

您的潜力，我们的动力

**Microsoft**  
微软(中国)有限公司

# 使用SQL Server 2005 Data Mining 构建智能应用

徐晓卓

Microsoft MVP

萌微科技有限公司首席顾问

# 本次课程内容包括

- SQL Server 2005商务智能概述
- 什么是数据挖掘
- 在SQL Server 2005中使用数据挖掘
- 示例

您的潜力. 我们的动力

**Microsoft**  
微软(中国)有限公司

# 收听本次课程需具备的条件

- 了解SQL Server 2005

## Level 300

# 如何.....

您的潜力. 我们的动力

**Microsoft**  
微软(中国)有限公司

- 从企业数据中获取信息
- 使商务智能成为企业业务平台的一部分
- 获取数据并为数据建立模型
- 与商务流程集成
- 通过单一的维度模型同时实现关系数据报表和OLAP

您的潜力. 我们的动力

**Microsoft**  
微软(中国)有限公司

# 定义商务智能

## 支持平台

- 多维分析（比如OLAP）
- 数据挖掘
- 预测
- 业务分析
- 查询、报告和图表
- 知识管理
- 企业门户
- 数字仪表板（Digital Dashboard）
- 数据仓库
- .....

您的潜力. 我们的动力

**Microsoft**  
微软(中国)有限公司

# 商务智能面临的挑战

- 众多的数据模型
- 众多的数据源
- 众多的编程接口
- 数据的冗余

# SQL Server 2005 商务智能概述

完整的商务智能平台

您的潜力, 我们的动力

**Microsoft**  
微软(中国)有限公司

- 完整的商务智能 (BI) 平台
  - 为用户提供了可用于构建典型和创新的分析应用程序所需的各种特性、工具和功能
- 构建分析程序
  - 新增功能使复杂 BI 系统的构建和管理比以往更加轻松

# 什么是数据挖掘

使用数据挖掘解决问题

- 消费者将购买什么产品？哪些产品会一起销售？
- 标识将流失的消费者
- 市场状况如何，将会如何发展？
- 分析网站
- 确定营销活动是否成功
- 文本分析

# 什么是数据挖掘

## 同传统统计分析的比较

---

### 统计分析

---

- 统计分析需要先有假设
- 统计分析需要根据假设列出方程
- 统计分析仅能分析数字
- 统计分析的结果可以直接表述

---

### 数据挖掘

---

- 数据挖掘不依赖于任何假设
- 数据挖掘算法能够自动列出方程
- 数据挖掘能够处理多种数据, 包括声音、文本等
- 数据挖掘的结果需要经过统计分析后才能形成直观的可表述的内容, 比如图表

# 什么是数据挖掘

## 同传统统计分析的比较（举例）

---

### 统计分析

---

- 市场人员想根据不同顾客的消费行为设计不同的市场活动。传统分析的做法会首先将顾客分成“学生”、“老师”.....而后用已经区分完成的类别进行行为分析。如果采用这样的分类得到的结果并不令人满意，则换一种分类方法，如“10-20岁”、“20-25岁”.....直到满意为止。

### 数据挖掘

---

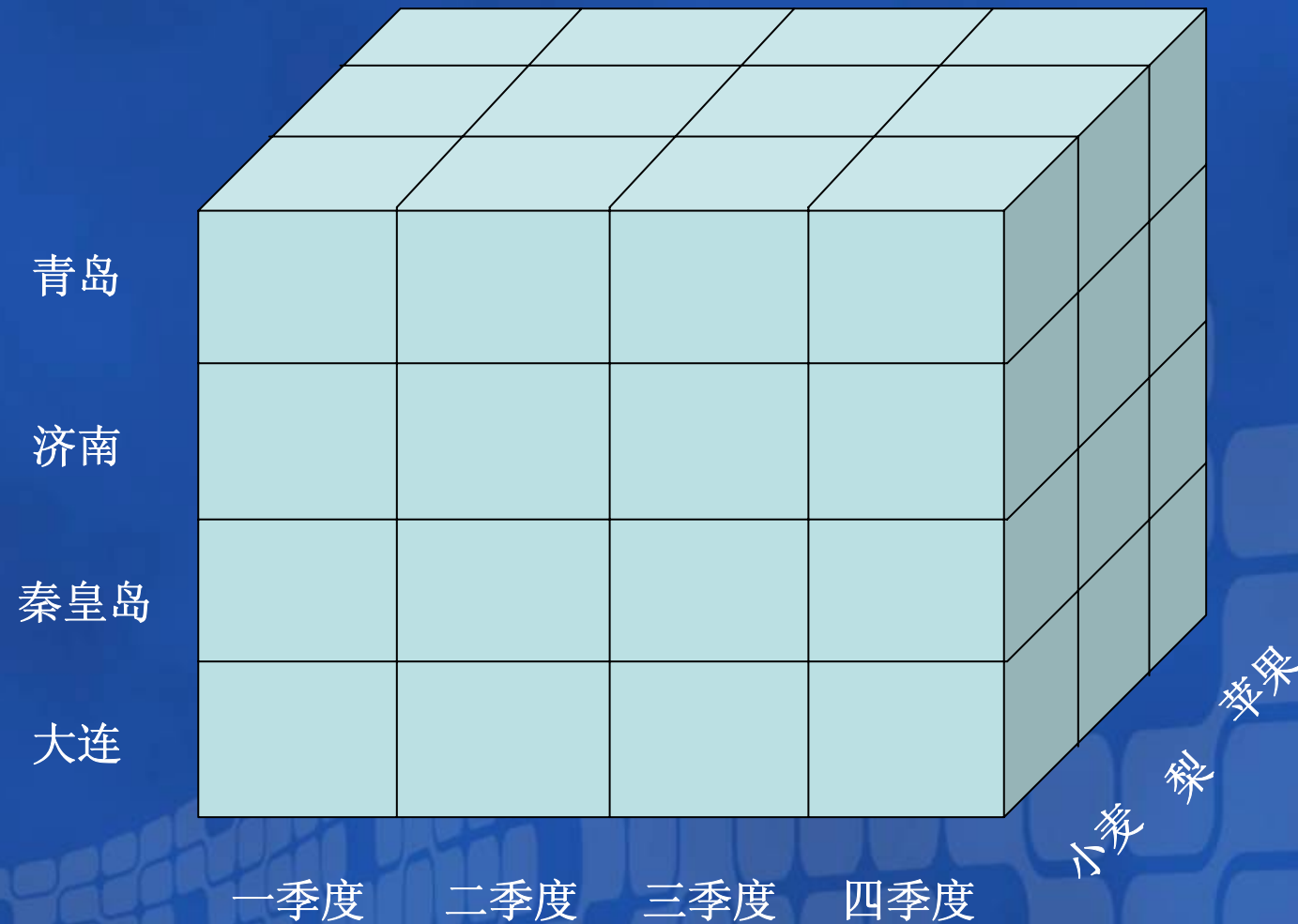
- 数据挖掘工具将对全部的数据进行“学习”，发现顾客的不同行为模式并按照该模式给顾客分类从而得到最适当的分组方法。

# 数据挖掘的基本概念

## Cube

您的潜力. 我们的动力

**Microsoft**  
微软(中国)有限公司



# 数据挖掘的基本概念

## Cube

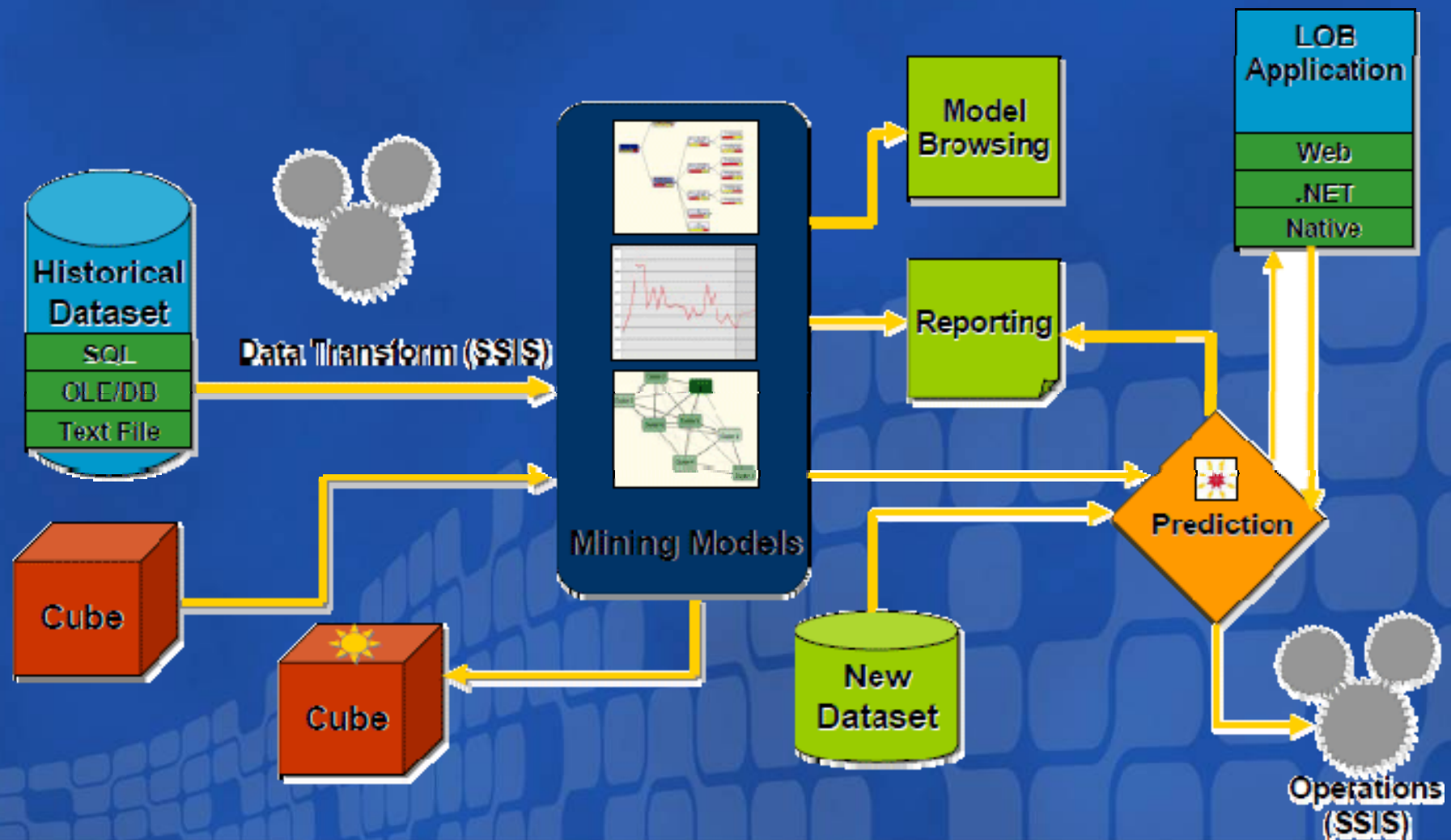
- 没有个数限制（仅仅受可创建对象上限限制）
- 使用XML进行存储
- 维（Dimensions）
- 度量（Measures）
- 属性（Attributes）
- 层次（Hierarchies）

# SQL Server 2005中的数据挖掘

架构

您的潜力，我们的动力

**Microsoft**  
微软(中国)有限公司

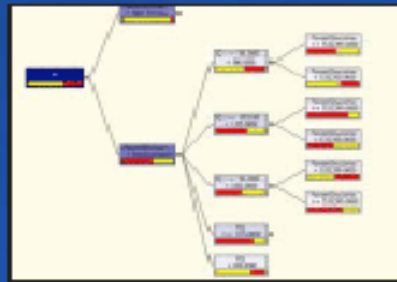


# SQL Server 2005中的数据挖掘

算法

您的潜力，我们的动力

**Microsoft®**  
微软(中国)有限公司

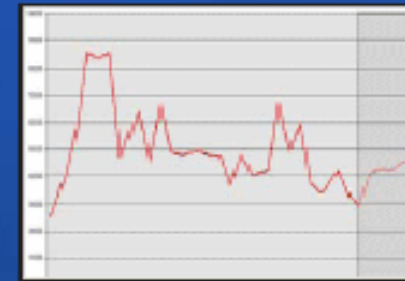


**Decision Trees**

*Introduced in SQL Server 2000*



**Clustering**



**Time Series**



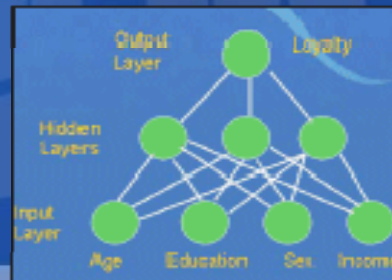
**Sequence Clustering**



**Association**

Attribute	Value	From: Police and Editor	From: Sales
Education Year	5-9		
Education Year	0-4		
Education Year	10+		
value: info:LINE-40: THE 101	Missing		
value: info:LINE-40: THE 101	Missing		
value: info:8: THE WORLD TURN	Missing		
value: info:8: THE WORLD TURN	Missing		

**Naïve Bayes**



**Neural Net:**

**MSDN Webcasts**

您的潜力，我们的动力

**Microsoft**  
微软(中国)有限公司

# DEMO

## 在SQL Server 2005中构建数据 挖掘程序

# 本次课程内容回顾

- SQL Server 2005商务智能概述
- 什么是数据挖掘
- 在SQL Server 2005中使用数据挖掘
- 示例

# 知道更多有关**DMX**的知识


- **SQL Server 2005**中文网站  
<http://www.microsoft.com/china/sql/2005/>
- 下载**SQL Server 2005**联机丛书  
<http://www.microsoft.com/downloads/details.aspx?FamilyID=be6a2c5d-00df-4220-b133-29c1e0b6585f&DisplayLang=en>
- 联系讲师  
[xiaozhuo\\_xu@msn.com](mailto:xiaozhuo_xu@msn.com)

# 获取更多MSDN资源


- **MSDN中文网站**  
<http://www.microsoft.com/china/msdn>
- **MSDN中文网络广播**  
<http://www.msdnwebcast.com.cn>
- **MSDN Flash**  
<http://www.microsoft.com/china/newsletter/case/msdn.aspx>
- **MSDN开发中心**  
<http://www.microsoft.com/china/msdn/DeveloperCenter/default.mspx>

# Question & Answer

如需提出问题，请单击“提问”按钮并在随后显示的浮动面板中输入问题内容。一旦完成问题输入后，请单击“提问”按钮。

 **问题和解答 (无问题)** ▲ ×

在此会议中尚未解答任何问题。

要向演示者提问，请在此处键入问 

提问(A)

删除(D)

问题管理器(Q)

您的潜力，我们的动力

**Microsoft®**  
微软(中国)有限公司

**Microsoft®**

msdn  


**MSDN Webcasts**