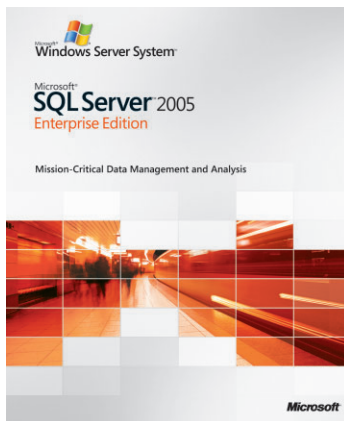


# Microsoft® SQL Server™ 2005

## Analysis Services – Data Mining

### Data Mining – die Plattform für intelligente Anwendungen



In den letzten zwei Jahrzehnten haben Organisationen eine enorme Menge an geschäftlichen Daten gesammelt. Über viele Daten zu verfügen bedeutet jedoch nicht, über umfangreiches geschäftliches Wissen zu verfügen. Data Mining ist der nächste Schritt im Bereich Business Intelligence. Data Mining beschäftigt sich mit der Erkundung von Daten, dem Erkennen von Mustern und der Anwendung dieser Muster auf das operative Geschäft.

Microsoft SQL Server 2005 Analysis Services etabliert eine neue Grundlage für das Data Mining. Durch eine einfach zu nutzende, erweiterbare, verfügbare und flexible Plattform machen die Data Mining-Möglichkeiten von SQL Server 2005 Analysis Services das Data Mining auch für die Unternehmen nutzbar, die sich bis jetzt noch keine Gedanken über eine entsprechende Lösung gemacht haben.

Microsoft SQL Server 2005 bietet viele neue Eigenschaften, mit deren Hilfe eine völlig neue Generation intelligenter Anwendungen geschaffen wird. Zu diesen neuen Eigenschaften zählen eine verbesserte, unternehmensfähige Architektur, eine tief greifende Verbindung mit den Business Intelligence-Features der SQL Server-Familie sowie umfangreiche Tools, APIs und Algorithmen. So können daten-

bezogene Lösungen entwickelt werden, die die Unternehmensproduktivität verbessern, den Gewinn erhöhen und die Kosten senken. Diese Lösungen lassen sich zum Beispiel auf die folgenden Geschäftsfelder anwenden:

- **Warenkorbanalyse**  
Feststellen von zusammen verkauften Waren und Analysen der Ergebnisse. Nutzen der Ergebnisse zur Erstellung eines Online-Empfehlungssystems oder um zu erkennen, welche Produktkombinationen zu Gewinnen beitragen.
- **Abwanderungsanalyse**  
Erstellen von Berichten bezüglich des Risikos von Kundenabwanderungen.
- **Marktanalyse**  
Definieren von Marktsegmenten durch die automatische Gruppierung von Kunden und Nutzung dieser Daten für Trendanalysen als Grundlage für Marketingkampagnen.
- **Prognose**  
Voraussage von Verkaufs- und Lagerdaten und Erkennen ihrer Beziehungen zueinander.
- **Datenerkundung**  
Verbessertes Verständnis der Kunden durch Abfragen von relevanten Datenmustern. Vergleichen der Unterschiede zwischen gewinnbringenden und weniger gewinnbringenden Kunden oder von Kunden, die unterschiedliche Marken des gleichen Produktes bevorzugen.
- **Websiteanalyse**  
Erkennen, wie die Unternehmenswebsite von Kunden genutzt wird. Gruppierung von gleichen Nutzungsmustern.
- **Kampagnenanalysen**  
Effizientere Ausrichtung der Marketingausgaben durch eine bessere Ausrichtung an den Zielgruppen.
- **Datenqualität**  
Erkennen von fehlenden Werten bzw. Anomalien in den Daten bei deren Eingabe oder Implementierung in das Data-Warehouse.
- **Textanalyse**  
Analyse von Feedback, um häufig auftretende Themen und Trendanalysen zu finden, die Kunden oder Mitarbeiter beschäftigen.

### Integration

SQL Server Data Mining ist in eine Familie von Business Intelligence-Technologien integriert, die zusammen zur Erweiterung und Entwicklung einer neuen Generation von intelligenten Anwendungen genutzt werden können.

- **SQL Server 2005 Integration Services**  
Die durch die Arbeit mit SQL Server 2005 verbesserte Datenverbindung ermöglicht Ihrer Organisation, „Ausreißer“ zu erkennen, Daten zu separieren und fehlende Werte auf Basis von Prognoseanalysen aufzufüllen.
- **SQL Server 2005 Analysis Services (OLAP)**  
Mit der neuen Data Mining-Funktion, die Ihre Daten nach inherenten Mustern aufteilt, schaffen Sie ein umfassendes Unified Dimensional Model.
- **SQL Server 2005 Reporting Services**  
Übersichtlichere, aufschlussreiche Berichte, mit den richtigen Informationen für alle Mitarbeiter oder einfache Selektion einzelner, dedizierter Empfänger.

### Einfache Nutzung

Mit dem standardmäßigen Data Mining-Wizard und -Designer können hoch entwickelte Modelle mit nur wenigen Mausklicks erstellt werden. Das SQL Server Data Mining-Toolset ist direkt in Microsoft Visual Studio® integriert und ermöglicht die Anzeige und Bearbeitung von Daten und das Design und die Bearbeitung von Modellen. SQL Server Data Mining stellt **über ein Dutzend interaktiver Visualisierungen** zur Verfügung, die beim Verständnis der vom Data Mining zu erkennenden Muster helfen. Zusätzlich gibt es Diagramme, mit denen die Qualität der Modelle vor der Bereitstellung verglichen und gegenübergestellt werden können.

### Einfache und umfangreiche API

Bei der Anwendung der Modelle beginnt mit SQL Server ein neues Kapitel im Bereich Data Mining.

### **Data Mining Extensions for SQL (DMX)**

macht es Entwicklern und Datenbank-Administratoren einfach, Anwendungen mit Data Mining-Funktionen zu erstellen. Eine Prognose gegen ein Data Mining-Modell ist einfach nur ein Join in einer bekannten SQL-Abfrage. Zum ersten Mal haben die Personen, die für die Erstellung von Anwendungen und das Handling von Daten verantwortlich sind, die Möglichkeit, Data Mining-Technologien mit bereits bekannten Tools zu nutzen.

### **Algorithmen**

Aufgrund der gemeinsamen Entwicklung durch Microsoft Research und das SQL Server-Produktteam beinhaltet SQL Server Data Mining aktuellste Algorithmen (hierzu gehören zum Beispiel Decision- und Regression-Trees, Time-Series, Clustering und Sequence-Clustering, Association-Rules, Naïve Bayes, Neuronale Netze und Text-Mining).

### **Erweiterbarkeit**

SQL Server Data Mining ist über .NET™-Prozeduren und **Plug-in-Algorithmen und -Viewer** vollständig erweiterbar. Die Adaption von SQL Server Data Mining als Plattform bedeutet, dass die bereits mitgelieferten Funktionen des Data Mining Systems jederzeit entsprechend den Bedürfnissen angepasst und erweitert werden können.

### **Architektur**

Durch Data Mining entstehen für Organisationen jeder Größe neue Herausforderungen. Bereitstellung, Skalierbarkeit und Sicherheit werden zu wichtigen Faktoren. SQL Server Data Mining ist ein Teil der SQL Server Analysis Services, die sich auszeichnen durch:

- **Bereitstellung**

SQL Server Data Mining basiert auf einer Client-Server-Architektur, mit der ein Zugriff auf Modelle über LAN, WAN oder das Internet möglich ist. Standard-APIs wie OLE DB, ADOMD.Net und sogar XML for Analysis (XML/A) stellen unabhängig vom Standort oder von der Client-Plattform einen Zugriff zur Verfügung. Einmal verarbeitet sind die Modelle durch das Transaktionssystem des Servers immer verfügbar – auch während Updates.

- **Skalierbarkeit**

SQL Server Data Mining ist von Grund auf mit einer parallelen Architektur entworfen worden. So sind umfangreiche Datenmengen und Tausende von gleichzeitigen Benutzern mit Millionen von Abfragen pro Tag möglich.

- **Verwaltbarkeit**

SQL Server Data Mining ist in das neue SQL Management Studio integriert und stellt so einen zentralen Punkt zur Verwaltung aller Eigenschaften der SQL Server-Familie zur Verfügung.

- **Sicherheit**

SQL Server Data Mining bietet eine rollenbasierte Sicherheit und stellt so den Schutz von geistigem Eigentum sicher.