**Fill with the Mean or the Median**

* **If the data were missing completely at random**, then mean /median imputation might be suitable. You might also want to capture if the data was originally missing or not by creating a “missing indicator” variable.

Both methods are extremely straight forward to implement.

* **If a variable is normally distributed**, the mean, median, and mode, are approximately the same. Therefore, replacing missing values by the mean and the median are almost equivalent.

Replacing missing data by the mode is not appropriate for numerical variables.

* **If the variable is skewed**, the mean is biased by the values at the far end of the distribution.

Therefore, the median is a better representation of the majority of the values in the variable.

Having said that, you should avoid filling with mean, if you observe and increasing or decreasing trend in your data, in which case you might want to consider ‘interpolation’ and [predicting the missing value using ML approach].