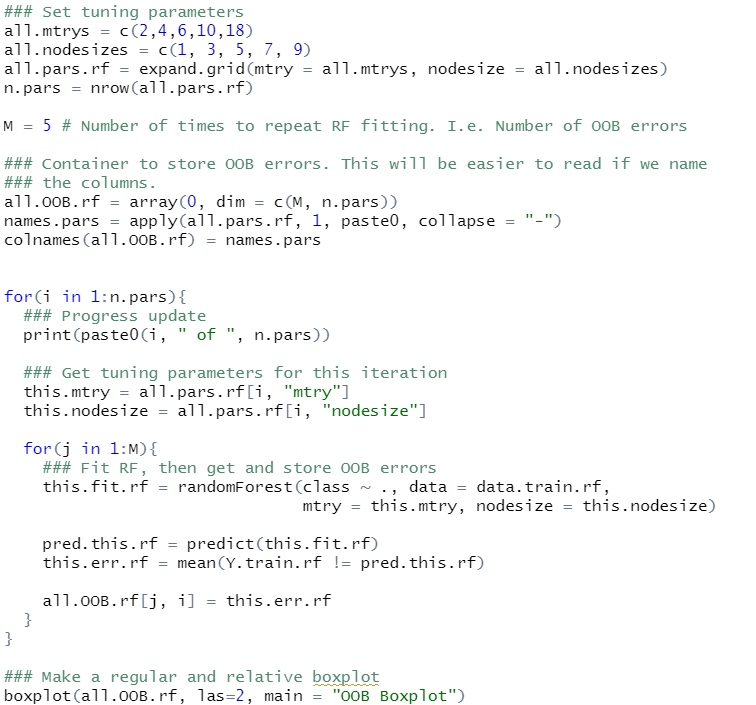
3. Tune the number of variables and the node sizes for the random forest using its internal

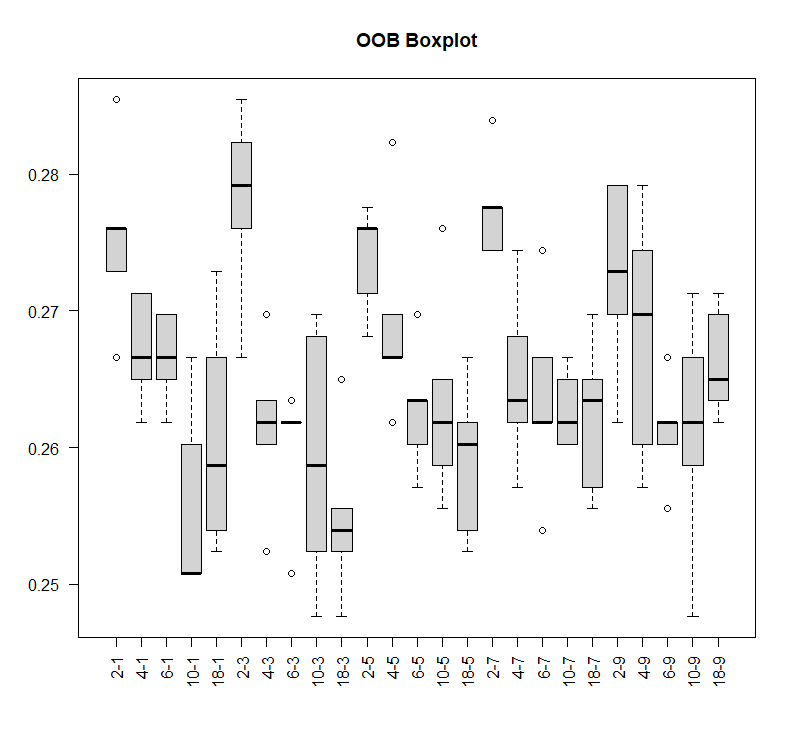
out-of-bag error. Since *p* = 18 and the default is mtry= p18 \_ 4, use a grid of

2,4,6,10,18. For node size, default is 1, so try 1,3,5,7,10. Rerun 5 random forests of

500 trees each.

(a) **Report the average misclassification rate for each combination**



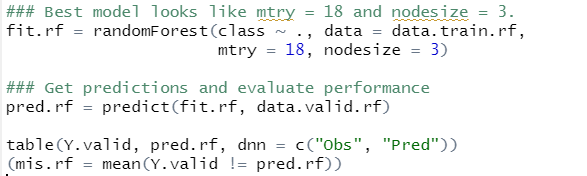


(c) **Which combination do you think is best?**

**->** 18-3 combination looks best

(d) **Report test error for that combination. Compare it to the default RF:**

**did it help? If so, how does it compare to other methods?**





It shows normal performance, but I don’t know why, it doesn’t show better result than default RF.