

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

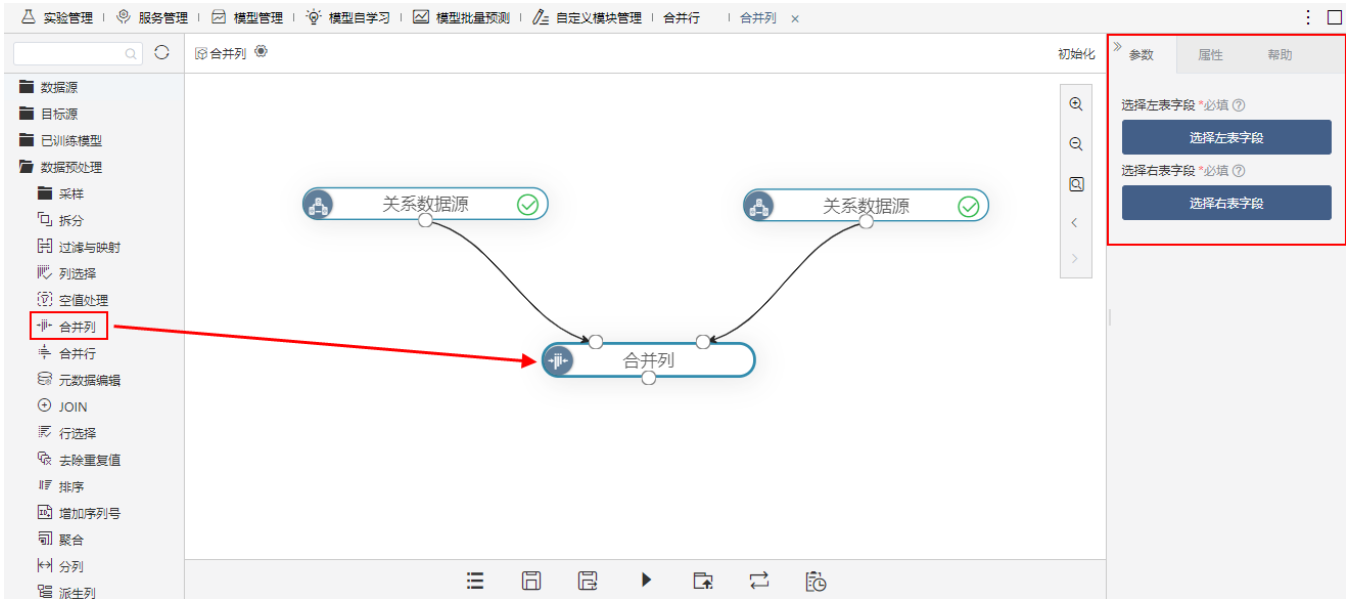
设置说明如下：

参数	说明
----	----

选择左表字段	<p>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：</p> <div><div><div>✦ 选择左表字段</div><div><div>源数据列表0/12</div><div><div>Q 请输入搜索内容</div><div><div><div><input type="checkbox"/> # 订单编号</div><div><input type="checkbox"/> A₀ 客户编号</div><div><input type="checkbox"/> # 雇员编号</div><div><input type="checkbox"/> A₀ 订单日期</div><div><input type="checkbox"/> A₀ 到货日期</div><div><input type="checkbox"/> # 运货商</div></div></div></div><div><div>< 到左边</div><div>到右边 ></div></div><div><div>已选字段列表0/3</div><div><div>Q 请输入搜索内容</div><div><div><input type="checkbox"/> A₀ 发货日期</div><div><input type="checkbox"/> A₀ 发货人</div><div><input type="checkbox"/> A₀ 发货城市</div></div></div><div><div>确定</div><div>取消</div></div></div></div></div></div>
选择右表字段	<div><div><div>✦ 选择右表字段</div><div><div>源数据列表0/3</div><div><div>Q 请输入搜索内容</div><div><div><input type="checkbox"/> # 订单编号</div><div><input type="checkbox"/> # 产品编号</div><div><input type="checkbox"/> # 折扣</div></div></div></div><div><div>< 到左边</div><div>到右边 ></div></div><div><div>已选字段列表0/2</div><div><div>Q 请输入搜索内容</div><div><div><input type="checkbox"/> # 单价</div><div><input type="checkbox"/> # 数量</div></div></div><div><div>确定</div><div>取消</div></div></div></div></div>

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”。



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 100 条 / 总共有 2154 条数据 提示:点击单元格可查看超出的内容

ShipName	ShipCity	ShippedDate	UnitPrice	Quantity
余小姐	北京	2016-07-16 00:00:00.0	18.6000	9
谢小姐	济南	2016-07-10 00:00:00.0	42.4000	40
谢小姐	秦皇岛	2016-07-12 00:00:00.0	7.7000	10
陈先生	南京	2016-07-15 00:00:00.0	42.4000	35
刘先生	长春	2016-07-11 00:00:00.0	16.8000	15
谢小姐	长治	2016-07-16 00:00:00.0	16.8000	6
林小姐	武汉	2016-07-23 00:00:00.0	15.6000	15
方先生	北京	2016-07-15 00:00:00.0	16.8000	20
何先生	济南	2016-07-17 00:00:00.0	64.8000	40
王先生	上海	2016-07-22 00:00:00.0	2.0000	25
王先生	济南	2016-07-23 00:00:00.0	27.2000	40

注意: 表头中*表示特征列, #表示标签列

表头真名 ☒ 表头别名

合并行

概述

合并行是将两张表的数据按行合并，合并成为一个新表，左表及右表选择输出的字段个数、类型应保持一致。



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

选择左表字段 *必填 ②

选择左表字段

选择右表字段 *必填 ②

选择右表字段

合并方式

并集

设置说明如下：

参数	说明
选择左表字段	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：

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：

选择左表字段

源数据列表0/0

Q 请输入搜索内容

无数据

< 到左边

到右边 >

已选字段列表0/6

Q 请输入搜索内容

☐ #id

☐ #Sepal_Length

☐ #Sepal_Width

☐ #Petal_Length

☐ #Petal_Width

☐ Aspecies

确定

取消

选择右表字段

源数据列表0/0

Q 请输入搜索内容

无数据

< 到左边

到右边 >

已选字段列表0/6

Q 请输入搜索内容

☐ #id

☐ #Sepal_Length

☒ #Sepal_Width

☐ #Petal_Length

☐ #Petal_Width

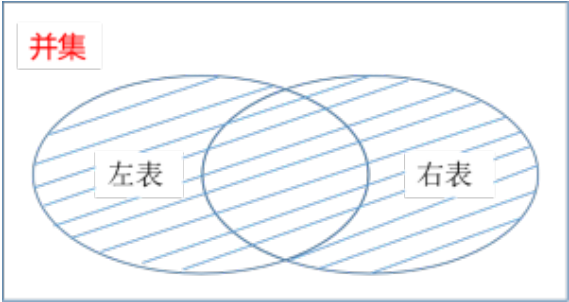
☐ Aspecies

确定

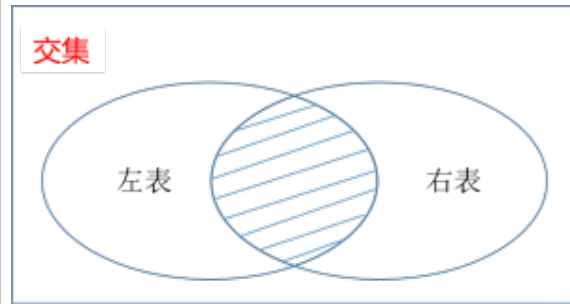
取消

合并方式

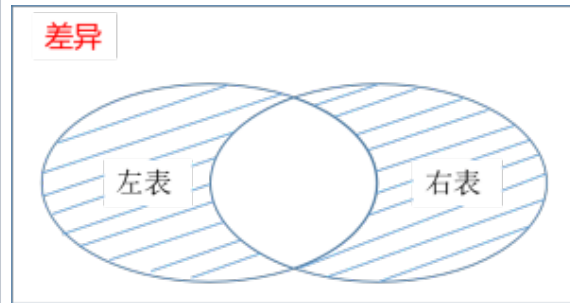
并集：子结果集所有的记录组合在一起形成新的结果集。如下图阴影部分。



交集：返回左边结果集和右边结果集中都有的记录。如下图阴影部分。



差异：返回左表和右表并集除去两表交集，如下图阴影部分。



示例

1、对两个数据源进行行合并，选择左表字段“id、number”，选择右表字段“id、number、count”，合并方式为“并集”。

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

合并方式

并集

2、合并时以左表字段为基准，将右表字段按顺序合并到左表字段中。当左表没有的字段“count”时，直接填充空值。

左表没有的字段，直接填充空值

id	number	count
7	4	null
8	null	null
9	6	null
10	27	null
1	1	100
2	2	null
3	3	50
4	3	null
5	4	30
6	4	null
7	4	null

左表字段

右表字段