

## REGRESSION TREES:

It is worked when we have dependent variable is continuous data.  
We cannot calculate Miss classification Error.

So, How to calculate Error?

How to split your dependent variable when it is in continuous in nature?

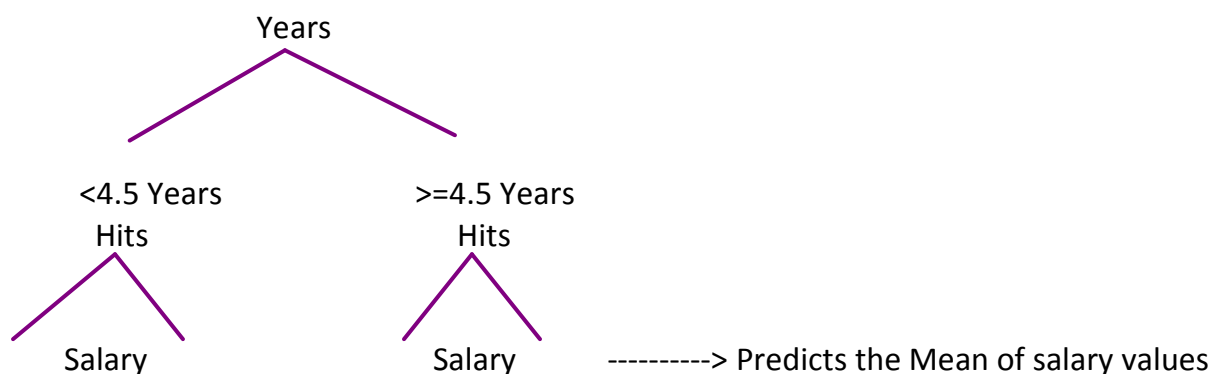
For what value of split where I am getting the minimum classification error that particular value going to finalized to split. This process will be repeated for all root nodes and intermediate nodes as well wherever your continuous data variable will be available.

### Example:

We use the Hitters data set to predict a baseball player's Salary based on Years, Hits  
So to predict Salary which is continuous variable we need to start split either from Years or Hits.

Let us choose Years as first variable

The predicted salary for these players is given by the mean response value for the players in the data set with Years<4.5.



So, Finally we have our own Actual salary values and Calculated Predicted values.

$$\text{Error} = \text{Actual} - \text{Predicted}$$