Logistic Regression: Heart Attack prediction (filling the NA/NaN values)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Column | No. of NaNs | Fill missing data |  |  |  | Remarks |
| 0 | 0 | NA |  |  |  |  |
| 1 | 0 | NA |  |  |  |  |
| 2 | 105 | Data is categorical; hence, replacement is “mode” replacement |  |  |  |  |
| 3 | 0 | NA |  |  |  |  |
| 4 | 29 | Data is highly skewed; hence, NaN-replacement is “median” replacement |  |  |  |  |
| 5 | 53 | Data is categorical; hence, replacement is “mode” replacement |  |  |  |  |
| 6 | 0 | NA |  |  |  |  |
| 7 | 0 | NA |  |  |  |  |
| 8 | 0 | NA |  |  |  |  |
| 9 | 50 | data is skewed; hence, NaN-replacement is “median” replacement |  |  |  |  |
| 10 |  |  |  |  |  |  |
| 11 |  |  |  |  |  |  |
| 12 |  |  |  |  |  |  |
| 13 | 1 | data is skewed; hence, NaN-replacement is “median” replacement |  |  |  |  |
| 14 | 388 | data is skewed; hence, NaN-replacement is “median” replacement |  |  |  |  |
| 15 | 0 | NA |  |  |  | This is the output; y data |