

- 1- ReplaceMissingValues replaced missing values of numerical data by using mean values of the dataset. For qualitative missing values, ReplaceMissingValues replaced missing values by using mode of the dataset.
- 2- For numeric attributes, only standard deviation was changed. For nominal attributes, weight of most common attribute value was changed.
- 3- In “duration” attribute, there was only one missing value and it replaced with 2.1607142857142856 since the mean of this attribute was about 2.161 . Because data points tends to be close to the mean, standard deviation was decreased.
In “standby-pay” attribute, there was 48 missing values which was 84% of all values. When it was raw statistics, values were distributed symmetrically on maximum and minimum values. After I applied ReplaceMissingValues, values were distrubed on mean. That resulted in huge decreasing on standard deviation from 5.028 to 1.9.
In “wage-increase-third-year” attribute, same things happened with “standby-pay” attribute, only values were different.
In “wage-increase-first-year” attribute, there was only one missing value also and it replaced with 3.803571428571428 where the mean was 3.804. In this attribute, same things happened with “standby-pay” attribute.
- 4- It was suitable since “class” attribute was not effected by replaced missing values. Also it increased the distinctness of attributes.