**Missing value:**

Calculated missing value percentages for each row and column.

Rows which are having missing value percentage greater than 50% was flagged. 6 rows found.

Columns which are having missing value percentage greater than 50% was flagged. 0 rows found.

**Imputation method:**

1. Simple imputer
   1. The SimpleImputer class provides basic strategies for imputing missing values.
2. Iterative imputer
   1. machine-learning based approach by imputing missing values as a function of remaining features using IterativeImputer()

**Checking validity:**

Indices which are having AGE not in range [65,89]

Indices which are having SEX not as 1 or 2

Indices which are having Laboratory not in range [1,6]

Indices which are having invalid AGE\_GROUP

Dropped all above indices from dataframe.

**Useless columns:**

Dropped all useless columns which are having all values as NaN

**Duplicate rows:**

All duplicate rows were deleted

**Box Plot for a factor SEX:**

Median age of male is greater than median age of female

All other cases female is dominant

**Box Plot for a factor Laboratory:**

Lab 4 there are more aged patients

Lab3 there are more inorganic phosphorous.