

SMARTBI 思迈特软件 让数据为客户创造价值

Smartbi数据模型 在业务应用上的实践

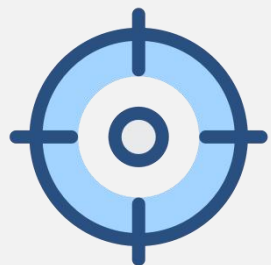
刘怡霞
高级产品培训师

直播时间: 2月23日 (周四) 19:00

- ▶ 五个维度分析企业为什么需要构建数据模型
- ▶ 三步手把手教你如何构建企业数据模型
- ▶ 企业数据建模工具及实践方法论

▶ 扫码预约直播
前10名可免费获得1对1辅导





游戏规则：组成语

最快答对的麦粉就可以获得88麦豆.

麦豆用途：观看教程、在线学习文档资料、兑换实物（麦豆商城即将上线.....）

组成语

今年是兔年，请大家列出一个带“兔”字的成语

答案：守株待兔、动若脱兔.....



游戏规则：猜名字

最快答对的麦粉就可以获得**88麦豆**。

麦豆用途：观看教程、在线学习文档资料、兑换实物（麦豆商城即将上线.....）

猜名字

最近非常火的一款由OpenAI研发的聊天机器人程序，它的名字是？

答案：ChatGPT

Smartbi 数据模型在业务应用上的实践

Smartbi 更聪明的大数据分析软件

目 录

CONTENTS

01 为什么需要数据模型

标题简介

02 什么是Smartbi数据模型

标题简介

03 Smartbi数据模型功能详解

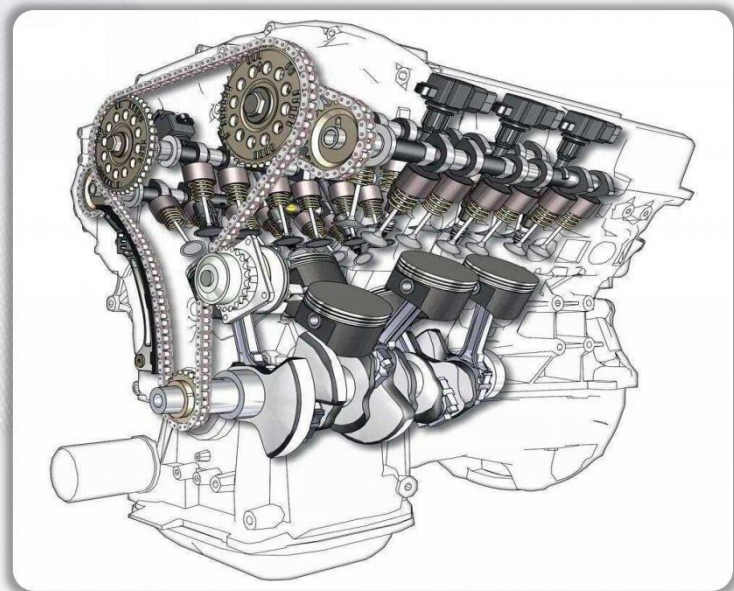
标题简介

04 使用规范和推荐用法

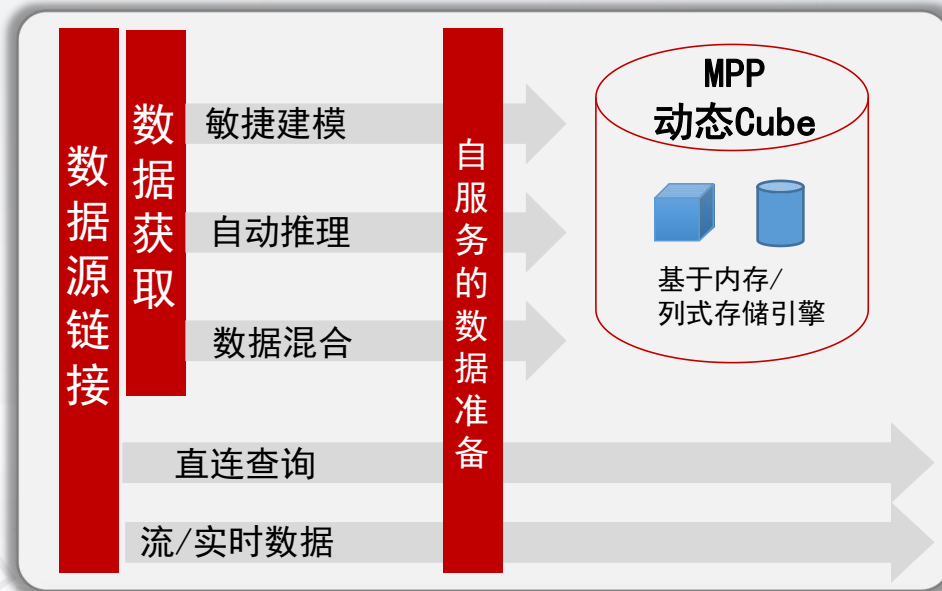
标题简介

01 为什么需要数据模型

对于“数据模型”的重要性与作用，就像一台汽车最重要的部件就是引擎，同样在一个BI产品里面也有一个引擎--数据建模功能。V10.5提供的“数据模型”相当于把汽车的发动机换掉了，从原来的V4发动机换成了 V12发动机，这是一个质的飞跃。



汽车引擎



BI引擎

增强“一站式”的架构能力

数据模型让“多维建模”能力融入BI产品，用户无论是否有数据仓库或是ODS，只需使用Smartbi，都可以敏捷建模实现数据集市。从而使得Smartbi具备一站式BI能力，即可一站完成数据连接、数据处理、数据建模、数据分析、数据分享等全流程功能。

增强“数据准备”的敏捷能力

现代化BI是以自助敏捷为理念，在数据分析及建模操作都更加简单敏捷。Smartbi数据模型通过更简单灵活的建模操作，更动态的Cube模式，让用户能更快速地响应需求、更敏捷地进行数据建模等。

增强“模型计算”的基础能力

Smartbi数据模型提供了丰富、强大的计算能力，不仅支持SQL计算、ETL分布式计算、MDX计算；还支持快速实现同环比时间计算、占比、排名、累计等；同时支持在计算中使用参数，动态获取结果。

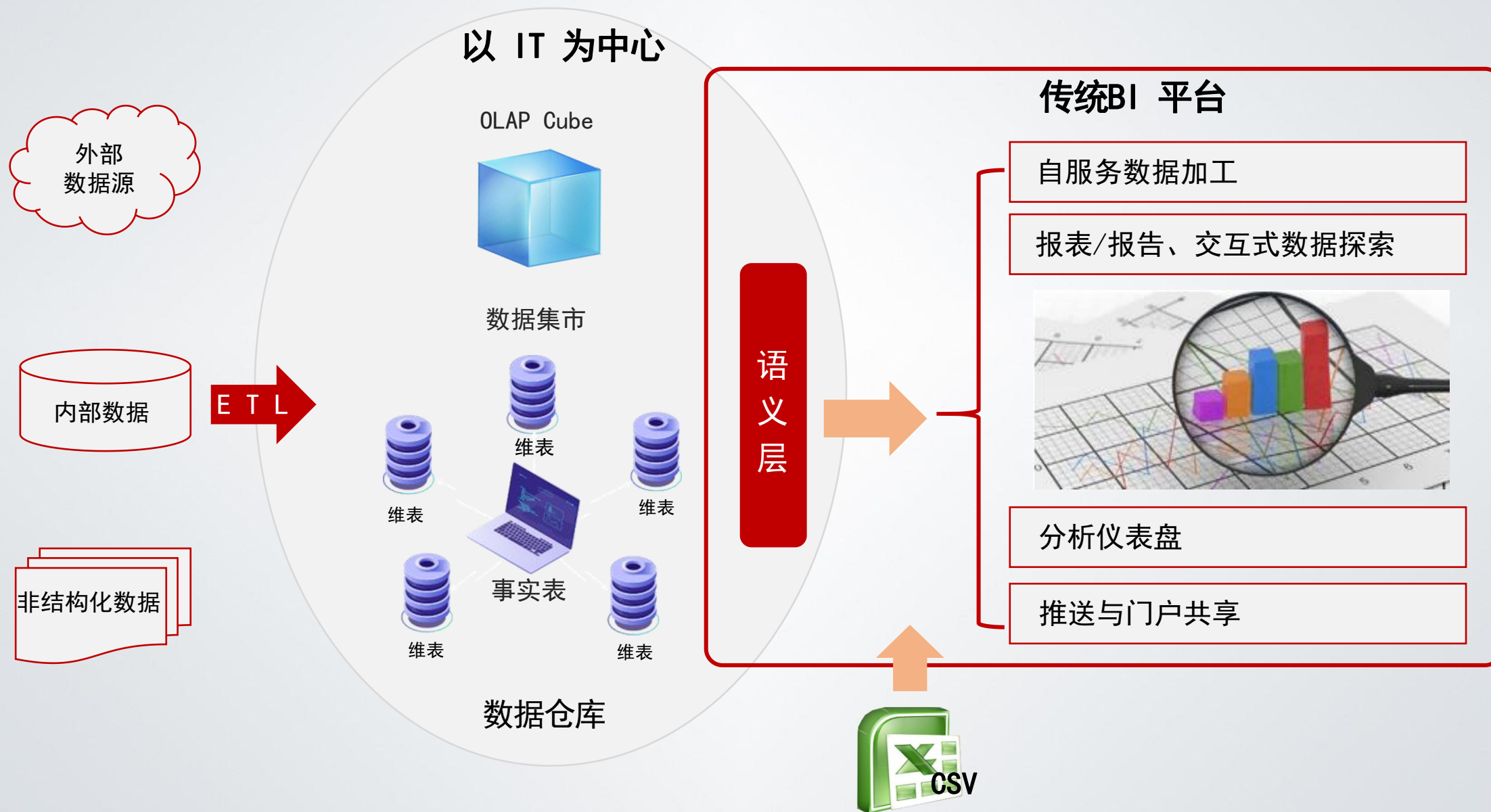


整合分散的业务数据，提升关联分析的效率

Smartbi数据模型整合多种数据集类型，融合线上线下数据，包括数据源表、即席查询、SQL查询、导入文件等，并且支持跨库数据整合，可以直接连接Oracle、MySQL、Excel文件等各类数据关联分析，满足不同用户的数据准备需求，提升数据分析的效率。

解决“网状路径”的选择问题

数据模型对Smartbi产品数据准备和数据分析之间的关系进行了重新梳理，统一了表现层和数据层，从而让表现层和数据层关系变得简单，用户一眼即可看清功能逻辑，学习和使用成本进一步降低。





按项目集中开发

提需求、提需求、提需求！

等待审批、等待立项、等待开发！

沟通业务需求、沟通数据来源、沟通测试方法！

数据处理、报表设计、核对数据……

慢

要报表，找科技部

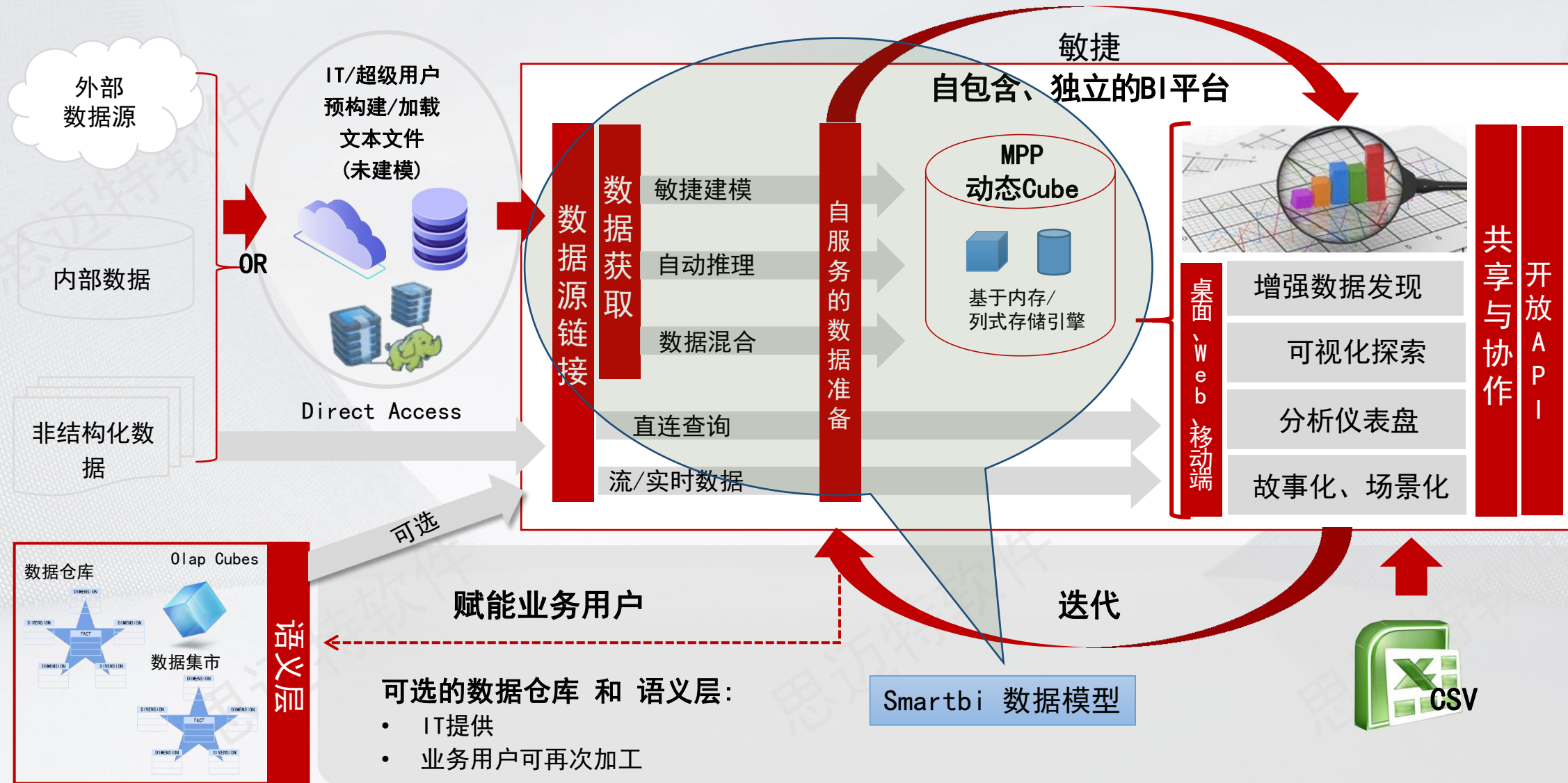
重

相似需求不断重复

难

BI工具使用难度高

以自助为理念的现代化BI带来了“敏捷建模”的新思路



查询场景：一名业务人员需要分析销售数据，但是有关数据都放在不同地方，部分数据存放在个人电脑的Excel上，如何进行数据分析工作？

数据表	存放方式	查询说明
产品类别表	MySQL数据库	
订单表	Oracle数据库	只获取订单表部分列数据
订单明细表	Excel文件	Excel需处理掉空数据，且需去除重复数据。



数据抽取(中间库)

用于将多来源数据汇聚，构建ODS贴源层



构建数据仓库(事实表维表)

用于加工、聚合数据，得到事实表、维表



数据集(SQL查询)

用于获取到需求报表需要的数据查询结果

业务人员能否独立完成此数据分析需求？



数据集市(大宽表)

用于整合汇总成某分析主题域的报表数据

“数据集”难以承担“敏捷建模”的使命



02 什么是Smartbi数据模型

应用商店

金融
银行|保险|证券...

政府
政数局|财政|工商...

教育
高校|高职...

科技
互联网|软件...

能源
石油|电力|电网...

制造
制药|汽车|冶金|...

运营商
电信|联通|移动|广电...

零售
线上|线下

...

平台产品

一站式BI大数据分析平台

嵌入式BI

SaaS BI云平台

传统BI

现代BI (自助BI+智能BI)

分析工具



电子表格

Excel电子表格
Web电子表格



分析报告

Word分析报告
PPT分析报告



即席查询

自助式
明细数据查询



透视分析

无需建模的OLAP
分析



交互仪表盘

可视化探索
数据大屏



Excel融合分析

面向Excel用户的
探索分析



对话式分析

基于自然语言的
智能数据分析



数据挖掘

预测性分析
机器学习



分享协同

帮助构建
数据文化

数据管理



数据接入

丰富的数据接入



数据采集

在线填报
批量导入



自助ETL

可视化的数据处理



数据模型

多源的数据整合
OLAP多维建模



指标管理

统一指标口径
提供指标自增长能力



高速缓存

提高数据查询性能



安全管理



运维监控



集群管理



元数据管理



用户行为分析



Smartbi的数据模型可灵活创建表间关系，将关联输出的结果集基于CUBE模型以“维度”和“度量”重新构建数据结构，实现数据模型灵活构建的同时，也整体提高了数据的处理效率与计算能力。



一体化建模能力，多种“表”类型

- 多种查询类型：表、SQL、脚本、**即系查询**、**ETL查询**等；
- 简易操作即可生成度量、时间维、地理维等；
- 支持多事实表，星型模型→雪花模型→**星座模型**等。

强大计算能力（库内计算，ETL计算，多维数据计算）

- 除了SQL计算能力，还支持ETL分布式计算能力，超大表关联；
- 快速使用智能计算同环比、排名等函数；
- 支持编写MDX函数进行复杂计算。

高性能：高速缓存MPP

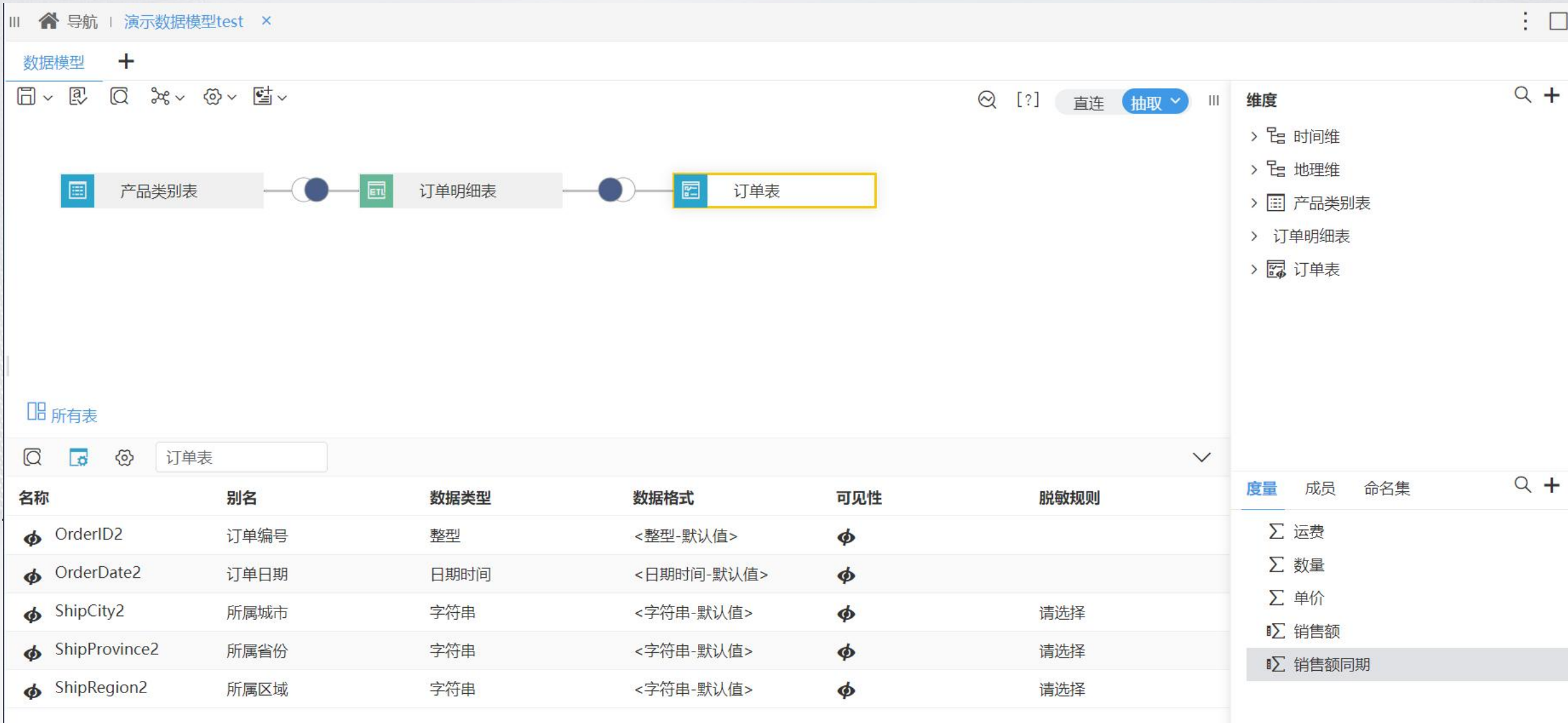
- 直连模式 & 抽取模式
- 数据抽取到分布式加速库中，提升查询速率。

将数据引入模型中，这些数据的来源和结构不尽相同，可能来源于数据库的其中一张表，或关联多张表查询，或本地的Excel文件等等，这些数据只能被当前的数据模型所使用。

私有查询	功能	适用情况
即席查询	提供自助化的操作界面，用户基于业务主题、数据源通过拖拽的方式，筛选、组合数据	0基础的业务人员
数据源表	基于业务库已建好的结果表，不需要再做其它数据预处理工作	数据已经处理好或者已有集市
SQL查询	支持自由书写SQL语句去查询数据	熟悉SQL的技术人员
导入文件	导入线下处理好的Excel、CSV、TXT等文件	存在线下数据需要合并分析
ETL高级查询	基于封装好的算法可视化组件，拖拉拽完成数据处理	进行高级数据加工、大数据量加工
脚本查询	支持编写代码脚本去查询数据	各种非结构化数据，特殊数据
存储过程查询	基于数据库中已编写好的存储过程获取数据结果	已存在的老数据系统
Java查询	通过用户自定义的类读取数据并输出的数据结果	不支持或不开放数据源取数的情形
生成日期表	基于用户设置的起止时间段及日期格式字段而自动生成的日期表结构和日期数据。	需系统生成日期时间段

数据模型—多种查询联合使用的实际场景

查询场景：一名业务人员需要分析销售数据，但是有关数据都放在不同地方，部分数据存放在个人电脑的Excel上，如何进行数据分析工作？



The screenshot displays the SMARTBI data model interface. At the top, there is a navigation bar with a home icon and the text '演示数据模型test'. Below this is a toolbar with various icons for file operations and a search bar. The main area shows a data model diagram with three tables: '产品类别表' (Product Category Table), '订单明细表' (Order Detail Table), and '订单表' (Orders Table). The '订单表' is highlighted with a yellow border. To the right of the diagram is a '维度' (Dimension) panel with a search icon and a plus sign. Below the diagram is a '所有表' (All Tables) section with a search bar containing '订单表'. Below this is a table with columns for '名称' (Name), '别名' (Alias), '数据类型' (Data Type), '数据格式' (Data Format), '可见性' (Visibility), and '脱敏规则' (Desensitization Rule). To the right of this table is a '度量' (Measure) panel with a search icon and a plus sign, containing a list of measures.

名称	别名	数据类型	数据格式	可见性	脱敏规则
OrderID2	订单编号	整型	<整型-默认值>	☐	
OrderDate2	订单日期	日期时间	<日期时间-默认值>	☐	
ShipCity2	所属城市	字符串	<字符串-默认值>	☐	请选择
ShipProvince2	所属省份	字符串	<字符串-默认值>	☐	请选择
ShipRegion2	所属区域	字符串	<字符串-默认值>	☐	请选择

维度

- > 时间维
- > 地理维
- > 产品类别表
- > 订单明细表
- > 订单表

度量 成员 命名集

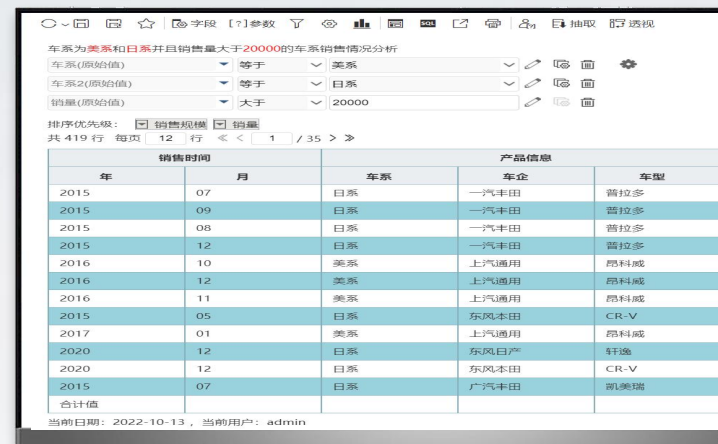
- Σ 运费
- Σ 数量
- Σ 单价
- Σ 销售额
- Σ 销售额同期

数据模型作为**数据来源**，可以直接应用在交互式仪表盘、固定格式报表、即席查询、对话式分析（NLA）。



交互式仪表盘

所见即所得、炫酷美观、丰富模板复用



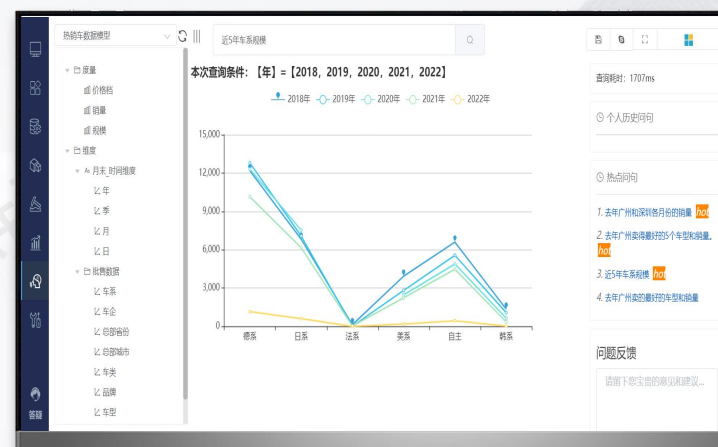
即席查询

秒级响应各种数据查询及分析请求

区域	产品	点心		谷类/薯片		肉/家禽				
		桂花糕	薯条	小米	鸡肉	盐焗鸡	小计			
东北	长春	2462	0	2462	0	2073	1291	3364		
	大连	0	0	0	390	4456	1514	6270		
	小计	2462	0	2462	390	6529	3105	9634		
华北	长治	0	640	640	0	0	0	0		
	北京	388	500	888	315	984	1299	3608	786	4394
	小计	388	1524	1912	315	984	1299	3608	879	4487
华东	上海	0	0	0	31	1856	0	1856		
	济南	1296	480	1776	0	477	477	0	393	393
	南京	1093	668	1761	422	339	761	2463	1441	3904
小计	2389	1148	3537	422	847	1269	4319	1834	6153	
华南	深圳	0	0	0	0	0	0	0	0	0
	厦门	1814	0	1814	0	249	249	566	590	1156
	海口	0	285	285	0	0	0	4640	0	4840
小计	1814	285	2099	0	249	249	5406	590	5996	
西南	昆明	0	240	240	504	78	582	2896	393	3289
	重庆	0	0	0	0	0	0	4539	328	4867
	成都	0	0	0	0	780	780	504	0	504
小计	0	240	240	504	858	1362	7939	721	8660	

固定格式报表

真Excel、支持Web版



对话式分析

PC/移动端，分析结果所问即所答

规则：发布题目后，大家可以将答案在聊天室刷起来，我们通过**随机截图**的方式送出礼品。（一共2道题）

奖品：台历+随行茶具

问题1、我们从几个维度说明为什么需要Smartbi 数据模型

答案：5

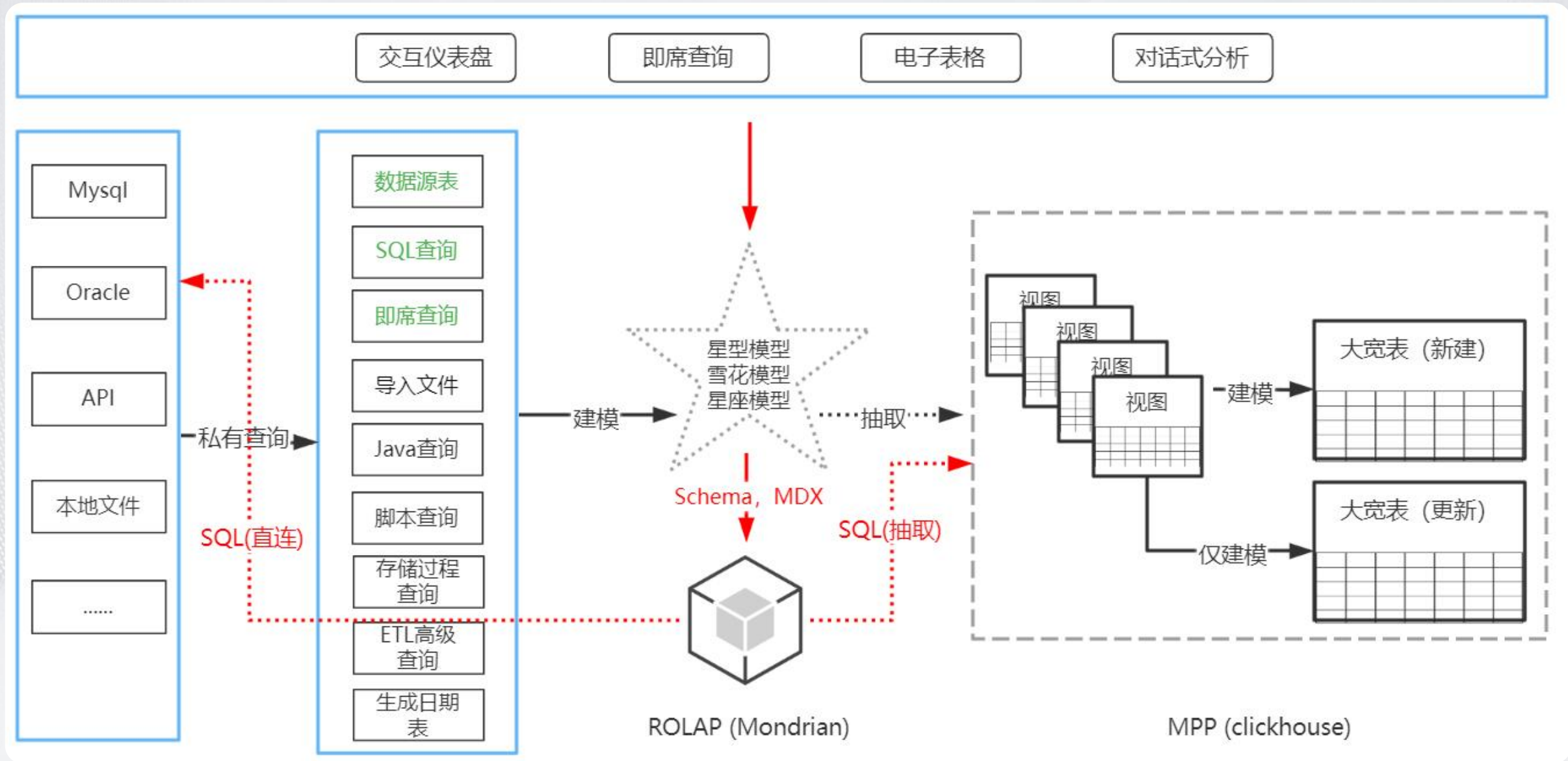
问题2、 Smartbi 数据模型提供了一体化的什么能力？

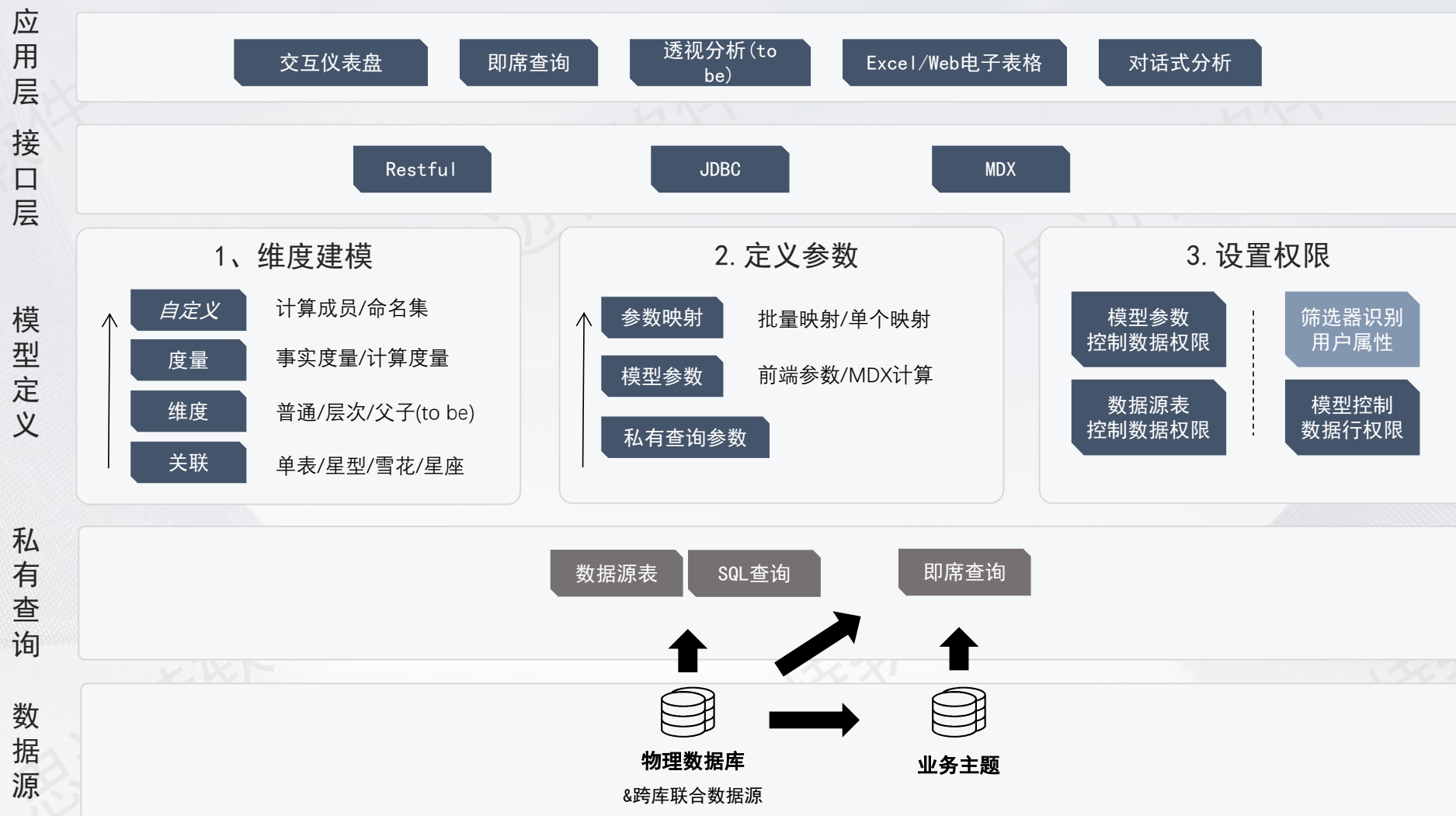
答案：建模



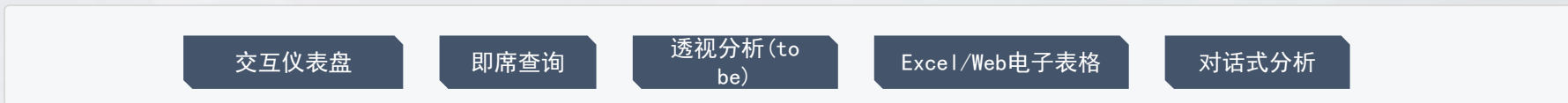
有奖竞答

03 Smartbi 数据模型功能详解





应用层



接口层



模型数据



模型定义



私有查询



数据源



什么场景适合用直连？

- 企业出于数据安全考虑，不允许数据加载到其他系统
- 数据频繁变化，又需要几乎实时的报表
- 数据量大且客户本身的存储性能很高

直连的注意事项

- 只支持SQL查询、数据源表、即席查询
- 直连时计算依赖原始库，要求数据库性能良好
- 个别数据库不支持直连：（由于mondrian的原因）
 - Smartbi JDBC for Excel
 - MongoDB
 - Tinysoft
 - Smartbi Jdbc4Olap

什么场景适合用抽取？

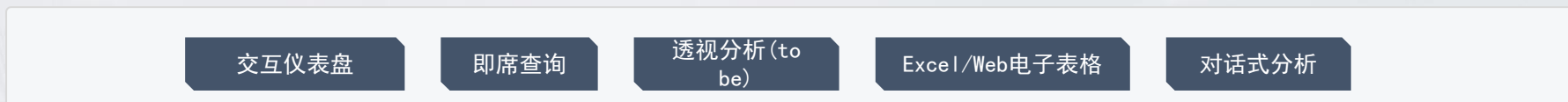
- 导入文件、脚本查询、ETL高级查询、存储过程查询必须用抽取模式，数据源类型多
- 原始数据库性能差，又难以优化性能，希望提升查询速度
- 数据变化频率不高

抽取的注意事项

- 需配置高速缓存库。
- 慎用参数映射
- 依赖数据源控制数据权限的场景慎用抽取

直连式数据模型的功能详解

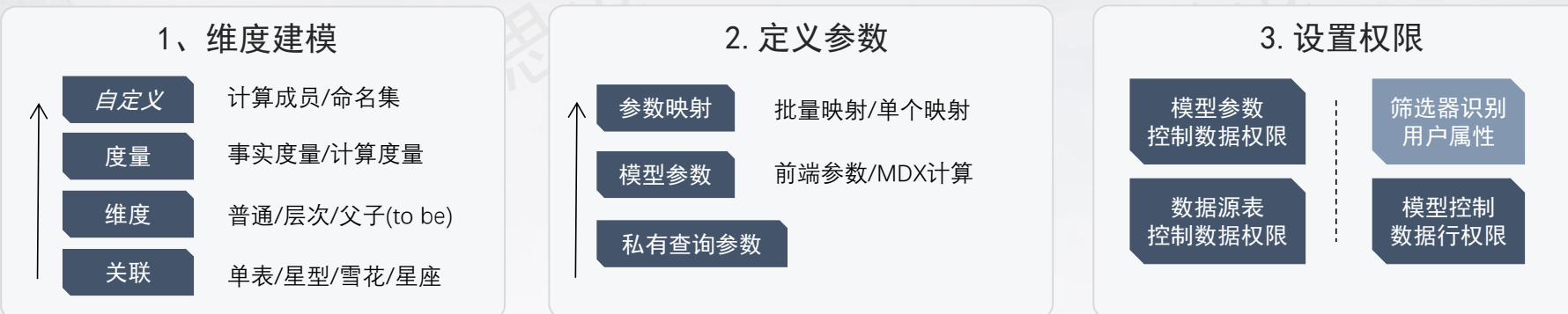
应用层



接口层



模型定义



私有查询



数据源



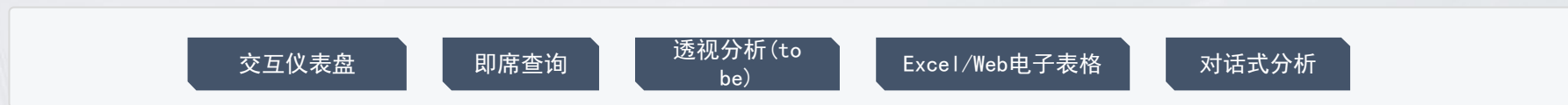
直连模式：报表直接从原业务库中查询取数。

私有查询	功能	适用情况
数据源表	基于业务库已建好的数据表	数据已经处理好或者已有集市
SQL查询	支持自由书写SQL语句去查询数据	熟悉SQL的技术人员
即席查询	基于业务主题、数据源直接拖拽字段	0基础的业务人员

注意事项

- 数据模型中仅包含**SQL查询、数据源表、即席查询**三种类型的私有查询时支持直连，且所有私有查询来源于**相同的数据源（或跨库联合数据源）**；
- 直连模式要求原业务库必须**支持SQL子查询**；
- 由于Mondrian的原因，个别数据库不支持直连（Smartbi JDBC for Excel、MongoDB、Tinysoft、Smartbi jdbc40lap）；

应用层



接口层



模型定义



私有查询



数据源



维度表(dimension table)

表示对分析主题所属类型的描述。比如“昨天晚上张三在京东花费200元购买了一个皮包”。那么以购买为主题进行分析，可从这段信息中提取三个维度：时间维度(昨天晚上)，地点维度(京东)，商品维度(皮包)。通常来说维度表信息比较固定，且数据量小。

维度：字符串类型的字段一般是维度

事实表(fact table)

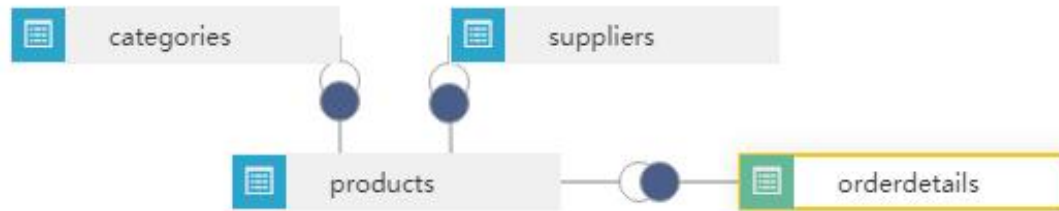
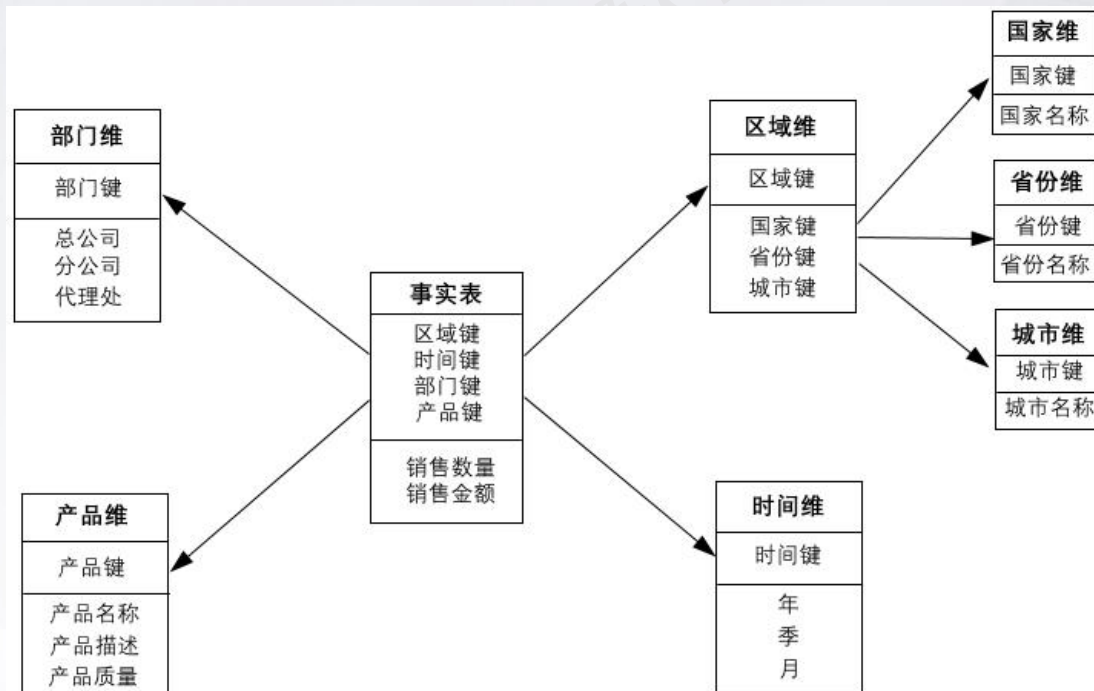
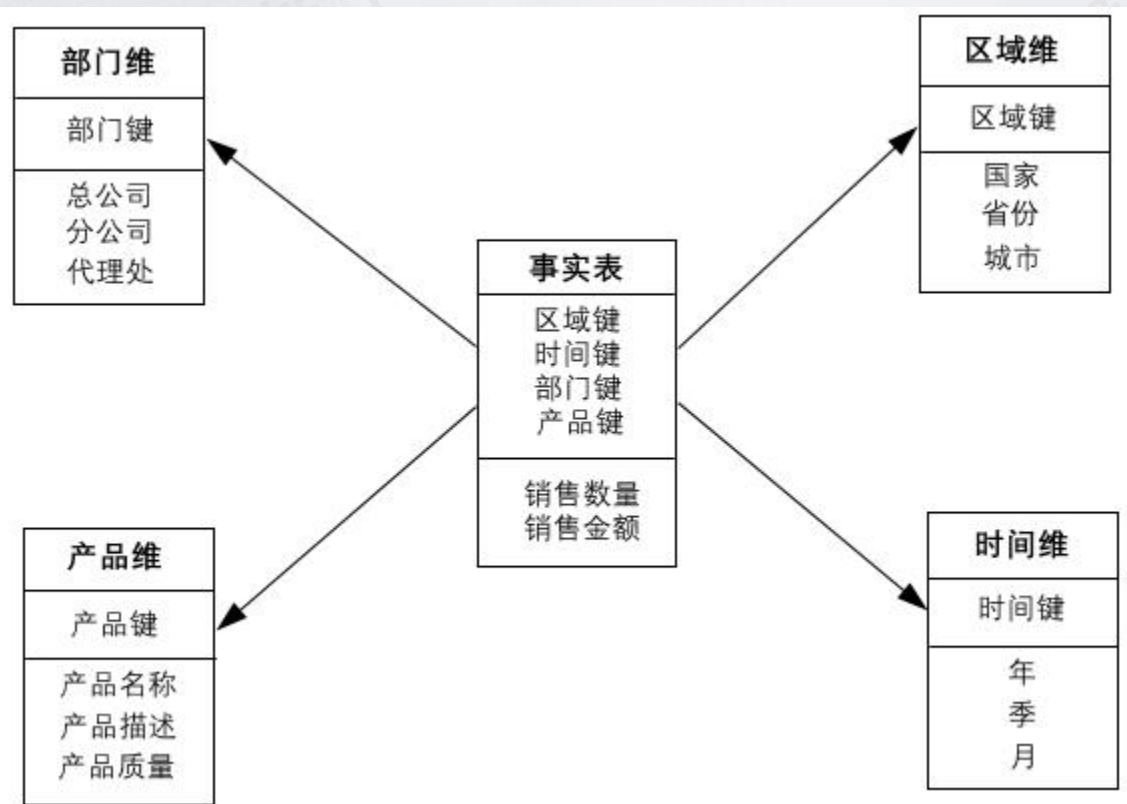
表示对分析主题的度量。比如上面那个例子中，200元就是事实信息。事实表包含了与各维度表相关联的外码，并通过JOIN方式与维度表关联。事实表的度量通常是数值类型，且记录数会不断增加，表规模迅速增长。

度量：数值类型的字段一般是度量

数据模型是围绕事实表和维度表的关系而进行模型的构建。模型架构：星型模型、雪花模型、星座模型、单表模型

星型模型：所有维表都直接连接到事实表上，整个视图就像星星一样

雪花模型：一个或多个维表没有直接连接到事实表上，而是通过其他维度表连接到事实表，视图就像多个雪花连接在一起



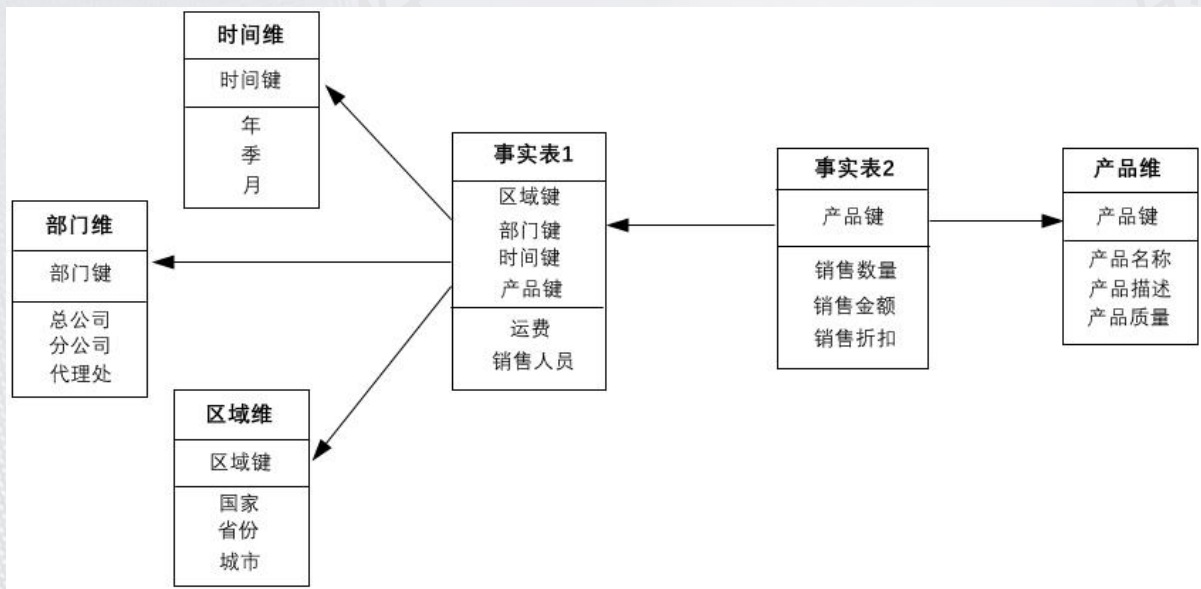
模型表左侧绿底表示事实表，蓝底表示维表

模型表左侧绿底表示事实表，蓝底表示维表

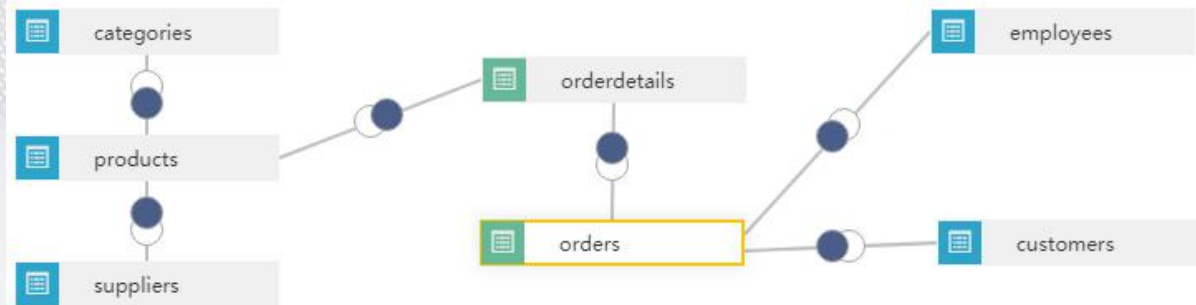
数据模型是围绕事实表和维度表的关系而进行模型的构建。模型架构：星型模型、雪花模型、星座模型、单表模型

星座模型：包含多个事实表，维表是公共的，可以共享

单表模型：指标、维度、属性关联在一起的一张表



事实表1	
区域键	部门键
时间键	产品键
运费	
销售人员	



模型表左侧绿底表示事实表，蓝底表示维表

订单	订单	订单	订单	订单	订单	订单	订单	订单	订单	订单	订单	订单	订单	订单	订单
订单 Id	订...	发货日期	邮寄...	...	细分	城市	省...	国...	...	类别	子...	产品名称	销售额	数量	折扣
US-20...	2019...	2019/4/29	二级	管惠	公司	杭州	浙江	中国	华东	办公用品	用品	Fiskars 剪刀, ...	¥ 130	2	40%
CN-20...	2019...	2019/6/19	标准级	许安	消费者	内江	四川	中国	西南	办公用品	信封	GlobeWeis 搭...	¥ 125	2	0%
CN-20...	2019...	2019/6/19	标准级	许安	消费者	内江	四川	中国	西南	办公用品	装订机	Cardinal 孔加...	¥ 32	2	40%
US-20...	2019...	2019/12/13	标准级	宋良	公司	镇江	江苏	中国	华东	办公用品	用品	Kleencut 开信...	¥ 321	4	40%
CN-20...	2018...	2018/6/2	二级	万兰	消费者	汕头	广东	中国	中南	办公用品	器具	KitchenAid 搅...	¥ 1,376	3	0%
CN-20...	2017...	2017/10/31	标准级	俞明	消费者	景德镇	江西	中国	华东	技术	设备	柯尼卡 打印机...	¥ 11,130	9	0%
CN-20...	2017...	2017/10/31	标准级	俞明	消费者	景德镇	江西	中国	华东	办公用品	装订机	Ibico 订书机, ...	¥ 480	2	0%

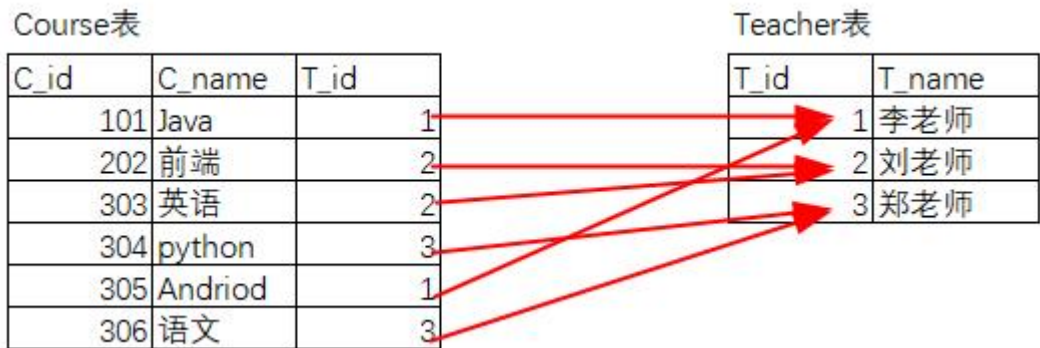
维度

度量

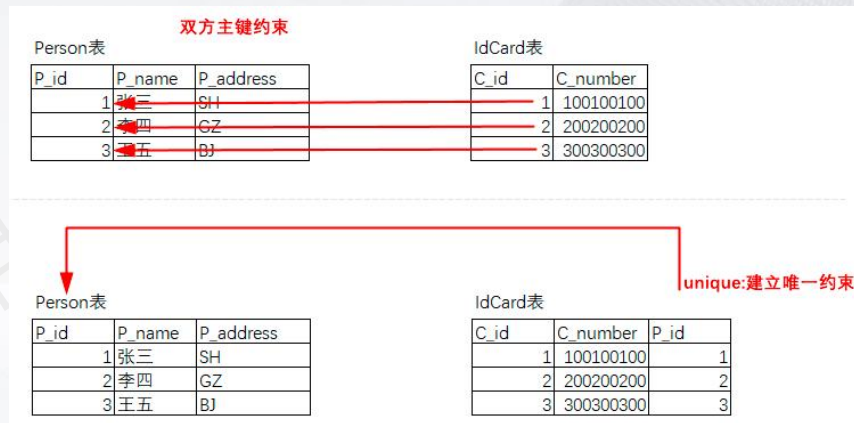
关系的确立需要通过匹配键列中的数据（通常是两表中同名的列）。在大多数情况下，该关系会将一个表中的主键（它为每行提供了唯一标识）与另一个表的外部键中的某项相匹配。例如，通过创建 orderdetails表中的orderid（主键）与orders表中的orderid列（外部键）之间的关系，则销售量就与订单相关联了。

表关系类型：一对多(多对一)关系、一对一关系、多对多关系

一对多（多对一）关系（一对多和多对一是相对关系）



一对一关系



数据模型表关系类型：一对多、多对一、一对一(内连接)、一对一(外连接)

编辑关系

关系* 产品表 = 请选择 +

基数* 请选择

- 多对一
- 一对多
- 一对一 (内连接)
- 一对一 (外连接)

温馨提示：只有数据类型以及字段名称一致的字段才能自动匹配关联关系

确定 取消

一对一（内连接）/一对一（外连接）：

关联字段匹配后，内连接关系将会删除未完全匹配数据，外连接会将未完全匹配数据以”空“补充。

如下我们有两张数据表“家庭表”和“学生表”：

ID	StudentID	fatherName	mortherName
1	1	wulei	dafang
2	2	sundawei	mei
3	3	zhong	xiaocui
4	4	pang	jiali
5	5	lizi	ping
6	6	zhao	mingmingzi

StudentID	Name	Sex
1	fangxiang	女
2	梅子	女
3	cuicui	男
4	佳佳	男
5	平子	女
6	明仔	男
7	明仔2	男
8	明仔3	男
9	明仔4	男
10	明仔5	男

StudentID	fatherName	mortherName	Name	Sex
1	wulei	dafang1	fangxiang	女
2	sundawei	mei001	梅子	女
3	zhong	xiaocui	cuicui	男
4	pang	jiali	佳佳	男
5	lizi	ping	平子	女
6	zhao	mingmingzi	明仔	男

StudentID	fatherName	mortherName	Name	Sex
1	wulei	dafang1	fangxiang	女
2	sundawei	mei001	梅子	女
3	zhong	xiaocui	cuicui	男
4	pang	jiali	佳佳	男
5	lizi	ping	平子	女
6	zhao	mingmingzi	明仔	男
7			明仔2	男
8			明仔3	男
9			明仔4	男
10			明仔5	男

维度建模--维度、度量、自定义

系统自动获取私有查询的输出结果，并将所有返回字段看作维度字段，我们可以在此基础上重新进行目录归集，层次划分，以及将表示度量的字段进行标记。

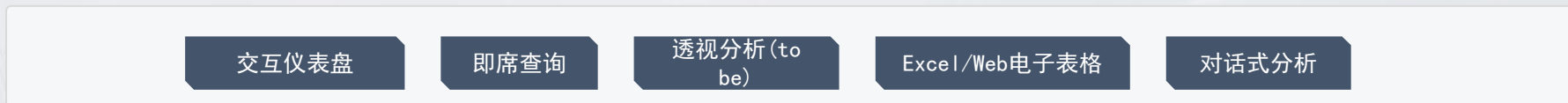
The image illustrates the process of building a data model in SMARTBI through several screenshots:

- 层次维度 (Hierarchical Dimension):** A screenshot of the '维度' (Dimension) panel showing a tree view of '订单表' (Orders). The '订单日期' (Order Date) field is highlighted with a red box. A context menu is open, with '创建时间层次结构' (Create time hierarchy) selected.
- 普通维度 (Standard Dimension):** A screenshot of the '维度' panel showing a tree view of '地理维' (Geography). The '区域' (Region), '省份' (Province), and '城市' (City) fields are highlighted with a red box.
- 事实度量 (Fact Measure):** A screenshot of the '维度' panel showing a list of fields. The '数量' (Quantity) field is highlighted with a red box, and a '标记为度量' (Mark as measure) button is visible.
- 计算度量 (Calculated Measure):** A screenshot of the '度量' (Measure) panel showing a list of measures. The '销售量' (Sales Volume) measure is highlighted with a red box, and a context menu is open with '新建计算度量' (New calculated measure) selected.
- 计算成员 (Calculated Member):** A screenshot of the '度量' (Measure) panel showing a list of members. The '一线城市' (First-tier cities) member is highlighted with a red box.
- 命名集 (Named Set):** A screenshot of the '命名集' (Named Set) panel showing a list of named sets. The '销量排行前三城市' (Top 3 cities by sales volume) named set is highlighted with a red box.

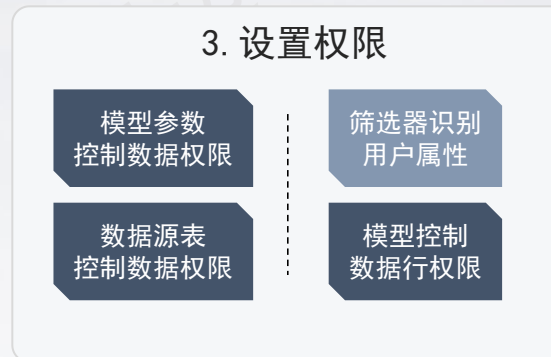
Labels for each screenshot are: 层次维度, 普通维度, 事实度量, 计算度量, 计算成员, 命名集.

- 简易操作即可生成度量、时间维、地理维等；
- 自定义：支持托拉拽，支持算术运算符，支持MDX表达式、MDX函数，实现丰富、复杂的计算场景；

应用层
接口层



模型定义



私有查询



数据源



直连式数据模型 -- 定义私有查询参数

私有查询中的私有参数，这类参数可以直接应用于报表数据的筛选及联动。

```
1 select
2     T1.`ProductName` as `产品名称`,
3     T2.`Quantity` as `销量`
4 from
5     `products` T1
6
7
8 where
9     {[T3
```

SQL查询

即

ShipRegion	ShipProvince	ShipCity	Freight
华东	山东省	济南	11.61
华东	山东省	济南	11.61
华北	河北省	秦皇岛	65.83
华北	河北省	秦皇岛	65.83
华北	河北省	秦皇岛	65.83
华东	江苏省	南京	41.34
华东	江苏省	南京	41.34

省份	产品名称	销量	平均运费
北京市			101.45
北京市	柳橙汁	580	108.02
北京市	绿茶	623	192.52
北京市	蜜桃汁	883	45.52
北京市	牛奶	1,057	118.74
北京市	浓缩咖啡	1,155	54.91
北京市	啤酒	506	8.19
北京市	苹果汁	828	98.05
北京市	汽水	1,125	58.10
北京市	苏打水	819	61.45
北京市	运动饮料	793	257.62
福建省			55.22
福建省	柳橙汁	580	43.51
福建省	绿茶	623	101.96

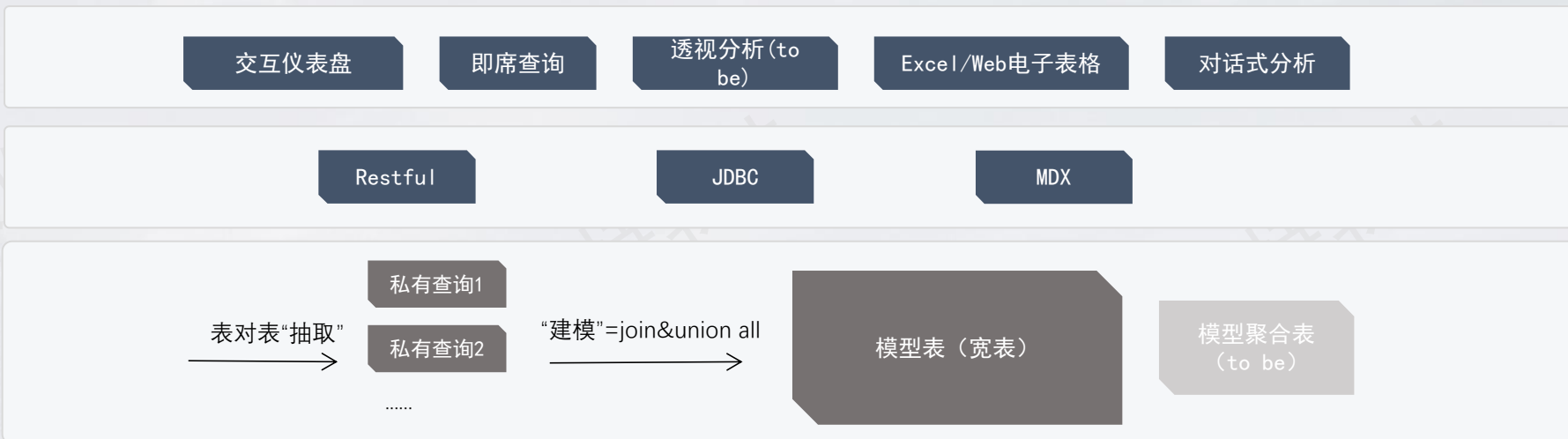
筛选效果

发货区域	省份	发货城市	平均运费
华北	北京市	北京	99.91

联动效果

抽取式数据模型的功能详解

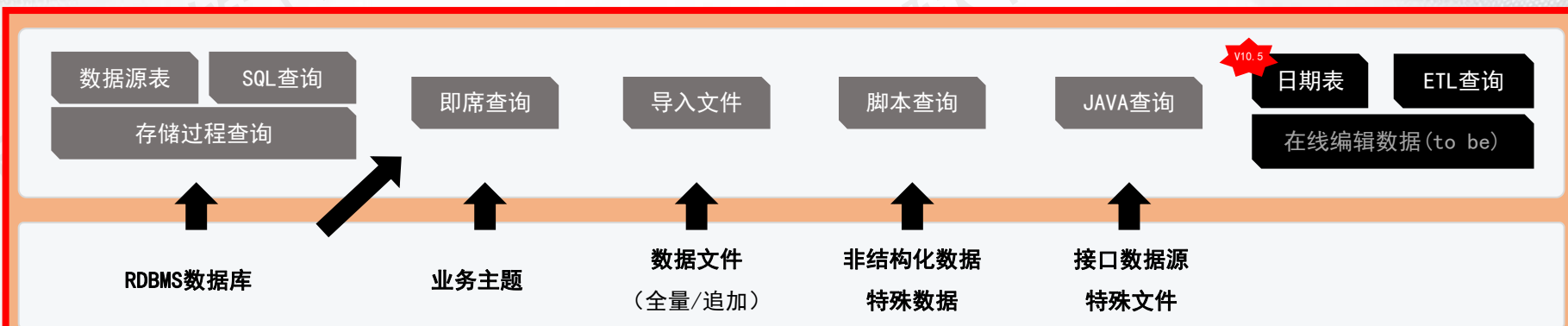
应用层
接口层
模型数据



模型定义

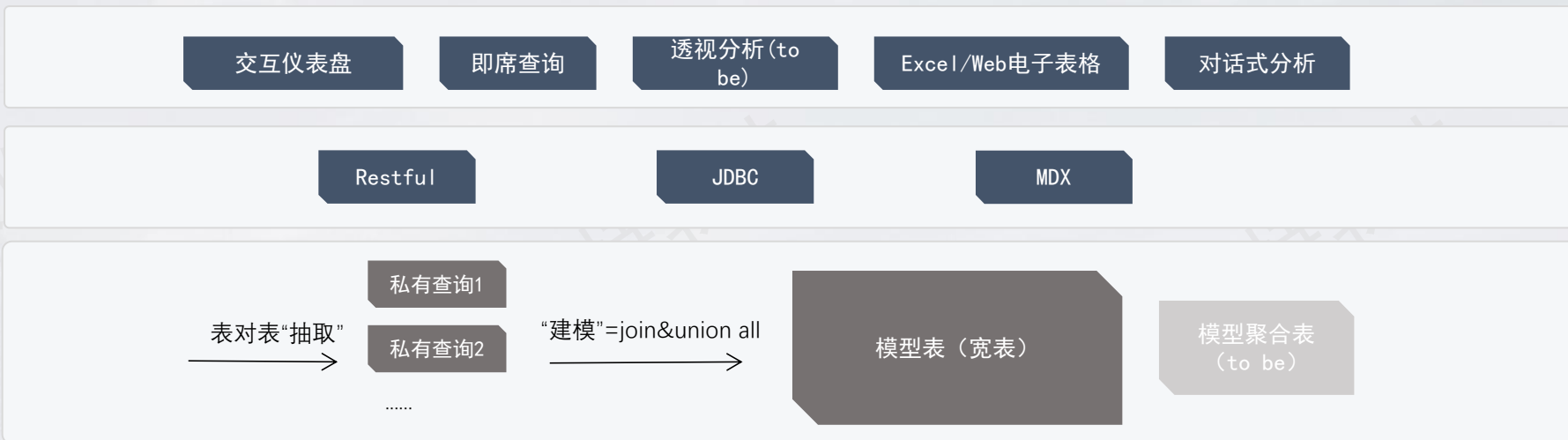


私有查询
数据源



私有查询	功能	适用情况
即席查询	基于业务主题、数据源直接拖拽字段	0基础的业务人员
数据源表	基于业务库已建好的结果表	数据已经处理好或者已有集市
SQL查询	支持自由书写SQL语句去查询数据	熟悉SQL的技术人员
导入文件	导入线下处理好的Excel、CSV、TXT等文件	存在线下数据需要合并分析
ETL高级查询	基于封装好的算法可视化组件，拖拉拽完成数据处理	进行高级数据加工、大数据量加工
脚本查询	支持编写代码脚本去查询数据	各种非结构化数据，特殊数据
存储过程查询	基于数据库已编写好的存储过程	已存在的老数据系统
Java查询	通过用户自定义的类读取数据并输出的数据结果	不支持或不开放数据源取数的情形
生成日期表	基于用户设置的起止时间段及日期格式字段而自动生成的日期表结构和日期数据。	需系统生成日期时间段

应用层
接口层
模型数据



模型定义



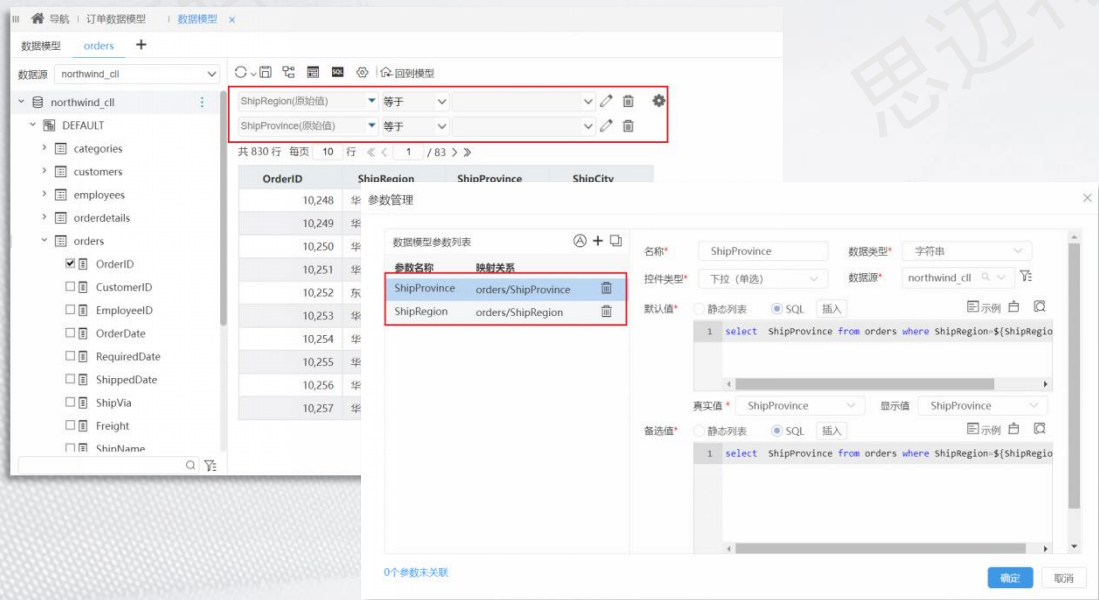
私有查询
数据源



过滤数据

私有参数：过滤数据

在私有查询中定义筛选器，可对业务库数据过滤及用户数据权限的控制。



注意事项：在数据模型中定义参数，一般用于过滤静态数据；若想要在报表展现层通过参数动态过滤数据，则建议在报表展现层定义参数。

动态结果应用

数据模型参数：主要用于计算度量、计算成员和命名集的表达式，实现动态获取结果的效果；（同直连模式）

动态获取结果 - 计算度量

发货区域	累计销售量
东北	2,468.00
华北	3,797.00
华东	3,460.00
华南	830.00
西南	1,917.00

动态获取结果 - 计算成员

发货区域	销量
排除所选区域销量汇总	29,086
东北	4,973
华北	22,206
华东	13,518
华南	4,686
华中	107
西北	592
西南	5,210

显示销量大于10000的区域

动态获取结果 - 命名集

发货区域	销量
华北	22,206
华东	13,518

全量

需求场景：

- 数据变化频率不高
- 业务库性能较差，提高
- 只需要抽取部分指定的

后台实现：

每次抽取前先清空之前抽取
若设置了抽取参数，则按照

The screenshot shows the SMARTBI data model interface. At the top, there's a toolbar with icons for home, refresh, search, and settings. A dropdown menu is open, showing '直连' (Direct) and '抽取' (Extract) options, with a red box and arrow highlighting the '抽取' option. Below the toolbar is a star schema diagram with four tables: '区域维' (Area Dimension), '城市维' (City Dimension), '产品维' (Product Dimension), and '销售事实表' (Sales Fact Table). The '区域维' table is highlighted with a yellow border. Below the diagram is a '所有表' (All Tables) list. In the foreground, an '抽取设置' (Extraction Settings) dialog is open. It has a table for '查询列表' (Query List) and two radio buttons for '抽取方式' (Extraction Method): '全量抽取' (Full Extraction) and '按次抽取' (Incremental Extraction). The '全量抽取' option is selected. Below the radio buttons, it says '抽取参数值: 无参数' (Extraction Parameter Value: No Parameters). At the bottom of the dialog are '确定' (Confirm) and '取消' (Cancel) buttons.

查询名称	抽取方式
区域维	全量
城市维	全量
产品维	全量
销售事实表	全量

抽取方式： 全量抽取 按次抽取

抽取参数值：无参数

用直连方式，

，同时又不

储。当这些发
定数据。

+ 【数据模型】数据模型新增SQL引擎

数据模型新增SQL引擎，实现MDX&SQL双引擎，类似汽车的“DM-i超级混动”，使大数据量清单表查询也能轻松获取。

+ 【数据模型】支持向导方式创建计算度量

在数据模型支持向导式创建计算度量，支持四则运算、时间计算、文本转换、命名集集合等，且计算类型支持无编码扩展。



数据 模型

一体化超强计算的数据建模

↑ 性能提升

↑ 更便捷更易用

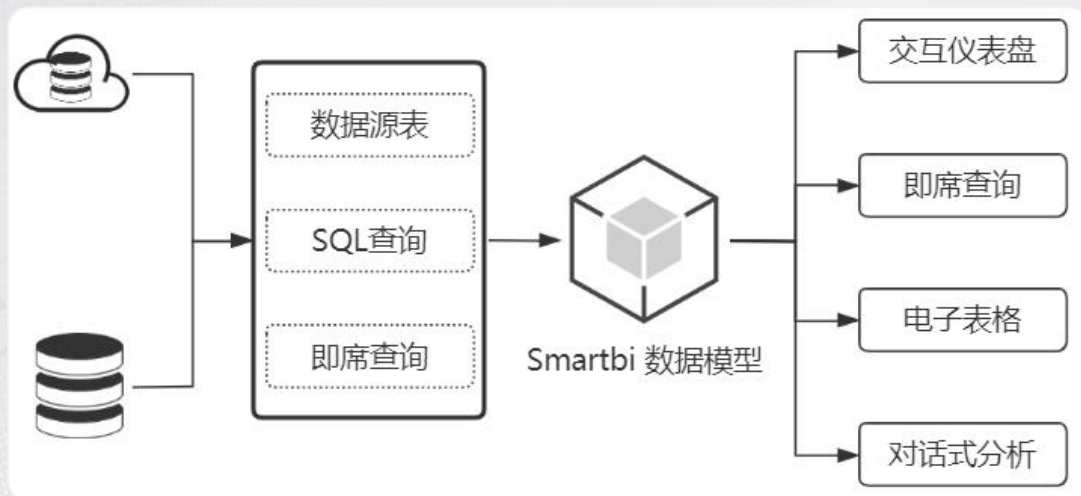
The screenshot displays the SMARTBI data modeling interface. On the left is a navigation tree with categories like '指标模型', '自助ETL', '作业流', 'ETL自动化', '业务主题', '数据集', and '订单模型'. The main area shows a data model diagram with tables: '订单表' (highlighted), '订单明细表', '产品表', '顾客表', and '产品类别表'. Below the diagram is a table titled '所有表' showing the schema for '订单表'.

名称	别名	数据类型	数据格式	可见性	脱敏规则
# OrderID2	订单编号	整型	<整型-默认值>	👁	
👁 CustomerID	顾客编号	字符串	<字符串-默认值>	🔒	请选择
👁 OrderDate	订单日期	日期时间	<日期时间-默认值>	👁	
👁 EmployeeID	雇员编号	整型	<整型-默认值>	🔒	
👁 RequiredDate	交货日期	日期时间	<日期时间-默认值>	👁	
👁 ShippedDate	发货日期	日期时间	<日期时间-默认值>	👁	
👁 ShipVia	经销途径	整型	<整型-默认值>	🔒	

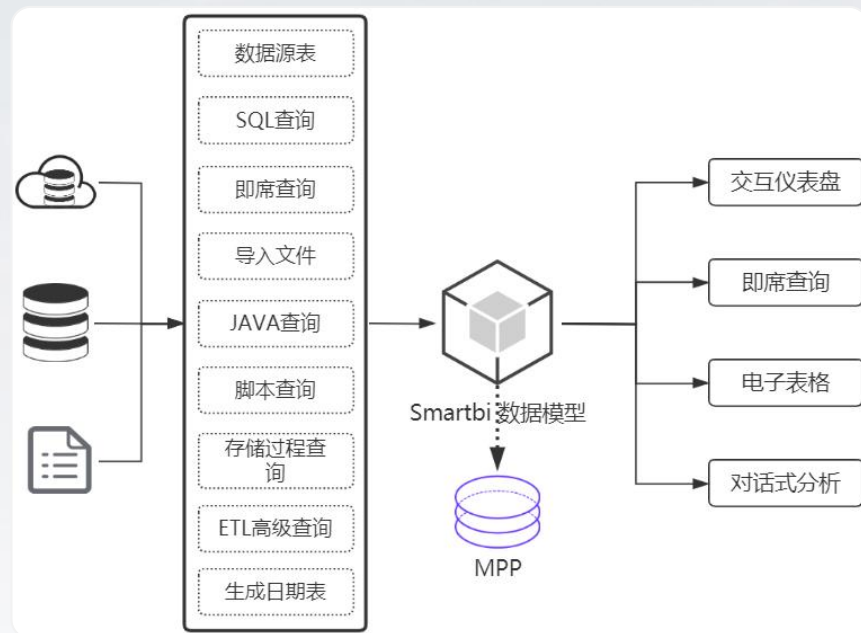
On the right side, there are panels for '维度' (Dimensions) and '度量' (Measures). The '度量' panel shows options like '新建计算度量', '新建计算成员', '新建命名集', and '新建目录'.

04 数据模型使用规范和推荐用法

模式	私有查询	参数	权限	使用规范
直连模式	3 种	前端参数、动态结果应用 (MDX计算)	推荐: 数据模型右键设置数据权限	<ol style="list-style-type: none">1、不使用抽取开关2、不允许使用其它私有查询3、不使用生成日期表 (除非数据源可写)
抽取模式 (约3万条/秒)	9 种	过滤数据、动态结果应用 (MDX计算)	必须: 数据模型右键设置数据权限	<ol style="list-style-type: none">1、不使用参数映射, 筛选条件在展现层拖拽字段实现2、依赖数据源控制数据权限时慎用抽取3、ETL查询的使用需要相关许可



直连模式



抽取模式

- 在直连模式下，做数据模型参数映射后，才可在前端数据分析实现参数效果。
- 在抽取模式下，私有查询参数做参数映射时，会以数据模型参数的默认值过滤数据，过滤后的数据抽取到MPP中。此时前端数据分析时，只有过滤后数据，且无论参数有一个或多个默认值，都只作为一个参数选项，同时没有备选值；所以，抽取模式下，私有查询参数不能做参数映射。建议以私有查询参数默认值的所有值，抽取全部私有查询结果数据，前端数据分析通过筛选器实现参数效果。

分类	详细说明
创建模型	模型架构，以事实表为中心，围绕维表；
	维度尽可能设计层次，并且层次有顺序；
	隐藏事实表、部分维表或维表字段；
	表、字段都应有相关的描述信息；
设置参数	在直连模式下，私有查询参数做数据模型参数映射后，才可在前端数据分析实现参数效果；
	在抽取模式下，私有查询参数不能做参数映射。建议抽取私有查询全部数据，前端数据分析通过筛选器实现参数效果；
设置权限	数据模型设置数据权限；
性能优化	如果通过单表日期字段生成年、月，将需要很多计算资源，降低性能。建议事实表准备好单独的年、月字段。
	创建筛选器时设置默认值；
	建议事实表所有字段不要有空值；
	日期字段用日期类型，不要用字符串类型；
	表关系不建议做一对一关联，一对一关联时性能差；
	表连接查询时，大表在前小表在后；因为表连接查询时后面的表必然会进行全表扫描。



填写调查问卷, get 抽奖机会!



填写问卷, 参加抽奖



公众号



小麦微信

更聪明的大数据分析软件，

快速挖掘企业数据价值！



广州思迈特软件有限公司



愿景：让数据为客户创造价值



北京、上海、深圳等办事处



www.smartbi.com.cn



sales@smartbi.com.cn



公众号：思迈特Smartbi