Model Comparison and Selection

JMP PRO Use the **Model Comparison** platform to compare competing statistical models and select the best performing model. For details on fitting predictive models in JMP Pro see the guides and videos at jmp.com/learn under *Data Mining*.

Model Comparison – Continuous Response

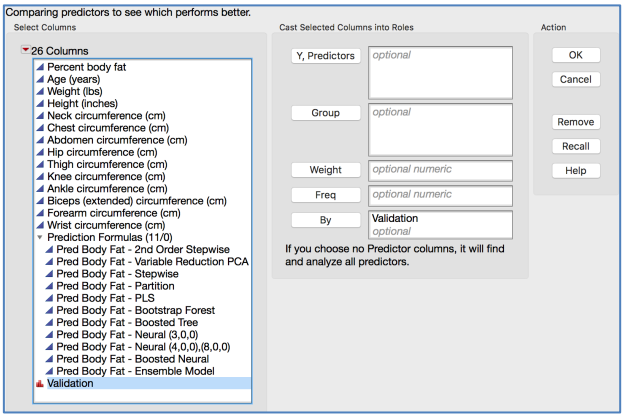
Example: We use the *Body Fat.jmp* data to predict Percent body fat. Formulas for several models, saved to the data table, are grouped under **Prediction Formulas** in the column panel. The models were built using the Validation column, which in this example partitions the data into Training and Validation data.

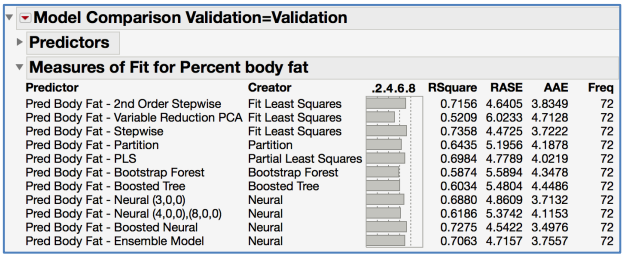
1. Open the dataset, and select **Analyze > Predictive Modeling > Model Comparison**.
2. Select the prediction formula columns from **Select Columns**, and click **Y,Predictors**. If no columns are selected JMP will use all saved prediction formulas.
3. Select **Validation** from the **Select Columns** list, and click **By**. Then, click **OK**.

JMP provides the following statistics for comparing the performance of the different models (statistics for the Validation set are displayed):

- RSquare (higher is better)
- RASE: Root Average Squared Error (like RMSE, but can be compared for competing models)
- AAE: Average Absolute Error (a measure of the magnitude of the errors)

Body Fat.jmp (Help > Sample Data)





Model Comparison Validation=Validation

Predictors

Measures of Fit for Percent body fat

Predictor	Creator	.2.4.6.8	RSquare	RASE	AAE	Freq
Pred Body Fat - 2nd Order Stepwise	Fit Least Squares		0.7156	4.6405	3.8349	72
Pred Body Fat - Variable Reduction PCA	Fit Least Squares		0.5209	6.0233	4.7128	72
Pred Body Fat - Stepwise	Fit Least Squares		0.7358	4.4725	3.7222	72
Pred Body Fat - Partition	Partition		0.6435	5.1956	4.1878	72
Pred Body Fat - PLS	Partial Least Squares		0.6984	4.7789	4.0219	72
Pred Body Fat - Bootstrap Forest	Bootstrap Forest		0.5874	5.5894	4.3478	72
Pred Body Fat - Boosted Tree	Boosted Tree		0.6034	5.4804	4.4486	72
Pred Body Fat - Neural (3,0,0)	Neural		0.6880	4.8609	3.7132	72
Pred Body Fat - Neural (4,0,0),(8,0,0)	Neural		0.6186	5.3742	4.1153	72
Pred Body Fat - Boosted Neural	Neural		0.7275	4.5422	3.4976	72
Pred Body Fat - Ensemble Model	Neural		0.7063	4.7157	3.7557	72

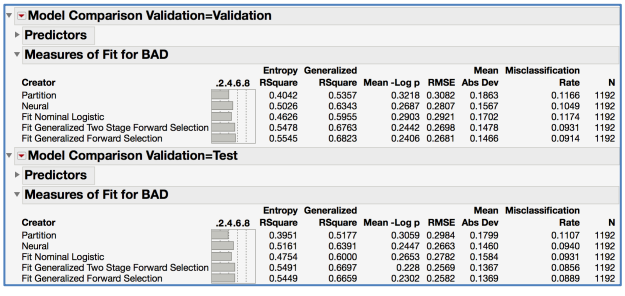
Model Comparison – Categorical Response

Example: In the *Equity.jmp* data, we predict Bad as a function of several predictors. In this example, the data were partitioned into Training, Validation and Test sets.

Create the predictive models of your choice (with Validation), then follow steps 1-3 above to generate model statistics for each value of the Validation column.

- The misclassification rate and other measures of performance are provided.
- ROC curves, lift curves, the confusion matrix and other options are available from the red triangles.

Equity.jmp (Help > Sample Data Library)



Model Comparison Validation=Validation

Predictors

Measures of Fit for BAD

Creator	.2.4.6.8	Entropy RSquare	Generalized RSquare	Mean -Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Partition		0.4042	0.5357	0.3218	0.3082	0.1863	0.1166	1192
Neural		0.5026	0.6343	0.2687	0.2807	0.1567	0.1049	1192
Fit Nominal Logistic		0.4626	0.5955	0.2903	0.2921	0.1702	0.1174	1192
Fit Generalized Two Stage Forward Selection		0.5478	0.6763	0.2442	0.2698	0.1478	0.0931	1192
Fit Generalized Forward Selection		0.5545	0.6823	0.2406	0.2681	0.1466	0.0914	1192

Model Comparison Validation=Test

Predictors

Measures of Fit for BAD

Creator	.2.4.6.8	Entropy RSquare	Generalized RSquare	Mean -Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Partition		0.3951	0.5177	0.3059	0.2984	0.1799	0.1107	1192
Neural		0.5161	0.6391	0.2447	0.2663	0.1460	0.0940	1192
Fit Nominal Logistic		0.4754	0.6000	0.2653	0.2782	0.1584	0.0931	1192
Fit Generalized Two Stage Forward Selection		0.5491	0.6697	0.228	0.2569	0.1367	0.0856	1192
Fit Generalized Forward Selection		0.5449	0.6659	0.2302	0.2582	0.1369	0.0889	1192

Notes: Additional options are available from the red triangles. **Model Comparison** is also available from the **Formula Depot**. For more details on model comparison and the **Formula Depot**, see the book ***Predictive and Specialized Modeling*** (under **Help > Books**) or search for "Model Comparison Platform" in the JMP Help.