

# Classification Trees (Partition)

Use this data mining technique to predict a categorical (nominal or ordinal) response as a function of potential predictor variables using recursive partitioning.

## Classification Trees

1. From an open JMP® table, select **Analyze > Predictive Modeling > Partition**.
2. Select a nominal or ordinal response variable from **Select Columns** and click **Y, Response**.
3. Select explanatory variables and click **X, Factor**.
4. If desired, enter the **Validation Portion** (a proportion, as shown) or select a validation column and click **Validation (JMP Pro only)**.
5. In **JMP Pro only**, select the tree **Method: Decision Tree** (the default), **Bootstrap Forest**, **Boosted Tree**, **K Nearest Neighbors** or **Naive Bayes**.
6. Click **OK**. JMP displays:

- A graph, with horizontal lines drawn at the proportion of observations in each response level.
- Statistics for the training and validation set(s). Note that results will vary if **Validation Portion** is used.
- A summary of **All Rows**. Click on the **gray triangle** next to **Candidates** to view split statistics.

7. Click the **Split** button. The original observations will be split into two nodes, or leaves (as shown).

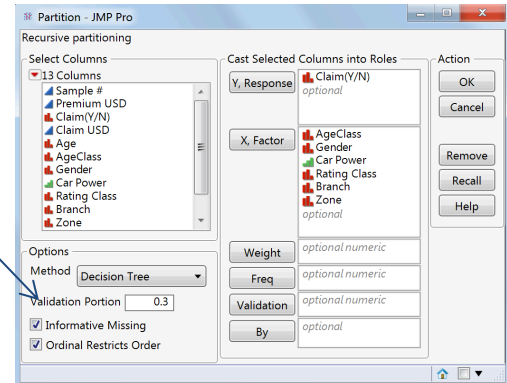
Note: Click on the **top red triangle** and select **Display Options > Show Split Prob** to show **Rates** (proportion of observations) and **Probs** (predicted probabilities) for the response levels in each leaf.

Interpretation (the response, in this example, is Claim (Y/N)):

- There are 1,165 rows in the left leaf, corresponding to AgeClass(Young). The response rate (predicted probability) for Claim(Y/N) = Y is 0.361.
- There are 12,840 rows in the right leaf, corresponding to AgeClass(Elder). The response rate (predicted probability) for Claim(Y/N) = Y is 0.097.

8. Click **Split** to make an additional split. Click **Prune** to remove a split. If a validation portion or validation column is used, click **Go** to perform automatic splitting.

Auto Raw Data.jmp (Help > Sample Data)



Notes:

For additional options, such as **Column Contributions**, **ROC** and **Lift Curves**, click the **top red triangle**. Other options, such as **Save Prediction Formula** and **Make SAS® DATA Step**, are available from the **top red triangle > Save Columns**. For split options for a particular node, click on the **red triangle for that node**.

For more information on fitting and evaluating classification trees, including **Validation**, **Bootstrap Forest** and **Boosted Trees**, see the book *Predictive and Specialized Models* (under **Help > Books**) or search for “partition trees” in the JMP Help.