

多维数据透视分析

01

多维数据模型概述

04

基本透视规则

02

多维数据模型创建方法

05

透视规则扩展

03

5W2H思维模型

06

多维透视分析应用

多维数据模型概述

01

以字段或记录作为数据的引用、操作及计算的基本单位的数据

- 字段：整列数
- 记录：整行数
- 维度：业务角度
- 度量：业务行为结果
- 维度字段：文本型
- 度量字段：数值型

表名：订单表

表名，用来区分识别订单表

维度字段

度量字段

订单ID	用户ID	付费时间	支付状态	订单金额	运费	用券抵扣金额
a001	Y0234	2020/11/11 20:00	已支付	1000	8	30
a002	Y0235	2020/11/11 20:01	已支付	600	8	60
a003	Y0236	2020/11/11 20:02	未支付	430	8	15
...
...

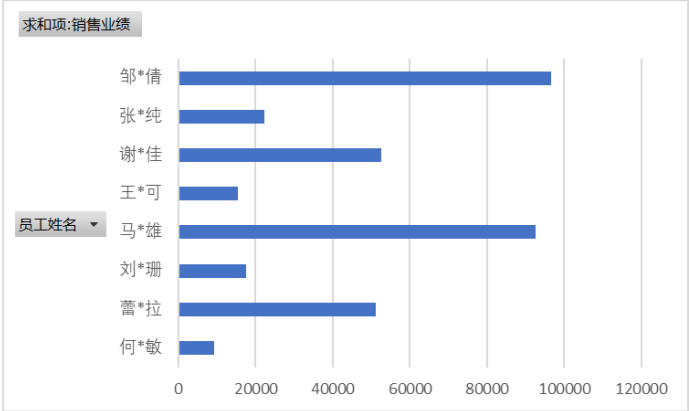
记录：一行记录
一笔不同的交易

数据透视 -- 对零散数据进行汇总分析

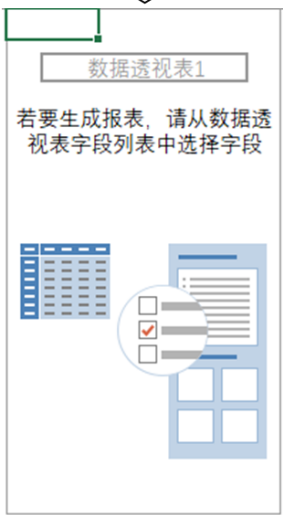
订单ID	员工姓名	销售业绩	销售提成
a001	邹*倩	75894	3794.7
a002	马*雄	92522	4626.1
a003	王*可	15457	772.85
a004	邹*倩	20698	1034.9
a005	刘*珊	17616	880.8
a006	张*纯	22371	1118.55
a007	谢*佳	52560	2628
a008	何*敏	9213	460.65
a009	蕾*拉	51142	2557.1

数据源：销售业绩表

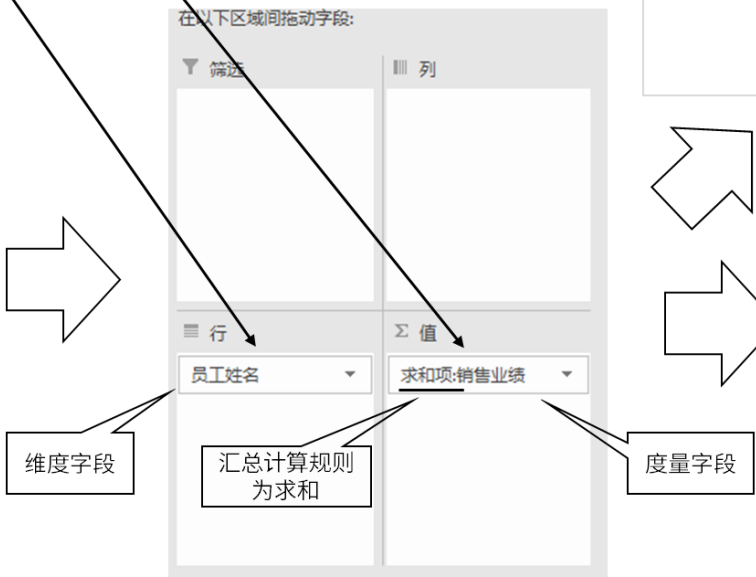
维度字段：员工姓名
度量字段：销售业绩
汇总计算规则：求和



数据透视图



数据透视表模板



设置维度、度量及汇总计算规则

行标签	求和项:销售业绩
何*敏	9213
蕾*拉	51142
刘*珊	17616
马*雄	92522
王*可	15457
谢*佳	52560
张*纯	22371
邹*倩	96592
总计	357473

数据透视表

复习 - 事实表及维度表

维度表：只包含维度信息的表

事实表：既包含维度信息又包含度量信息的表

事实表：销售表

订单号
销售日期
客户ID
销售金额
.....

事实表：采购表

进货单号
进货日期
进货金额
.....
.....

事实表：库存表

库存编码
产品编码
库存日期
库存金额
.....

维度表：产品表

产品编号
产品名称
品牌编号
.....
.....

维度表：品牌表

品牌编号
品牌名称
.....
.....
.....

维度表：客户表

客户编号
客户名称
公司规模
所属行业
.....

一个表中有且只有一个主键

物理意义

单字段主键：由一个字段构成的主键

多字段联合主键：由多个字段构成的主键

非空不重复

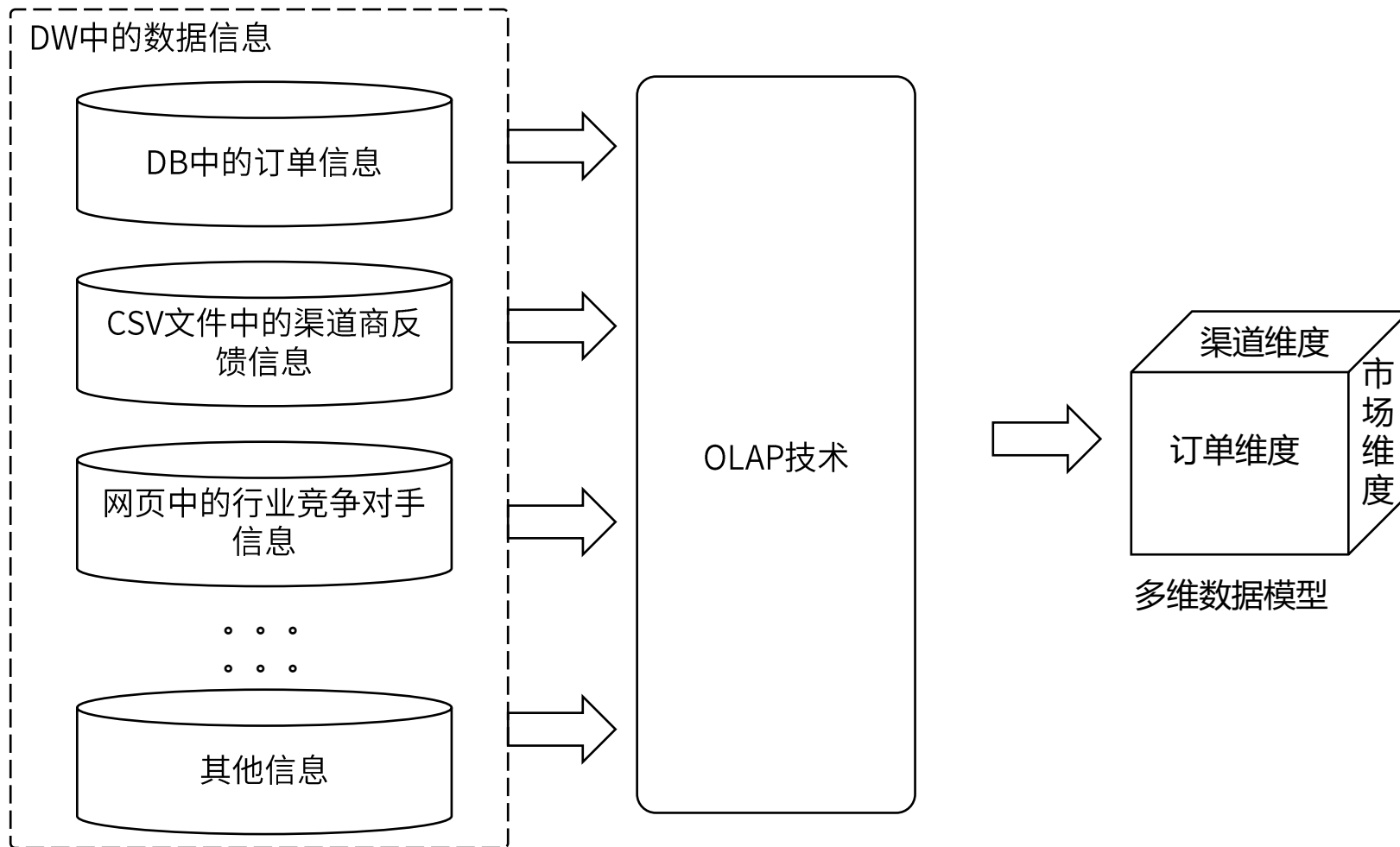
定位记录行、字段名+主键值定位具体数值

多以“xxID”、“xxNo”、“xx编号”等名称命名

业务意义

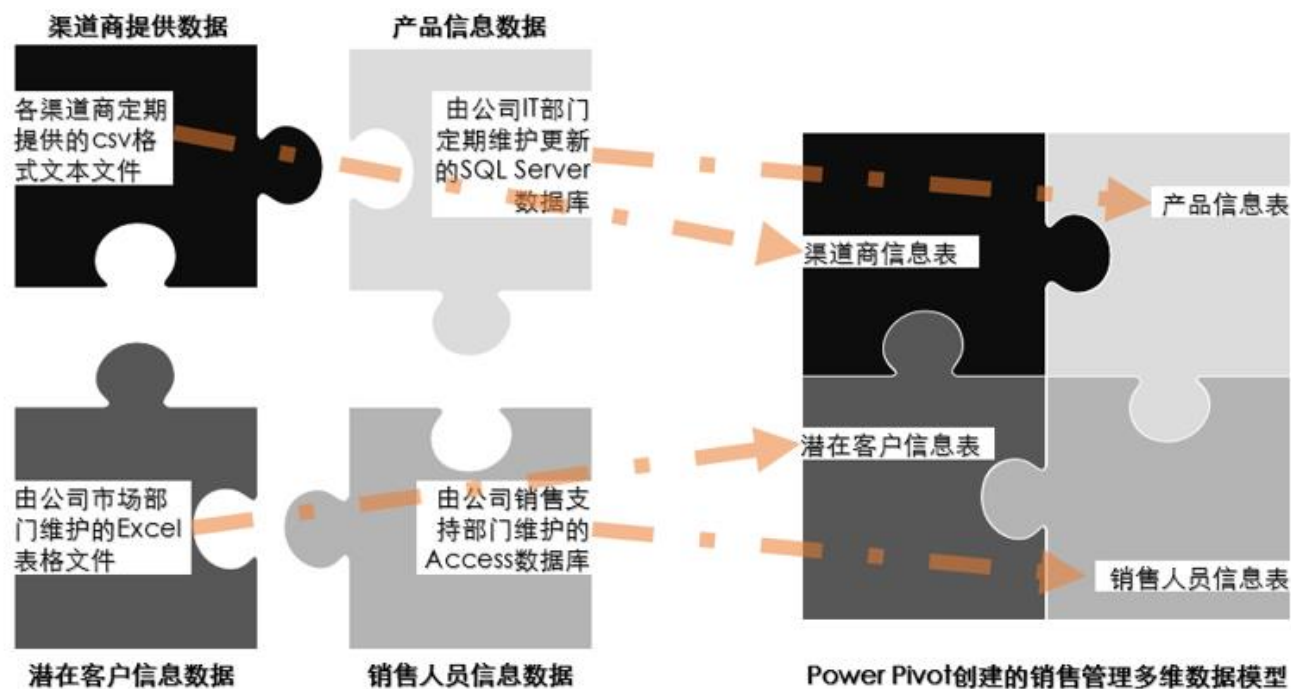
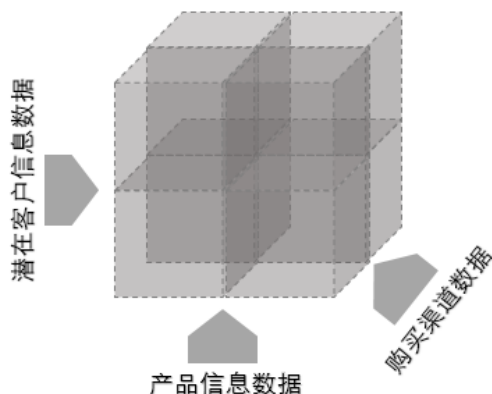
表的业务记录单位。在一个数据表中的所有非主键字段都要围绕主键展开

连接信息孤岛、创建多维数据模型



多维数据模型又叫多维数据集、立方体，指的是相互间通过某种联系被关联在一起的不同类别的数据集合

多维数据模型：在咨询公司以及BI工具厂商的介绍性资料中又被称为“立方体（Cube）”，在这些资料中常以一个立体正方形的形式出现。多维数据集可以从多角度用数据全面映射某种业务的实际状况。



多维数据模型 创建方法

02

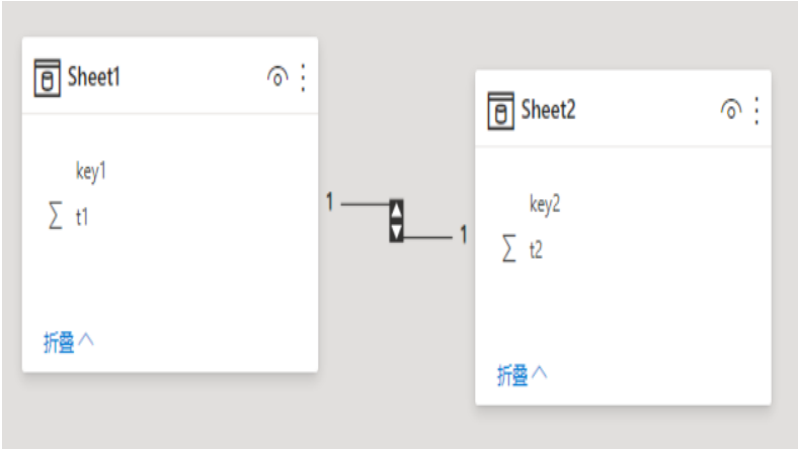
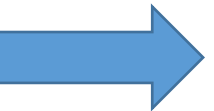
相邻两表间连接汇总

通过公共字段连接两表，选择不同表中字段分别做为维度、度量，选择汇总计算规则

key1	t1
a	1
b	2
c	3



key2	t2
a	4
b	5
c	6



key1	t2
c	6
b	5
a	4
总计	15

影响连接汇总的三要素

筛选器方向、对应关系、汇总角色

筛选器方向

分为单向及双向两种、筛选器方向决定维度与度量的出处

对应关系

分为三类，一对一、多对一与多对多，决定连接汇总的结果

汇总角色

维度、度量

筛选器方向

筛选器方向决定维度与度量的出处、不可逆筛选器方向做筛选

单向

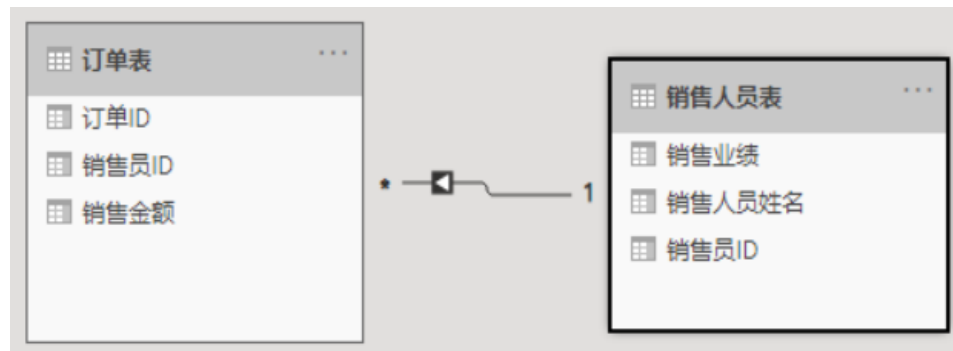
箭头出发一侧为维度、指向一侧为度量

订单表

订单ID	销售员ID	销售金额
a	S1	1
b	S2	2
d	S3	3
e	S3	4

销售人员表

销售员ID	销售人员姓名	销售业绩
S1	赵大	1
S2	王二	2
S3	张三	7



正确

销售人员姓名	销售金额
王二	2
张三	7
赵大	1
总计	10

错误

销售员ID	销售业绩 的总和
S1	10
S2	10
S3	10
总计	10

双向

两表间互为筛选

OLAP与数据库连接汇总的逻辑区别(1)

OLAP连接汇总时，哪个表提供度量，哪个表为主表

订单表

订单ID	销售员ID	销售金额
a	S1	1
b	S2	2
d	S3	3
e	S3	3

销售人员表

销售员ID	销售人员姓名
S1	赵大
S2	王二
S3	张三
S4	李四



```
select 销售人员姓名, sum(销售金额) as 销售金额
from 订单表 left join 销售人员表 on 订单表.销售人员ID = 销售人员表.销售人员ID
group by 销售人员姓名
```

OLAP与数据库连接汇总的逻辑区别(2)

OLAP连接汇总时，哪个表提供度量，哪个表为主表

订单表

订单ID	销售员ID	销售金额
a	S1	1
b	S2	2
d	S3	3
e	S5	3

销售人员表

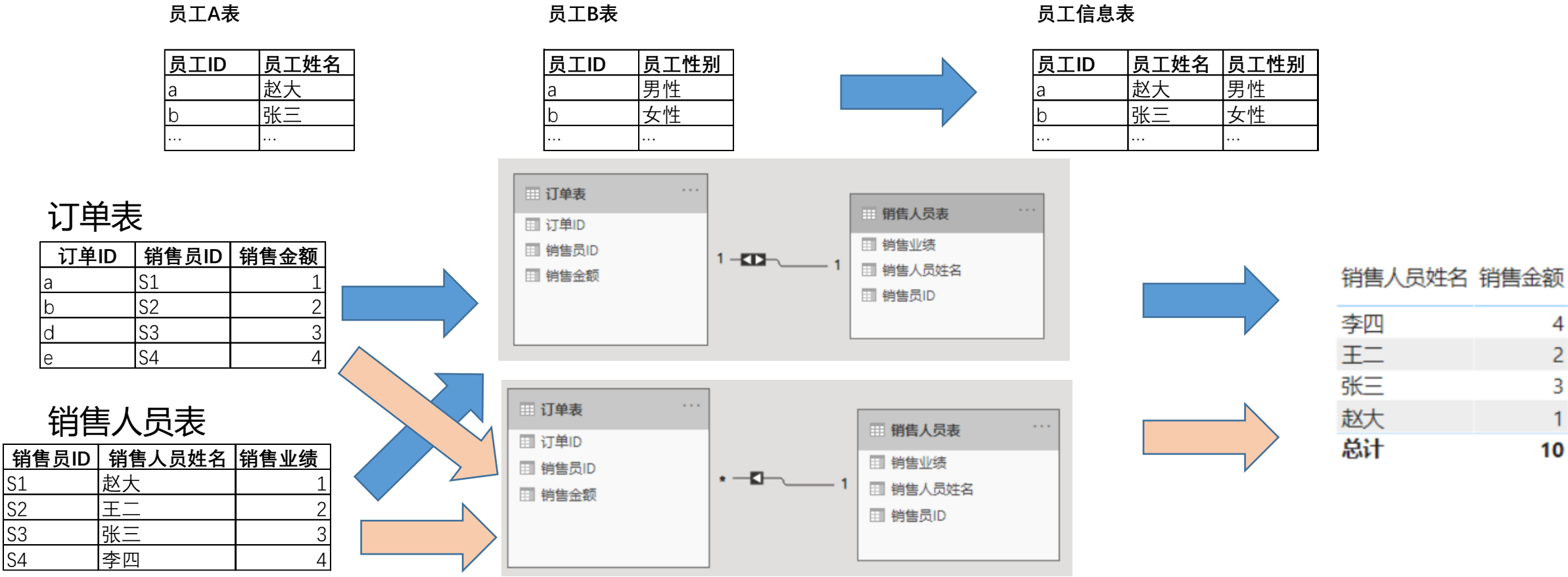
销售员ID	销售人员姓名
S1	赵大
S2	王二
S3	张三
S4	李四



```
select 销售人员姓名, sum(销售金额) as 销售金额
from 订单表 left join 销售人员表 on 订单表.销售员ID = 销售人员表.销售员ID
group by 销售人员姓名
```

三种对应关系 - 1 对 1

主键与主键相连、两表具有相同主键，这种情况在实际场景中几乎不会出现，如果出现物理层面上的1对1关系连接，应该结合字段属性进行确认，如果不是主键与主键相连的一对一关系应该及时进行调整



三种对应关系 - 多对多

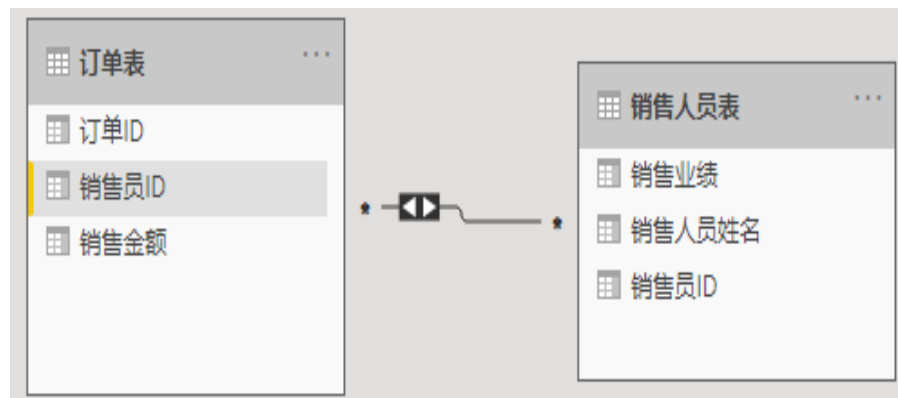
非主键连接非主键，虽然会出现这种连接情况，但是会造成度量值在求和、计数等常用汇总规则下翻倍，应尽量避免使用多对多的连接关系

订单表

订单ID	销售员ID	销售金额
a	S1	1
b	S2	2
d	S3	3
e	S3	4

销售人员表

销售员ID	销售人员姓名	销售业绩
S1	赵大	1
S2	王二	2
S3	张三	3
S3	李四	4



销售人员姓名	销售金额
李四	7
王二	2
张三	7
赵大	1

三种对应关系 - 多对一（单向）

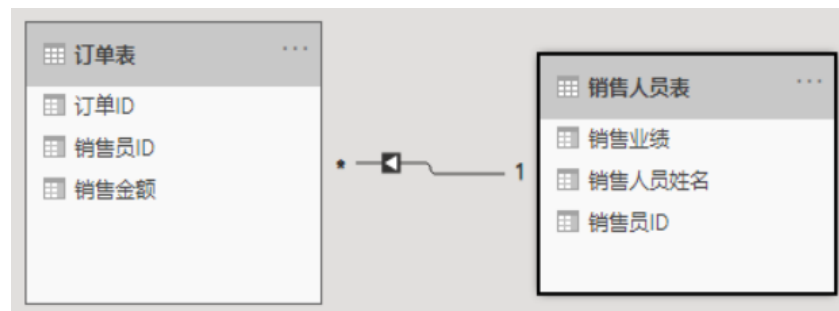
相邻两表连接时应尽量使用多对一的关系，单向筛选方向时，一表筛选多表

订单表

订单ID	销售员ID	销售金额
a	S1	1
b	S2	2
d	S3	3
e	S3	4

销售人员表

销售员ID	销售人员姓名	销售业绩
S1	赵大	1
S2	王二	2
S3	张三	3



销售人员姓名 销售金额

王二	2
张三	7
赵大	1

三种对应关系 - 多对一（双向）

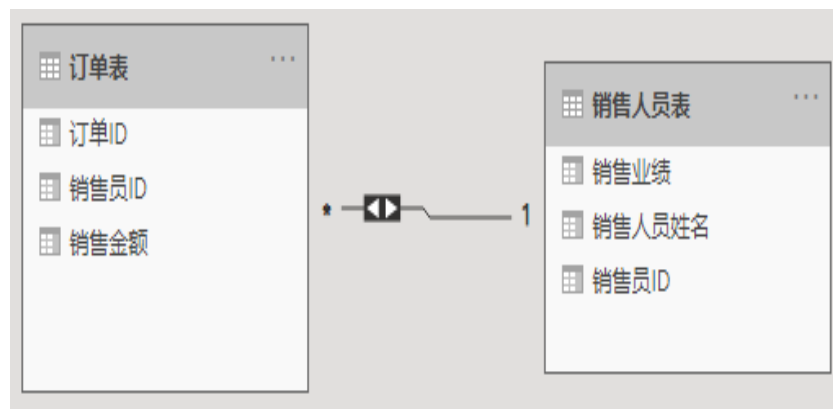
相邻两表连接时应尽量使用多对一的关系，双向筛选方向时，多表可以筛选一表，但是筛选方式不同于一表筛选多表的筛选方式

订单表

订单ID	销售员ID	销售金额
a	S1	1
b	S2	2
d	S3	3
e	S3	4

销售人员表

销售员ID	销售人员姓名	销售业绩
S1	赵大	1
S2	王二	2
S3	张三	3



销售员ID 销售业绩 的总和

S1	1
S2	2
S3	3

两种汇总方式

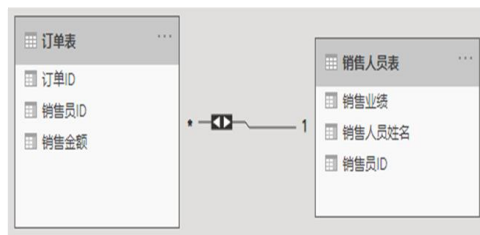
类型一：合并维度，汇总度量、维度筛选度量

订单表

订单ID	销售员ID	销售金额
a	S1	1
b	S2	2
d	S3	3
e	S3	4

销售人员表

销售员ID	销售人员姓名	销售业绩
S1	赵大	1
S2	王二	2
S3	张三	3



销售人员姓名 销售金额

销售人员姓名	销售金额
王二	2
张三	7
赵大	1

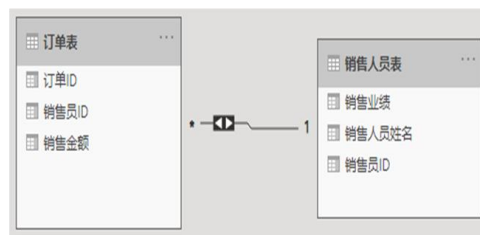
类型二：合并维度，汇总维度下不同公共字段对应的度量、维度选取公共字段，公共字段筛选度量

订单表

订单ID	销售员ID	销售金额
a	S1	1
b	S2	2
d	S3	3
e	S3	4

销售人员表

销售员ID	销售人员姓名	销售业绩
S1	赵大	1
S2	王二	2
S3	张三	3



销售员ID 销售业绩 的总和

销售员ID	销售业绩 的总和
S1	1
S2	2
S3	3

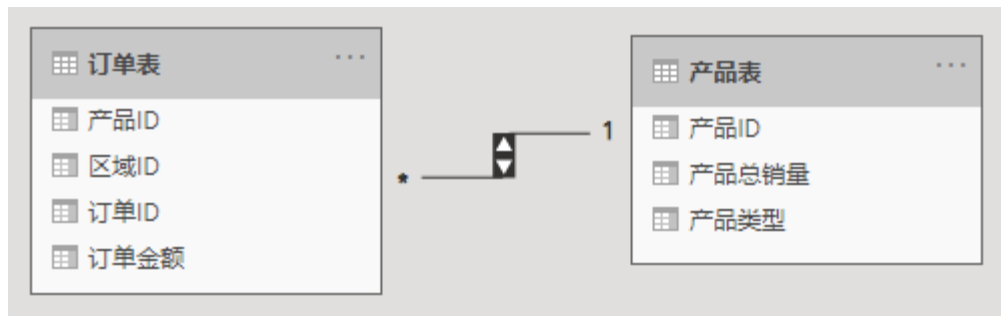
类型二进一步说明

订单表

订单ID	产品ID	区域ID	订单金额
101	a	e01	1000
102	a	e02	1000
103	b	e03	1000
104	b	e04	1000
105	b	e05	1000
106	c	e06	1000
107	d	e07	1000
108	d	e08	1000
109	d	e09	1000
110	e	e10	1000
111	e	e11	1000
112	f	e12	1000
113	f	e13	1000
114	f	e14	1000

产品表

产品ID	产品类型	产品总销量
a	D型	100
b	D型	100
c	S型	100
d	S型	100
e	S型	100
f	O型	100



区域ID	产品总销量的总和
e01	100
e02	100
e03	100
e04	100
e05	100
e06	100
e07	100
e08	100
e09	100
e10	100
e11	100
e12	100
e13	100
e14	100
总计	600

相邻两表间应尽量使用一对多的连接关系，汇总时应遵循一表出维度，多表出度量的规则

	双向		单向	
	维度	值	维度	值
多表	●	X	△	■
一表	X	●	■	△

-
- 值的结果为多表下一表的不同情况的汇总值
- X
- 值的结果为多表下值字段的汇总值
- △
- 无法正确汇总值
-
- 与X相同，为多表下值字段的汇总值

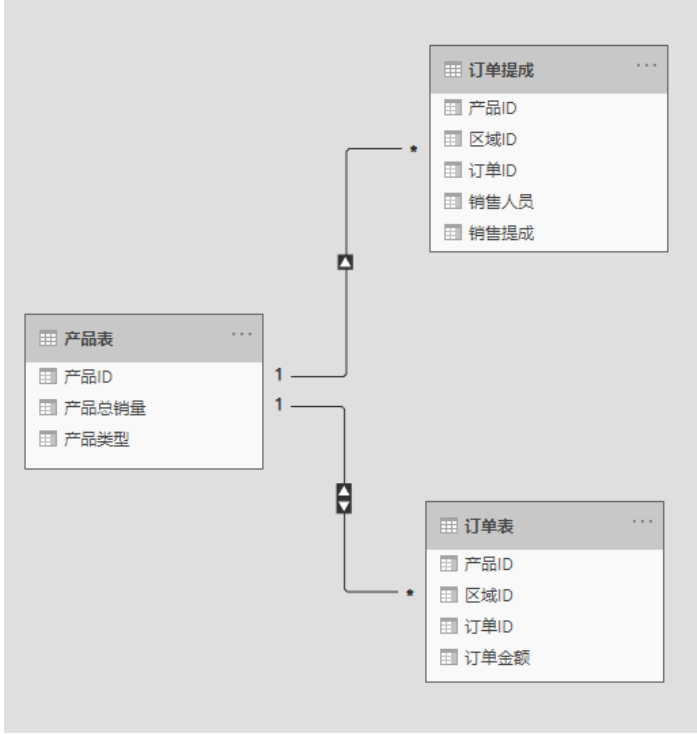
跨表筛选（筛选路径不通时无法正确汇总）

订单表

订单ID	产品ID	区域ID	订单金额
101	a	e01	1000
102	a	e02	1000
103	b	e03	1000
104	b	e04	1000
105	b	e05	1000
106	c	e06	1000
107	d	e07	1000
108	d	e08	1000
109	d	e09	1000
110	e	e10	1000
111	e	e11	1000
112	f	e12	1000
113	f	e13	1000
114	f	e14	1000

订单提成表

订单ID	销售人员	产品ID	区域ID	销售提成
101	赵大	a	e01	50
102	赵大	a	e02	50
103	赵大	b	e03	50
104	赵大	b	e04	50
105	赵大	b	e05	50
106	赵大	c	e06	50
107	赵大	d	e07	50
108	王二	d	e08	50
109	王二	d	e09	50
110	王二	e	e10	50
111	王二	e	e11	50
112	孙三	f	e12	50
113	孙三	f	e13	50
114	孙三	f	e14	50

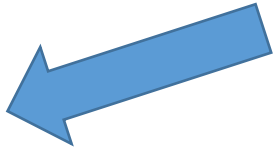


产品表

产品ID	产品类型	产品总销量
a	D型	100
b	D型	100
c	S型	100
d	S型	100
e	S型	100
f	O型	100

销售人员 订单金额

孙三	14000
王二	14000
赵大	14000
总计	14000



跨表筛选（筛选路径通时汇总规则为类型二）

订单表

订单ID	产品ID	区域ID	订单金额
101	a	e01	1000
102	a	e02	1000
103	b	e03	1000
104	b	e04	1000
105	b	e05	1000
106	c	e06	1000
107	d	e07	1000
108	d	e08	1000
109	d	e09	1000
110	e	e10	1000
111	e	e11	1000
112	f	e12	1000
113	f	e13	1000
114	f	e14	1000

订单提成表

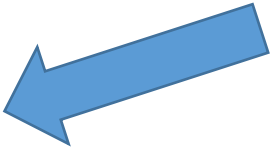
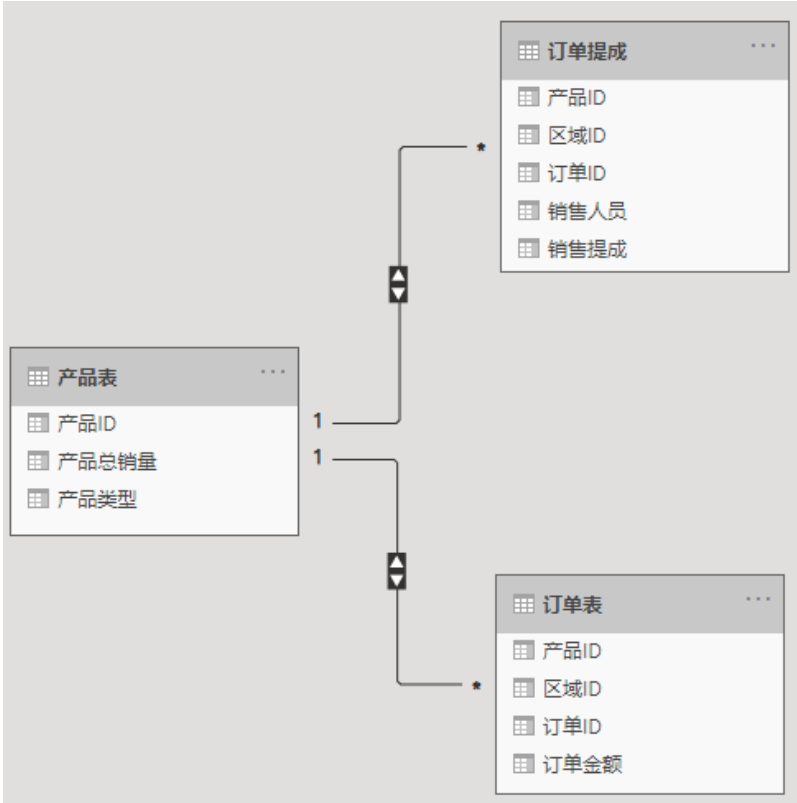
订单ID	销售人员	产品ID	区域ID	销售提成
101	赵大	a	e01	50
102	赵大	a	e02	50
103	赵大	b	e03	50
104	赵大	b	e04	50
105	赵大	b	e05	50
106	赵大	c	e06	50
107	赵大	d	e07	50
108	王二	d	e08	50
109	王二	d	e09	50
110	王二	e	e10	50
111	王二	e	e11	50
112	孙三	f	e12	50
113	孙三	f	e13	50
114	孙三	f	e14	50

产品表

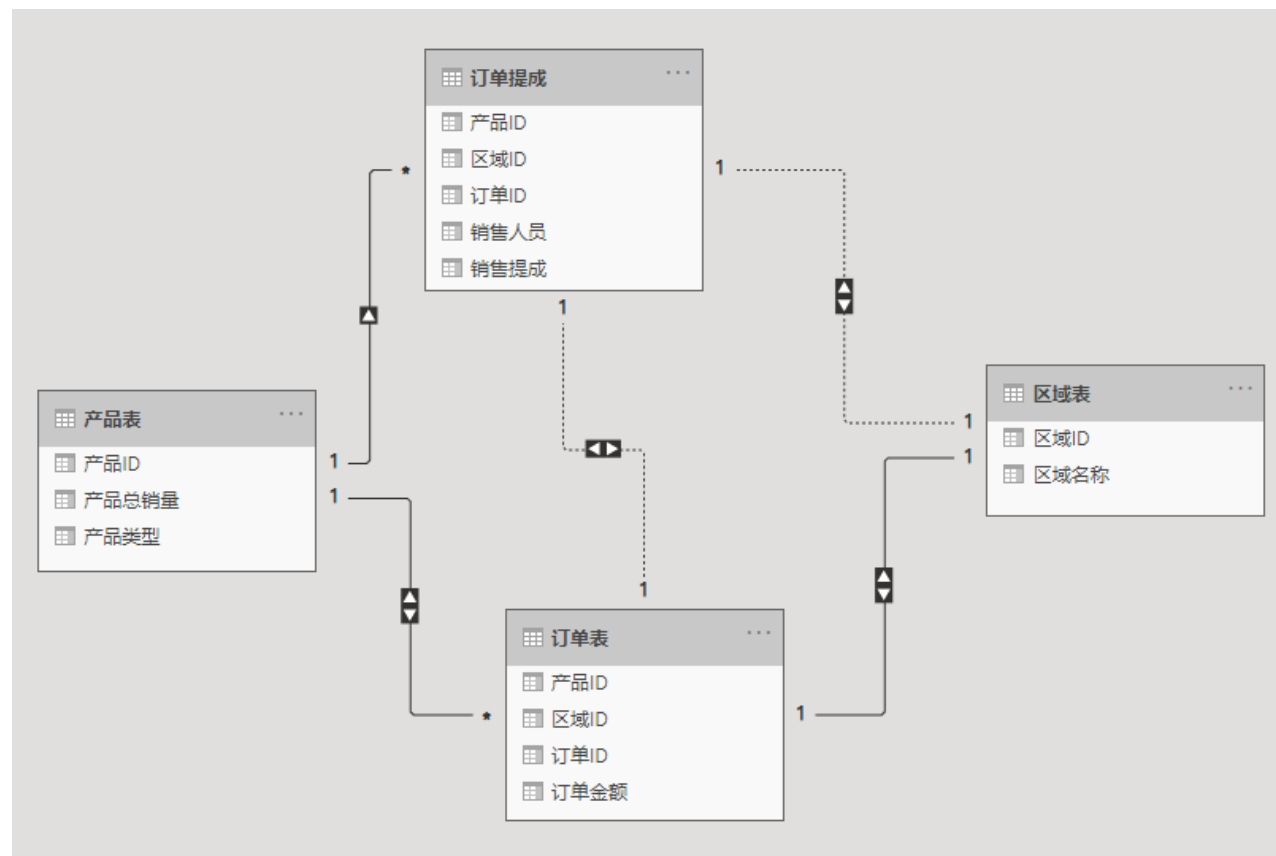
产品ID	产品类型	产品总销量
a	D型	100
b	D型	100
c	S型	100
d	S型	100
e	S型	100
f	O型	100

销售人员 订单金额

孙三	3000
王二	5000
赵大	9000
总计	14000



交叉连接下只有一条路径为有效路径

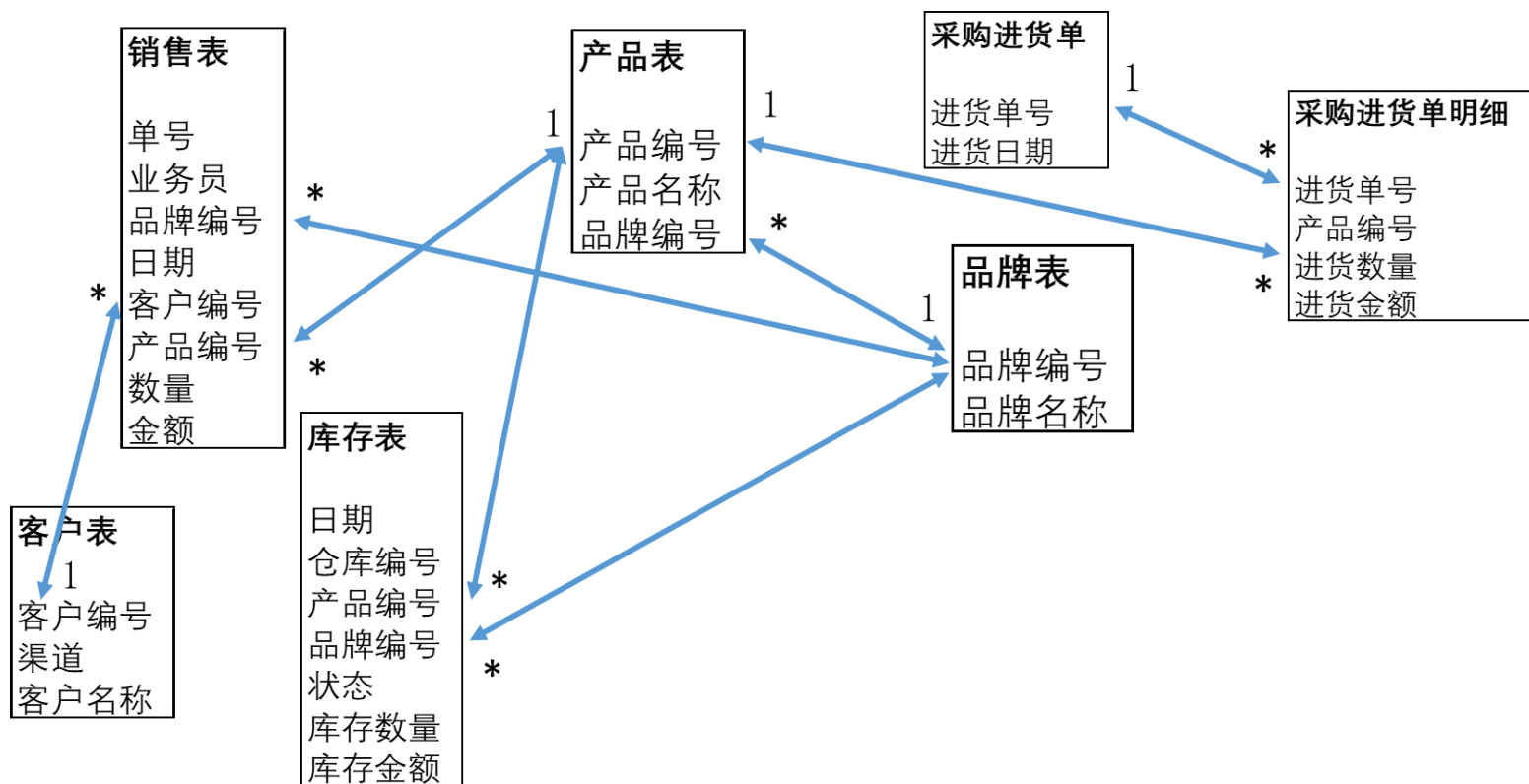


三种模式

星型模式：一个事实表和多个
维度表相连

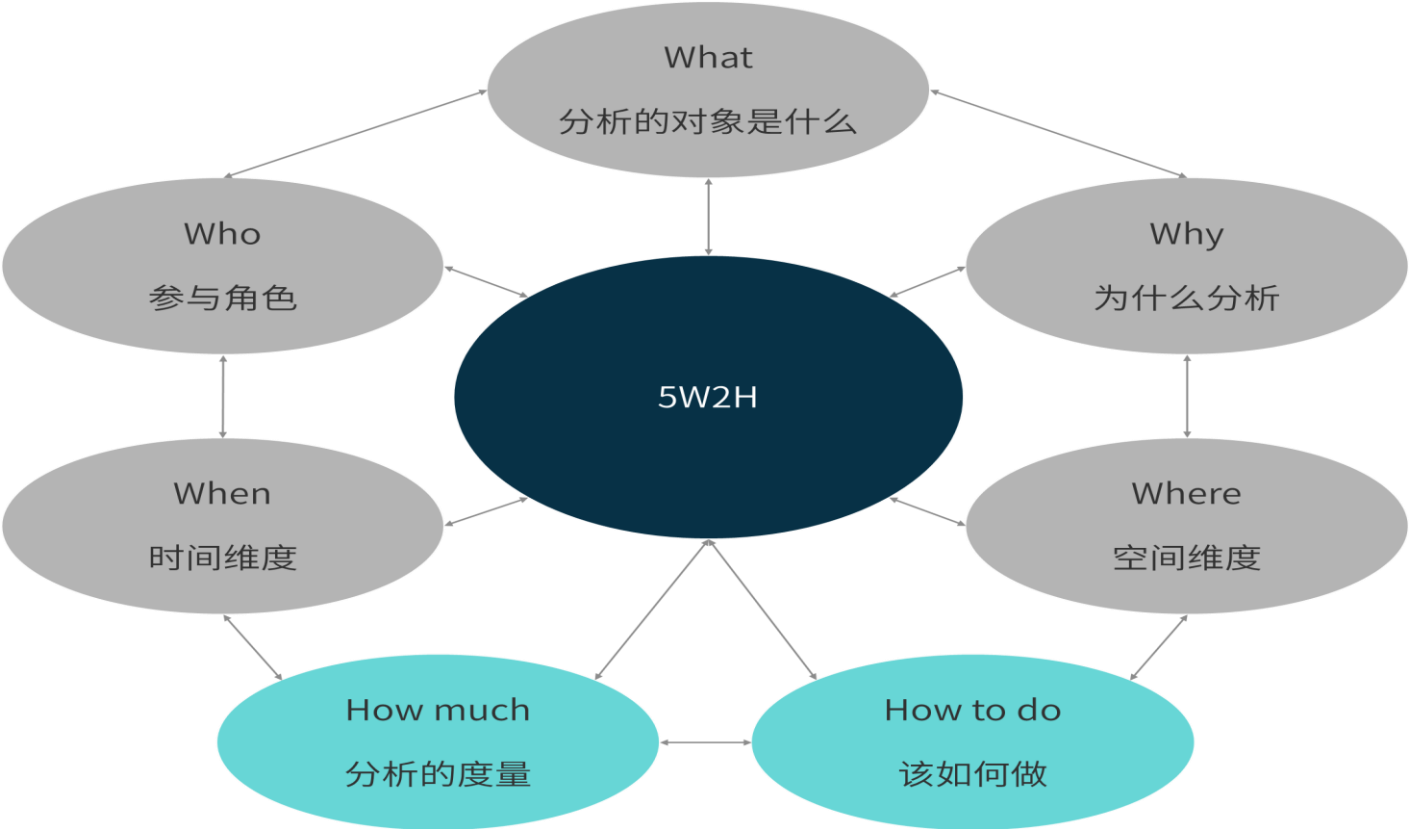
雪花模式：维度表与维度表相
连，进行维度的扩展

星座模式：多个事实表共用某
些维度表

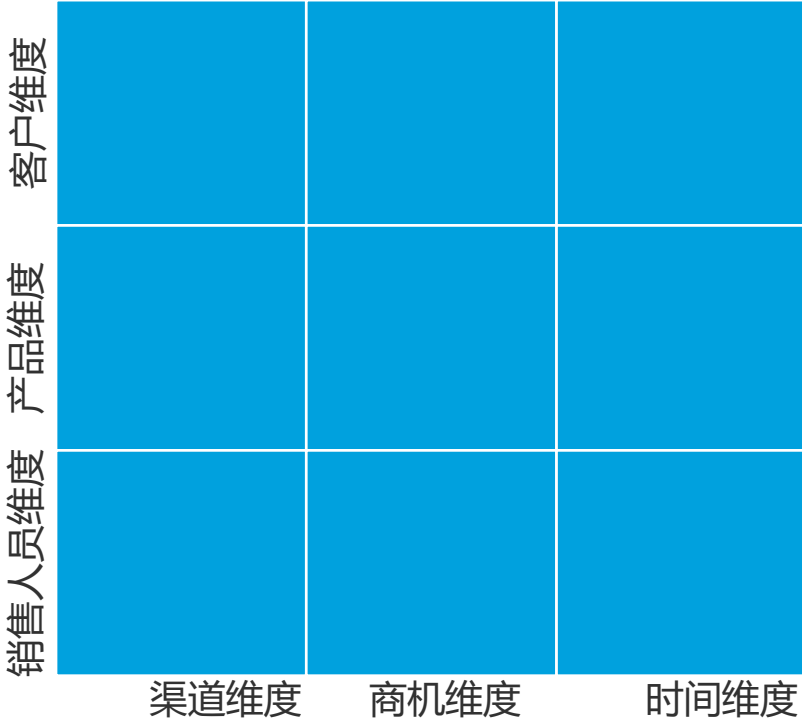
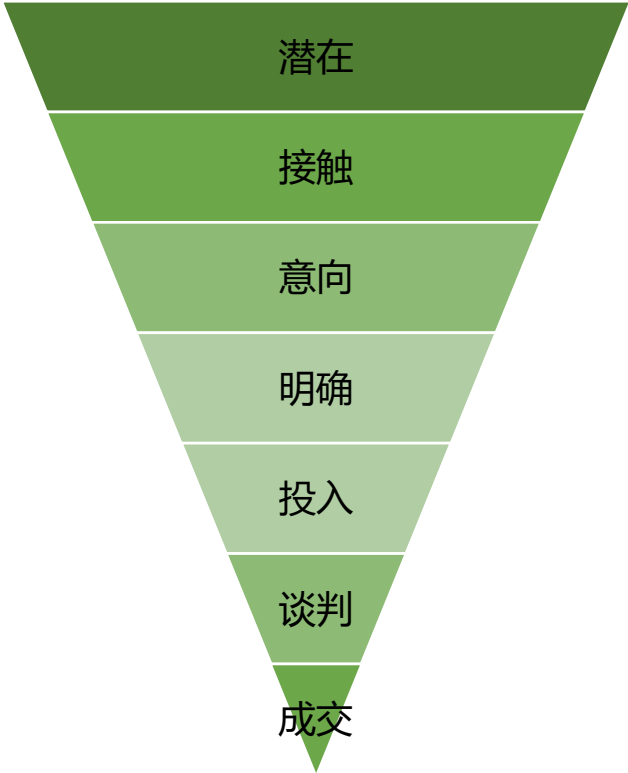


5W2H思维 模型

03



销售漏斗模型



5W2H分析与销售管理分析

What

销售进度管理

Why

发现并控制销售阶段风险

Where

销售地点

Who

销售人员、客户

When

时间维度

How Much

商机金额、商机个数、商机规模等

How to do

如何制定销售策略

5W2H分析与销售管理分析

What	销售进度管理 → 商机维度（商机规模、商机号、商机来源等）
Why	发现并控制销售阶段风险 → 商机维度（销售阶段、上周销售阶段、赢单率等）
Where	销售地点 → 销售大区、销售城市、销售区域等
Who	销售人员、客户 → 商机发现者、商机管理者、销售人员能力、销售人员成本等，客户负责人、与客户以往交易情况、客户需求等
When	时间维度 → 创建商机日期、预计签约日期、商机停留时间、商机状态变化时间、客户需求变更日期等
How Much	商机金额、商机个数、商机规模等
How to do	如何制定销售策略