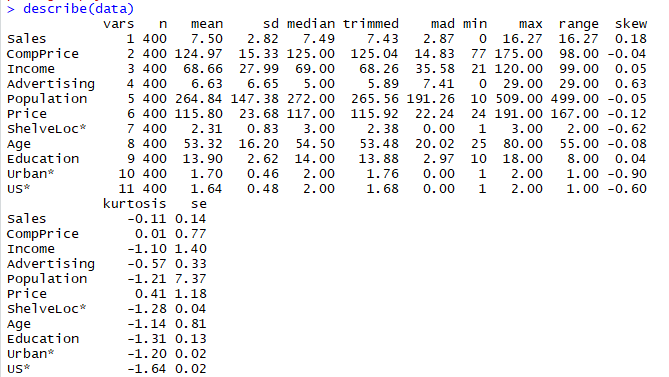
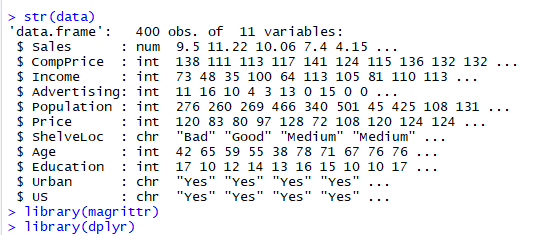
# Decision tree and random forest assignment on Company Data:

* Business Problem: To make a decision tree and random forest model for company data.
* Data Collection: Company data.csv
* EDA:

1. Using the describe() function, we can find that there are no missing values.
2. All the variables are approximately normally distributed, as viewed by the skewness and kurtosis value.





1. Using str() function, we can see that Shelveloc, urban and US variables are character datatype. We will coerce them into factors.
2. Now we will split our dataset into training and test data.

* Data Mining:

1. Using rpart() function, we will make a regression tree model using Sales as dependent variable and other variables as predictors.
2. Then we will predict the test dataset and find the rmse value. The rmse value obtained is 2.24.
3. We will make another model using random forest and find that the rmse value obtained is 1.76.
4. We can conclude that random forest model is predicting better accurate values.