

# 数据挖掘-合并列/合并行

将两张表的数据按列或按行合并，组成新表。

- 合并列
- 合并行

## 合并列

### 概述

合并列是将两张表的数据按列合并，组成新表，行数不同时，缺失部分则自动填充空值。



### 输入/输出

输入	两个输入端口，用于接收两个数据集。
输出	一个输出端口，用于输出两个数据集按列合并后的结果。

### 参数设置

设置合并列的参数：

>> 参数

属性

帮助

选择左表字段 \*必填 ?

选择左表字段

选择右表字段 \*必填 ?

选择右表字段

设置说明如下：

参数	说明
选择左表字段	<div>用于选择通过左侧输入端口连接的前置节点资源输出的数据集的列：<div><div>选择左表字段</div><div><div>源数据列表0/12</div><div>已选字段列表0/3</div></div></div></div>

选择右表字段

用于选择通过右侧输入端口连接的前置节点资源输出的数据集的列：

✦ 选择右表字段

源数据列表0/3

Q 请输入搜索内容

☐ #订单编号

☐ #产品编号

☐ #折扣

全部

字符

数字

到右边>

< 到左边

已选字段列表0/2

Q 请输入搜索内容

☐ #单价

☐ #数量

确定

取消

示例

1、将两个关系数据源的列进行合并，选择左表字段“ShipName、ShipCity、ShippedDate”，选择右表字段“UnitPrice、Quantity”。

实验管理 | 服务管理 | 模型管理 | 模型自学习 | 模型批量预测 | 自定义模块管理 | 合并并 | 合并并 x

数据源

目标源

已训练模型

数据预处理

采样

拆分

过滤与映射

列选择

空值处理

**合并并**

合并并

元数据编辑

JOIN

行选择

去除重复值

排序

增加序列号

聚合

分列

派生列

关系数据源

关系数据源

合并并

初始化

参数

属性

帮助

选择左表字段 \*必填 ?

选择左表字段

选择右表字段 \*必填 ?

选择右表字段

Q

Q

Q

<

>

初始化

保存

删除

上一步

下一步

刷新

帮助

搜索

文档

2、来源于左表和右表的字段按顺序组成一张新的表。

当前显示 5 列 / 总共 5 列， 100 条 / 总共有 2154 条数据

列筛选 请选择

表头真名 ☒ 表头别名

ShipName	ShipCity	ShippedDate	UnitPrice	Quantity
周先生	上海	2013-10-15 00:00:00	8.0000	15
鍾小姐	天津	2013-10-21 00:00:00	27.2000	60
陈先生	厦门	2013-10-20 00:00:00	28.8000	20
方先生	天津	2013-11-07 00:00:00	77.6000	20
王先生	常州	2013-11-18 00:00:00	4.8000	2
王俊元	石家庄	2013-10-21 00:00:00	12.0000	8
王先生	成都	2013-10-23 00:00:00	12.0000	20
方先生	张家口	2013-11-05 00:00:00	7.3000	4
刘先生	上海	2013-11-20 00:00:00	24.9000	30

提示：点击单元格可查看超出的内容。注意：表头中 表示特征列， \* 表示标签列

下载预览数据

## 合并行

### 概述

合并行是将多张表的数据按行合并成一个新表，缺失的列以“null”值填充，如示例效果。



### 输入/输出

输入	六个输入端口，用于接收多个数据集。
输出	一个输出端口，用于输出两个数据集按行合并后的结果。

### 参数配置

设置合并行的参数：

» 参数

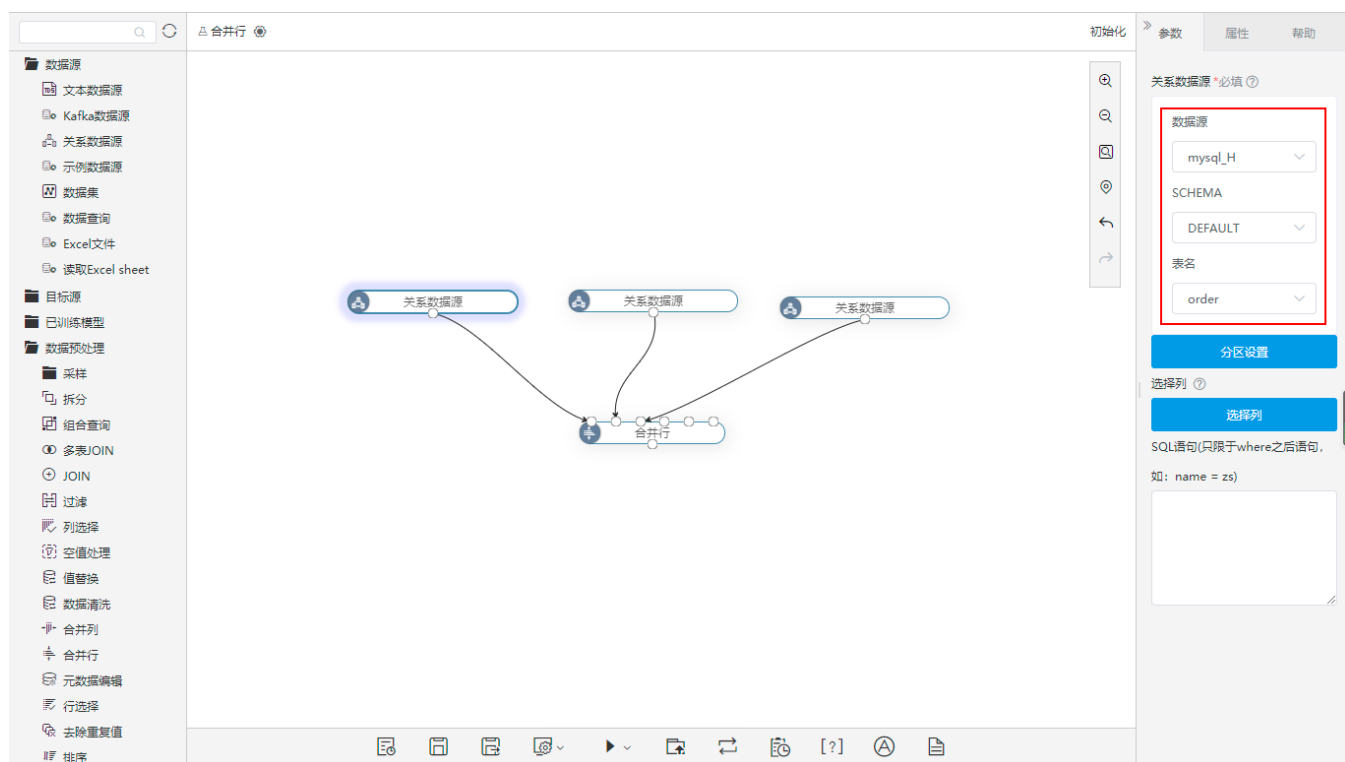
属性帮助

选择列 \*必填

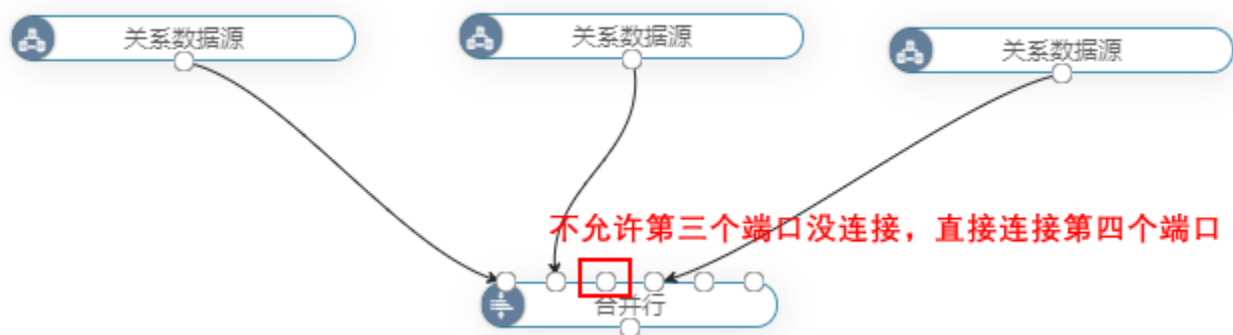
选择列

参数面板用于设置各表的字段对应关系：单击 选择列 按钮，在弹出的窗口进行对应关系处理。





⚠ 多个数据源必须从“合并行”输入端口第一个依次连接，不允许跨端口，例如以下设置是错误示范：



3、在画布中分别选定三个“关系数据源”节点，从左至右依次在参数面板中选择表“order”、“product”、“cust”。

4、依次执行这三个“关系数据源”节点，执行成功后，选定“合并行”节点，在参数面板中单击 **选择列**，弹出“合并行”设置窗口，按照如下进行设置：



5、设置完成后，执行“合并行”节点，执行成功后，可以通过查看输出，浏览效果：

查看输出

当前显示 3 列 / 总共 3 列。13 条 / 总共有 13 条数据

列筛选 请选择

表头真名 ☒ 表头别名

# 编码	名称	# 销量
11	油条	315.0
12	炒粉	46.0
13	炒面	53.0
14	馄饨	73.0
15	酸奶	56.0
1	包子	321.0
2	饺子	426.0
3	面包	65.0
4	豆浆	62.0
5	牛奶	79.0
21	粽子	null
22	月饼	null
23	蛋糕	null

提示：点击单元格可查看超出的内容。注意：表头中 表示特征列， 表示标签列

下载预览数据