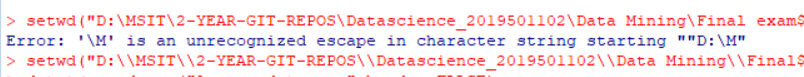
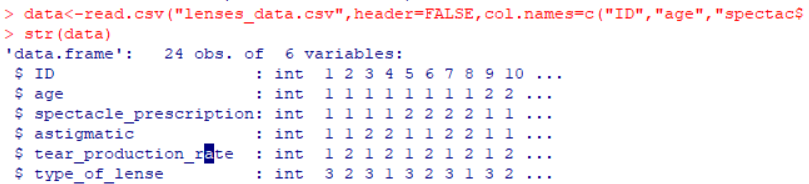
Q3. Build Decision Trees by using i) information gain and ii) misclassification error rate for Lenses Data Set provided at <http://archive.ics.uci.edu/ml/datasets/Lenses>. In terms of tree size what do you conclude comparing these two? (10M)

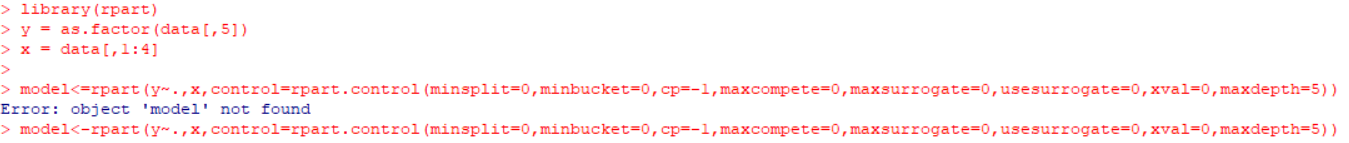
Solution:



Setting up the environment where our datasets are present.



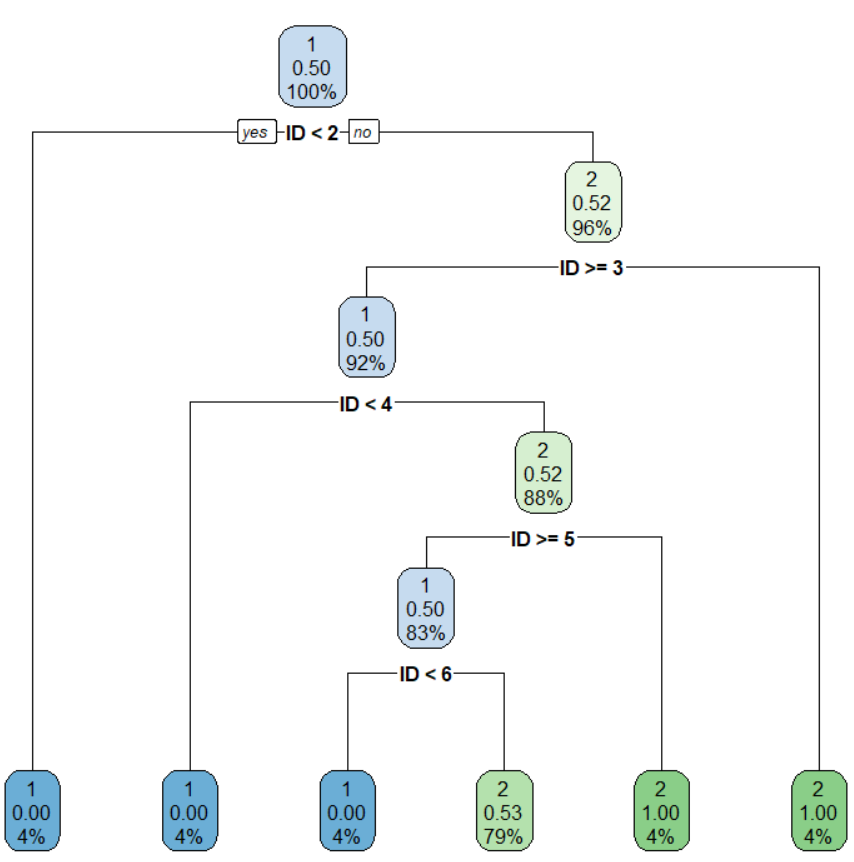
Reading the data and giving the data column names and printing them.



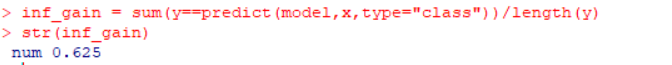
* Importing rpart library which is used for partitioning a dataset into decision trees for classification purposes.
* Taking x and y as training set and testing sets.
* For model, giving default values for the variables that should be passed to rpart function.



Importing rpart.plot for plotting the decision tree by passing the model as argument for plot function.



In the above graph, the decision tree is giving us values for each value from the dataset their confidence and support values.



Calculated information gain for comparison.



Calculated misclassification error as m\_error.

By comparing the above two, we conclude that the error difference between the two are so small, so neglible, so the decision tree is almost accurate by point of 0.2.