

MineSet™ 3.0 Enterprise Edition

Visual Insight on Every Desktop

Features

- MineSet server on Windows NT® and IRIX, MineSet client on Windows NT, Windows 95, and Windows 98
- Ability for Windows based clients to connect to both Windows NT and IRIX servers
- ODBC connectivity; provides access to SQL Server and other ODBC-compliant databases
- ActiveX visual components
- Software Development Kit providing an application programming interface to major MineSet components (will be available as an option)

Visual Insight on Every Desktop

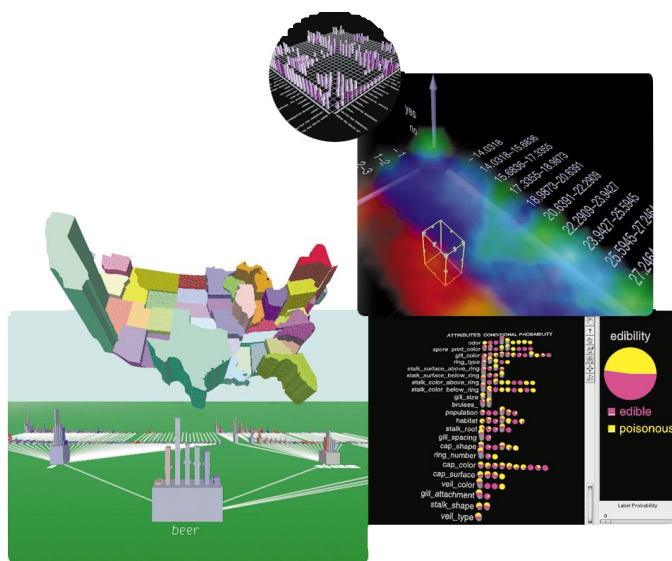
Buried within your vast store of transaction logs and legacy data, subtle correlations and elusive relationships hold the secrets to enhance your business. To reveal the hidden value in your data warehouse, SGI developed MineSet, an integrated suite of software tools for data mining and data visualization. Powered by industry-leading Silicon Graphics® visual workstations and servers, MineSet revolutionizes the way you use data mining for strategic business analysis. With the introduction of MineSet 3.0 Enterprise Edition, Silicon Graphics extends the power of visual data mining, previously available only to UNIX® users, to the Windows® desktop. By providing the same functionality on both IRIX® and Windows, SGI allows business users to enjoy visual interpretation of complex data mining algorithms on PCs as well as take advantage of parallelization features on Origin™ servers, famous for industry-leading scalability.

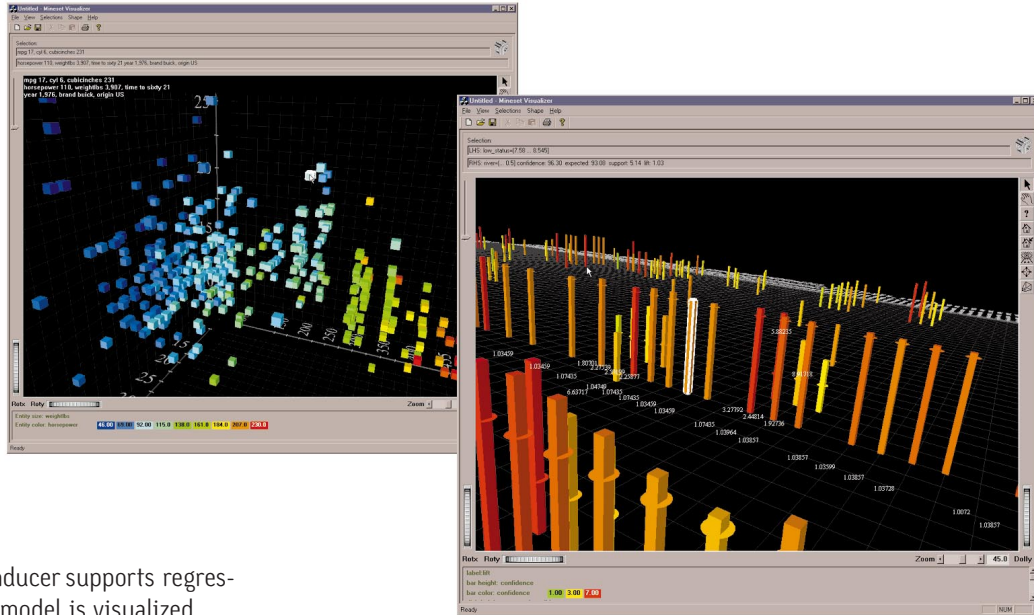
Intuitive Visual Data Mining Tools

MineSet enables interactive exploration of data through an advanced suite of visual tools for faster discovery of meaningful trends and relationships. The *Splat Visualizer* and the *Scatter Visualizer* represent complex data in up to eight dimensions. The *Map Visualizer* displays data with geographical relationships by using a map metaphor. Animation and view synchronization techniques are used to reveal patterns over critical dimensions such as time. The *Tree Visualizer* depicts data with hierarchical relationships utilizing a fly-through technique set in a 3D landscape. The *Statistics Visualizer* presents a visual summary of basic statistical information. Advanced drill-through techniques give you fast access to the original records that created entities within your visualization for additional analysis. The *Record Viewer* allows quick access to the original data with column sorting and HTML output capabilities.

Powerful Analytical Data Mining Algorithms with Matching Visualizers

Below the elegant interfaces of the visual tools, state-of-the-art analytic data mining algorithms build comprehensive models. MineSet includes multiple classifier inducers, including a *Decision Tree*, an *Option Tree*, *Evidence*, and *Decision Tables*. Boosting allows MineSet to further improve the accuracy of these classifiers. *Decision Tree* and *Option Tree* classification models are viewed through the *Tree Visualizer*. An associated tool, the *Evidence Visualizer*, displays the structure and properties of the *Evidence Classifier* and supports instantaneous what-if analysis. The *Decision Table Visualizer* displays multidimensional data on important columns automatically chosen by the matching inducer. The visual-OLAP paradigm supports multilevel drill-through capability.





The *Regression Tree* inducer supports regression, and the resulting model is visualized using the *Tree Visualizer*. The *Association Rules Generator* analyzes data to discover product affinities and relationships between data entities. This feature can extract both one-to-one and multiway associations. The resulting one-to-one rules are depicted by the *Scatter Visualizer* in an easy-to-understand graphical format. Multiway rules are summarized in a text report. The *Column Importance* feature enables automatic or user-guided selection of important columns for use with the MineSet visualizers. The k-means-based Clustering algorithm is used for segmentation analysis and is visualized using the Cluster Visualizer.

Intuitive User Interface and Innovative Enterprise Solutions

MineSet is designed to fit the needs of users at several levels of technical ability. For technical users, MineSet offers unlimited potential with a complete suite of data mining tools, database integration, and scalable performance. To accommodate business users, data mining results and visualizations are easily deployed across corporate networks through point-and-click access. Analytical models developed in MineSet can be applied to any data set with the touch of a button. This easy scoring capability requires no programming or running of complicated batch programs. The MineSet Application Interface allows developers to embed MineSet tools within their customized solutions.

Integrating MineSet into Your Computing Environment

With seamless support for ODBC-compliant databases including SQL Server and DB2 and popular databases such as Oracle®, Informix®, and Sybase®, MineSet makes it easy to add new capabilities to your data warehouse. Queries to your relational database are supported through graphical or SQL commands. An open architecture allows MineSet to coexist with other data mining and visual tools, including SAS® software. Hot links help MineSet users share discoveries and results through Web pages, launching appropriate MineSet visualization tools for further analysis.

The capability to connect to MineSet server on both IRIX and Windows NT from a MineSet client on Windows provides unparalleled flexibility for business users who want to protect their investment in software as well as allow for explosive data growth.

Industry-Leading Scalability and Large Memory Support*

MineSet is optimized for the ultimate scalable performance on the Origin server. In addition to parallelization, MineSet offers very large memory support [over 2GB] and data handling capability using the 64-bit implementation.

MineSet Enterprise Edition allows business users unequalled flexibility and investment protection by providing the same back-end analytical processing capability on the popular NT platform. Parallelization and 64-bit support are offered on the highly scalable SGI Origin servers.

Using MineSet as a Platform for Developing Applications

MineSet 3.0 Enterprise Edition constitutes the ideal data mining application development platform through an advanced API and ActiveX visualization components. Business users and software vendors alike can reuse MineSet components or plug in custom algorithms.

Data Access and Transformation Features

- Support for ODBC-compliant databases, including SQL Server and DB2
- Support for access to Oracle, Informix, and Sybase running on all major UNIX platforms, including IBM, Sun, HP, and Digital
- Support for access to flat files (ASCII or binary)
- Data transformation history and graphical editing facility
- SAS file import/export utility
- Data transformation support for:
 - Automatic or user-specified binning of data
 - Data aggregations with indexed arrays using average, minimum, maximum, sum, and count
 - Data distribution (transpose)
 - New column creation using expressions
 - Data type conversion
 - Sampling
- Save and restore session management
- Client/server architecture
- Statistical tool for determining minimum, maximum, means, median, standard deviation, histograms, and quartiles

Analytical Data Mining Features

- | | |
|---------------------------|---------------------------------|
| Classification | Regression |
| • Decision Tree | • Regression trees |
| • Option Tree | |
| • Evidence (simple Bayes) | • Associations |
| • Decision Table | • One-to-one |
| | • Multiway |
| Clustering | Column Importance |
| • K-means | • Automatic attribute selection |
| • Iterative k-means | |

Advanced Analytical Options

User-specified parameters for control of classifier induction, including:

- Boosting
- Node-splitting criteria for decision tree induction
- Pruning factors for decision tree induction
- Laplace correction for evidence induction
- Automatic feature selection for evidence induction
- Data holdout percentage for automatic accuracy estimation (holdout and cross-validation)

Advanced features for classification, including:

- | | |
|-------------------|-------------------|
| • Loss matrices | • Record weights |
| • Lift curves | • Learning curves |
| • Binning options | • Backfitting |
| • Scoring | |

Visual Data Mining Features

- User-defined mapping of data to visualization components
- Visual and numeric normalization of data across visual components such as height, color, and size
- Discrete or continuous mapping of data to colors
- Web launching support
- Visual drill-up, drill-down, and drill-through to source data
- Animation of dependent data by up to two user-defined independent variables for trend analysis
- VCR-style animation playback
- Visual filtering and querying of data
- Synchronous animation of multiple visualizations
- 3D fly-through

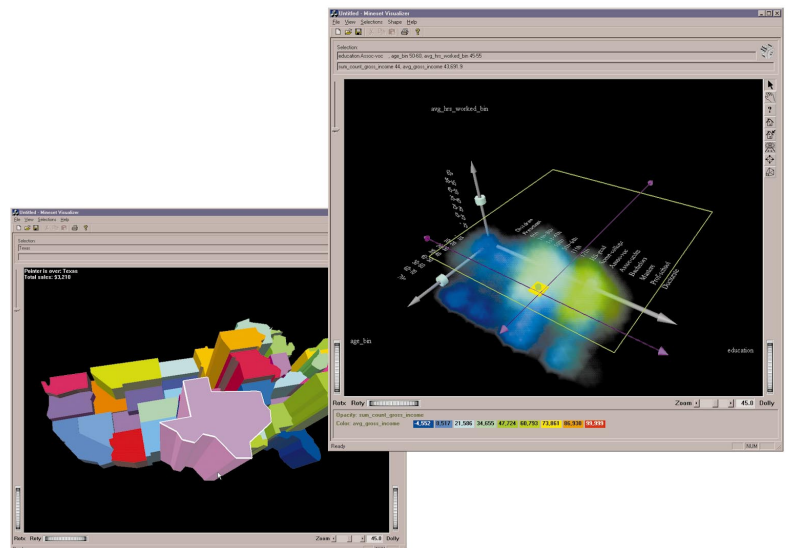
Integration with Windows Applications


Visualizers are provided as ActiveX components. This allows embedding of visualizations into Windows applications. Application vendors can embed industry-leading MineSet visualizations into business reports generated by their decision support tools.

Product Requirements

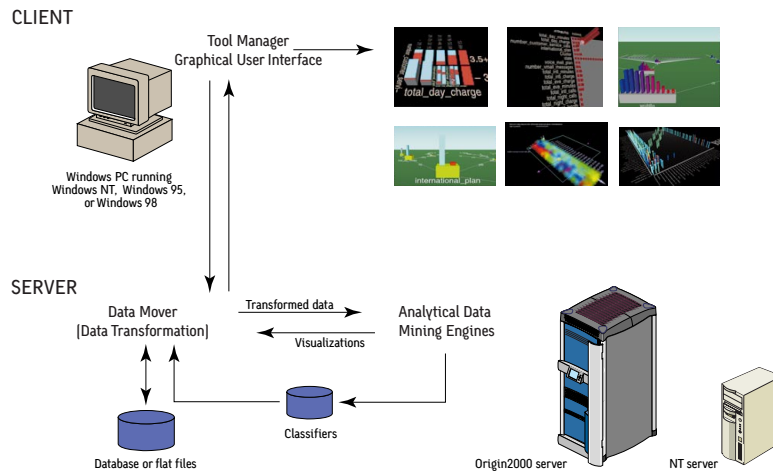
MineSet server and client are supported on all PC-compatible systems running on Windows NT 4.0. The client can also run on Windows 95 and Windows 98. Memory requirements for the server vary with the size of the data. Clients must have at least 64MB of memory; 96MB is recommended. Graphics accelerator cards enhance visualization performance; 1024x768 resolution with 65K colors is the minimum recommended.

IRIX 6.4 or higher is required for 64-bit support and parallelization on the server. Silicon Graphics platforms must be running IRIX 6.2 and above.



	Silicon Graphics 320™	Silicon Graphics 540™	NT Server	Origin2000™
				
Sample Configurations	1 user entry level client/server	1 user heavy graphics client/server	1-4 users heavy analytics server only	5 or more users heavy analytics, big data server only
Parallelized	NA	NA	NA	Yes
Model	Single Pentium® II 500 MHz	Single Pentium II 550 MHz	Depends on the platform	8-CPU 250 MHz R10000®
Memory	128MB	256MB	Depends on the platform	4GB
Resolution	1900x1200	1900x1200	NA	NA
Color Depth	32-bit	32-bit	NA	NA
Maximum Memory	1GB	2GB	Depends on the platform	256GB
Maximum CPUs	2	44	8/32	256

MineSet Client/Server Architecture



Corporate Office
1600 Amphitheatre Pkwy.
Mountain View, CA 94043
(650) 960-1980
www.sgi.com

North America (800) 800-7441
Latin America (650) 933-4637
Europe (44) 118.925.75.00
Japan (81) 3.5488.1811
Asia Pacific (65) 77.10.290

www.sgi.com/go/mineset/

© 1999 Silicon Graphics, Inc. All rights reserved. Specifications subject to change without notice. Silicon Graphics and IRIX are registered trademarks, and MineSet, Origin, Origin2000, Silicon Graphics 320, Silicon Graphics 540, and the Silicon Graphics logo are trademarks, of Silicon Graphics, Inc. UNIX is a registered trademark in the U.S. and other countries, licensed exclusively through X/Open Company Limited. Oracle is a registered trademark of Oracle Corporation. Sybase is a registered trademark of Sybase, Inc. The following are worldwide trademarks of Informix Software, Inc., Informix Corporation, or their subsidiaries, registered in the U.S. as indicated by "™" and in numerous countries worldwide: Informix, Windows and Windows NT are registered trademarks of Microsoft Corporation. SAS is registered by SAS Institute, Inc. IBM is a registered trademark of International Business Machines. Sun and Java are trademarks of Sun Microsystems, Inc. HP is a registered trademark of Hewlett-Packard. R10000 is a registered trademark of MIPS Technologies, Inc. Pentium is a registered trademark of Intel Corporation. All other trademarks mentioned herein are the property of their respective owners. Image credits: Underwater Camera courtesy of Katz Design. Brake disc assembly designed with CATIA/Dassault Systemes.