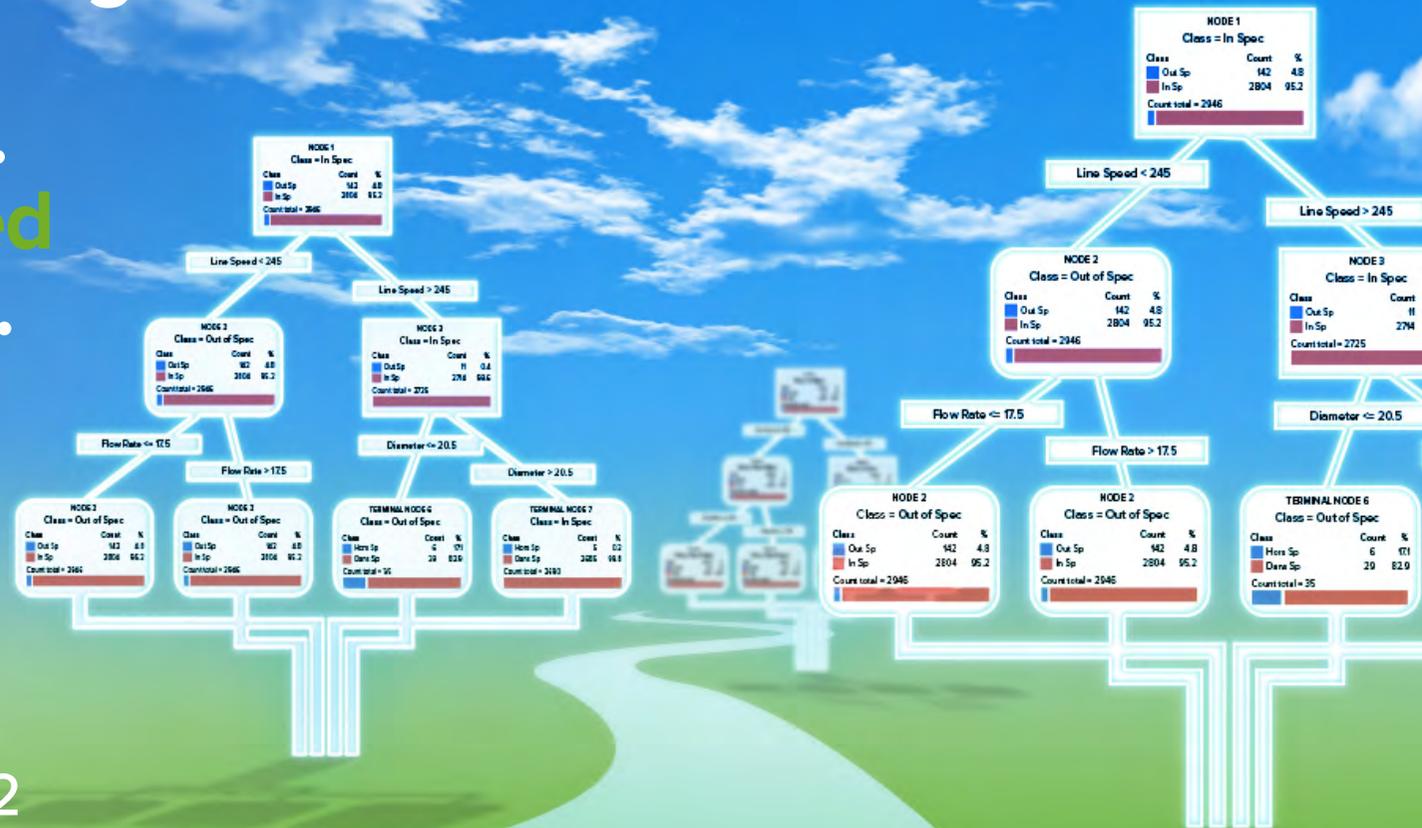


Better decision making.
Faster insights.
Easier on the cloud.
 Your path to **boosted** predictive analytics.



What's New in Minitab 20.2

Contents

Visualizations

02

New: Correlogram

Update: By Variables for Binned Scatterplot,
Parallel Coordinates Plot, and Heat Map

Improvement

04

Lag

Additions

05

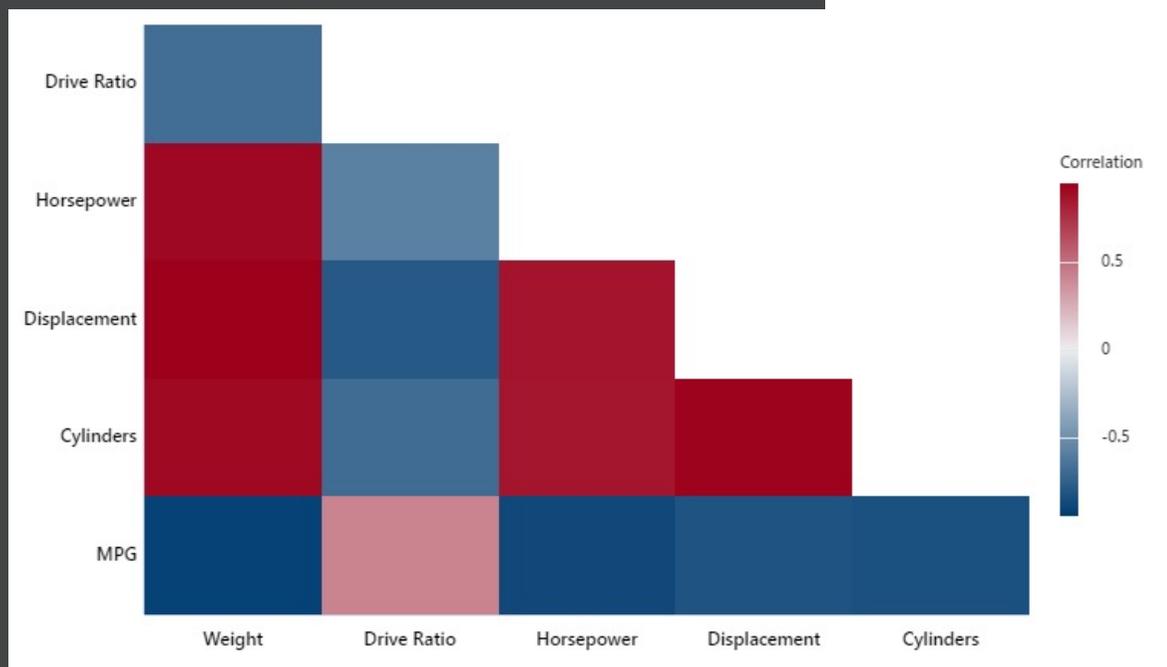
R Integration

Predictive Analytics Module

Random Forests

TreeNet





Visualization

New: Correlogram

Improvement: Correlograms added to the graph menu. A correlogram is a visualization that displays correlation coefficients with variables names defining the rows and columns. Optionally, users can display the correlation statistics.

User Benefit: Correlograms are useful for finding important correlations when faced with many variables. Viewing correlations as a color gradient is an alternative approach to displaying a matrix plot or a table of correlation statistics.

Summary: The correlogram makes it easy to visualize a matrix or correlations, particularly when the number of variables is large.

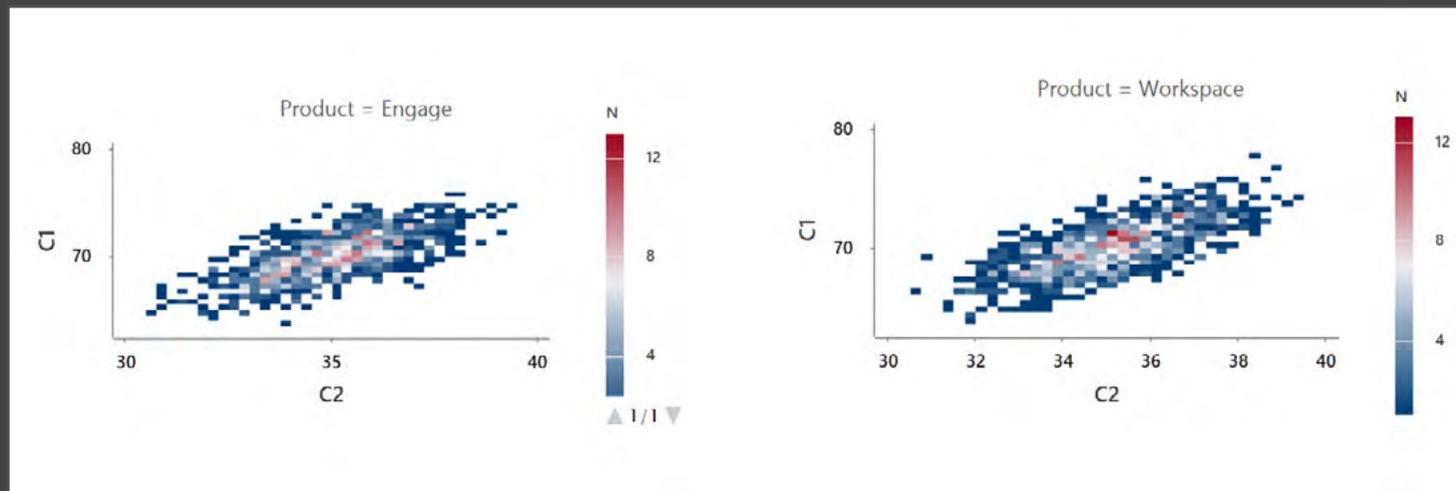
Visualization

Updated: Variables for Binned Scatterplot, Parallel Coordinates Plot, and Heat Map

Improvement: The binned scatterplot, parallel coordinates plot and the heat map can accommodate by variables.

User Benefit: Graphs can be split by variables faster, removing the need to first subset the data before creating these graphs.

Summary: In each of these dialogs, grouping variables can be designated using the By Variable section.



Improvement: Lag

Improvement: Previous versions of Minitab Statistical Software performed lags for 1 variable at a time. Users can now create lags for up to 12 variables (series) at a time. The dialog can also accommodate multiple lags (e.g., 1, 2, 3, 4 and 5), as shown.

User Benefit: Lagged columns are commonly used in time series modeling and supervised machine learning, such as CART, TreeNet, and Random Forests. This improvement allows for faster data preparation.

Summary: Numerous lag columns can be easily generated for one or multiple time series columns. This command is located in Stat-Time Series-Lag.

Specify root series and lags:

Root Series	Lags
Measures	1:5



C1	C2	C3	C4	C5	C6
Measures	Measures_Lag1	Measures_Lag2	Measures_Lag3	Measures_Lag4	Measures_Lag5
20	*	*	*	*	*
32	20	*	*	*	*
22	32	20	*	*	*
27	22	32	20	*	*
7	27	22	32	20	*
12	7	27	22	32	20
18	12	7	27	22	32
9	18	12	7	27	22
11	9	18	12	7	27
34	11	9	18	12	7
29	34	11	9	18	12
10	29	11	9	18	12

Addition

R Integration

Addition: Call R scripts from Minitab Statistical Software. R is a language and environment for statistical computing and graphics.

R scripts can run in 3 ways:

- Run the RSCR in the command line pane
 - Run a Minitab exec that includes the RSCR command
 - Customize the interface to run a Minitab exec that includes the RSCR command
-

Summary:

Execute external R scripts that use Minitab Statistical Software variables (columns, constants, matrices) as inputs. Results are returned to Minitab and displayed in the output navigator and output pane.



Addition

Minitab's Predictive Analytics Module*

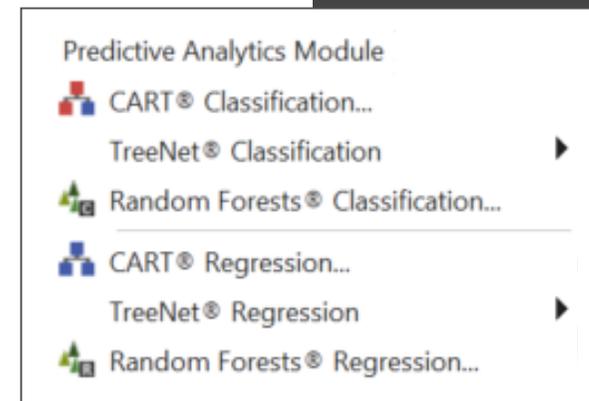
Improvement: Tree-based machine learning algorithms in an easily accessible module.

User Benefit: Tree-based methods empower predictive analytics with not only speed to answer, but also remarkable accuracy and ease of interpretation. Users can quickly understand the key drivers of a process.

Summary: Our proprietary, best-in-class, tree-based machine learning algorithms not only have the power to provide deeper insights and visualize multiple complex interactions with decision trees but are equipped to handle larger data sets with more variables, messy data, missing values, random outliers, and non-linear relationships. These methods are now available in a module that you easily add to Minitab Statistical Software.

*Available in Minitab Statistical Software for an additional fee

[Learn more about Minitab's Predictive Analytics Module](#)



Addition

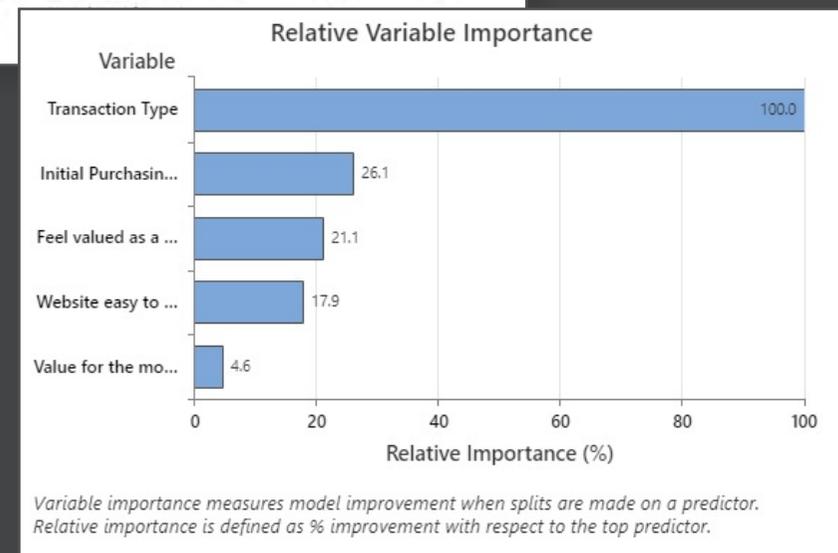
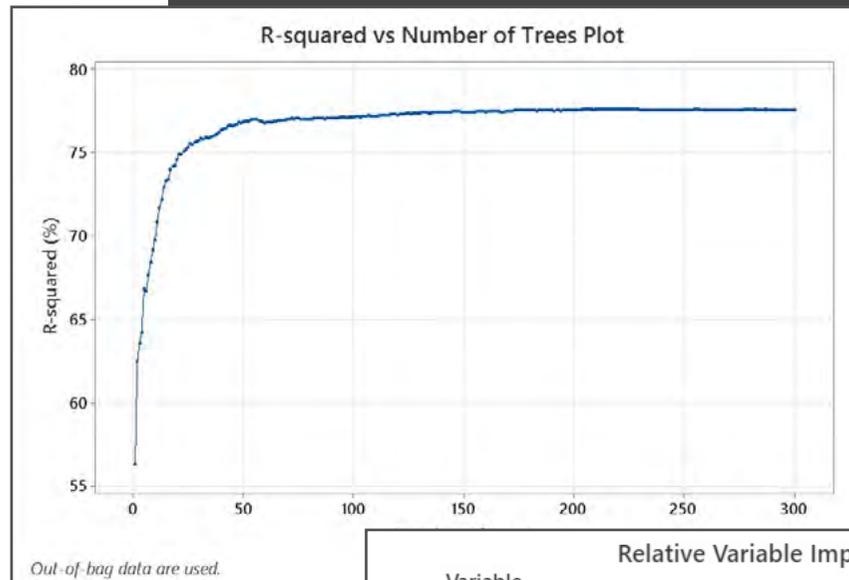
Random Forests®

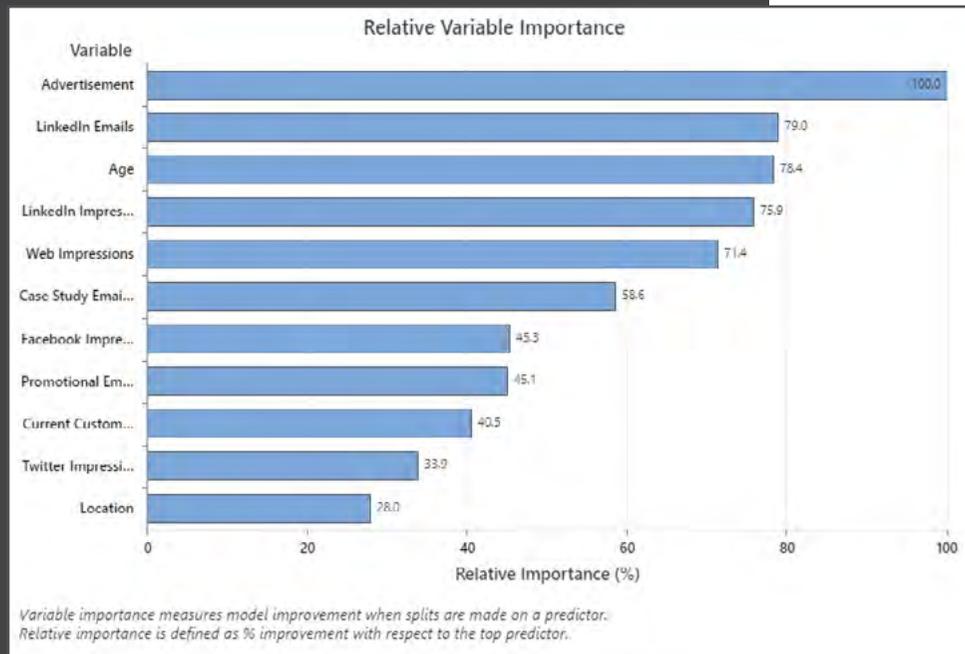
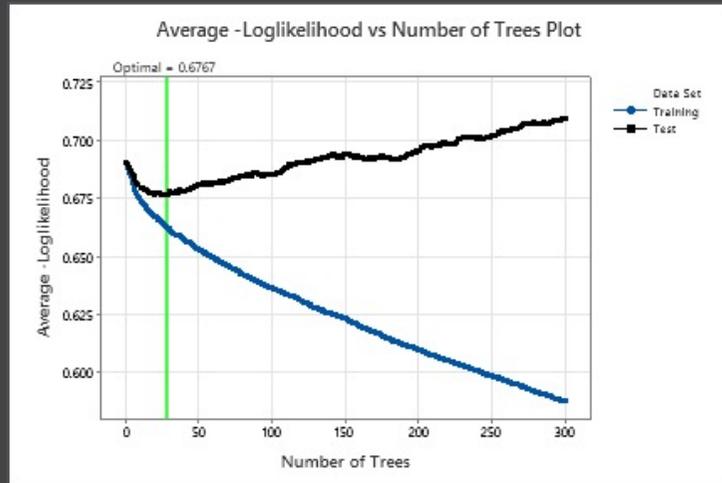
New Feature: Random Forests consists of many individual decision trees that operate as an ensemble.

User Benefit: Random Forests generally provides better predictive power than a single decision tree.

Summary: Based on a collection of CART Trees, Random Forests leverages repetition, randomization, sampling, and ensemble learning in one convenient place that brings together independent trees and determines the overall prediction of the forest.

[Learn more about Random Forests in Minitab's Predictive Analytics Module](#)





Addition

TreeNet® (Gradient Boosting)

New Feature: TreeNet Classification and TreeNet Regression. Includes Fit Model and Discover Key Predictor submenus.

User Benefit: Gradient boosting can deliver optimal prediction accuracy and unique insights.

Summary: Our most flexible, award-winning and powerful machine learning tool, TreeNet Gradient Boosting, is known for its superb and consistent predictive accuracy due to its iterative structure that corrects combined errors of the ensemble as it builds.

[Learn more about TreeNet in Minitab's Predictive Analytics Module](#)