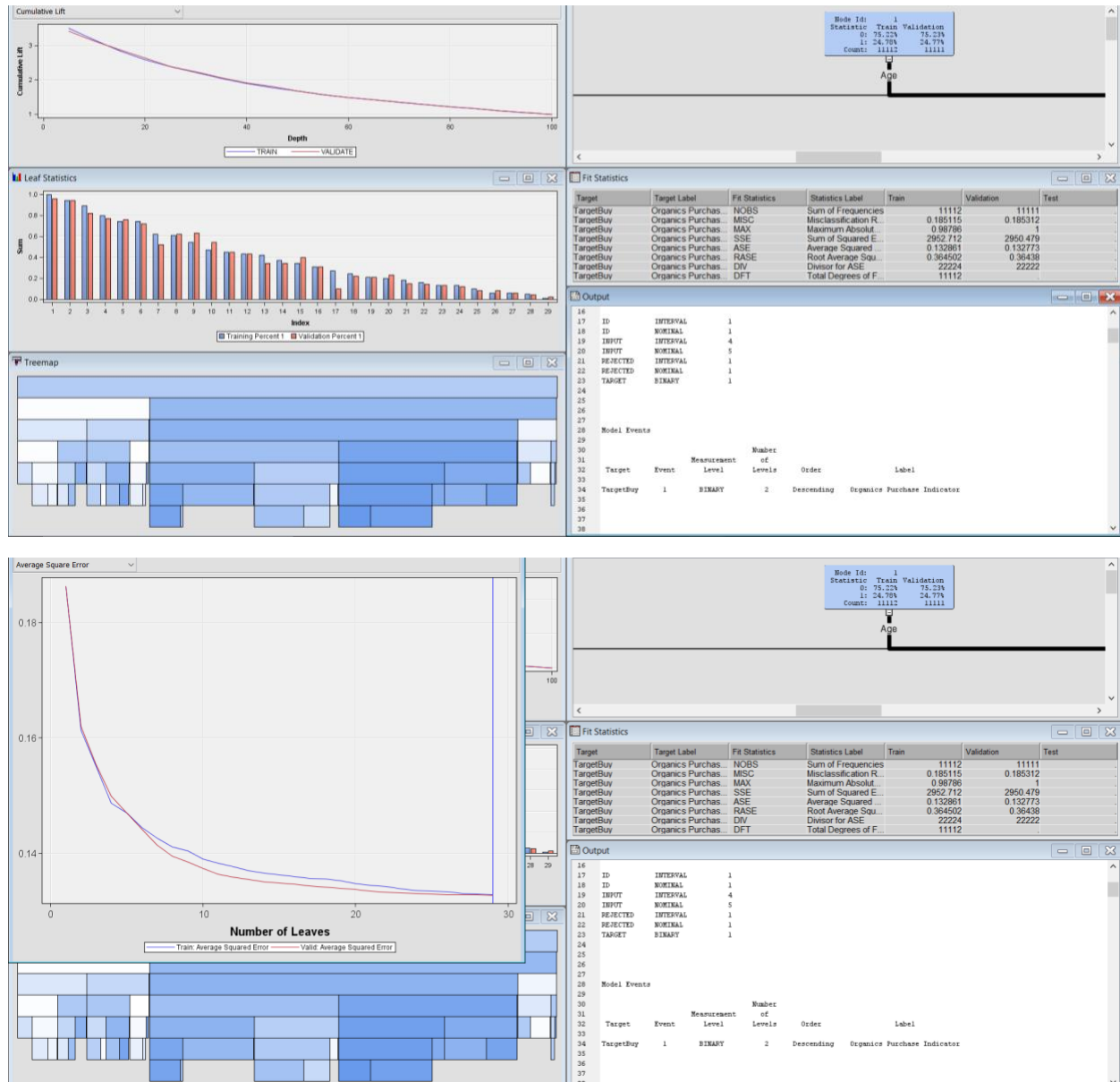


Data Exploration – Decision Trees

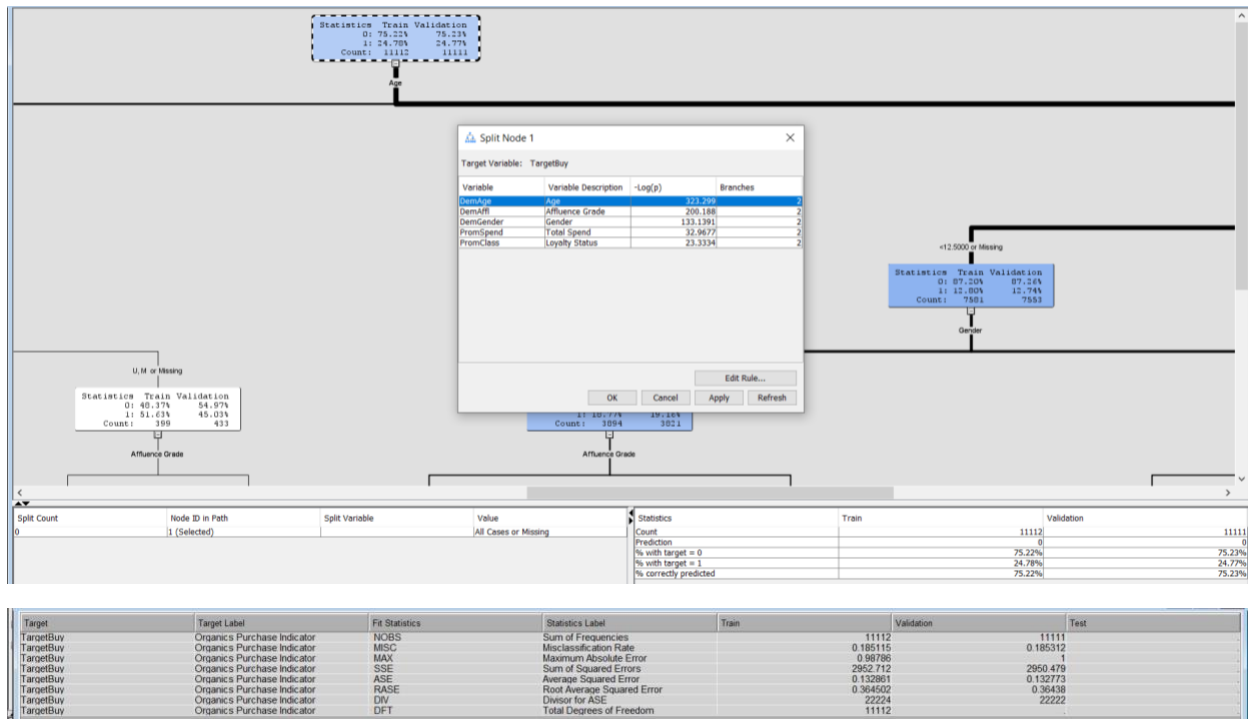
The following are screenshots of the steps that are performed on the dataset :

1. How many leaves are in the optimal tree?

Decision Tree (1)

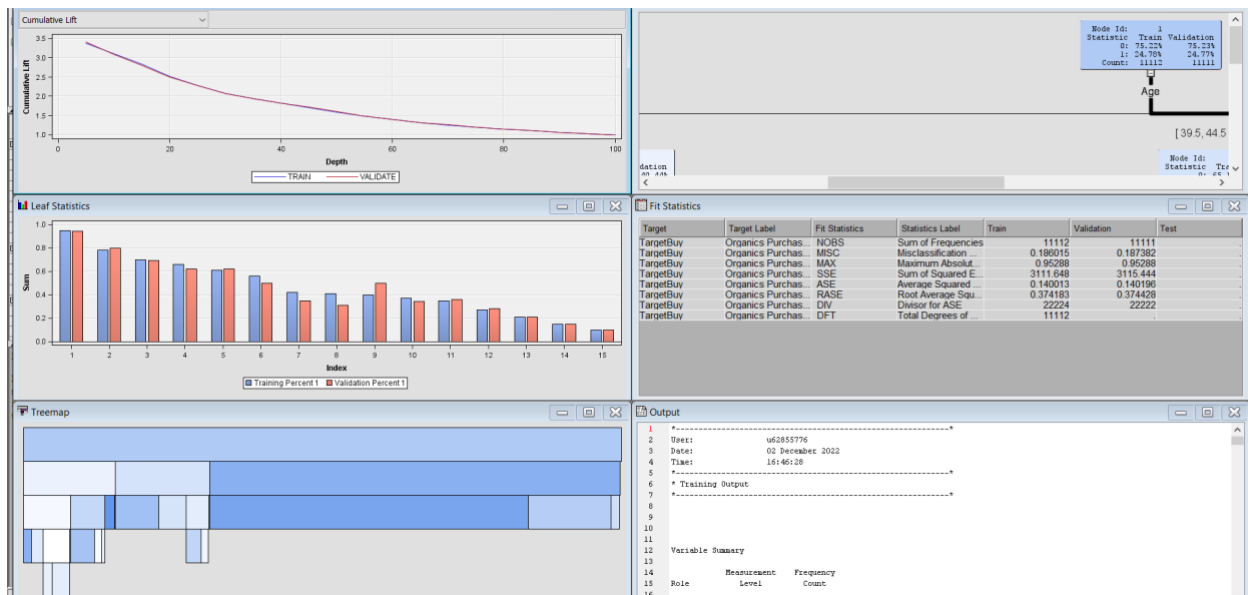


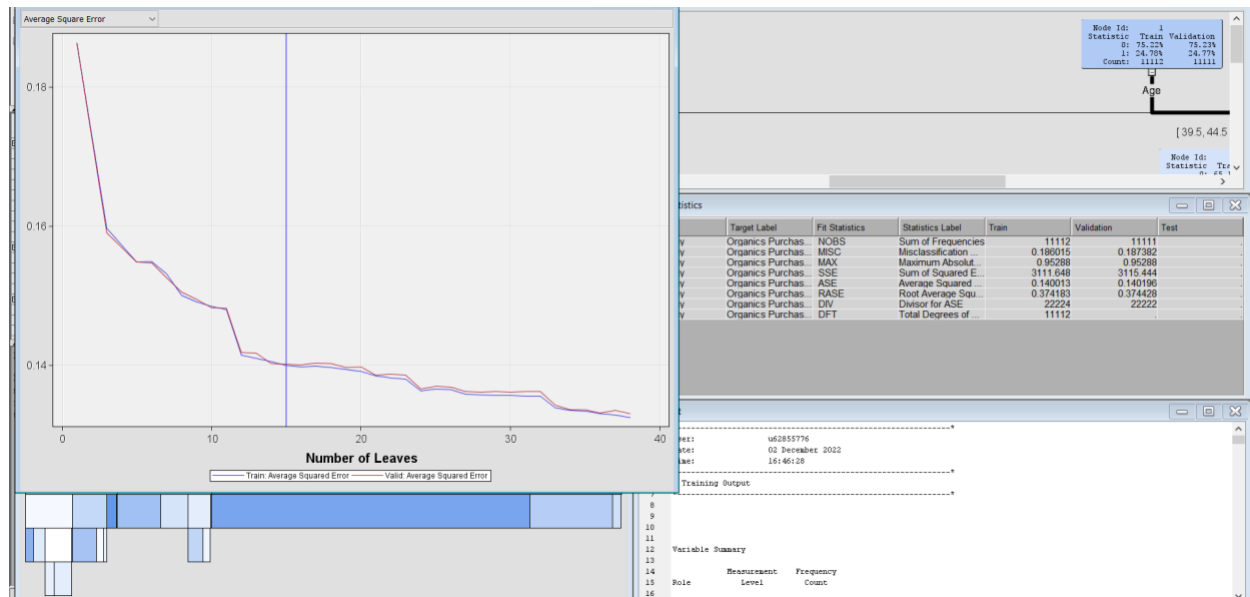
2. Which variable was used for the first split? What were the competing splits for this first split?



3. Add a second Decision Tree node to the diagram and connect it to the Data Partition node.

Desition tree (2)

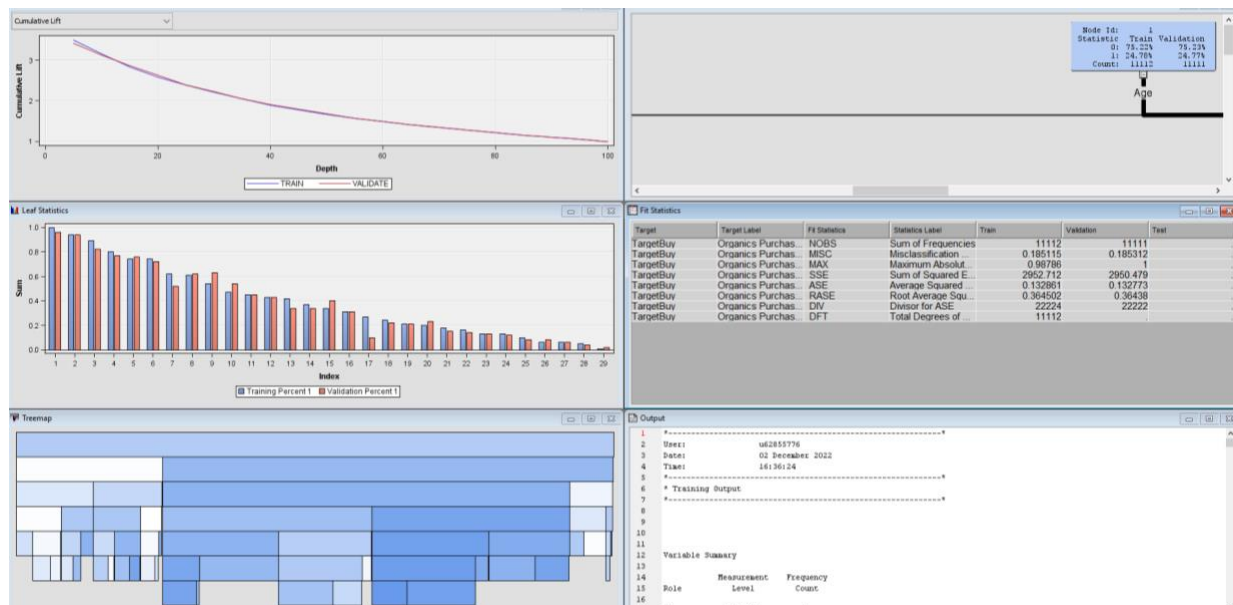




Set the model roles for the analysis variables as shown above.

3) How many leaves are in the optimal tree?

33 leaves



Result – Decision Tree SAS organic Diagram

h. Based on the average square error, which of the decision tree models appears to be better?

Decision tree 2 appears to a better model

