

Wrangle report

The given task is to wrangle the data given

- Data wrangling, which consists of:
 - Gathering data (downloadable file in the Resources tab in the left most panel of your classroom and linked in step 1 below).
 - Assessing data
 - Cleaning data
- Storing, analyzing, and visualizing your wrangled data
- Reporting on 1) your data wrangling efforts and 2) your data analyses and visualizations

- ✓ Now we have jupyter notebook to complete the task
- ✓ To complete the task, I gathered data from twitter account “We Rate Dogs”.
- ✓ After reading the data gathered, found some issues to correct before finding insights .

After gathering each of the above pieces of data, assess them visually and programmatically for quality and tidiness issues. Detect and document at least **eight (8) quality issues** and **two (2) tidiness issues** in your `wrangle_act.ipynb` Jupyter Notebook. To meet specifications, the issues that satisfy the Project Motivation (see the *Key Points* header on the previous page) must be assessed.

- ✓ We going to start with tidiness

- 1 Tweet_df and image predications should be merged with twitterArchive
- 2 Columns puppo ,puppo, flppfer and doggo merged to be in one

column

- ✓ After tidiness we start with quality issues
1. Tweet_id is integer instead of object.
 2. Converte timestamp to date type
 3. Time and date are in one column (timestamp) in twitter archive.
 4. The columns in_reply_to_status_id, in_reply_to_user_id, retweeted_status_id and retweeted_status_user_id should be object (string)
 5. source column is not readable
 6. Remove the duplicated rows from jpg_url column
 7. The column 'name' has NAN values that assigned as None

8. Drop the columns that won't be used for analysis(`in_reply_to_status_id`, `in_reply_to_user_id`).

Then after storing it into `wrangled_df`

No, we have cleaned and addressed some key issues

It's time for us to move on to find insights and visualizing some columns and analyzing the insights from data

Reporting for this Project

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, which is now what is been doing

Create a **250-word-minimum written report** called `act_report.pdf` or `act_report.html` that communicates the insights and displays the visualization(s) produced from your wrangled data. This is to be framed as an external document, like a blog post or magazine article

After these the has come to end