一【2】.上机目标及内容

目标:

1. 懂得数据发掘的根本概念及其进程;

2. 懂得数据发掘与数据仓库. OLAP 之间的关系

3. 懂得根本的数据发掘技巧与办法的工作道理与进程,控制数据发掘相干对象的运用. 内容:

将创建一个数据发掘模子以练习发卖数据,并运用"Microsoft决议计划树"算法在客户群中找 出购置自行车模式.请将要发掘的维度(事例维度)设置为客户,再将客户的属性设置为数据发掘算 法辨认模式时要运用的信息.然后算法将运用决议计划树从中肯定模式.下一步须要练习模子,以便可 以或许浏览树视图并从中读取模式.市场部将依据这些模式选择潜在的客户发送自行车促销信息. 请求:

运用试验室和指点教师供给的试验软件,卖力完成划定的试验内容,真实地记载试验中碰到的各类问题息争决的办法与进程,并依据试验案例绘出模子及操作进程.试验完成后,应依据试验情形写出试验报告.

二.试验道理及根本技巧路线图(方框道理图或程序流程图)

联系关系剖析: 联系关系剖析是从数据库中发明常识的一类主要办法.

时序模式: 经由过程时光序列搜刮出反复产生概率较高的模式.

分类:分类是在聚类的基本上对已肯定的类找出该类别的概念描写,代表了这类数据的整体信息,既该类的内在描写,一般用规矩或决议计划树范式表示.

三. 所用仪器. 材料(装备名称. 型号. 规格等或运用软件)

1台PC及Microsoft SQL Server 套件

第1页,一共26页

四. 试验办法. 步骤(或: 程序代码或操作进程)及试验进程原始记载(测试数据. 图表. 盘 算等)

创建 Analysis Services 项目

- 打开 Business Intelligence Development Studio. 1.
- 在"文件"菜单上,指向"新建",然后选择"项目". 2.
- 确保已选中 模板"窗格中的"Analysis Services项目". 3.
- 在 "名称"框中,将新项目定名为 AdventureWorks. 4.
- 单击"肯定". 5.

项目类型 (£):		模板 (I).
商业智能	项目 类型	Visual Studio 已安装的模板 Analysis Services 项目 Integration Services 项目 服表服务器项目向导 服表模型项目 我的模板 證 搜索联机模板
创建新的 Ana 名称 00 :	lysis Services AdventureW	项目 orks

■更改存储数据发掘对象的实例

- 在 Business Intelligence Development Studio 的 项目"菜单中,选择 属性". 1.

在 属性页"窗格的左侧,单击 部署". 2.

在'目标"选项部分,验证数据库名称是否为 localhost.假如运用的是其他实例,请键入该实例的 3.

名称.单击"肯定".

第2页,一共26页

- 配寄属性		
一内部版本	服务器	lecalhost
■ 配置原性 - 内部版本 - 调试 - 問言	数据库	AdventureTorks
	□ 洗項	
	处理)先项	默认值
	事务部署	False
	部署模式	仅部署更改
	服务器	1727

创建数据源

1. 在解决计划资本治理器中,右键单击 数据源"文件夹,然后选择"新建数据

欢迎使用数据源向导
使用此向导创建新数据源。 数据源表示到数据的连接。 数据源不提供缓存元数据、添加关系、添加计算和添加批注等功能。若要对数据 源应用这些功能,请使用此向导来创建数据源,然后使用数据源视图向导来创建 包含相应功能的视图。

体系将打开数据源领导.

第3页, 一共26页

2. 在 迎接运用数据源领导"页面中,单击"下一步"按

数据连接 (E):	iteat (<u>C</u>)	数据连接属性 (T):	
127.0.0.1.AdventureWorksDW		属性	值	1.
		Initial Ca Integrated Provider	AdventureWorksDW SSFI SQLOLEDB. 1	

3. 在选择若何界说衔接"页上,单击"新建"向 Adventure Works 数据库中添加衔接.

体系将打开 衔接治理器"对话

2	教援源向导
	莫拟信息 可以定义 Analysis Services 使用何种 Windows 凭据来连接到数据源。
	⊙ 使用特定 Windows 用户名和密码 (≦)
	用户名创: 3a
	密码 (2):
	○1使用服务帐户(2)
Let	○使用当前用户的凭据 (2)
框.	

- 4. 在 衔接治理器"的 供给程序"列表中,选择 "本机 OLE DB\Microsoft OLE DB Provider for SQL Server".
- 5. 在 "办事器名称"列表中,键入或选择承载 AdventureWorksDW 的办事器的名称.
- 6. 在 登录到办事器"组中,选择身份验证办法,并输入凭证.
- 7. 在选择或输入数据库名称"列表中,选择 AdventureWorksDW, 再单击"肯定"按钮.

第4页,一共26页

- 8. 单击"下一步"按钮进入领导的下一页.
- 9. 在 模仿信息"页中,选择 运用办事帐户",再单击"下一步".
- 10. 请留意,在 完成领导"页中,数据源名称默以为 Adventure Works

	数据源向号
	完成向导 话提供一个名称,然后单击"完成"以创建新数据源。
	数据源名称 @):
	Adventure Works DW
	预览 (2):
	连接字符串: Provider=SQLOLEDB.1;Data Source=127.0.0.1;Integrated Security=SSPI;Initial Catalog=AdventureWorksDW
DW.	

11. 单击 完成".

新的数据源 Adventure Works DW 将显示在解决计划资本治理器的 数据源"文件夹中.



创建数据源视图

1. 在解决计划资本治理器中,右键单击 数据源视图",选择 新建数据源视图".

体系将打开数据源视图领导.



第5页, 一共26页

- 2. 在 迎接运用数据源视图领导"页上,单击"下一步".
- 3. 在选择数据源"页的"关系数据源"下,体系将默认选中您在上一个义务中创建的 Adventure Works DW 数据源. 单击"下一步".

若要创建新数据源,请单击"新建数据源",启动数据源领导.

华择教准通	
选择现有的关系数据源,或新建一	个关系数据源。
关系数据源(医):	数据源属性(2):
关系数据源(B): Adventure Works DW	数据源属性 (2): 属性 值
关系数据源(图): Adventure Works DW	数据源属性 (2): 属性 值 Data Source 127.0.0.1
关系数据源(图): Adventure Works DW	数据源属性 (2): 属性 值 Data Source 127.0.0.1 Initial Ca AdventureWorksDW
关系数据源(E): Adventure Works DW	数据源属性 (2): 属性 值 Data Source 127.0.0.1 Initial Ca AdventureWorksDW Integrated SSPI

- 4. 在"选择表和视图"页上,选择下列各表,然后单击右箭头键,将这些表包括在新数据源视图中:
 - dbo. ProspectiveBuyer
 - dbo.vAssocSeqLineItems
 - dbo.vAssocSeqOrders
 - dbo.vTargetMail
 - dbo.vTimeSeries

第6页, 一共26页

选择表和视图 从要包含在数据源视图中的关系数据库中选择对象。



可用对象(A):

包含的对象(I):

名称	类型	~		名称	类型
🔢 DimReseller (dbo)	表			TrospectiveBuyer (dbo)	表
🔝 DimSalesReason (dbo)	表			🔄 vAssocSeqLineItems (dbo)	视图
🔟 DimSalesTerritory (dbo)	表			vAssocSeqOrders (dbo)	视图
🗾 DimScenario (dbo)	表			🔄 vTargetMail (dbo)	视图
🗾 DimTime (dbo)	表		<	vTimeSeries (dbo)	视图
🔢 FactCurrencyRate (dbo)	表			-	
🔢 FactFinance (dbo)	表				
🔢 FactInternetSales (dbo)	表		[]		
🔟 FactInternetSalesReaso	表	(E)	\rightarrow		
🔢 FactResellerSales (dbo)	表				
🔟 FactSalesQuota (dbo)	表				
🛒 vDMPrep (dbo)	视图		55		

5. 单击"下一



步".

在 完成领导"页上,默认情形下,体系将数据源视图定名为 Adventure Works DW. 单击 完成". 6.

体系将打开数据源视图设计器,显示 Adventure Works DW 数据源视图.

第7页,一共26页



创建用于目标邮件计划的发掘构造

1. 在解决计划资本治理器中,右键单击"发掘构造"并选择"新建发掘构造"启动数据发掘领导.



- 2. 在 迎接运用数据发掘领导"页上,单击"下一步".
- 3. 在"选择界说办法"页上,确保已选中"从现有关系数据库或数据仓库",再单击"下一步".

《数据按据同导	
选择定义方法 选择创建挖掘结构定义时要使用的方法。	
使用何种方法定义挖掘结构?	
④(从现有关系数据库或数据仓库页)	
○ 从现有多维数据集 (2)	
说明:	

第8页,一共26页

4. 在 创建数据发掘构造"页的 您要运用何种数据发掘技巧?"下,选择"Microsoft决议计划



5. 单击"下一步".

致语活症回守 选择数据源视图 选择为挖掘结构提供数据的数据源视图。	
可用数据源视图(A): Adventure Works DW	□ 表: ProspectiveBuyer vAssocSeqLineItems vAssocSeqOrders vTargetMail vTimeSeries

6. 在 选择数据源视图"页上,请留意已默认选中 Adventure Works DW. 在数据源视图中,单击 浏 览"查看各表,然后单击 對闭"返回该领导.

第9页, 一共26页

WEAdventure Works DW	🔚 🗖 🔀
3宿前(2) +	
vTime vTargetMail ■	Prospectiv
	关闭 (二)

- 7. 单击"下一步".
- 8. 在 指定表类型"页上,选中 vTargetMail 表旁边 "事例"列中的复选框,再单击"下一步".

指分	2 表 类 型 指定分析时要使用的表类型。		l
前入事	長(江):		
	表	爭例	嵌套
	ProspectiveBuyer		
	vAssocSeqLineItems		
	VA ssocSeqOrders		
5	vTargatMail	~ .	

9. 在"指定定型数据"页上,确保已选中 CustomerKey 列旁边 Key 列中的复选框.

假如数据源视图中的源表表示一个键,则数据发掘领导将主动选择该列作为模子的键.

第10页,一共26页

指是苏	, 府中所用的列。			
田桓型结构	ធ (ភ្លា) :			
	表/列	键	输入	🗌 可预测
	CommuteDistance			
	CustomerAlternateKey			
M 20	CustomerKey	-		
	DateFirstPurchase			
	EmailAddress			
	EnglishEducation			
	EnglishOccupation			
_	and the second			

10. 选中 BikeBuyer 列旁边的 输入"和 '可猜测".

指定定型数 指定分析	汉据 中所用的列。			C
挖掘模型结构	©): 表/列	键	■ 输入	■ 可预测
	Age			
V 🖏	BikeBuyer			
	BirthDate			
(C)	C 1. D. 1	E CONTRACTOR OF CONTRACTOR OFO		

11. 单击 建议"打开"供给相干列建议"对话框.

只要选中至少一个可猜测属性,即可启用 建议"按钮. 供给相干列建议"对话框将列出与可猜测列联 系关系最亲密的列,并按照与可猜测属性的互相关系对属性进行排序.值大于 0.05 的列将被主动选 中,以包括在模子中.

12. 浏览建议,然后单击 撤消"疏忽建议并保留领导设置的原始值.

13. 选中以下各列旁边的 输入"复选框:

Age ٠

第11页,一共26页

- CommuteDistance
- EnglishEducation
- EnglishOccupation
- FirstName
- Gender
- GeographyKey
- HouseOwnerFlag
- LastName
- MaritalStatus
- NumberCarsOwned
- NumberChildrenAtHome
- Region
- TotalChildren
- YearlyIncome

第12页,一共26页

	表/列	键	■ 输入	
	FrenchOccupation			[
V 🐖	Gender		~	[
 • 	GeographyKey			[
🗹 🖅	HouseOwnerFlag			[
9 🖅	LastName		~	[
V 🖅	MaritalStatus		V	[
	MiddleName			[
	NameStyle			[
V 🖅	NumberCarsOwned			[
-	NumberChildrenAtHome			[

挖掘模型结构(S):

	表/列	键	■ 输入	
V 41	NumberCarsOwned			
🗹 📲	NumberChildrenAtHome		~	
	Phone			
🗹 🖅	Region		~	
	SpanishEducation			
	SpanishOccupation			
	Suffix			
	Title			
V 🖅	TotalChildren		~	
	YearlyIncome			

14. 单击"下一步".

模型	2結构(2):	
	列。内容类型	数据类型
	Age Continuou	s Long
=	Bike Buyer Continuou	s Long
E	Commute Distance Discrete	Test
2	Customer Key Key	Long
Ξ	English Education Discrete	Text
Ξ	English Occupation Discrete	Text
Ξ	First Name Discrete	Test
=	Gender Discrete	Text
2	Geography Key Continuou	s Long
	House Owner Flag Discrete	Text
E	Last Name Discrete	Text

第13页,一共26页

以上内容仅为本文档的试下载部分,为可阅读页数的一半内容。如 要下载或阅读全文,请访问: <u>https://d.book118.com/60514334321</u> <u>3012001</u>