Data Wrangling -Milestone Report

1. Heart Disease UCI dataset is fetched from Kaggle.

Reference URL: <https://www.kaggle.com/ronitf/heart-disease-uci>

1. The Cleveland database is the only one that has been used by ML researchers to this date. The "goal" field refers to the presence of heart disease in the patient. It is integer valued from 0 (no presence) to 4.
2. The dataset is model ready as they are no missing or Null values found.
3. We were able to determine the outliers for the cholestrol and Resting blood pressure variable However we need to perform EDA on the column to understand the impact of the outliers on the data