Project Details

Your tasks in this project are as follows:

- Data wrangling, which consists of:
  - Gathering data
    - Download any preferred datasets from **UCI repository** or **Kaggle**.
  - Assessing data.
    - After gathering your 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**.
  - Cleaning data.
    - Clean each of the issues you documented while assessing.
- Storing, analyzing, and visualizing your wrangled data
  - Store the clean DataFrame(s) in a CSV file. If additional files exist because multiple tables are required for tidiness, name these files appropriately.
  - Use *matplotlib* and *seaborn* packages for visualizing.
- Apply any classification/clustering algorithm on your data and visualize your results
  - Hint: use *scikit-learn* Python package for classifying or clustering: https://scikit-learn.org/stable/
- Reporting on 1) your data wrangling efforts and 2) your data analyses and visualizations
  - Create a **300-600 word** written report called **wrangle_report.pdf** that briefly describes your wrangling efforts (i.e., define and describe briefly **at least eight (8) quality issues** and **two (2) tidiness issues**).
  - Create a **250-word-minimum** written report called **act_report.pdf** that communicates the insights and displays the visualization(s) produced from your wrangled data (i.e., put graphs plotted by *matplotlib* package and describe them briefly).
  - Create a **100-word-minimum** written report called **prediction_results.pdf** that briefly describes the used classification/clustering algorithm and concludes the output.