

The Power of JMP[®] Pro

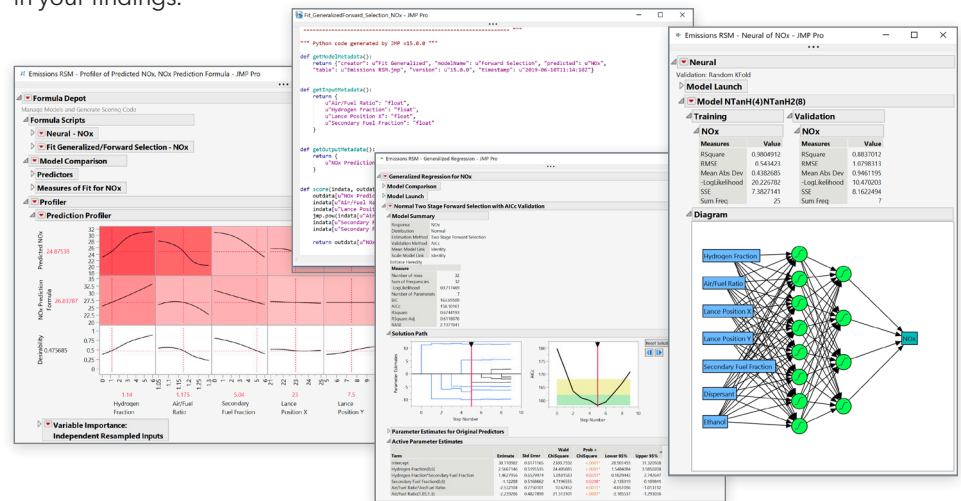
Visual interactive statistics joined with the most sophisticated modeling techniques



JMP Pro offers all the capabilities of JMP plus features for next-level analysis, including predictive modeling and cross-validation techniques, all in an easy-to-use interface.

Scientists, engineers and data explorers in a variety of industries find everything they need in JMP. It is, after all, the most visual and interactive way to discover what is hidden in your data.

However, we know that sometimes you need to put especially sophisticated statistical techniques behind your data exploration. JMP Pro is for people who are problem solving to inform major decisions. It is how you will multiply your impact and be more confident in your findings.



JMP Pro predictive analytics software from SAS lets you harness the power of the supercomputer on your desktop for exploring and understanding data with an easy-to-use interface.

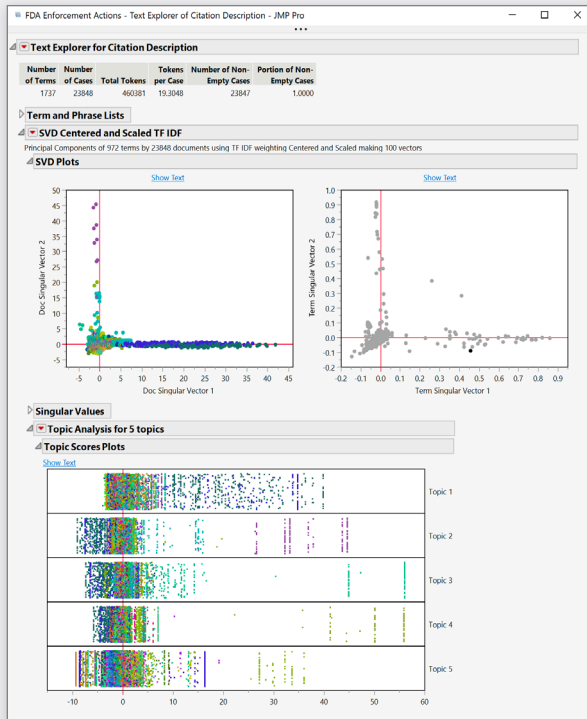
JMP Pro is a collection of the most relevant statistical instruments for analyzing data without being overwhelmed by choices and without needing to program.

Who uses the power of JMP Pro?

- Biologists at a major global pharmaceutical company reduced processing time for biomarker analysis from 20 hours to 30 minutes using a customized application developed with JMP Pro.
- Research scientists at one of the world's largest consumer products companies use design of experiments capabilities in JMP Pro to improve product quality while cutting costs across its R&D operations.
- Statisticians at a leading chemical company use JMP Pro for its advanced capabilities with, for example, the Partial Least Squares platform. The group now uses the Custom Designer and Prediction Profiler to provide support to R&D and manufacturing teams on a daily basis.
- Manufacturing engineers at an international semiconductor corporation use the regularization techniques in JMP Pro to deal with their highly dimensional, highly correlated data.
- Researchers at an international animal nutrition technology company use advanced predictive modeling tools in JMP Pro to combine data from tests conducted around the world into a single database, validate it and use decision trees to analyze the data in confidence.

And so can you.





Let JMP Pro process your unstructured text data, reduce dimensionality and generate data that can be easily consumed by the other predictive modeling tools. This helps you understand and model the latent information present in text data such as engineering reports, surveys or other free text documents.

JMP Pro has the tools to handle the complexity of various types of data.

• Text Exploration

If you have unstructured text data, Text Explorer organizes and shapes that data into a usable form for your predictive models, and the latent information you've collected in the form of repair logs, engineering notes or customer survey response comments is now available to boost the power of your models.

Text Explorer in JMP Pro turns the difficulty of dealing with text data into a standard data analysis problem. It handles latent class analysis, latent semantic analysis, topic analysis, and cluster terms and documents. With the enhanced capabilities of the Text Explorer platform, text data analysis becomes attainable, like any other data exploration.

• Functional Data Explorer

Data collected from batch processes and sensor streams require specialized data cleaning and modeling techniques. Functional Data Explorer in JMP Pro simplifies the complex problems of managing and cleaning up this data. Furthermore, with Functional Data Explorer, perform the challenging tasks of handling messy data, removing outliers, aligning curves, building surrogate models and creating features to better understand data coming from your streams or process measures (for use as inputs, outputs or both) from sensor streams. Streamlined workflows simplify the process of creating, reshaping and joining tables, especially for functional experiment design (functional DOE).

Key features

Predictive modeling

Anyone can do a fair job of describing last year's performance. But without the right tools and the most modern techniques, building a model to predict what will happen with new customers, new processes or new risks becomes more difficult. Some of the most useful techniques for predictive modeling are available in JMP Pro:

- The Partition platform in JMP Pro automates the decision tree-building process.
- The Bootstrap Forest platform, which uses a random-forest technique, grows dozens of decision trees using random subsets of the available data and averages the computed influence of each factor in these trees.
- The Boosted Tree platform builds many simple trees, repeatedly fitting any residual variation from one tree to the next.
- The advanced Neural platform lets you build one- or two-layer neural networks with your choice of three activation functions and automatic model construction using gradient boosting. The platform automatically handles missing values and transformation of continuous X's, which saves time and

effort, includes robust fitting options and guards against overfitting without having to mentally tune parameters.

- The Generalized Regression platform lets you perform variable selection, generalize a response to a number of distribution and performs regularization using penalized methods, forward selection as well as has a rich set of diagnostics; it's an all-in-one linear modeling platform in JMP Pro.
- K-Nearest Neighbors and Naive Bayes models.
- Support Vector Machine (SVM) classifiers.

Each of these platforms in JMP Pro uses cross-validation described in the next section. In addition, stepwise regression, logistic regression (both nominal and ordinal), PLS and discriminant analysis in JMP Pro can take advantage of using a validation column allowing you to use these modeling methods in your predictive modeling workflow.

Using diverse types of data

It's great when your data is straightforward, with columns of numbers, dates or categories. But data, in its many forms and from its many sources, isn't always so clear. When your data gets complicated,

"Data is worthless if we don't have the right tools to work with it. So we need to be using the newest and most innovative means to manage this data. JMP Pro uses the most innovative and effective methods, and that's very important."

Corinne Bergès, PhD
Advanced Automotive Analog
Department, NXP

Partial least squares

Are you trying to model data that is wider than it is tall? Traditional techniques won't work, but partial least squares (PLS) does. PLS is a powerful modeling technique to have in your toolbox, especially when you have more X variables than observations, highly correlated X variables, a large number of X variables, or several Y variables and many X variables. All of these are situations where ordinary least squares would produce unsatisfactory results.

PLS modeling fits linear models based on factors, namely linear combinations of explanatory variables (the X's). The factors are obtained in a way that attempts to maximize the covariance between the X's and the response or responses (the Y's). In JMP Pro, you can build PLS models with either continuous or categorical responses (PLS-DA), specify curvature terms or interaction effects and perform missing value imputation.

Cross-validation

For effective predictive modeling, you need sound ways to validate your model, and with a large model, you can easily get into trouble over-fitting. Large models should always be

cross-validated, and JMP Pro does this through data partitioning, or holdback. The cross-validation technique helps you build models that generalize well to tomorrow's data – about new customers, new processes or new risks – so you can make data-driven inferences about the future.

Dividing the data into training, validation and test data sets has long been used to avoid over-fitting, ensuring that the models you build are not reliant on the properties of the specific sample used to build them. The general approach to cross-validation in JMP Pro is to use a validation column. You can easily split your data into different sets for different purposes using the validation column utility (either with a purely random sample stratified random or using a time cut-point).

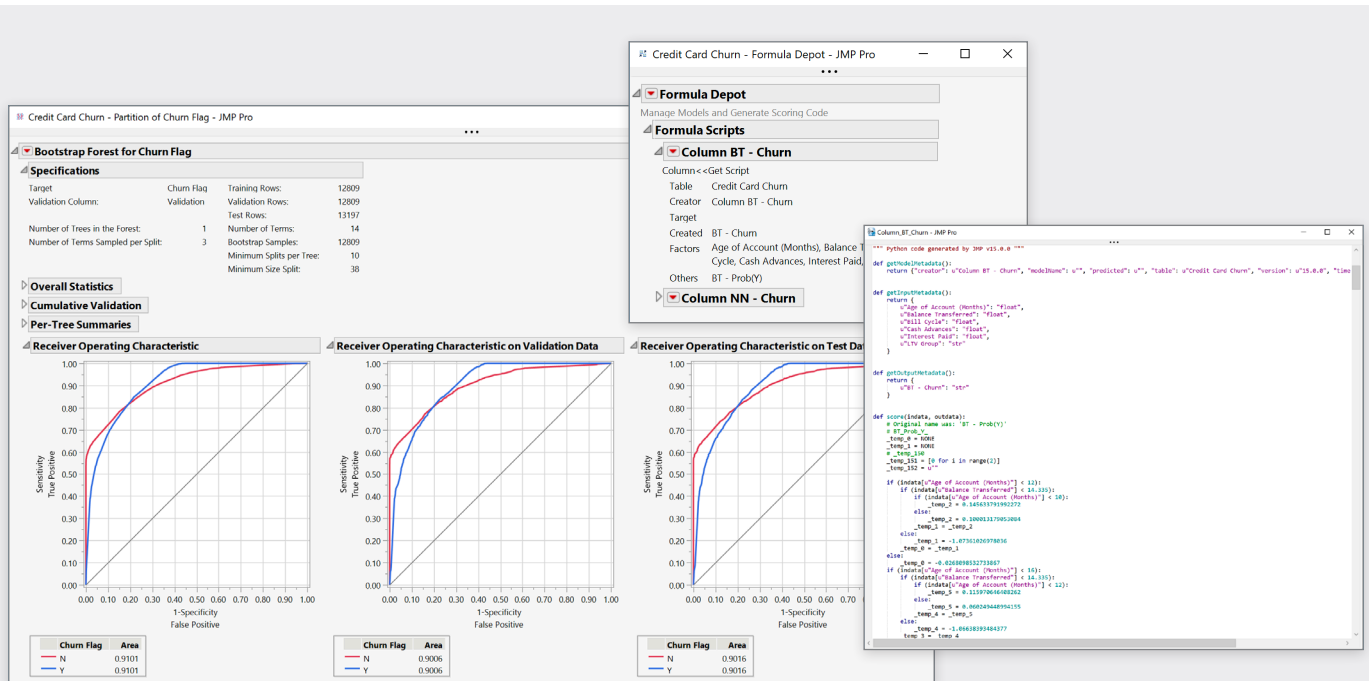
The training set is used to build the model(s); the validation set is used in the model-building process to help choose how complex the model should be. Finally, the test set is held out completely from the model-building process and used to assess the quality of the model(s). For smaller data sets, k-fold cross-validation also can be used. This process helps you build models that generalize to new data effectively.

It is important to consider that observational data can only take you so far. To truly understand cause and effect, many times you may wish to employ design of experiments (DOE).

JMP provides world-class tools for optimal DOE in a form you can easily use.

Structural Equation Modeling

When you need to examine relationships between measured variables and latent constructs, Structural Equation Modeling in JMP Pro is a sound, but flexible, modeling framework that allows specification of a variety of statistical models. You can explore competing theories about latent variables and fit any general linear model, including confirmatory factor analysis, path analysis and latent growth curve models. Additionally, you can use a graphical model builder for creating interrelated models, specify a model, duplicate it and switch among several models' results, or use a model specification repeatedly with different sets of data. The flexibility of Structural Equation Modeling lets you fit models where variables can be both inputs and outputs at the same time.



Formula Depot is the hub for organizing, comparing, profiling, storing and deploying models built in JMP Pro. You can convert models built in JMP Pro in a single click to score code in SAS, C, Python, JavaScript and/or SQL.

Model comparison

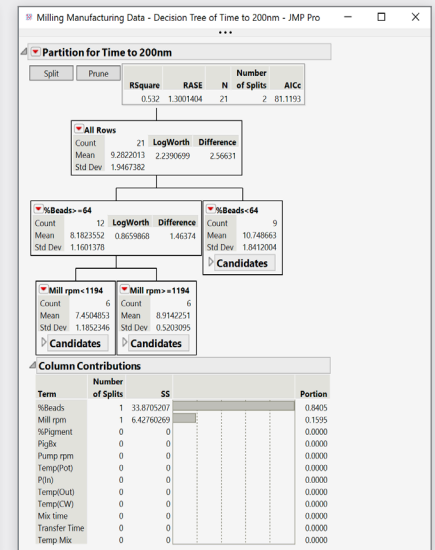
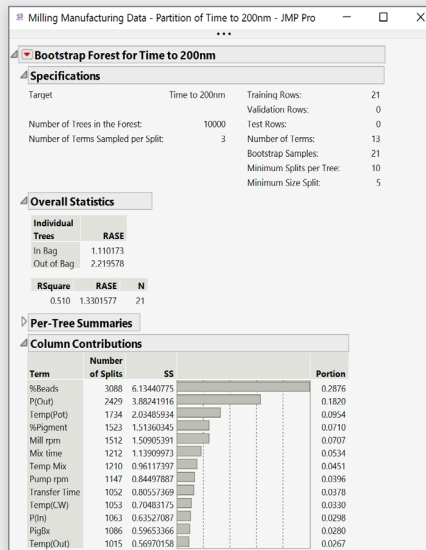
In the real world, some kinds of models fit well in certain situations but poorly in others. With JMP Pro, there are many ways to fit, and you need to find out which is most appropriate in a given situation. A typical approach to model building is that you will try many different models: models with more or less complexity, models with or without certain factors/predictors, models built using different kinds of methods or even averages of multiple models (ensemble models).

Each of these models will have common quality measures that can be used to assess the model: R2, misclassification rate, ROC curves, AUC, lift curves, etc.

Using model comparison in JMP Pro, you can compare all the saved prediction columns from various fits and pick the best combination of goodness of fit, parsimony and cross-validation. JMP Pro makes this comparison automatically. At the same time, you can interact with visual model profilers to see which important factors each model is picking up. Model comparison in JMP Pro makes it easy to compare multiple models at the same time, and also to do simple model averaging, if desired.

Formula Depot and score code

Managing your models doesn't have to be painful - the Formula Depot in JMP Pro organizes your work when dealing with many models. This central repository lets you store, profile, compare and selectively generate code to aid in deployment of JMP Pro models to other languages such as SAS®, Python, SQL and JavaScript. And even if you don't choose to deploy models outside of JMP, your data tables are no longer weighed down with numerous extra columns of prediction formulas needed to perform model comparison. The score code can be saved to the Formula Depot and applied to new data. The result is a central modeling hub for easy access to your models and simple deployment to other systems.



The bootstrap forest will identify all the potential variables that could be affecting your responses for further investigation, ideally with design of experiments. A standard decision tree model will not always do this, as shown here.

Connect to the richness of SAS®

As one of the SAS offerings for predictive analytics and data mining, JMP Pro easily connects to SAS including SAS® Viya®, expanding options and giving access to the unparalleled depth of SAS Analytics and data integration.

Modern modeling

The Generalized Regression personality in Fit Model is an all-inclusive approach to doing regression.

With Generalized Regression, you can build models for many kinds of data, including DOE data, observational data, data with categorical responses, messy data, text regression, highly correlated data and more.

In one location, with one tool, you can fit models, select variables, handle multicollinearity and investigate diagnostic reports. And it's only in JMP Pro.

Standard estimation techniques break down when you have predictors that are strongly correlated or more predictors than observations. And when there are many correlated predictors (as is often the case in observational data), stepwise regression

or other standard techniques can yield unsatisfactory results. Such models are often overfit and generalize poorly to new data. But how do you decide which variables to cull before modeling - or, worse, how much time do you lose manually preprocessing data sets in preparation for modeling?

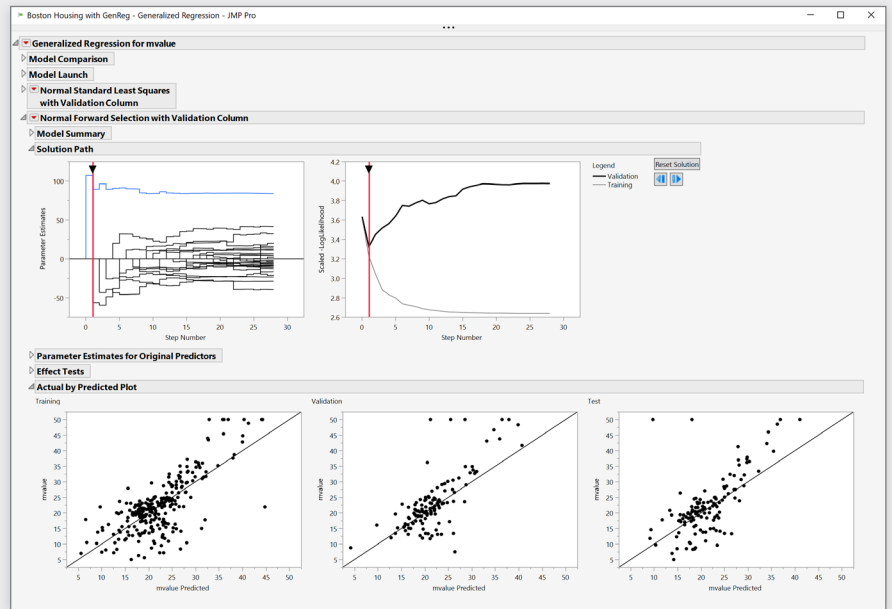
The Generalized Regression personality in Fit Model is an all-inclusive approach to doing regression. It's a complete modeling framework from variable selection through model diagnostics to LS means comparisons, inverse prediction and profiling. And it's only in JMP Pro.

The estimation methods in Generalized Regression include basic ones such as standard least squares, logistic regression and maximum likelihood (for a multinomial response) as well as selection techniques like best subset and pruned forward selection and regularization techniques such as Lasso, Double Lasso and the Dantzig Selector. Harnessing these techniques is as easy as any other modeling personality in Fit Model - simply identify your response, construct model effects and pick the desired estimation and validation method. JMP automatically fits your data, performs

“JMP continues to amaze me. There are always new modeling techniques and tools to make my life as a data analyst easier. The Generalized Regression platform is easy to use and fast. And the ability to create test/validation variables on the fly is great. Check it out!”

Dick De Veaux

C. Carlisle and Margaret Tippit
Professor of Statistics,
Williams College



Generalized Regression is a one-stop shop for interactive variable selection, modeling and generalizing data to diverse distributions, even with challenging data problems.

variable selection when appropriate, and builds a predictive model that can be generalized to new data.

Finally, Generalized Regression gives options to choose the appropriate distribution for the response you are modeling, letting you model more diverse responses such as counts, data with many outliers, or skewed data. And like all the advanced modeling platforms in JMP Pro, you have your choice of cross-validation techniques.

Reliability

Reliability Block Diagram

Often you are faced with analyzing the reliability of a more complex analytical system – a RAID storage array with multiple hard drives, or an airplane with four engines, for example. With JMP, you have many tools to analyze the reliability of single components within those systems. But with JMP Pro, you can take the reliability of single components, build a complex system of multiple components and analyze the reliability of the entire system. Using the Reliability Block Diagram, you can easily design and fix weak spots in your system – and be better informed to prevent future system failures.

With this platform, you can perform what-if analyses by looking at different designs and comparing plots across multiple system designs. You can also determine the best places to add redundancy and decrease the probability of a system failure.

Repairable Systems Simulation

Some systems or components of complex systems are too costly to have offline for very long. Maintaining the integrity of these systems requires you to schedule repairs for system components or maximize the benefit realized by an unplanned outage by completing additional repairs while the system is unavailable. With JMP Pro, you can use the Repairable Systems Simulation to determine how long a system will be unavailable and answer key questions of how many repairable events to expect in a given period of time and how much a repair event will cost.

Covering arrays

Covering arrays are used in testing applications where factor interactions may lead to failures where each experimental run may be costly. As a result, you need to design an experiment to maximize the probability of finding defects while also

minimizing cost and time. Covering arrays let you do just that. JMP Pro lets you design an experiment to test deterministic systems and cover all possible combinations of factors up to a certain order of interactions.

And when there are combinations of factors that create implausible conditions, you can use the interactive Disallowed Combinations filter to automatically exclude these combinations of factor settings from the design.

One of the advantages of covering arrays in JMP Pro is that JMP Pro is a statistical analysis tool, not just a covering arrays design tool. For example, you can design covering arrays and analyze the data using generalized regression. This is an advantage of JMP Pro over other tools that only design covering arrays.

Finally, JMP Pro allows you to import any covering array design – generated by any software – and further optimize it and analyze the results. You can design the arrays yourself without having to rely on others to build experiments for you. Test smarter with covering arrays in JMP Pro.

Mixed models

Mixed models contain both fixed effects and random effects in the analysis. These models let you analyze data that involve both time and space. For example, you might use mixed models in a study design where multiple subjects are measured at multiple times during the course of a drug trial, or in crossover designs in the pharmaceutical, manufacturing or chemical industries.

JMP Pro lets you fit mixed models to your data, letting you specify fixed, random and repeated effects; correlate groups of variables; and set up subject and continuous effects - all with an intuitive drag-and-drop interface.

In addition, you can calculate the covariance parameters for a wide variety of correlation structures. Such examples include when the experimental units on which the data is measured can be grouped into clusters, and the data from a common cluster is correlated. Another example is when repeated measurements are taken on the same experimental unit, and these repeated measurements are correlated or exhibit variability that changes.

It is also easy to visually determine which, if any, spatial covariance structure is appropriate to utilize in your model specification when building mixed models in JMP Pro.

Uplift models

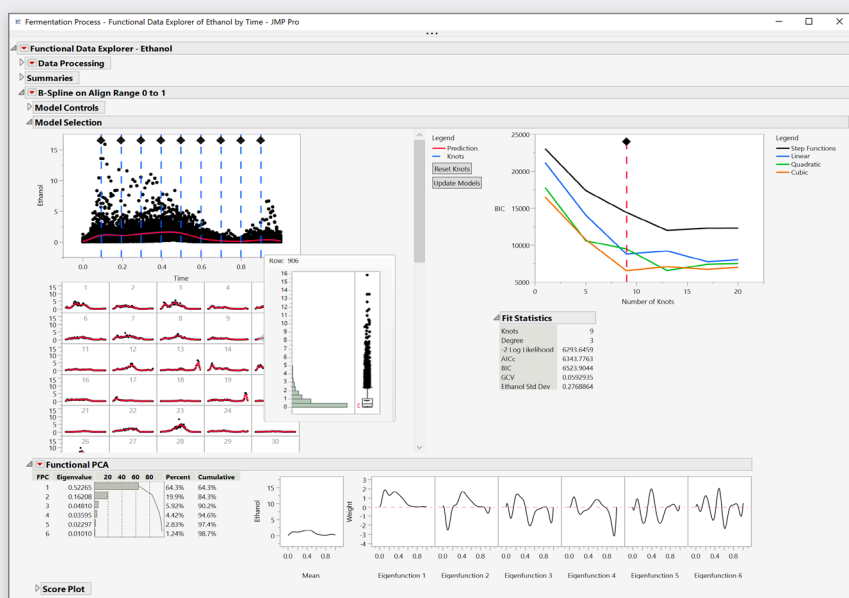
You may want to maximize the impact of a limited marketing budget by sending offers only to individuals who are likely to respond favorably. But that task may seem daunting, especially when there are large data sets and many possible behavioral or demographic predictors. This is where uplift models can help. Also known as incremental modeling, true lift modeling or net modeling, uplift models have been developed to help optimize marketing decisions, define personalized medicine protocols or, more generally, to identify characteristics of individuals who are likely to respond to some action.

Uplift modeling in JMP Pro lets you make these predictions. JMP Pro fits partition models that find splits to maximize a treatment difference. The models help identify groups of individuals who are most likely to respond favorably to an action; they help to lead to efficient and targeted decisions that optimize resource allocation and impact on the individual.

Advanced computational statistics

JMP Pro includes exact statistical tests for contingency tables and exact non-parametric statistical tests for one-way ANOVA. Additionally, JMP Pro includes general resampling methods that can be accessed through a right-click on most MP reports.

- **Bootstrap:** Provides one-click bootstrapping, a resampling method that approximates the sampling distribution of a statistic and is useful when textbook assumptions are in question or don't exist. For example, try applying bootstrapping techniques to nonlinear model results that are being used to make predictions or determining coverage intervals around quantiles. Also, you can use bootstrapping as an alternative way to gauge the uncertainty in predictive models.
- **Bagging:** Conducts model averaging using bootstrap samples from the training set in a cross-validated model. This technique is also known as bootstrap aggregating.
- **Simulate:** Provides simulated results for a column of statistics in a report and lets you answer challenging questions with parametric resampling. Use simulate to obtain power calculations in nonstandard situations, evaluate new or existing statistical methods or conduct permutation tests.



Functional Data Explorer is useful for understanding, cleaning up, aligning and building models from sensor streams or batch process data.

Share and communicate results

JMP has always been about discovery and finding the best way of communicating those discoveries across your organization. JMP Pro includes all the visual and interactive features of JMP, making your data accessible in ways you might never have experienced. Through dynamically linked data, graphics and statistics and easy to create dashboards, JMP Pro brings your investigation alive in an interactive profiler, geographic map or filtered report to show change over time, generating valuable new insights that inform both the model-building and explanation process.

Operating system guidelines

JMP Pro runs on Microsoft Windows and Mac OS. It includes support for 64-bit systems. See jmp.com/system for complete system requirements.

Key features exclusive to JMP® Pro

JMP Pro includes all the features of JMP, plus the additional capabilities for analytics listed below.

Platforms

Association Analysis

- Performs market basket analysis to identify items that frequently occur together.

Bootstrap Forest

- Constructs a predictive model by averaging predicted values from many decision trees constructed using randomly selected predictors and observations.

Boosted Tree

- Constructs a predictive model by adding a sequence of decision trees where each of the trees is fit on the residuals of the previous tree.

Covering Arrays

- Creates designs for testing deterministic systems for failures caused by interactions among components.
- Constructs highly efficient covering arrays.
- Constructs covering arrays that take into account disallowed (infeasible) combinations of factor levels.
- Supports importing of covering arrays created by other software.
- Calculates coverage metrics for evaluating covering arrays.
- Provides ability to further optimize a covering array once it is created.

Formula Depot

- A container for prediction models that is launched through the Publish commands in modeling platforms.
- Generate Score Code: SAS(DS2), C, Python, Javascript, SQL (with choice of syntax options for different destinations).

Functional Data Explorer

- Useful for dealing with sensor/signal/streaming or batch process data.
- Cleanup, align and conform data coming from sensors.
- Build surrogate models using a variety of techniques.
- Use output models and features in other predictive modeling platforms.
- Works with Custom Designer and Generalized Regression for functional DOE.

Generalized Regression

- General model fitting tool that supports variable selection.
- Fits penalized generalized linear models, including the lasso and elastic net.
- Provides forward selection and maximum likelihood.
- Includes many different cross-validation methods.
- Provides variable selection for censored data.
- Fits quantile regression models.

K-Nearest Neighbors

- Predicts a response based on the responses of the k-nearest neighbors in the space of the X's.

Mixed Model

- Fits linear models for continuous responses and complex covariance structures.
- Fits random coefficient (hierarchical) models.
- Fits repeated measures models.
- Fits spatial data models.
- Displays variograms.
- Fits models for multiple correlated responses.
- Conducts multiple comparisons.

Model Comparison

- Compares the predictive ability of candidate models.
- Includes fit statistics (R², misclassification rate, ROC curves, AUC, lift curves).
- Model averaging.

Naive Bayes

- Predicts group membership for a categorical variable based on the closeness of its predictor values to the predictor values for each group.

Reliability Block Diagram

- Graphically displays the relationships among the components of a system.
- Provides a drag-and-drop interface for viewing and editing the components of a reliability system.
- Provides profilers that enable you to explore the overall reliability of the system.

Repairable Systems Simulation

- Provides a graphical representation for system maintenance scheduling based on a reliability block diagram.
- Uses simulation to study and understand complex system behavior and to schedule desired maintenance.

SEM

- Graphical model builder that is useful for creating interrelated models.
- Can be used to build models where variables cannot be measured directly and where measurement error should be considered.
- Can be used to build a model specification before data is available.

SVM

- Classifier, similar to Neural Net, that provides additional options when building models.

Uplift

- Models incremental impact of an action on individuals.
- Identifies individuals who are likely to react positively to an action.
- Finds splits to maximize a treatment difference.

(continued to next page)

JMP® Pro Specific Platform Features

Choice

- Bayesian Random Effects (Hierarchical Bayes).
- Save subject estimates for Bayesian Random Effects.

Fit Y by X

- Contingency: Provides exact tests for association and trend.
- Oneway: Provides nonparametric exact tests.

Gaussian Process

- Analyze models with categorical factors.
- Ability to fit models with thousands of rows through FastGASP.

Neural

- Include rows with missing values in the model, instead of ignoring the rows.
- Enables you to fit a multi-layer perceptron with three different types of activation functions.
- Fits an additive sequence of models, also called boosting.
- Provides additional fitting options to transform covariates, minimize the impact of response outliers, and apply penalties to help avoid overfitting.

Partial Least Squares

- Conducts PLS Discriminant Analysis (PLS-DA).
- Enables use of Fit Model launch window path.
- Fits interaction and polynomial terms.
- Fits categorical effects.
- Includes K-Fold and Holdback validation methods.
- Imputes missing data with your choice of two methods.

Resampling Methods

- Bootstrap: Provides one-click bootstrapping, a resampling method that approximates the sampling distribution of a statistic.
- Bagging: Conducts model averaging using bootstrap samples from the training set. Also known as bootstrap aggregating.
- Simulate: Provides simulated results for a column of statistics in a report.

Text Explorer

- Fits latent class analysis models.
- Performs a singular value decomposition on the document-term matrix, also known as latent semantic analysis.
- Performs a varimax rotated singular value decomposition to help you find topics in your text data.
- Shows clusters of terms and clusters of documents.
- Provides scatterplot matrices of singular values.
- Support for discriminant analysis.

Validation Column

Platforms that support a validation column:

- Bootstrap Forest
- Boosted Tree
- Discriminant
- Explore Missing Values
- Fit Model: Standard Least Squares
- Fit Model: Stepwise
- Fit Model: Generalized Regression
- Fit Model: Nominal Logistic
- Fit Model: Ordinal Logistic
- Fit Model: Parametric Survival
- Fit Model: Partial Least Squares
- K-Nearest Neighbors
- Naive Bayes
- Neural
- Partial Least Squares
- Partition
- Text Explorer
- Uplift

Are you ready to take the next step?

Contact JMP Sales to ask questions, learn more and schedule a demo.

To contact your local JMP office, please visit: jmp.com/offices

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies. Copyright © 2019, SAS Institute Inc. All rights reserved. 105522_G99726.0819

