. This methodology helped me to got accurate data to the last item and to get the accurate visualizing data.

## 0.6 Storing, analyzing and visualizing your wrangled Data

The last step was to analyze the data cleaned and provide accurate visualization of the Data that gonna be explained in the act_report.pdf

[ ]: