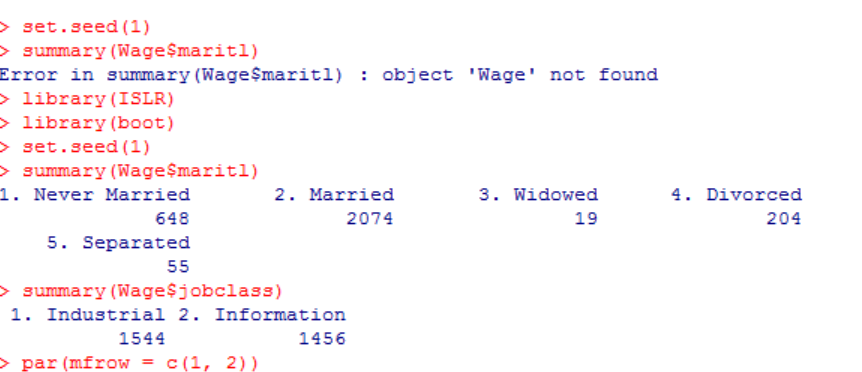
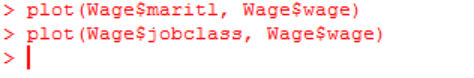
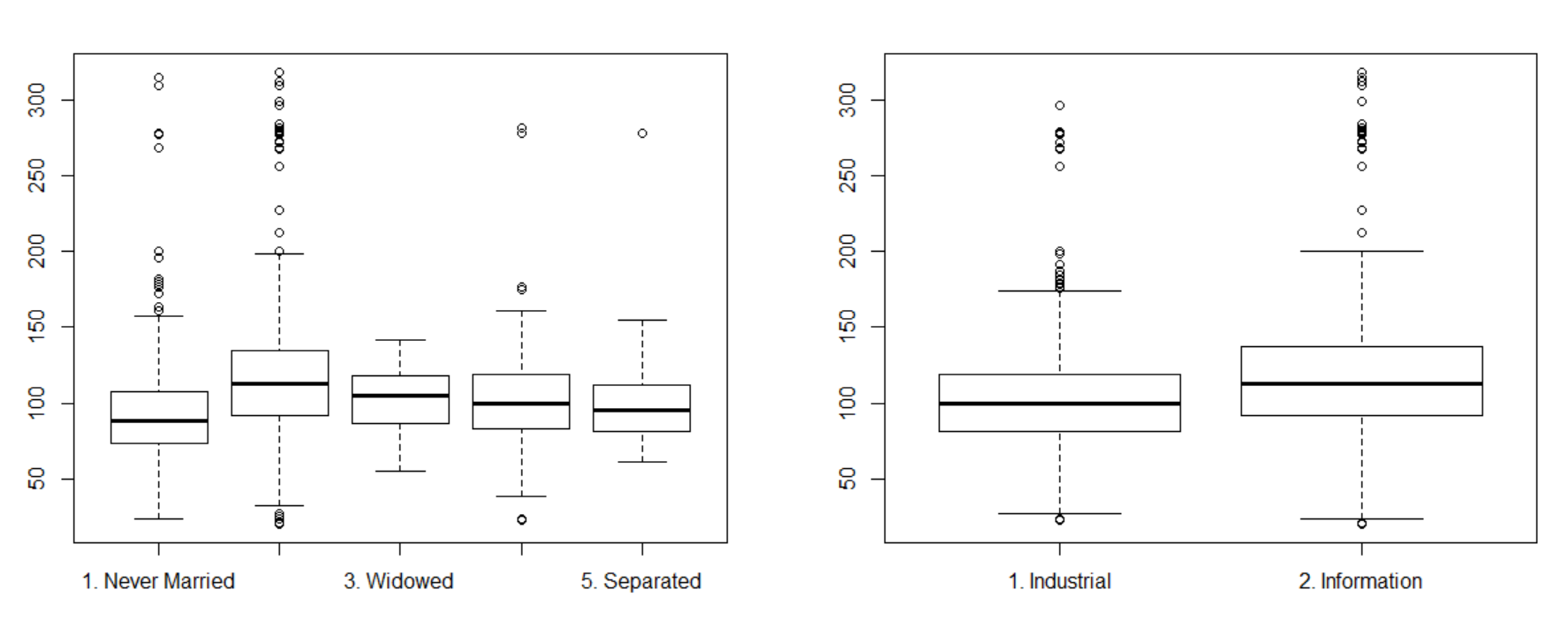
**HOME WORK FOUR Qian Yu**

**ST635 Intermediate Statistical Modeling for Business Fall 2017**

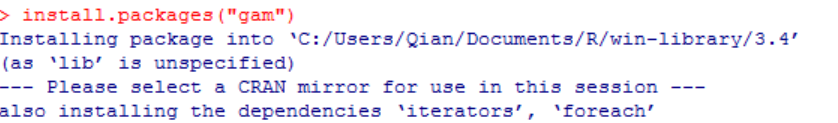
**Exercise 7.9: Problem 7**

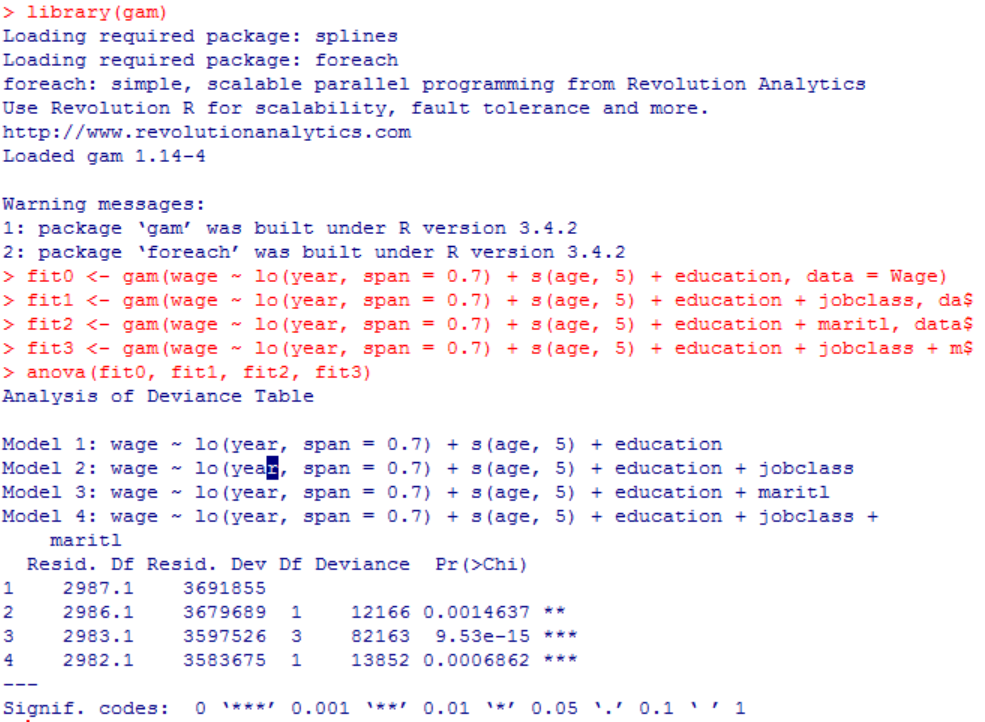


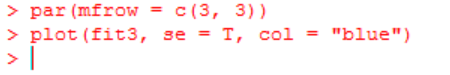


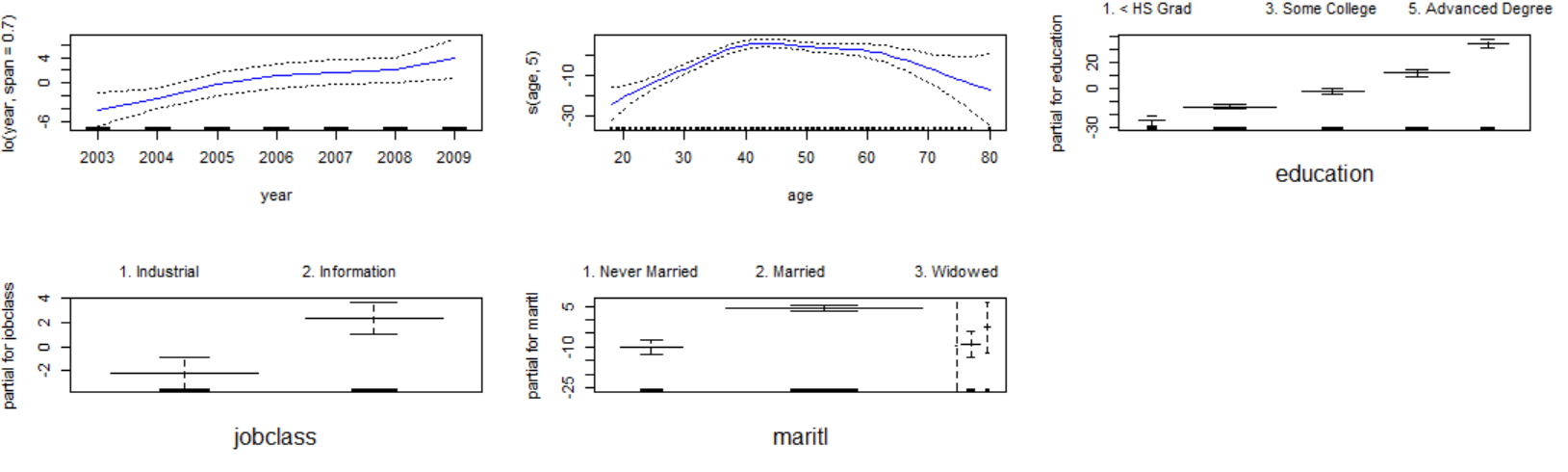


We can say that a married couple earns more money on average, and informational jobs earns more on average as well.





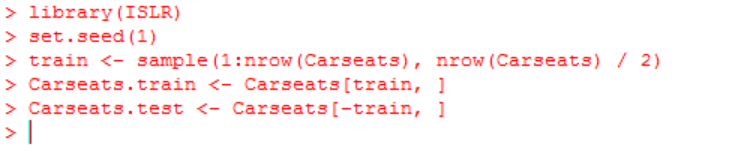


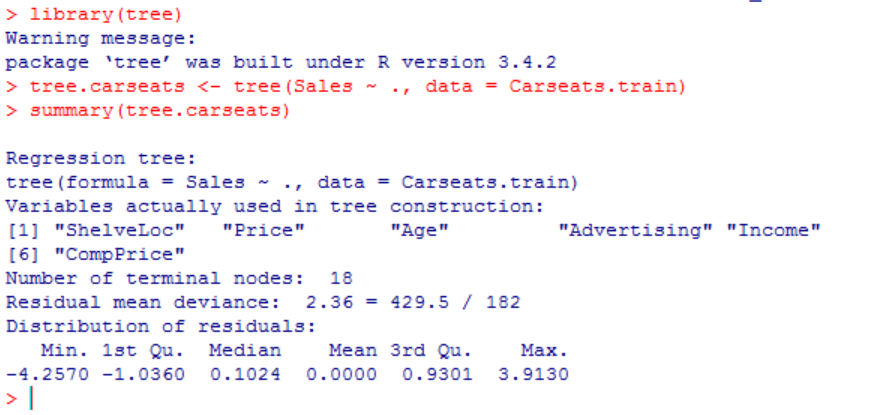


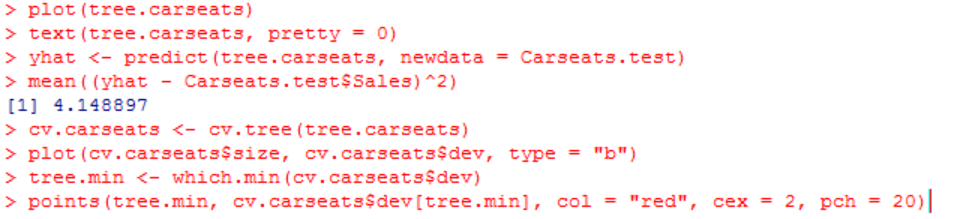
After use GAM to predict “wage” using natural spline functions of “year”, “age”, “education”, “jobclass” and “maritl”, it shows that the model “fit3” is significantly better.

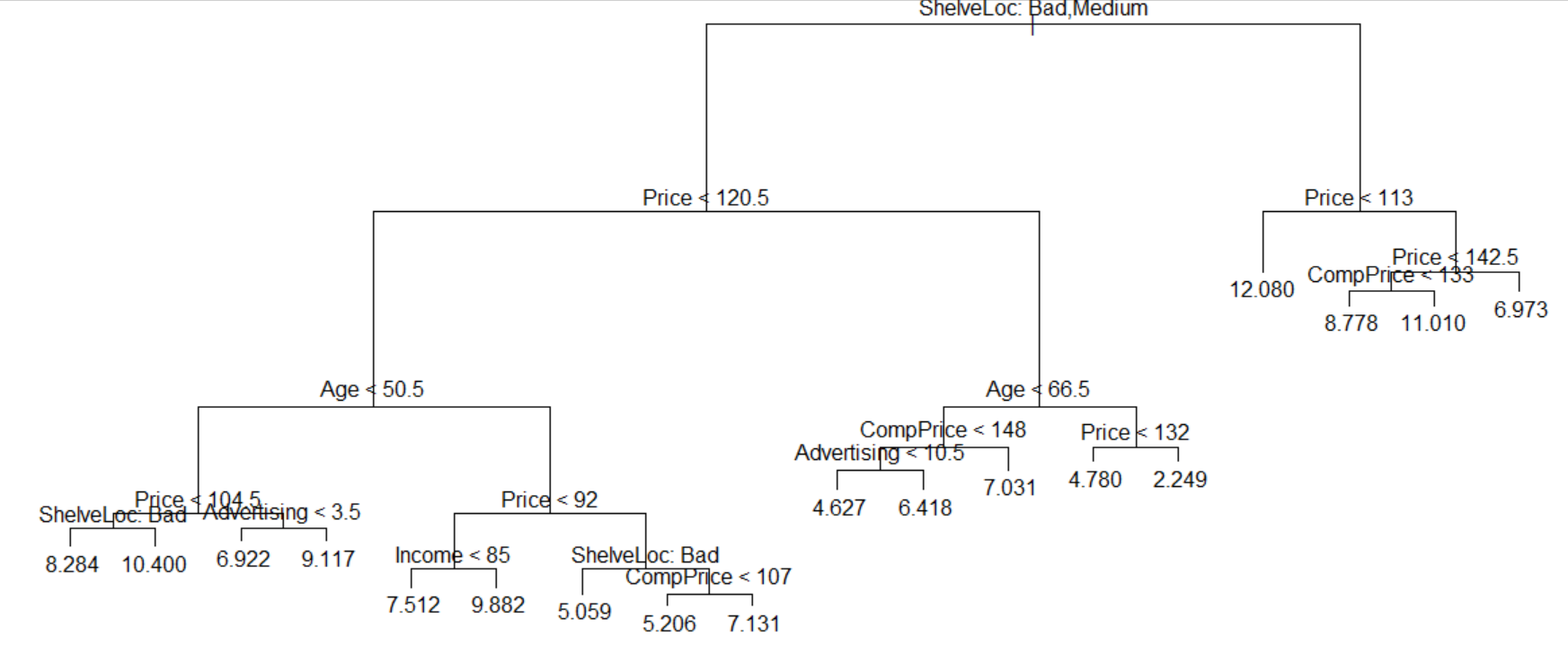
**Exercise 8.4: Problem 8**

**a&b.**



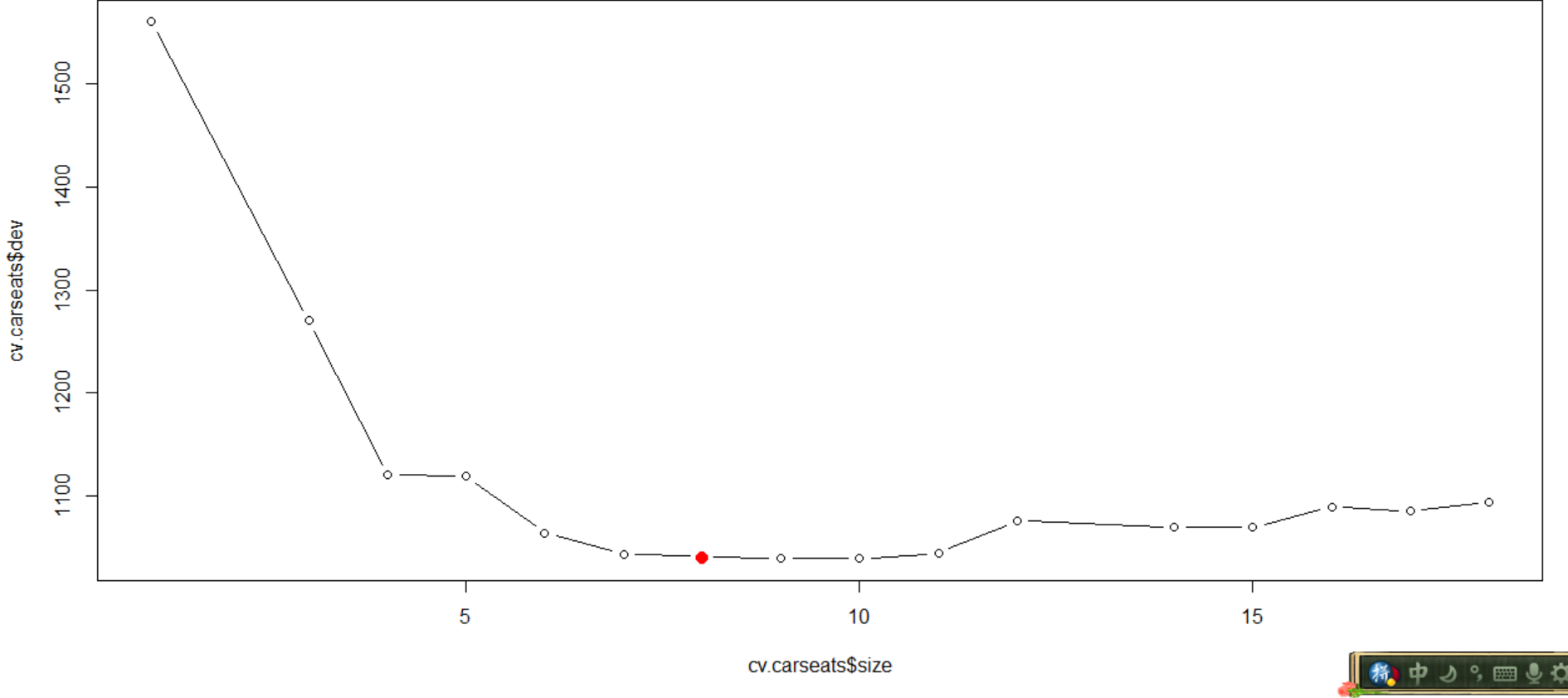


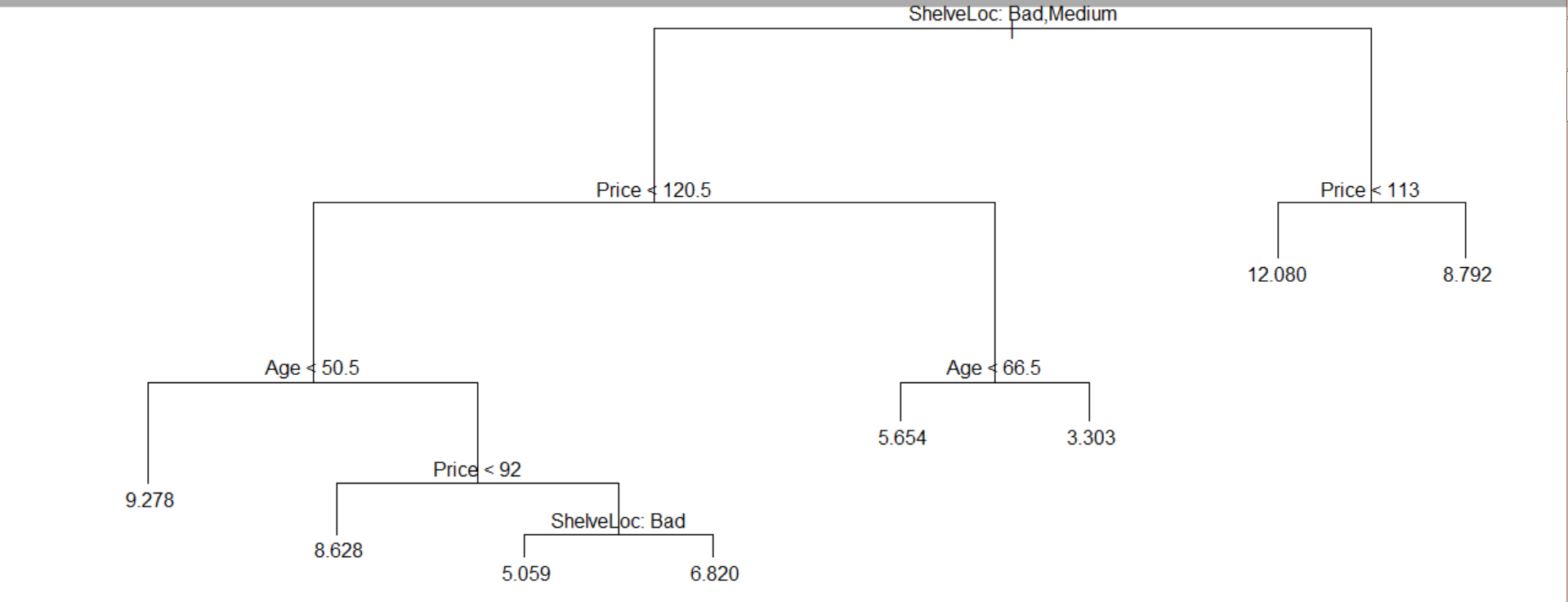


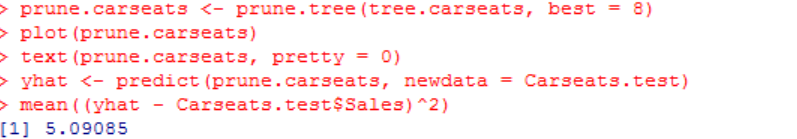


We may say that the Test MSE is about 4.149.

**c.**

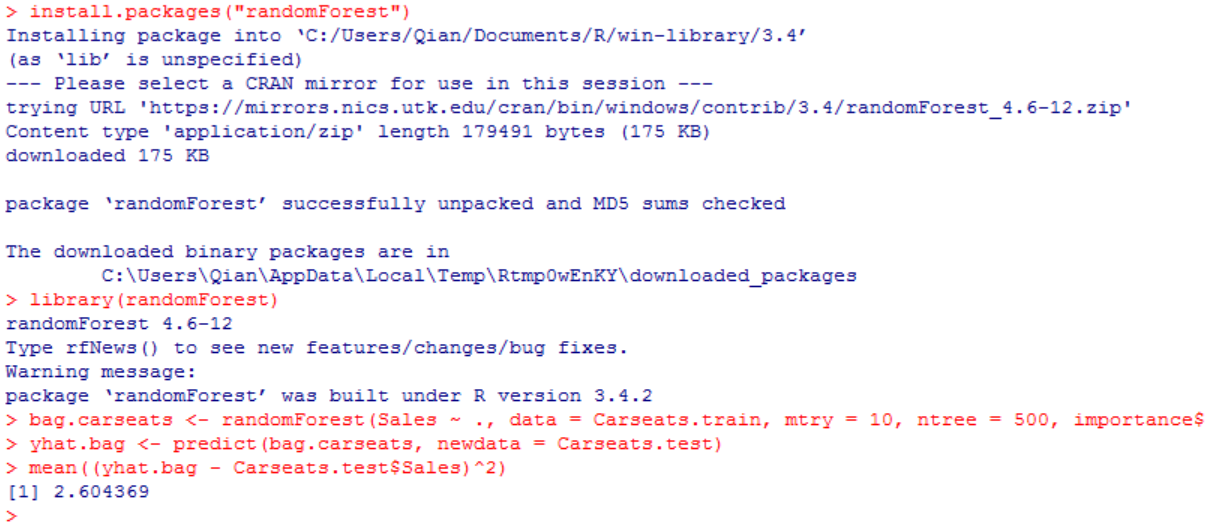


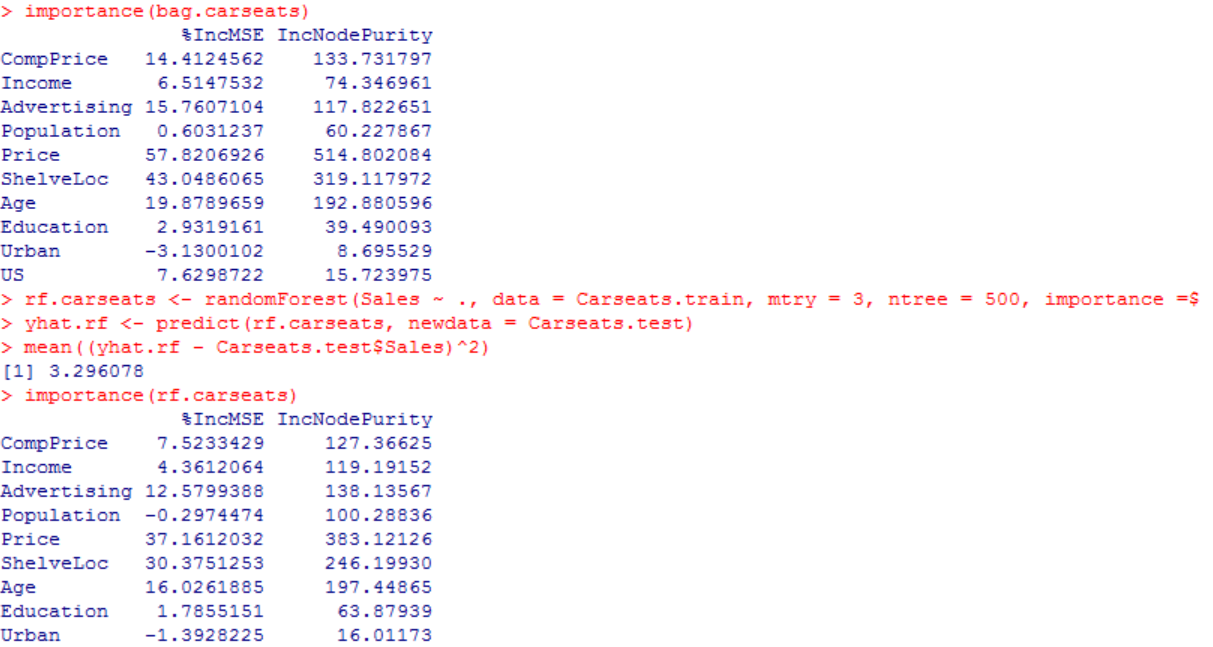




Yes, we can. The Test MSE increases to 5.091 after we pruned the tree to obtain the 8-node tree.

**d&e.**





Bagging decreases the Test MSE to 2.604 and price” and “ShelveLoc” are the two most important variables; while using random forests we have a Test MSE of 3.296 and “Price” and “ShelveLoc” are still the two most important variables.