



























Imputing Data

The **Impute** option allows you to perform in-place imputation by filling missing values with aggregates computed on the “na.rm’d” vector. Additionally, you can also perform imputation based on groupings of columns from within the dataset. These columns can be passed by index or by column name using the Group By option. Note that if a factor column is supplied, then the method must be Mode.

he following options can be specified when imputing dataset:

* **Frame**: The dataset containing the column to impute
* **Column**: A specific column to impute.
* **Method**: The type of imputation to perform. Mean replaces NAs with the column mean; Median replaces NAs with the column median; Mode replaces with the most common factor (for factor columns only).
* **Group By**: If the **Method** is either Mean or Mode, then choose the column or columns to group by.
* **Combine Method**: If the **Method** is Median, then choose how to combine quantiles on even sample sizes. Available **Combine Method** options include Interpolate, Average, Low, and High.

