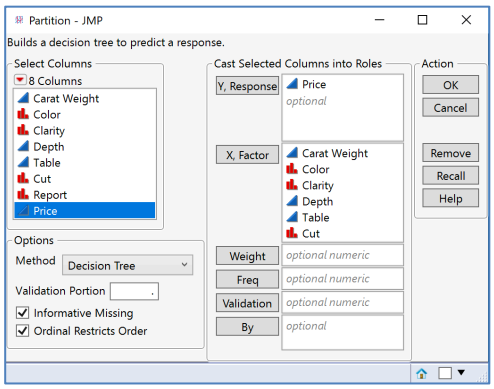
Regression Trees (Partition)

Use this data mining technique to predict a numeric (continuous) response as a function of potential predictor variables using recursive partitioning.

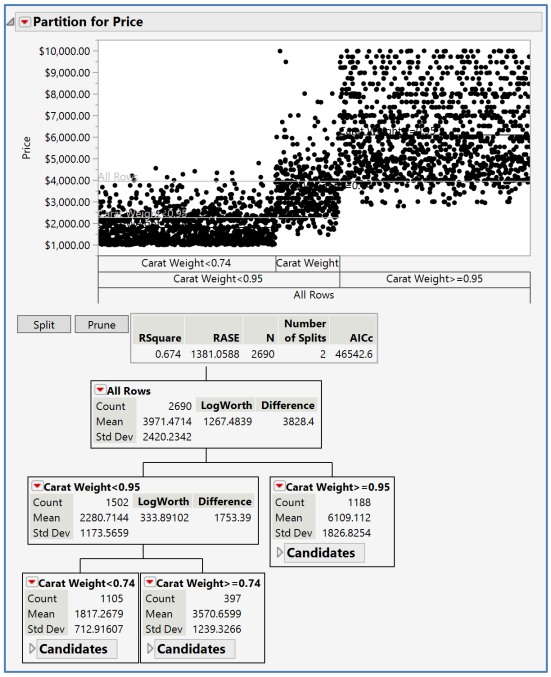
Regression Trees

Diamonds Data.jmp (Help > Sample Data Library)



1. From an open JMP® table, select **Analyze > Predictive Modeling > Partition.**
2. Select a continuous response variable from **Select Columns** and click **Y, Response**.
3. Select explanatory variables and click **X, Factor**.
4. If desired, enter the **Validation Portion** or select a validation column and click **Validation** (**JMP Pro only**). A validation set was not used in this illustration.
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 initially displays a graph of showing the response for all the observations and a line drawn at the overall mean ($3,971 in this example).
7. Click the **Split** button. The original observations will be split into two nodes, or leaves.

   In the top graph, horizontal lines are drawn at the mean response within each leaf and vertical lines depict the leaf's relative size.



> Interpretation for the first two splits (Response is Diamond Price in $):
>
> - There are 1,502 obs with Carat Weight < 0.95. The mean price of these obs is $2,281.
> - There are 1,188 obs with Carat Weight >= 0.95. The mean price of these obs is $6,109.
> - For the 1,502 obs in with Carat Weight <0.95, the second split, also based on the Carat Weight variable, is at a Carat Weight of 0.74.

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 and pruning optimizing the fit on the validation data.

Notes:

For additional options, such as **Leaf Report**, **Small Tree View**, **Column Contributions**, 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 regression trees, including **Validation**, **Bootstrap Forest** and **Boosted Trees**, see the book ***Predictive and Specialized Models*** book (under **Help > Books**) or search for "partition trees" in the JMP Help.