

# *STATISTICA*

## DATA MINER



### Verwandeln Sie Rohdaten in Wissens-Werte

- ✓ Intuitive Benutzerführung
- ✓ Automatisierung von Standardaufgaben
- ✓ Performante Analyse auch sehr großer Datenmengen
- ✓ Günstige Total Costs of Ownership
- ✓ Lizenzmodelle für kleine und große Unternehmen
- ✓ Großes Methodenspektrum, klare Ergebnispräsentation



## Data Mining = Wissens-Werte aus Daten gewinnen ...

Mit dem technischen Fortschritt wachsen auch die Datenmengen in jedem Unternehmen. Doch viele Daten werden nur für operative Zwecke gesammelt und gespeichert – aber nicht ausgewertet. Dabei verbergen sich hier Wissensschätze, die nur gehoben sein wollen!

Die Anwendungsgebiete von Data Mining sind ebenso vielfältig wie branchenübergreifend: Viele Banken, Versicherungen, Medien-, Handels-, Versand-, E-Commerce- oder Telekommunikationsunternehmen setzen es bereits erfolgreich ein. Aber auch in der industriellen Fertigung, Versorgung, Biochemie, Life Sciences, Medizin und Pharmaforschung wird zunehmend mit Data Mining gearbeitet, um Prozesse zu optimieren und Ursachen zu erforschen.

Profitieren Sie von einem der modernsten Data Mining Tools am Markt! Der *STATISTICA Data Miner* erkennt auch in Ihren Daten die entscheidenden Muster, Trends und Zusammenhänge.



*„Durch die Anwendung von Data-Mining-Methoden war es möglich, vorhersagende Modelle zur Prozessoptimierung zu erstellen. Maßgeblich hierfür war der STATISTICA Data Miner mit seiner einzigartigen Vielfalt an implementierten Algorithmen.“*

*Silvio Miccio,  
Process Development, Procter & Gamble Service GmbH*

## ... beim Einsatz in klassischen Bereichen ...



### Kundensegmentierung

Segmentieren Sie für eine gezielte Ansprache Ihren Kundenbestand in Gruppen mit ähnlichem sozio-demographischen Profil und Nachfrageverhalten.



### Risikoprososen

Leiten Sie aus Ihren Daten valide Prognosemodelle ab, um Ausfallrisiken zu minimieren. Erstellen Sie Modelle und „Was-wäre-wenn“-Szenarien zur Risikoabschätzung. Bewerten Sie Kreditrisiken nach Basel II.



### Marketing-Optimierung

Steigern Sie den Response und ROI Ihrer Direktmarketingaktionen: mit individualisierter Kundenansprache, Mailing-Frequenz und passgenauen Angeboten. Optimieren Sie Up- und Cross-Selling.



### Warenkorb- und Webanalysen

Welche Produkte und Leistungen sollen welchen Kunden angeboten werden? Und zu welchem Zeitpunkt und welchen Konditionen? Die Vorlieben Ihrer Kunden geben Ihnen Antworten.



### Weitere CRM-Szenarien

Setzen Sie im analytischen CRM Data Mining zur Betrugserkennung, für Kundenwertanalysen und im Beschwerdemanagement ein.



### OLAP

Sichten Sie Ihre Daten Schritt für Schritt mittels OLAP-Tools. Bewerten Sie z. B. Umsätze vor dem Hintergrund von PLZ-Bereichen oder Produktkategorien. Identifizieren Sie umsatzstarke Niederlassungen oder Problemregionen.



### Textanalyse

Mit Data Mining werten Sie auch Textinformationen ohne feste Struktur aus. So gewinnen Sie Informationen aus umfangreichen Dokumentensammlungen – wie beispielsweise Ihre Kundenkorrespondenz oder Online-Foren.

## ... sowie im industriellen Umfeld



### Prozessoptimierung

Automatisierte Prozesse speichern Daten in hoher Taktrate, insbesondere in der industriellen Fertigung oder der Energieerzeugung. Mit Data Mining identifizieren Sie selbst in Massendaten die entscheidenden Parameter. „Was-wäre-wenn“-Simulationen zeigen Ihnen, was Sie konkret verbessern können.



### Qualitätssicherung

Überwachen Sie mit multivariaten Regelkarten eine Vielzahl an Qualitätsmerkmalen simultan. Online-Monitoring und Echtzeitprognosen erlauben noch im laufenden Herstellungsprozess ein Nachjustieren.



### Fehleranalyse

Erkennen Sie Störgrößen und Fehlermuster bei Rohmaterialien, im Herstellungsprozess oder bei fertigen Produkten. Assoziationsregeln offenbaren das gemeinsame Auftreten verschiedener Qualitätsprobleme oder typische Fehlerabfolgen. Interaktive Entscheidungsbäume erlauben eine flexible und verständliche Ursachenanalyse.



### Forschung und Entwicklung

In Forschung und Entwicklung müssen zuweilen extrem große Datenmengen analysiert werden. Mit Data Mining erkennen Sie selbst im Terabyte-Bereich noch Muster und komplexe Zusammenhänge. Moderne Verfahren spüren beispielsweise auf, welche Gene für bestimmte Krankheiten verantwortlich sind.



### Bilddatenanalysen

Mit Data Mining lassen sich auch Bilder auswerten. So werden in Produktionsstraßen Oberflächen von Bauteilen gefilmt und Fehler mithilfe von Verfahren des Maschinellen Lernens automatisch erkannt und lokalisiert.

# Bedienung und Automatisierung

Der *STATISTICA Data Miner* passt sich Ihren Anforderungen und Kenntnissen an. Es existieren verschiedene Anwendungsmodi, die in Bedienerfreiheit und Automatisierungsgrad variieren:

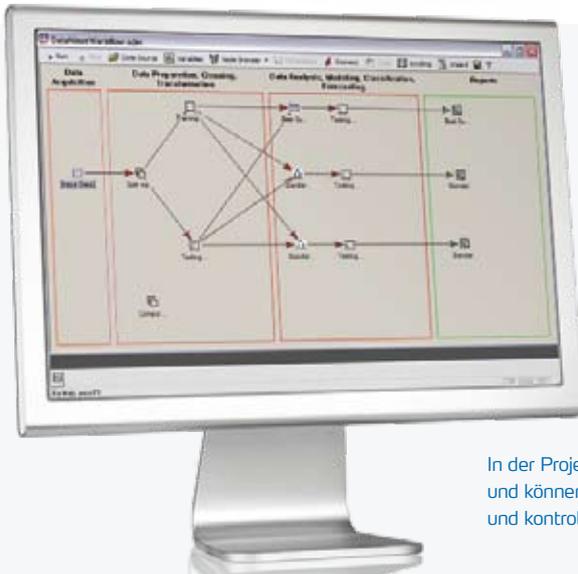
## Expertenmodus



Als Spezialist steht Ihnen über intuitiv strukturierte Menüs das gesamte Spektrum an Methoden und Optionen für alle Schritte des Data-Mining-Prozesses zur Verfügung.

Freier Zugriff auf alle Methoden und Einstellungen über eine moderne und übersichtliche Bedieneroberfläche – ideal für den Experten.

## Projektoberfläche



Alle Analyseschritte des Data-Mining-Prozesses – von der Datenübernahme, der Datentransformation, der Modellierung bis zur Ergebnisdarstellung und Rückschreiben von Scorewerten – lassen sich in einem Projekt abbilden und speichern. Die Prozessschritte werden übersichtlich durch Icons und Verbindungslinien dargestellt. Selbst hochkomplexe Projekte können so auf Knopfdruck ausgeführt, modifiziert und wiederverwendet werden.

In der Projektoberfläche werden alle Arbeitsschritte visualisiert und können per Knopfdruck ausgeführt werden. So strukturiert und kontrolliert man seinen Data-Mining-Prozess.

**voestalpine**

„Beim *STATISTICA Data Miner* wird ein breites Spektrum an Analysemethoden zur Verfügung gestellt, das uns ermöglicht, auch komplexe Zusammenhänge in unseren Prozessdaten zu identifizieren und entsprechende Klassifikations- und Prognosemodelle zu generieren. Auf diese Weise können die Prozesse und die Produktqualität noch weiter verbessert werden.“

Silvia Mitter  
Forschungsingenieur, voestalpine Stahl GmbH

## Assistentenführung

Ein Assistent führt Schritt für Schritt durch den gesamten Data-Mining-Prozess – wie ein Kochrezept. Alle Schritte werden in der klaren Benutzerführung erläutert, sinnvolle Voreinstellungen angeboten und Fehleingaben abgefangen. So kann auch der gelegentliche Anwender grundlegende Data-Mining-Analysen erfolgreich ausführen.



Data Mining auf Rezept: Der Assistent führt durch den Prozess, der Anwender bleibt nie auf sich allein gestellt.

## Einbettung

Data-Mining-Aufgaben wie Prognosen oder Segmentierungen laufen praktisch unsichtbar in einer Anwenderumgebung wie *STATISTICA Enterprise* oder anderen BI-Applikationen. Sie müssen keinerlei Einstellungen vornehmen, sondern lediglich Aufgaben aus der vertrauten Nutzeroberfläche starten. Die Ergebnisse werden gemäß den vorher festgelegten Einstellungen erstellt und angezeigt.

## Überzeugende Performance

Data Mining bewegt oft sehr große Datenbestände und führt aufwendige Modellberechnungen durch – eine gute Performance der Software ist daher Voraussetzung. Mit *STATISTICA Data Miner* sind Sie bestens gerüstet:

- Optimierte Algorithmen, die mit modernster Compiler-Technologie entwickelt wurden
- Eine 64-Bit-Version für schnellste Mikroprozessoren
- Die Verteilung von Prozessen auf zwei Prozessoren und die Übergabe von Abfragen auf einen Datenbankserver bereits in der Desktopversion des *STATISTICA Data Miner*
- Optimierter Lese- und Schreibzugriff auf große Datenbanken: Die IDP-Technologie (In-Place Database Processing) liest Daten asynchron direkt von Datenbankservern.
- Die Serverversion des *STATISTICA Data Miner* ermöglicht die Verlagerung aller Berechnungen auf leistungsstarke Multiprozessorsysteme. Die Serverversion kann außerdem von einem PC in einem Intra- oder Internet über eine Web-Oberfläche – d. h. ohne installierte *STATISTICA*-Komponenten – gesteuert werden.

**DEVK**  
VERSICHERUNGEN

„Die schnelle Einsatzbereitschaft und der Betrieb von *STATISTICA Data Miner* durch unsere Fachabteilung ermöglichten uns ein schlankes Projektvorgehen.“

Bruno Küpper,  
Referent für Kommunikationstechnik im Bereich Kundenbindung und Dialog-Marketing,  
DEVK Versicherungen

# Umfassendes Methodenspektrum

Die „beste Methode“ für Data Mining gibt es nicht. Studien belegen, dass abhängig von der Fragestellung und den Daten jeweils unterschiedliche Methoden die besten Ergebnisse liefern. Ein gutes Data-Mining-Werkzeug muss daher ein breites Methodenspektrum anbieten. Der *STATISTICA Data Miner* erfüllt diese Voraussetzung perfekt.

## Methoden des *STATISTICA Data Miner*:

- Feature Selection zur Identifikation relevanter Prädiktoren
- Klassifikations- und Regressionsbäume (CART, CHAID)
- Interaktive Entscheidungsbäume, Boosted Trees und Random Forests
- Verallgemeinerte Additive Modelle
- MAR Splines
- EM- und k-Means-Clusterverfahren
- Independent-Component-Analysen
- Einfache und sequenzielle Assoziationsregeln
- Neuronale Netzverfahren einschließlich Self-Organizing Maps (SOM)
- K-Nearest Neighbors
- Support Vector Machines
- Bayessche Verfahren
- Alle einfachen und höheren statistischen Methoden (wie Regressions-, Diskriminanz- und Zeitreihenanalysen)
- Exzellente interaktive Visualisierungswerkzeuge und vieles mehr

## Das Zusatzprodukt *STATISTICA Process Optimization* bietet darüber hinaus:

- Ursachenanalysen zur Bestimmung relevanter Einflussfaktoren (mit Wechselwirkungen) und Analysemethoden
- Methoden (Simplex, Rastersuche, Genetische Algorithmen) zur simultanen Optimierung von Data-Mining-Modellen oder beliebiger Funktionen
- Multivariate und Prädiktive Qualitätsregelkarten

Die Datenvorbereitung ist der erste und zeitraubendste Schritt beim Data Mining. Hier bietet der *STATISTICA Data Miner* zahlreiche Verfahren zum Filtern, Selektieren und Transformieren der Daten.

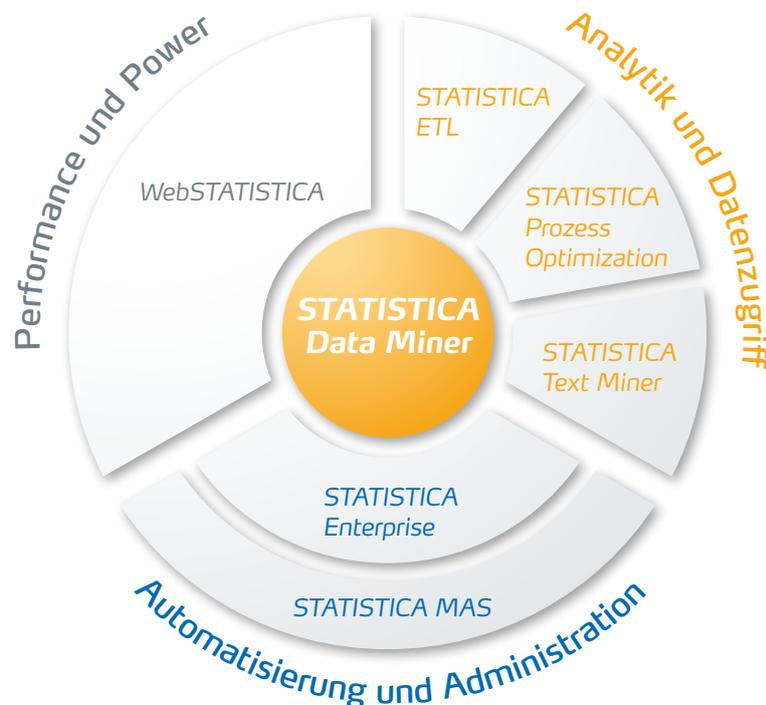


*„Dank der großen Auswahl an Data-Mining-Methoden in STATISTICA konnten wir ein geeignetes Verfahren für unsere Fragestellungen finden. Die Flexibilität von STATISTICA ermöglicht es uns, unsere Fertigungsprozesse schnell und effektiv abzubilden.“*

*Dr. Udo Kreißig,  
Leiter Technologieentwicklung für den Bereich Piezoaktor, Continental Automotive*

# Integration in die BI-Plattform von STATISTICA

Der *STATISTICA Data Miner* lässt sich entweder als Stand-Alone-Werkzeug betreiben oder nahtlos in die *STATISTICA* Business-Intelligence-Landschaft integrieren. Diese bietet folgende Komponenten:



## **STATISTICA Enterprise**

Mit dieser Komponente wird *STATISTICA* zu Ihrer unternehmensweiten Business-Intelligence-Plattform. Auf Knopfdruck werden vorgefertigte Analysen und Berichte erstellt und an die relevanten Konsumenten verteilt. Alle Ressourcen werden zentral gespeichert und verwaltet. Ein integriertes Dokumentenmanagement und Audit Trails erfüllen Ihre Compliance-Anforderungen.

## **STATISTICA MAS**

Mit dem *STATISTICA Monitoring and Alerting Server* lässt sich eine Vielzahl von Prozessen parallel auf leistungsstarken Servern im Hintergrund ausführen. Scorewerte und Analysen werden zeitgesteuert berechnet, Berichte generiert und Benachrichtigungen über relevante Ergebnisse automatisch versendet. Ein Dashboard informiert jederzeit über den Status der Analysen.

## **WebSTATISTICA**

Die Web-Lösung von *STATISTICA* bietet Client-Server-Power für Multiprozessorsysteme zur parallelen Verarbeitung hochkomplexer Berechnungen und kann über das Internet gesteuert werden.

## **STATISTICA ETL**

Die zu analysierenden Daten liegen oft in verschiedenen Datenquellen mit unterschiedlichen Strukturen. *STATISTICA ETL* extrahiert, transformiert und lädt Daten verschiedenster Quellen komfortabel und menügesteuert. Heterogene Datenquellen lassen sich nach analytischen Vorgaben vereinheitlichen und aggregieren. So stehen sie unmittelbar für eine weitergehende Analyse bereit.

## **STATISTICA Process Optimization**

Erweitert das Methodenspektrum des *STATISTICA Data Miner* um spezielle Methoden zur Ursachenanalyse und allgemeinen Optimierung.

## **STATISTICA Text Miner**

Eine Erweiterung des *STATISTICA Data Miner*, die unstrukturierte Textdaten aus unterschiedlichen Quellen einliest und in eine für die Datenanalyse geeignete Form umwandelt.

# Wenn es um Data Mining geht: StatSoft!

StatSoft zählt zu den weltweit führenden Herstellern von Datenanalyse-Software. Wir bieten Ihnen hohe Investitionssicherheit, kompetentes Engagement und exzellenten Service.

## ✓ Kontinuität und Internationalität

StatSoft hat sich seit seiner Gründung 1984 schnell als einer der weltweit führenden Hersteller von Datenanalyse-Software etabliert. Niederlassungen gibt es in über 20 Ländern, die Software ist in zehn Sprachen erhältlich.

## ✓ Moderne Software-Technologie

64-Bit-Architektur, Client-Server- und Multiprozessor-Lösungen, Web-Funktionalitäten, nahtlose Integration mit Datenbanken, STATISTICA Visual Basic, PMML-Unterstützung und interaktive Grafiktools sind nur einige Komponenten, die unsere technologische Führungsposition unterstreichen.

## ✓ Kundennähe

Das Feedback und die Wünsche unserer Kunden spielen von jeher eine zentrale Rolle bei der Entwicklung unserer Software. Diese werden regelmäßig analysiert und fließen in die Entwicklung ein, was uns eine einzigartige Praxisnähe sichert.

## ✓ Branchenübergreifende Kompetenz

Unsere Mitarbeiter sind hochqualifiziert und verfügen über fundiertes statistisches Know-how sowie langjährige Erfahrung in unterschiedlichsten Datenanalyseprojekten diverser Branchen.

## Software und Service



Viele Unternehmen sind sich der vielfältigen Möglichkeiten und Wettbewerbsvorteile durch Data Mining durchaus bewusst. Jedoch verfügen gerade mittelständische Unternehmen oft nicht über entsprechendes Know-how und die erforderlichen personellen Ressourcen. In Hinblick auf Ihre Data-Mining-Projekte bieten wir je nach vorhandenem Know-how und personellen Ressourcen folgende Optionen:

### Data Mining Training:

Data-Mining-Ressourcen sind in Ihrem Unternehmen vorhanden und Sie erwerben lediglich die Software sowie Training zur Toolbedienung und Methodik von uns.

### Data Mining Support:

Data-Mining-Know-how ist in Ihrem Unternehmen noch nicht vorhanden, soll aber – auch für künftige Projekte – erworben werden. Spezielle Trainings, Workshops und Consulting sorgen für den nötigen Wissenstransfer. Für künftige Projekte bieten wir auch die Überprüfung von Auswertungen auf statistisch korrektes Vorgehen an.

### Data Mining Turn-Key

Sie möchten Ihr Data-Mining-Projekt komplett outsourcen? In diesem Fall übernehmen wir das gesamte Data-Mining-Projekt für Sie.