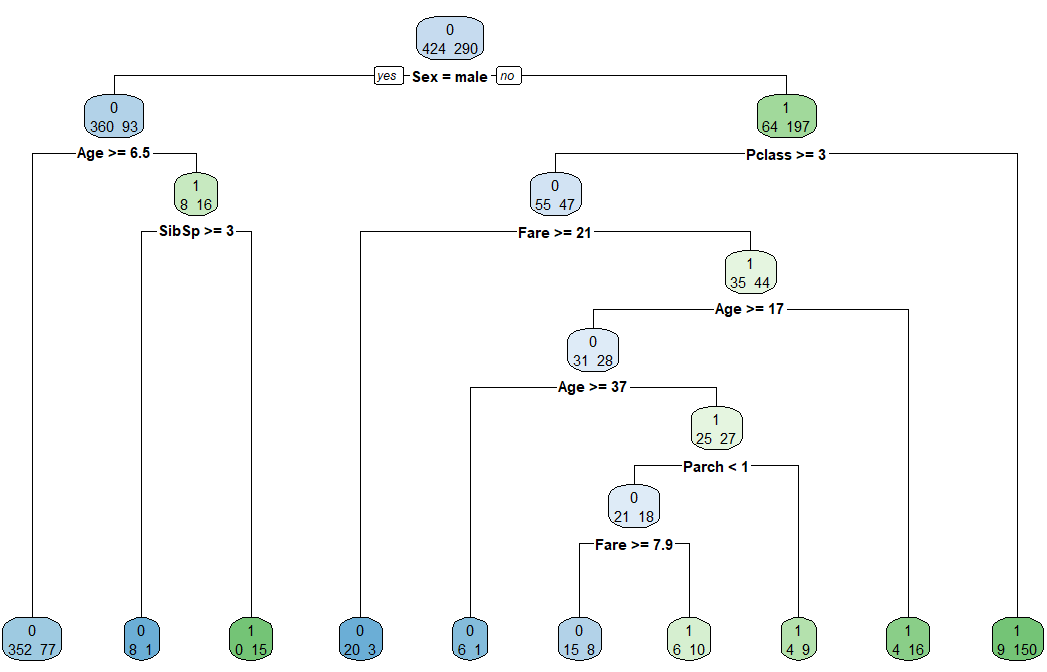
Decision Tree Classification on “Titanic” Report

In “Titanic train.csv” file, the model chooses columns “Pclass”, “Sex”, “Age”, “SibSp”, “Parch”, “Fare”, “Embarked” as variables and column “Survived” as target, while other columns are meta attributes or “PassengerId”. Also, the model removes the rows which have at least one null field. To determine if a person is survived, the model shows the most significant factor is “Sex”. The survival rate of the female is much higher than that of the male.

The model computed in R is plotted below:



Using this model to evaluate the “Titanic test.csv” file, the result is shown in “Prediction Titanic.xlsx” file.

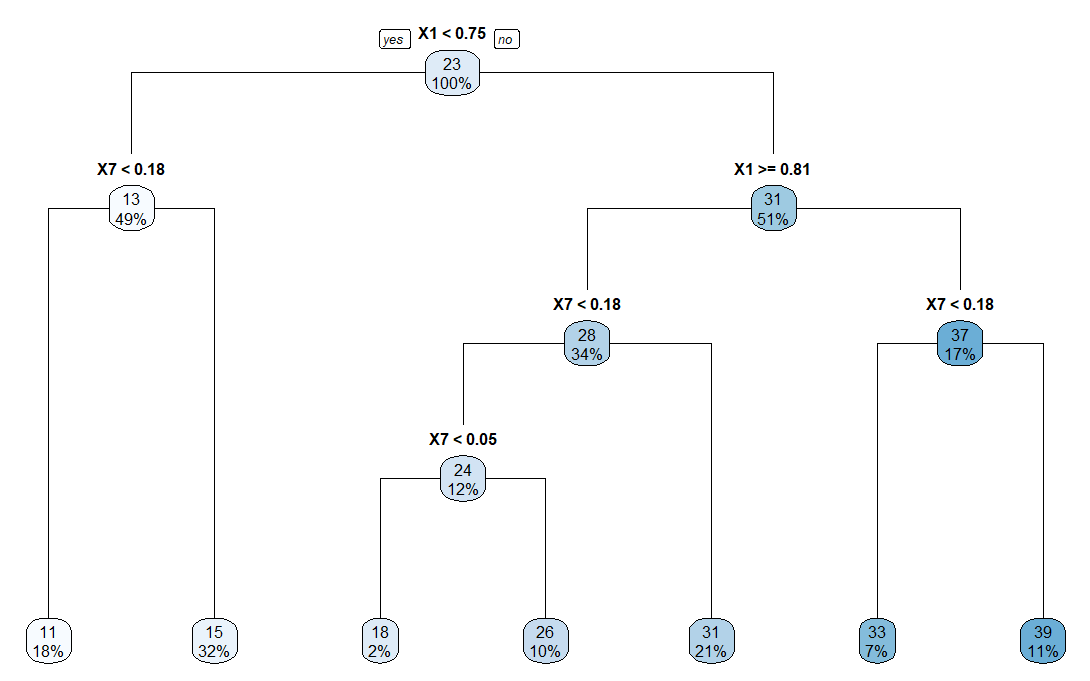
Decision Tree Regression on “Energy Efficiency” Report

The model removes the rows which have at least one null field. The model randomly chooses 650 rows for training and the rest is used to predict. For Y1, variable’s importance is shown below:

X1 X2 X4 X5 X3 X7

23 23 21 21 9 2

The model is plotted below:

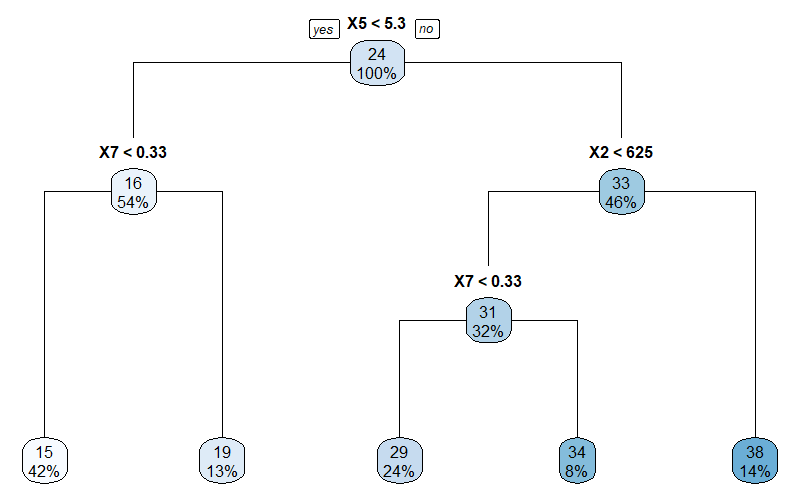


For Y2, variable’s importance is shown below:

X1 X2 X4 X5 X3

24 24 23 23 5

The model is plotted below:



The prediction is shown in “Prediction Energy Efficiency.xlsx”.