

## 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.

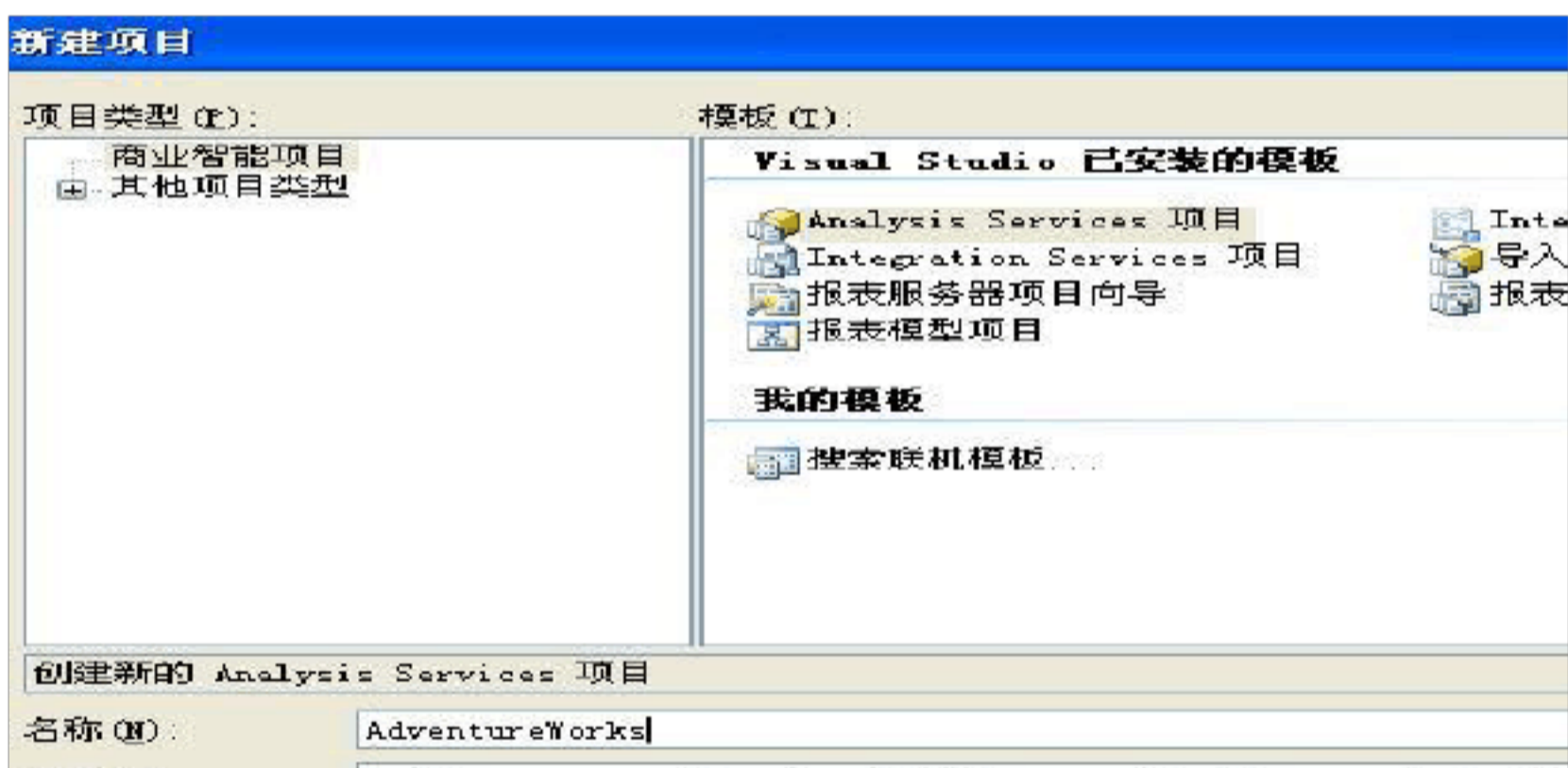
### 三. 所用仪器. 材料(装备名称. 型号. 规格等或运用软件)

1 台 PC 及 Microsoft SQL Server 套件

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）

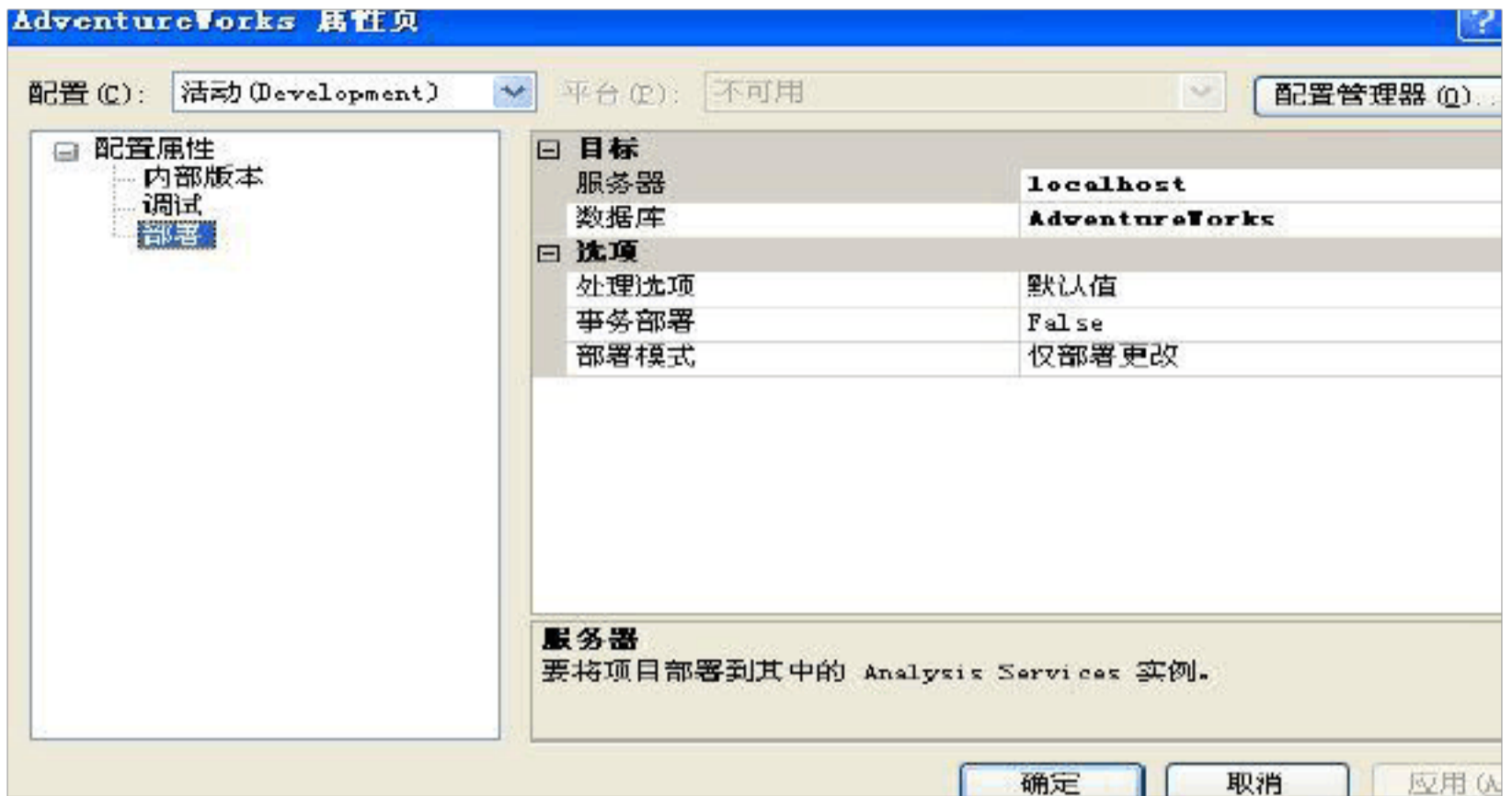
### 创建 Analysis Services 项目

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



### 更改存储数据发掘对象的实例

1. 在 Business Intelligence Development Studio 的“项目”菜单中, 选择“属性”.
2. 在“属性页”窗格的左侧, 单击“部署”.
3. 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, 请键入该实例的名称. 单击“肯定”.



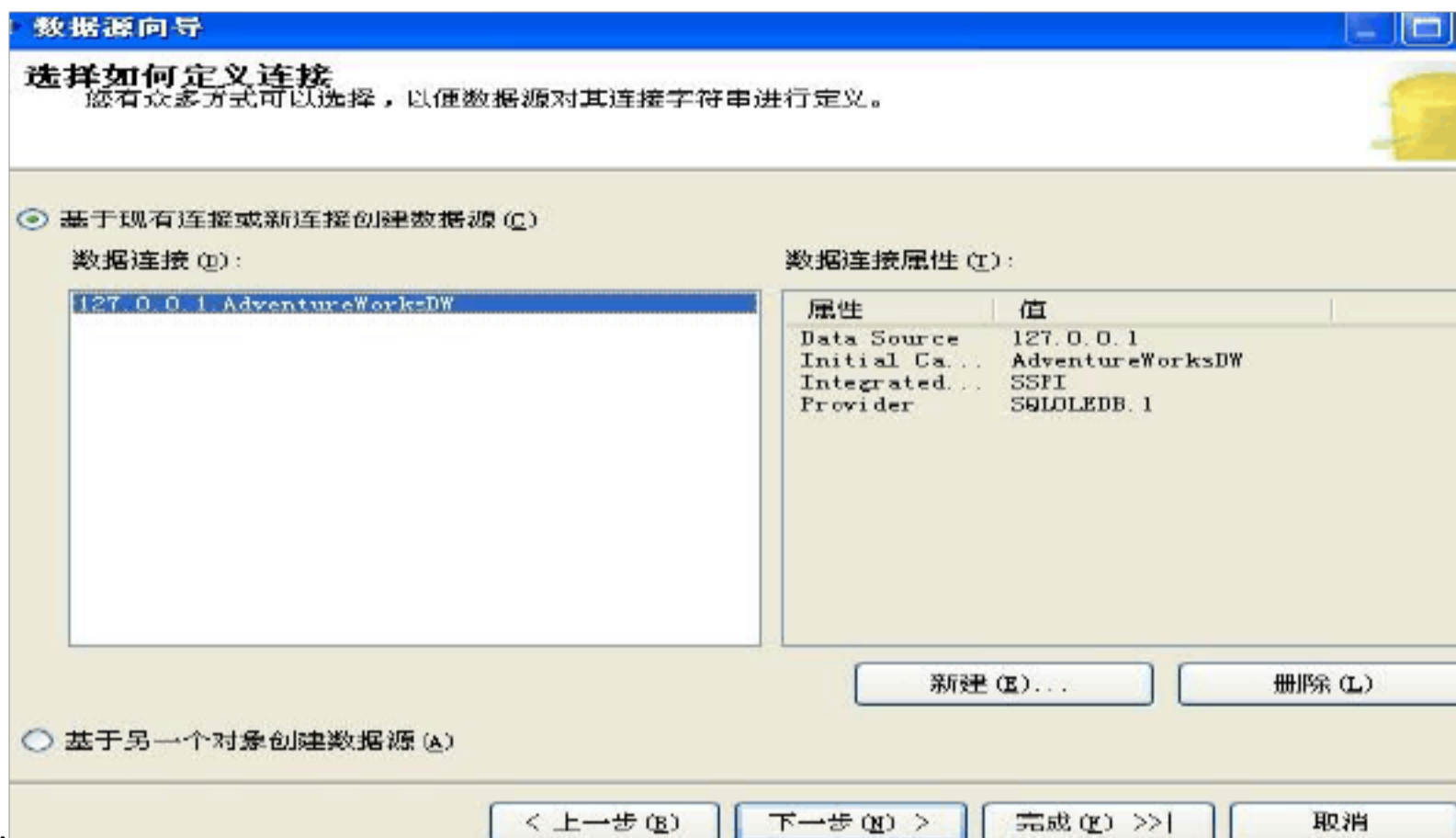
## 创建数据源

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



体系将打开数据源领导.

2. 在“迎接运用数据源领导”页面中,单击“下一步”按钮。



3. 在“选择若何界说衔接”页上,单击“新建”向 Adventure Works 数据库中添加入衔接。

体系将打开“衔接治理器”对话框。



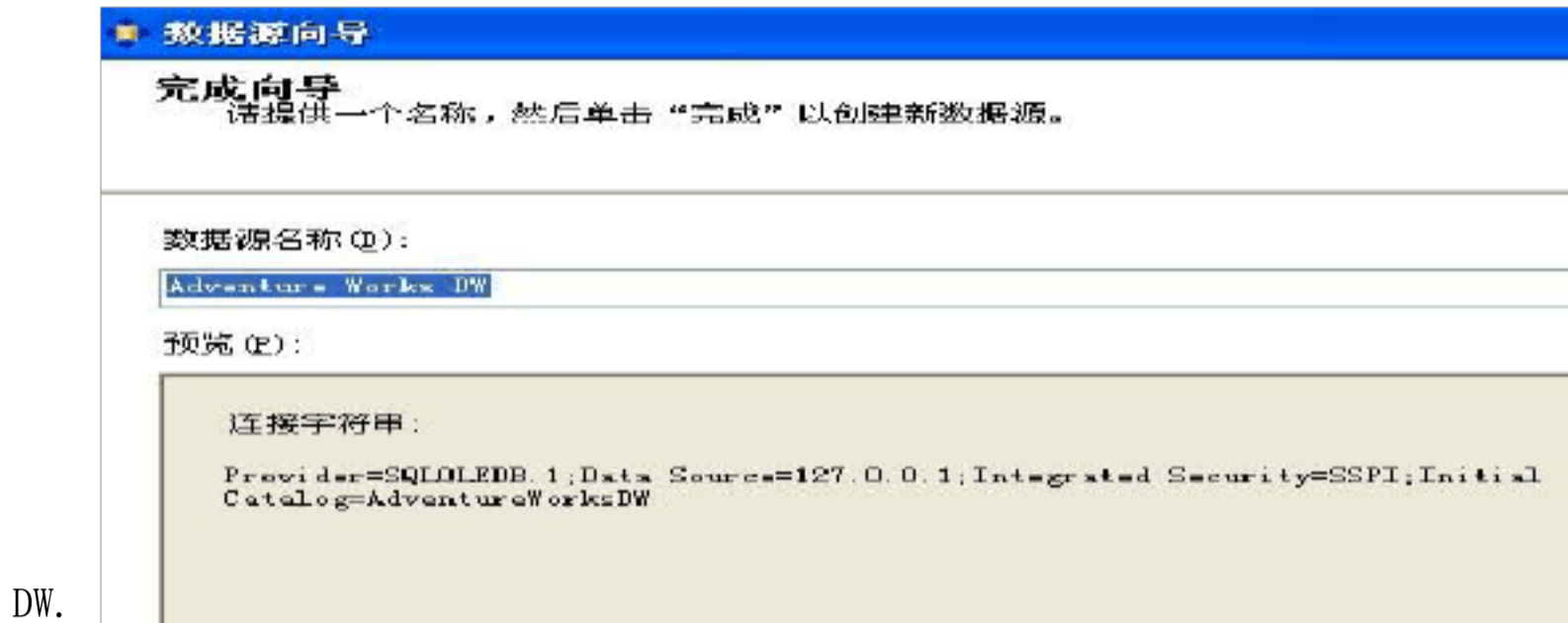
4. 在“衔接治理器”的“供给程序”列表中,选择“本机 OLE DB\Microsoft OLE DB Provider for SQL Server”。

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。

- 单击“下一步”按钮进入领导的下一页。
- 在“模仿信息”页中,选择“运用办事帐户”,再单击“下一步”。
- 请注意,在“完成领导”页中,数据源名称默以为 Adventure Works



- 单击“完成”。

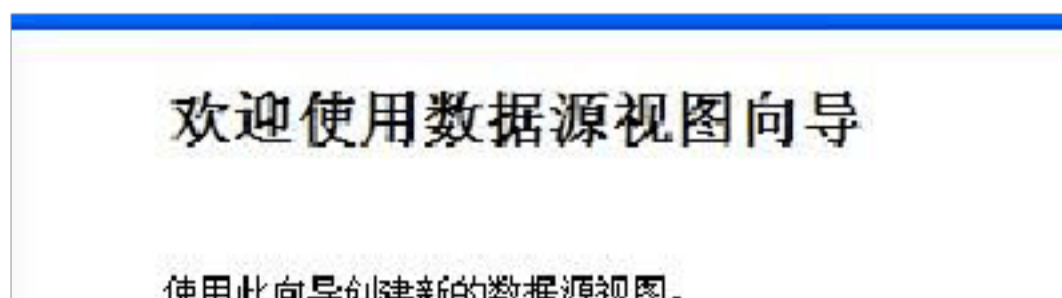
新的数据源 Adventure Works DW 将显示在解决计划资本治理器的“数据源”文件夹中。



## 创建数据源视图

- 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。



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”。

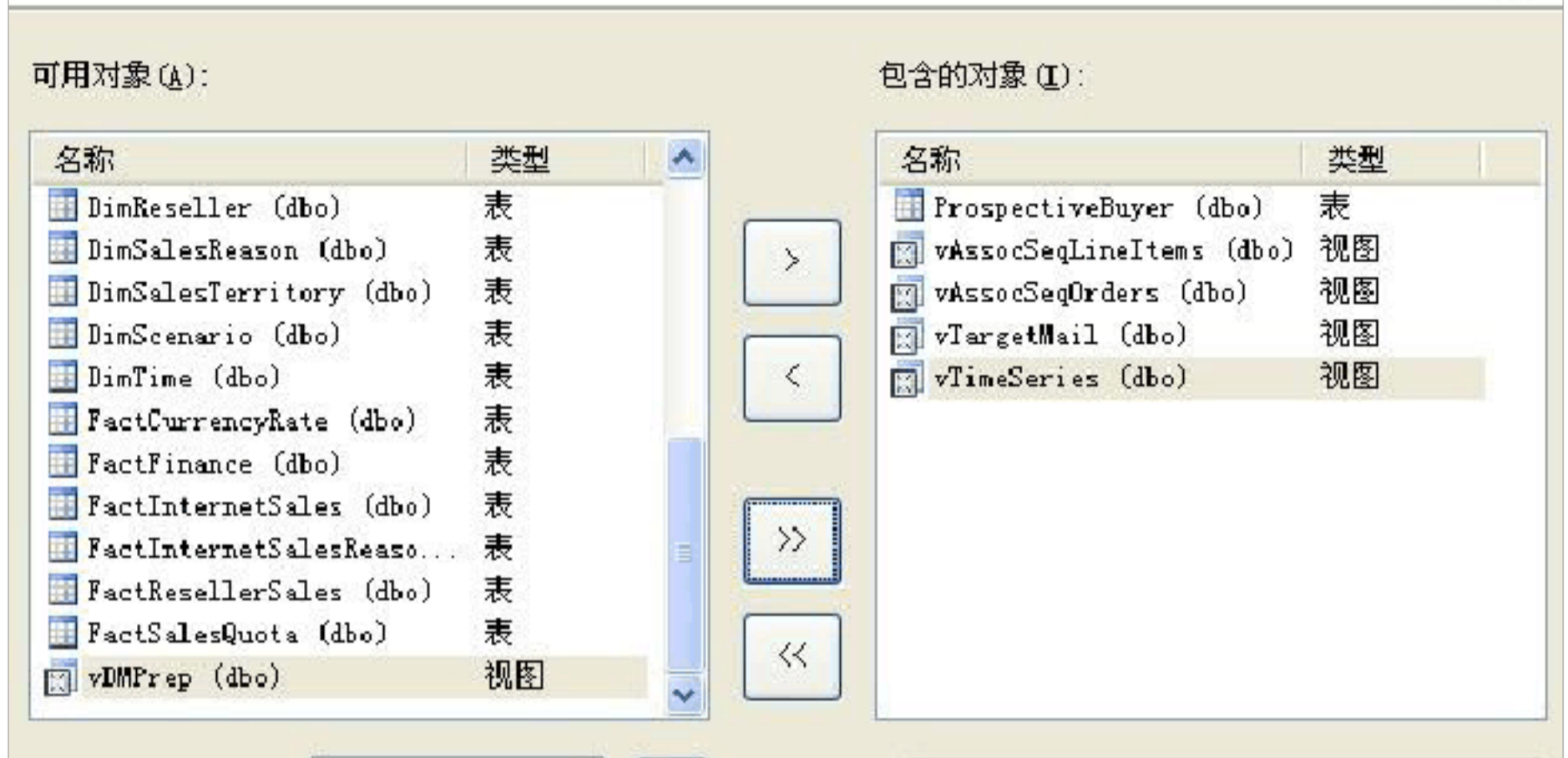
若要创建新数据源,请单击“新建数据源”,启动数据源领导。



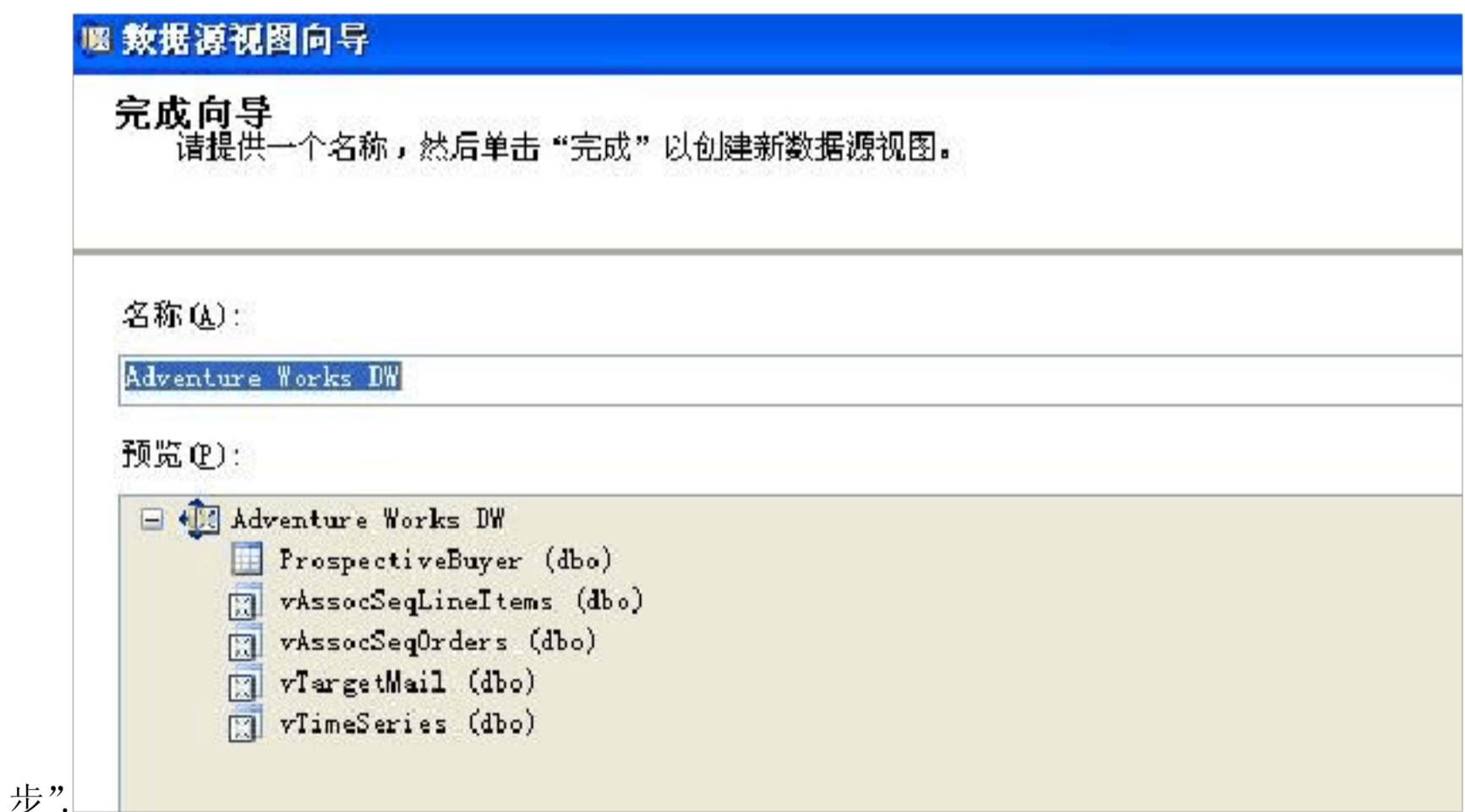
4. 在“选择表和视图”页上,选择下列各表,然后单击右箭头键,将这些表包括在新数据源视图中:
  - dbo.ProspectiveBuyer
  - dbo.vAssocSeqLineItems
  - dbo.vAssocSeqOrders
  - dbo.vTargetMail
  - dbo.vTimeSeries

## 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。



5. 单击“下一



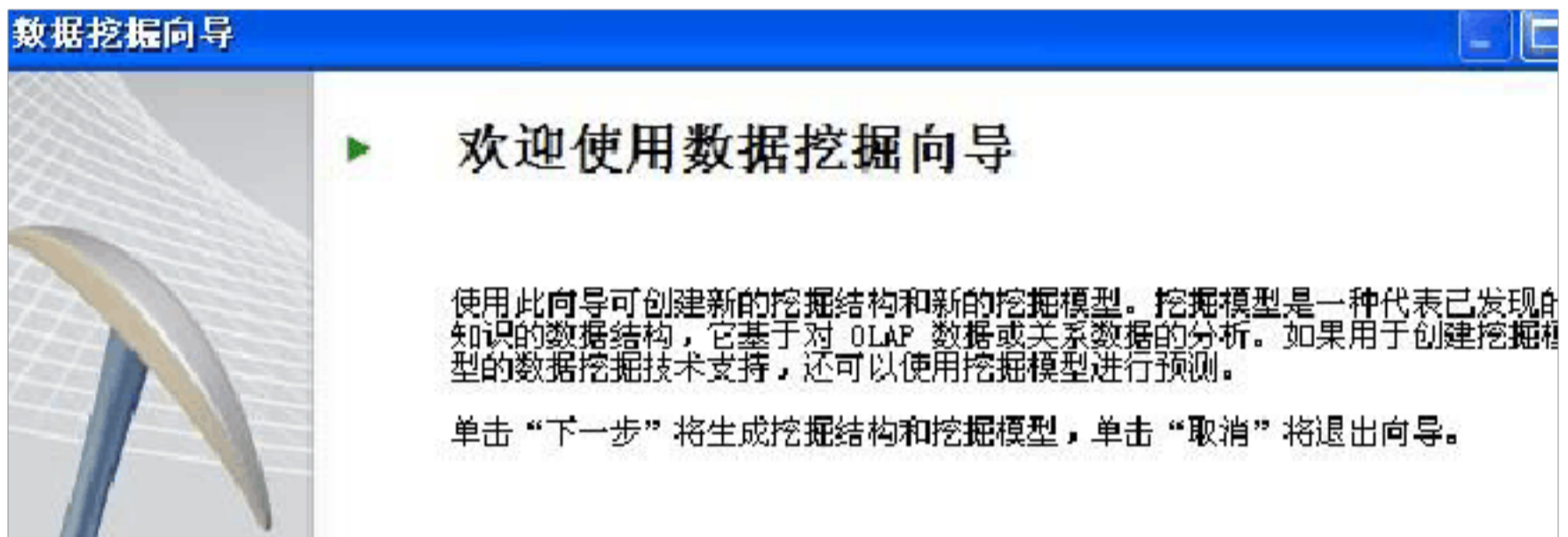
6. 在“完成向导”页上,默认情形下,体系将数据源视图定名为 Adventure Works DW. 单击“完成”.

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

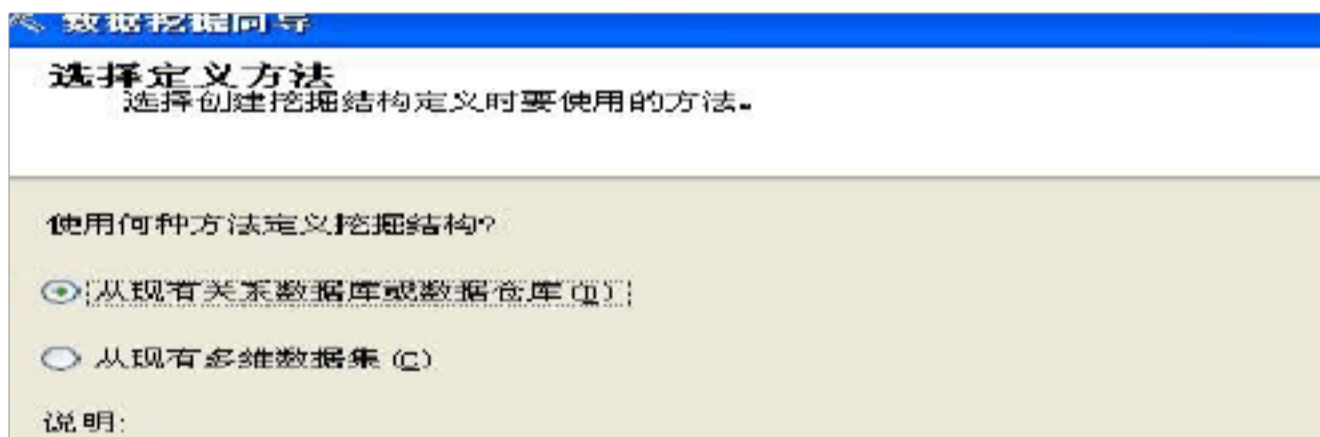


## 创建用于目标邮件计划的发掘构造

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”。



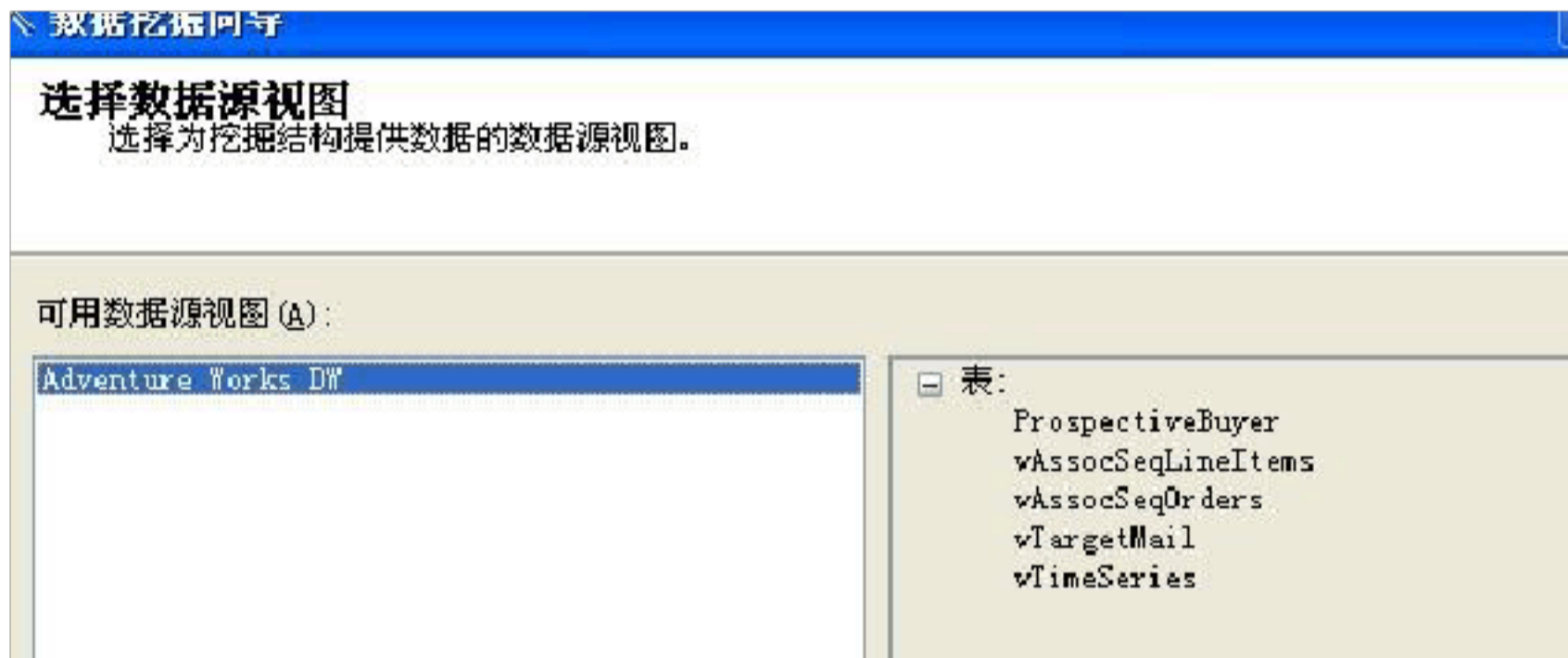


4. 在“创建数据发掘构造”页的“您要运用何种数据发掘技巧？”下,选择“Microsoft决议计划

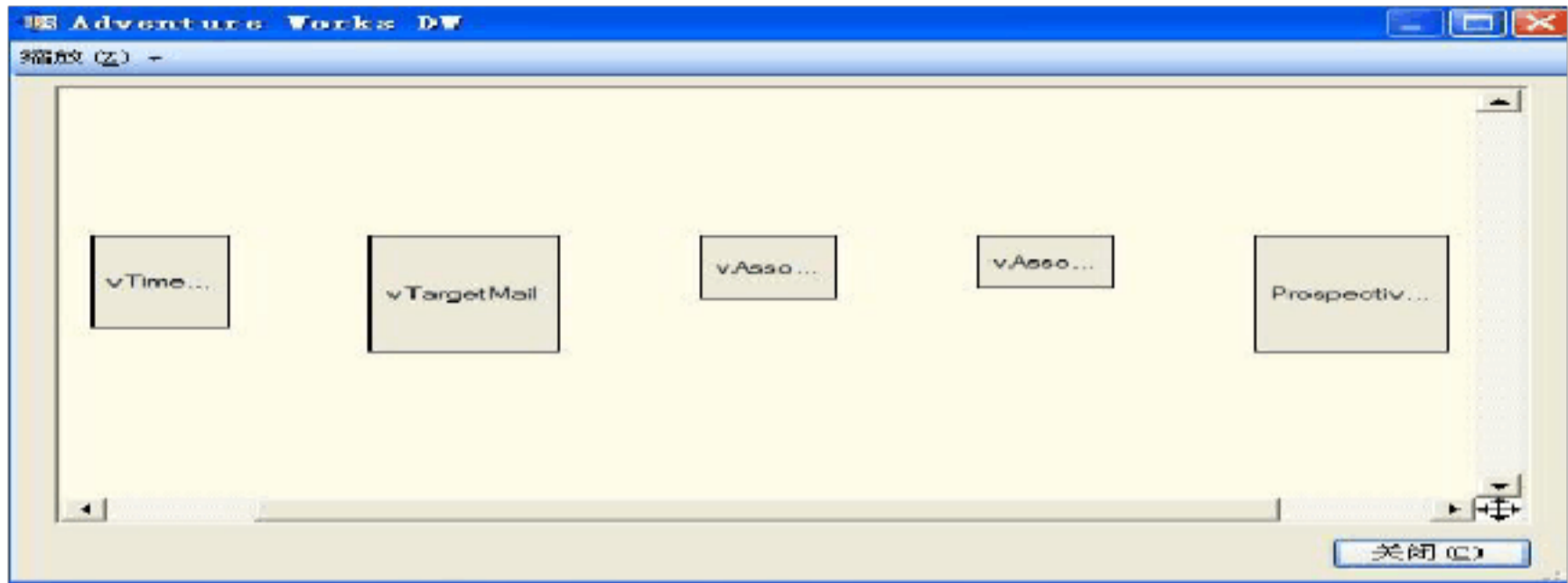


树”。

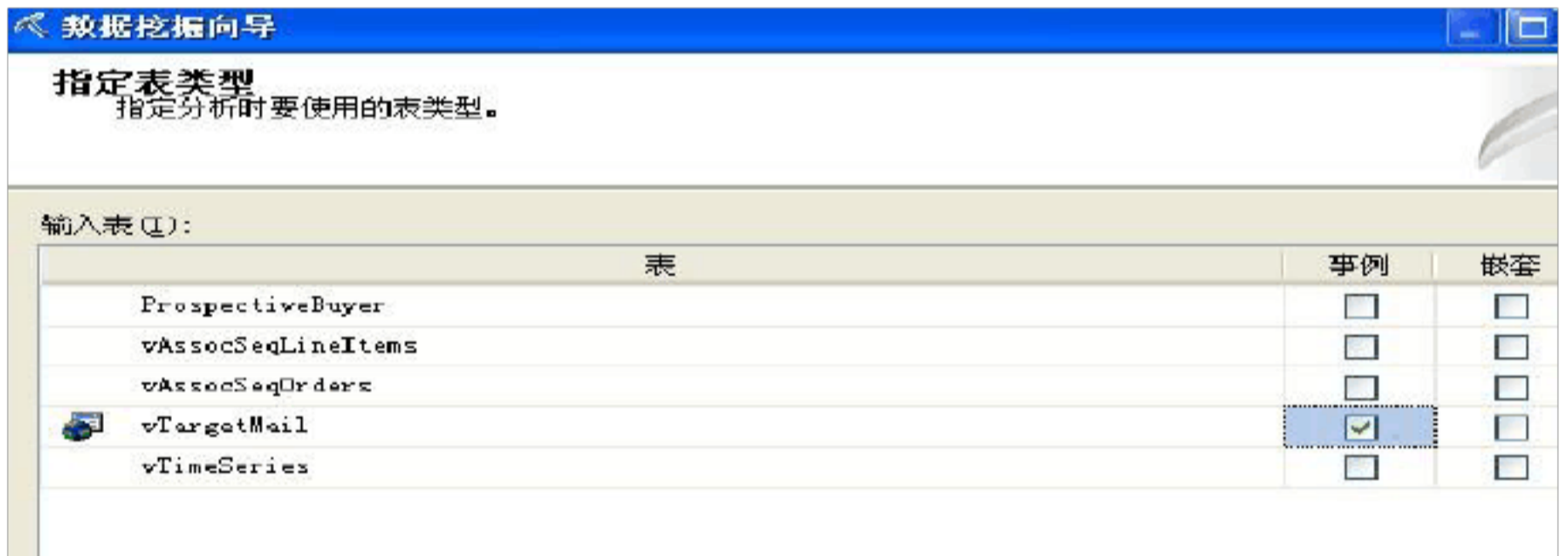
5. 单击“下一步”。



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“关闭”返回该领导。

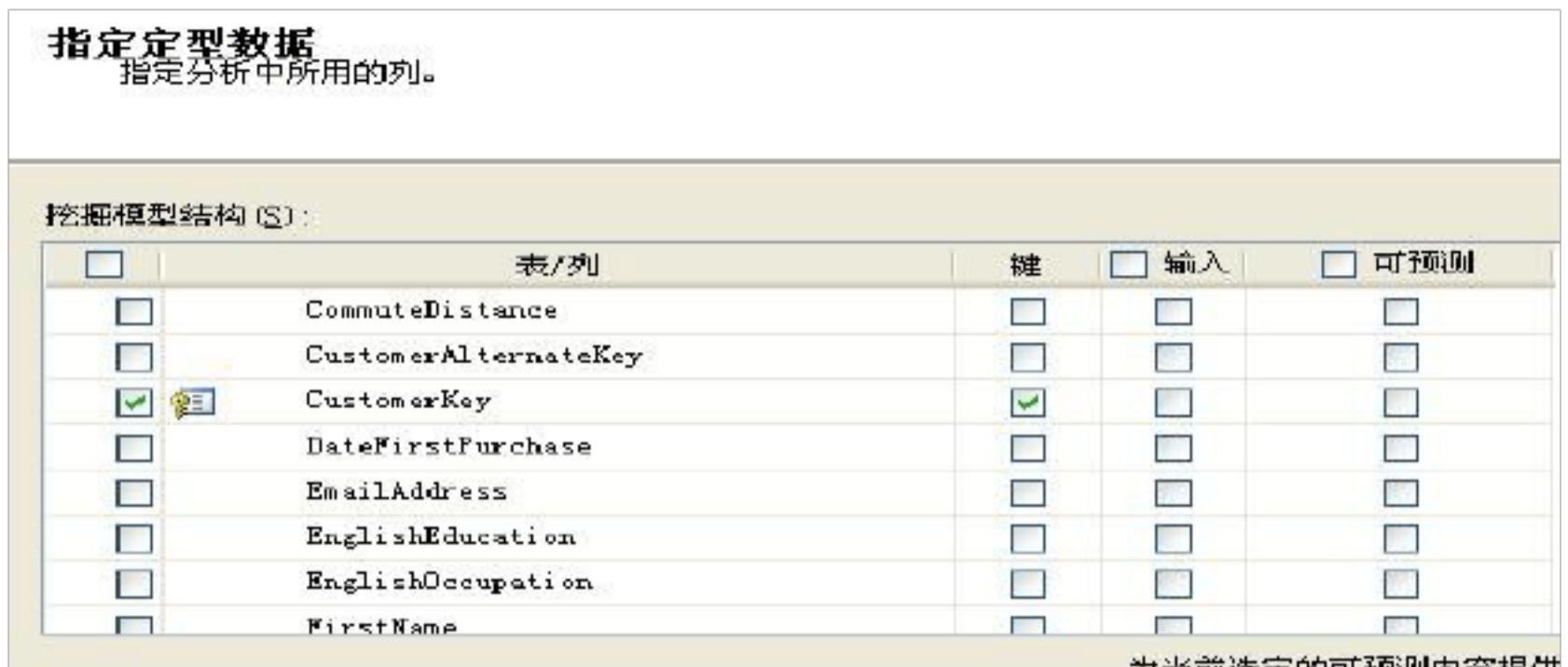


7. 单击“下一步”。
8. 在“指定表类型”页上,选中 vTargetMail 表旁边“事例”列中的复选框,再单击“下一步”。



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,则数据发掘向导将主动选择该列作为模型的键。



10. 选中 BikeBuyer 列旁边的“输入”和“可猜测”。



11. 单击“建议”打开“供给相关列建议”对话框。

只要选中至少一个可猜测属性,即可启用“建议”按钮。“供给相关列建议”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,以包括在模型中。

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

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

<input type="checkbox"/>	表/列	键	<input type="checkbox"/> 输入	<input type="checkbox"/> 可
<input type="checkbox"/>	FrenchOccupation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	Gender	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	GeographyKey	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	HouseOwnerFlag	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	LastName	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	MaritalStatus	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	MiddleName	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	NameStyle	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	NumberCarsOwned	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	NumberChildrenAtHome	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

挖掘模型结构 (S):

<input type="checkbox"/>	表/列	键	<input type="checkbox"/> 输入	<input type="checkbox"/> 可
<input checked="" type="checkbox"/>	NumberCarsOwned	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	NumberChildrenAtHome	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	Phone	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	Region	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	SpanishEducation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	SpanishOccupation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	Suffix	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	Title	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	TotalChildren	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	YearlyIncome	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

14. 单击“下一步”。

**指定列的内容和数据类型**  
指定挖掘结构列的内容和数据类型。

挖掘模型结构 (S):

列	内容类型	数据类型
Age	Continuous	Long
Bike Buyer	Continuous	Long
Commute Distance	Discrete	Text
Customer Key	Key	Long
English Education	Discrete	Text
English Occupation	Discrete	Text
First Name	Discrete	Text
Gender	Discrete	Text
Geography Key	Continuous	Long
House Owner Flag	Discrete	Text
Last Name	Discrete	Text

检测数值列的连续或离散数据:

< 上一步 (B)    下一步 (N) >    完成 (F) >> |    取消

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：<https://d.book118.com/605143343213012001>