

CHAPTER

1

Introduction to Enterprise Miner Software

<i>Data Mining Overview</i>	1
<i>Layout of Enterprise Miner</i>	2
<i>Using the Application Main Menus</i>	3
<i>Using the Toolbox</i>	8
<i>Using the Pop-up Menus</i>	9
<i>Unavailable and Dimmed Items</i>	9
<i>Enterprise Miner Documentation</i>	9

Data Mining Overview

SAS defines *data mining* as the process of *Selecting, Exploring, Modifying, Modeling, and Assessing (SEMMA)* large amounts of data to uncover previously unknown patterns that can be utilized as a business advantage. The data mining process is applicable across a variety of industries and provides methodologies for such diverse business problems as fraud detection, householding, customer retention and attrition, database marketing, market segmentation, risk analysis, affinity analysis, customer satisfaction, bankruptcy prediction, and portfolio analysis.

Enterprise Miner software is an integrated product that provides an end-to-end business solution for data mining. A graphical user interface (GUI) provides a user-friendly front end to the SEMMA data mining process:

- *Sample* the data by creating one or more data tables. The sample should be large enough to contain the significant information, yet small enough to process.
- *Explore* the data by searching for anticipated relationships, unanticipated trends, and anomalies in order to gain understanding and ideas.
- *Modify* the data by creating, selecting, and transforming the variables to focus the model selection process.
- *Model* the data by using the analytical tools to search for a combination of the data that reliably predicts a desired outcome.
- *Assess* the data by evaluating the usefulness and reliability of the findings from the data mining process.

You may or may not include all of these steps in your analysis, and it may be necessary to repeat one or more of the steps several times before you are satisfied with the results. After you have completed the assess phase of the SEMMA process, you apply the scoring formula from one or more champion models to new data that may or may not contain the target. Scoring new data that is not available at the time of model training is the end result of most data mining problems.

The SEMMA data mining process is driven by a process flow diagram, which you can modify and save. The GUI is designed in such a way that the business analyst who has

little statistical expertise can navigate through the data mining methodology, while the quantitative expert can go “behind the scenes” to fine-tune the analytical process.

Enterprise Miner contains a collection of sophisticated analysis tools that have a common user-friendly interface that you can use to create and compare multiple models. Statistical tools include clustering, self-organizing maps / Kohonen, variable selection, trees, linear and logistic regression, and neural networking. Data preparation tools include outlier detection, variable transformations, data imputation, random sampling, and the partitioning of data sets (into train, test, and validate data sets). Advanced visualization tools enable you to quickly and easily examine large amounts of data in multidimensional histograms and to graphically compare modeling results.

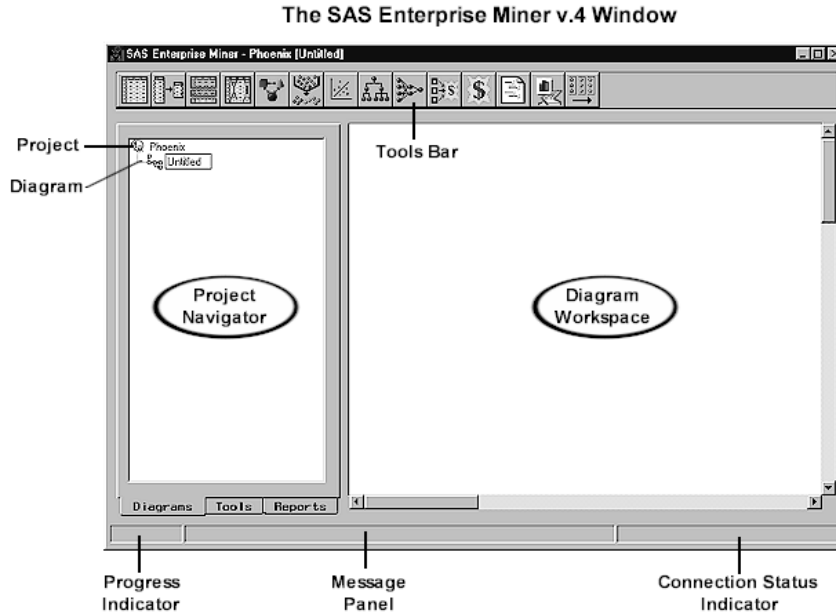
This documentation describes the core functionality of Enterprise Miner and shows you how to perform basic tasks. For more detail on the individual tool nodes, see the online Enterprise Miner Reference Help documentation that is available from the main menu:

Help ► Enterprise Miner Reference

Enterprise Miner is designed for PCs that are running Windows 95/98, Windows NT 4.0, or subsequent releases of those operating environments. The screen captures presented in this document were taken on a PC running Windows NT 4.0.

Layout of Enterprise Miner

To start Enterprise Miner, start up SAS and then type **miner** in the command dialog box of the SAS Display Manager. The SAS Enterprise Miner window opens:



The SAS Enterprise Miner window contains the following interface components:

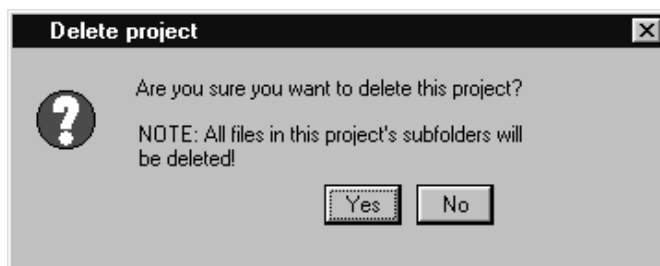
- *Project Navigator* — You use the Project Navigator to manage projects and diagrams, add tools to the Diagram Workspace, and view HTML reports that are created by the *Reporter* node. Note that once a tool is added to the Diagram Workspace, it is referred to as a node. The Project Navigator has three tabs:

- **Diagram tab** — lists the current project and the diagrams within the project. By default, Enterprise Miner creates a local project that uses your userid as the name (for this example, **Phoenix**) and a process flow diagram named **Untitled**. When you re-open Enterprise Miner, the last project you were working with is loaded.
- **Tools tab** — contains the Enterprise Miner tools pallet. The tools pallet presents all the data mining tools that are available for constructing process flow diagrams. The tools are grouped according to the SEMMA data mining methodology.
- **Reports tab** — contains the HTML report entries for the project. You generate reports using Enterprise Miner's *Reporter* node.
- **Diagram Workspace** — enables you to build, edit, run, and save process flow diagrams.
- **Tools Bar** — contains a customizable subset of Enterprise Miner tools that are commonly used to build process flow diagrams in the Diagram Workspace. You can add or delete tools from the Tools Bar.
- **Progress Indicator** — displays a progress indicator bar that indicates the execution status of an Enterprise Miner task.
- **Message Panel** — displays messages about the execution of an Enterprise Miner task.
- **Connection Status Indicator** — displays the remote host name and indicates whether or not the connection is active for a client/server project.

Using the Application Main Menus

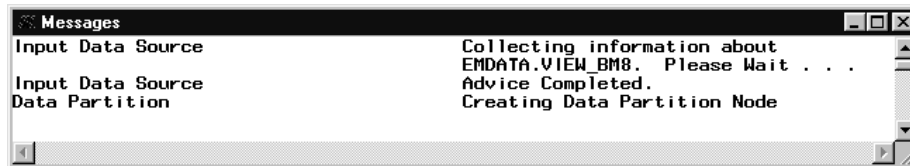
You use Enterprise Miner's application menus to quickly and efficiently perform common tasks. Different windows have different menus. Here is a brief summary of the SAS Enterprise Miner window pull-down menus:

- **File**
 - **New**
 - **Project** — creates a new project.
 - **Diagram** — creates a new diagram.
 - **Open** — opens a new or existing diagram within the current project.
 - **Import Version 2 project** — imports an Enterprise Miner 2.0x project.
 - **Save Diagram** — saves the current diagram to the project.
 - **Save Diagram as** — names and saves the current diagram to the project.
 - **Print Setup** — specifies printing options.
 - **Print** — prints the contents of the SAS Enterprise Miner window.
 - **Delete current project** — deletes the current project. If you select this menu item, then a message appears that asks you to verify the deletion.



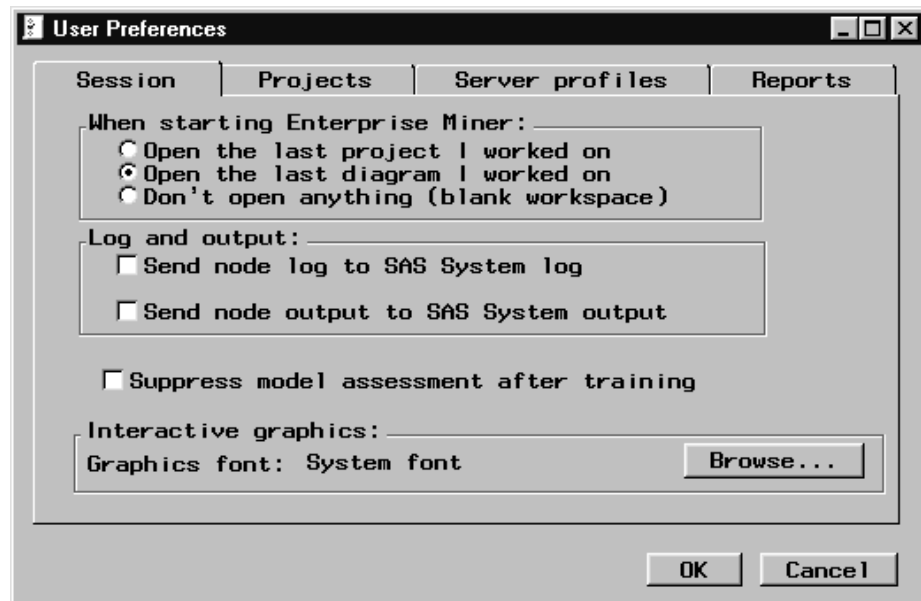
To delete the project, select **Yes**. All files in the project folder and any subfolders will be deleted. You cannot delete a project that someone else has opened.

- Close diagram** – closes and saves the diagram.
- Close project** – closes the project.
- Exit Enterprise Miner** – ends the Enterprise Miner session.
- Edit**
 - Copy diagram to clipboard** – copies the diagram to the clipboard.
 - Undelete** – undeletes the last deleted node.
 - Copy** – copies a node or object to the paste buffer. You must first select the node or object you want to copy.
 - Delete** – deletes the selected node or connection from the Diagram Workspace.
 - Clone** – clones the selected node (copies the current node and adds it to the Custom folder of the Tools Palette).
 - Paste** – pastes the node, diagram, or object from the clipboard.
 - Select all** – selects all nodes in the Diagram Workspace.
 - Create subdiagram** – creates a subdiagram that you use to collapse a set of selected nodes and connections into a single *Subdiagram* node icon.
- View**
 - Messages** – opens the Messages window that displays collected messages for this diagram.



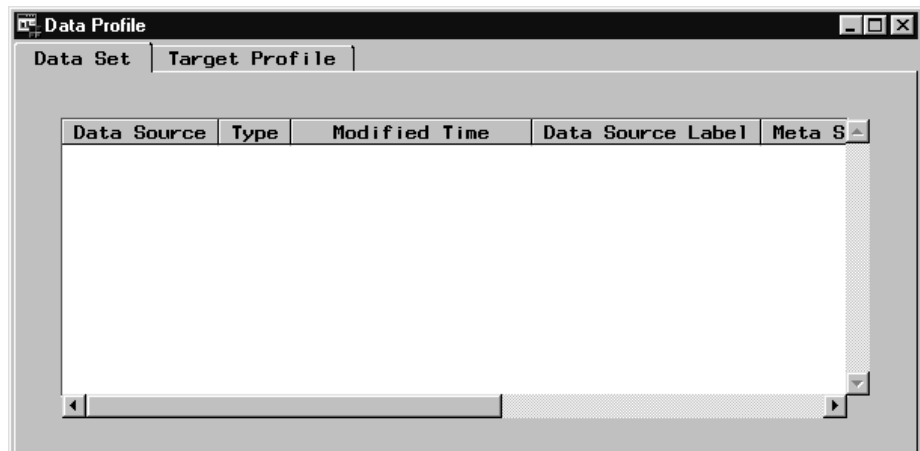
Use the window control buttons to close the Messages window.

- Refresh** – refreshes the Project Navigator and the Diagram Workspace view.
- Up One Level** – displays the next higher level of the process flow diagram. If there are no subdiagrams in your process flow diagram, then only one level can be displayed. If there are subdiagrams in your process flow diagram, then the subdiagrams can be displayed either in their condensed form (hiding the internal structure of the subdiagram) or in their expanded form (showing their internal structure). The higher level displays subdiagrams in condensed form.
- Top Level** – displays the process flow diagram in its most condensed form. All subdiagrams are condensed.
- Options**
 - User preferences** – opens the User Preferences window that enables you to specify start-up options, set the default directories for new projects, set up server profiles for client/server projects, and specify the HTML report items that you want to display when you run the *Reporter* node.



□ **Projects**

- **Data Profiles** – opens the Data Profile window that you use to define target profile information for a data set. To learn about defining target profiles, read the Target Profiler chapter in the online Enterprise Miner Reference, available from the application’s main Help menu.

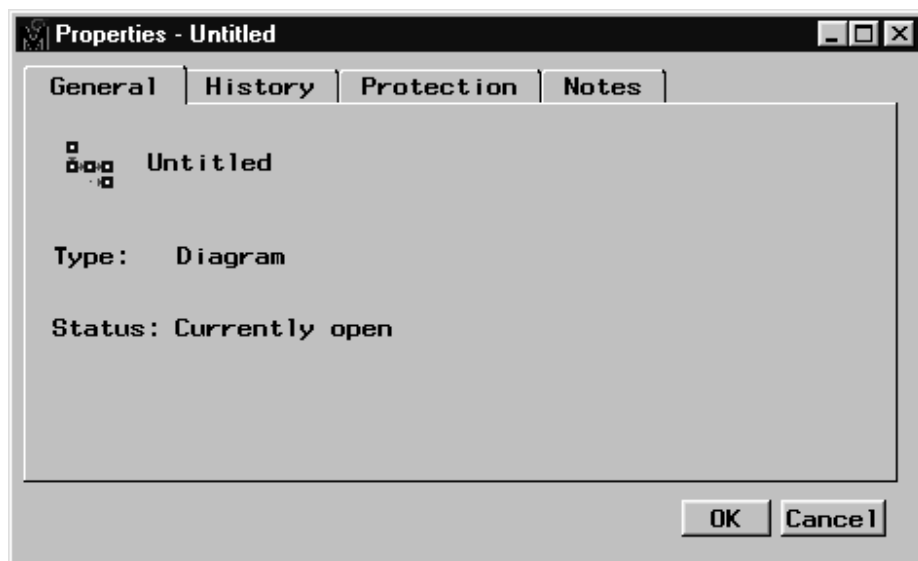


- **Properties** – shows project properties, such as the project name and type, share status, project location, server information, and a list of users that have the project open. You can also specify start-up and exit code that runs when the project is opened and closed, specify the SAS warehouse path, or update the project server profile. These settings can be specified if the project is opened by one user.



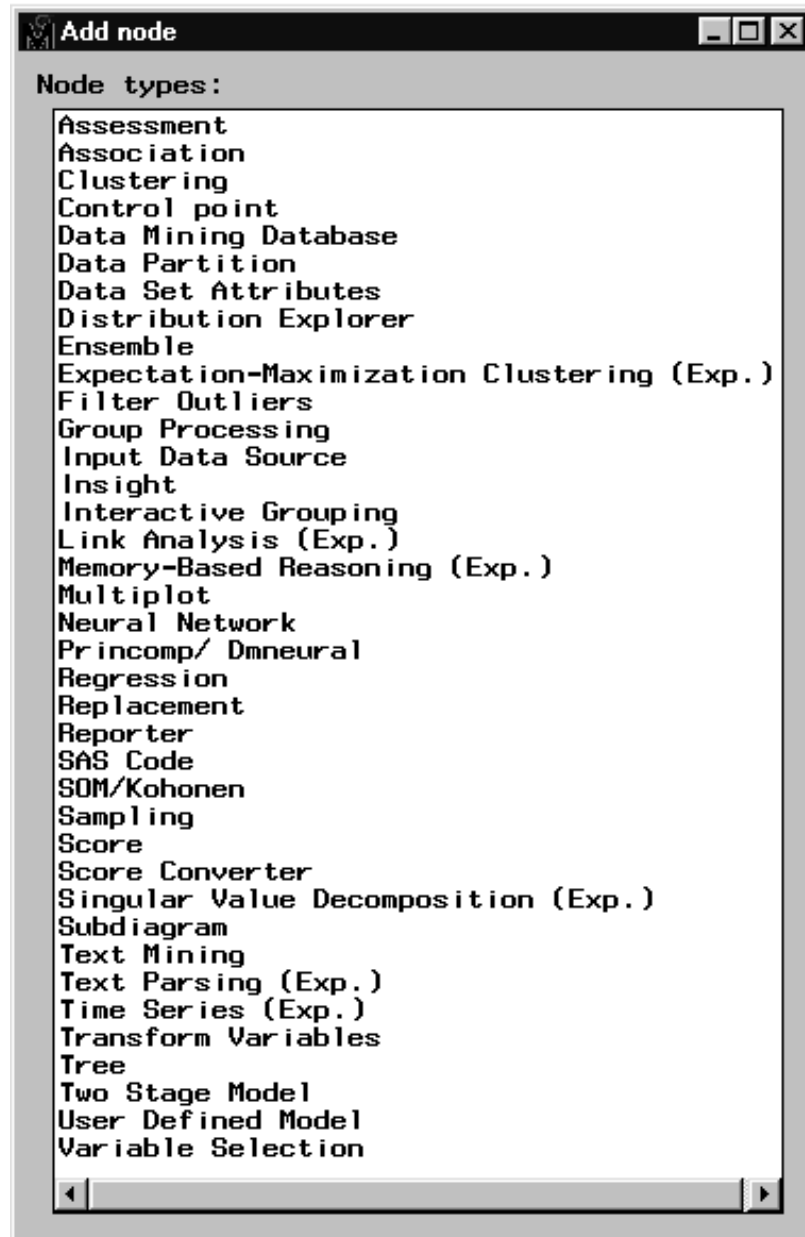
□ **Diagram**

- **Connect items** – a mode of editing where node icons are fixed in location, so that you can more easily make connections among them.
- **Move items** – a mode of editing where node icons cannot be connected, so that you may more easily move the icons into desired locations in the process flow diagram.
- **Move and Connect**– (default) a mode of editing where node icons can be moved and connected in the Diagram Workspace.
- **Large icons** – displays large node icons in the Diagram Workspace. This item is grayed out if a node is selected in the Diagram Workspace.
- **Small icons** – (default) displays small node icons in the Diagram Workspace. This item is grayed out if a node is selected in the Diagram Workspace.
- **Properties** – opens the Properties window that enables you to view diagram properties, such as the name, type, status, date created, and date last modified. You can also password protect a diagram in the Protection tab and type and store notes about the diagram in the Notes tab.



□ **Actions**

- **Open** – opens the selected node.
- **Run** – runs the selected node and any predecessor nodes in the process flow that have not been previously executed.
- **Results**– opens the results browser for those nodes that generate results.
- **Add Node** – opens the Add node window that enables you to add a new node to the Diagram Workspace.



- **Add Endpoints** – add endpoints to the process flow.
- **Help**
 - **Getting Started with Enterprise Miner** – this document in online Help format.
 - **About Enterprise Miner** – Production release information about your copy of Enterprise Miner.

- **Help** – Enterprise Miner Reference Help. This is a hierarchical help listing. Window frame-based help is also available using the “What’s This?” tool icon from the Enterprise Miner Toolbox.

Using the Toolbox

With Enterprise Miner, the toolbox at the top of the SAS window appears as follows:



Tool tips are displayed when you hold your mouse pointer over a tool icon. The icons of the Toolbox represent the following:

- 1 *Open* – opens a diagram within a project. You can also perform this task by using the main menu:
 - File ► Open
- 2 *Save diagram* – saves the current diagram. You can also perform this task by using the main menu
 - File ► Save diagram
- 3 *Delete Current Project* – deletes the current project. You can also perform this task by using the main menu:
 - File ► Delete current project
- 4 *User Preferences* – opens the User Preferences window that you use to specify start-up options, set the default directories for new projects, set up server profiles for client/server projects, and specify the HTML report items that you want to display when you run the *Reporter* node. You can also open the User Preferences window from the main menu:
 - Options ► User preferences
- 5 *Print diagram* – prints the current process flow diagram. You can also perform this task by using the main menu:
 - File ► Print
- 6 *Copy diagram to clipboard* – copies the diagram to the clipboard. You can also perform this task by using the main menu:
 - Edit ► Copy diagram to clipboard
- 7 *What’s This?* – turns the mouse pointer into a question mark, indicating that Enterprise Miner is in help mode. Click inside an Enterprise Miner window with the question mark mouse pointer to launch frame help. The window-based frame help provides information on objects in the window frame that was selected with the question mark mouse pointer.
- 8 *Getting Started* – launches the electronic version of this document from within Enterprise Miner.

Using the Pop-up Menus

Many of the functions found in the Enterprise Miner application main menus can be duplicated using pop-up menus. Pop-up menus are launched by right-clicking a GUI object within Enterprise Miner. You can use pop-up menus to perform the same tasks. This documentation frequently shows you how to perform tasks using the main menu, but many of these tasks can also be performed by using the pop-up menus.

For example, instead of using the main menu to insert a diagram, follow these steps:

- 1 Select the Diagram tab of the Project Navigator and right-click on either the Project folder, the diagram icon, or anywhere in the background of the tab. A pop-up menu appears.
- 2 Select the **New Diagram** pop-up menu item. A new diagram is automatically added to the project.

Unavailable and Dimmed Items

At times particular window objects and menu items are unavailable for use. When a window object or menu item is unavailable for use, it will appear dimmed and will not respond to mouse clicks. Some window objects are grayed dynamically; that is, when an object's use is contextually inappropriate for the current data mining settings, it appears dimmed. When the data mining settings are changed so that the object's use is contextually appropriate, then the objects become active and will accept user input.

For statistical functions, dimmed objects are used to prevent the selection of inappropriate combinations of options. In some cases, specifying particular options implies that additional options must also be selected to completely specify the method or technique. Be aware that not fully specifying options can cause errors.

Enterprise Miner Documentation

In addition to this document, Enterprise Miner contains the following documentation:

- *Enterprise Miner Reference* – describes the usage of each Enterprise Miner node. It also contains other chapters that describe how to use various features of Enterprise Miner, such as the Target Profiler. The Enterprise Miner Reference Help is available through the application's main menu:

Help ► EM Reference

Note: Users can search for text strings in the EM Reference Help, using the menu choices:

Edit ► Find

and typing in the desired search string. △

- *Frame Help* – Enterprise Miner nodes have Help available at the window frame level. For Help on a field or object contained in a particular window frame:

- 1 Click on the **What's This?** tool icon in the Enterprise Miner Toolbox. The mouse pointer will change into a question mark.



- 2 Click inside the window that you would like help on.
 - 3 The Enterprise Miner Help system will deliver Help on the contents of the window frame you selected.
- *Enterprise Miner add-ins to SAS/Warehouse Administrator* – This document describes how to use the Enterprise Miner add-ins to SAS/Warehouse Administrator. To download this document in HTML format, go to the SAS web site http://www.sas.com/software/data_warehouse, and select the **Demos/Downloads** link.