

1. **Complete Case Analysis**

Use when missing data is distributed at random in the dataset, then we can remove them otherwise our model fails to work.

A screenshot of a computer

Description automatically generated

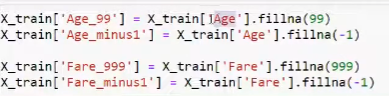
1. If data is numerical then plot histogram to check the distribution.
2. If data is categorical then check whether the ration of the categories are same before and after the CCA.
3. **Numerical Data | Simple Imputer**
4. **Mean – Normal Distribution**
5. **Median – Skewed Distribution**

A blackboard with white text

Description automatically generated

1. **Arbitrary Value Imputation**

Assign a random number like -1 or 0 or highest num into the missing number.



1. **End of Distribution Imputation**

Assign a (mean + 3 \* sd) or (mean – 3 \* sd) to the missing value

A blackboard with white text

Description automatically generated A blackboard with white text

Description automatically generated

Disadvantages

1. **Categorical Data | Simple Imputer**
2. **Random Sample Imputation**

Assign the missing value with the random value from the remaining values of the column

Disadvantage:

