Data Partitioning and Modeling

The data was partitioned into train and test datasets.
The train data set was used to create the decision tree model.
The trained model was then applied to the test dataset.
This is important because when we do data analysis, we should test our model on a data set that was not used to train the model. After a model has been processed by using the training set, you test the model by making predictions against the test set.

When partitioning the data using sampling, it is important to set the random seed to make sure the partition is the same every time you run the program. That is needed when you need a reproducible result.

A screenshot of the resulting decision tree can be seen below: