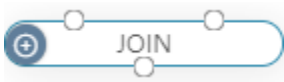


数据挖掘-JOIN

概述

JOIN是基于连接字段和给定的连接方式，进行两个数据集字段的组合后得到新的数据表。
支持两个数据表的单个或多个字段为连接字段，连接方式包括左连接、右连接、内连接、全连接。

- [概述](#)
- [输入/输出](#)
- [参数设置](#)
- [示例](#)



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，用于输出连接后的结果。

参数设置

设置JOIN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

连接类型

内连接

关联条件

请选择 = 请选择

+

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><div><div>✦选择左表字段</div><div><div><div>源数据列表0/26</div><div><div>Q 请输入搜索内容</div><div><div><div><input type="checkbox"/> #序号</div><div><input type="checkbox"/> #供应商编码</div><div><input type="checkbox"/> A#年</div><div><input type="checkbox"/> A#月</div><div><input type="checkbox"/> A#区域</div><div><input type="checkbox"/> A#省份</div></div></div><div><div>全部</div><div>字符</div><div>数字</div></div></div><div><div>到右边></div><div><到左边</div></div><div><div>已选字段列表0/6</div><div><div>Q 请输入搜索内容</div><div><div><div><input type="checkbox"/> A#顾客编号</div><div><input type="checkbox"/> A#产品大类名称</div><div><input type="checkbox"/> A#产品名称</div><div><input type="checkbox"/> #订单编号</div><div><input type="checkbox"/> #数量</div><div><input type="checkbox"/> #销售额</div></div></div><div><div>确定</div><div>取消</div></div></div></div></div></div></div></div>
选择右表字段	<div>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：</div> <div><div><div>✦选择右表字段</div><div><div><div>源数据列表0/10</div><div><div>Q 请输入搜索内容</div><div><div><div><input type="checkbox"/> A#公司类型</div><div><input type="checkbox"/> A#联系人头衔</div><div><input type="checkbox"/> A#顾客所在城市</div><div><input type="checkbox"/> A#区域</div><div><input type="checkbox"/> A#邮政编码</div><div><input type="checkbox"/> A#国家</div></div></div><div><div>全部</div><div>字符</div><div>数字</div></div></div><div><div>到右边></div><div><到左边</div></div><div><div>已选字段列表0/4</div><div><div>Q 请输入搜索内容</div><div><div><div><input type="checkbox"/> A#顾客编号</div><div><input type="checkbox"/> A#联系人</div><div><input type="checkbox"/> A#公司名称</div><div><input type="checkbox"/> A#顾客地址</div></div></div><div><div>确定</div><div>取消</div></div></div></div></div></div></div></div>
连接类型	<div>用于设置JOIN的连接类型：</div> <ul style="list-style-type: none">• 左连接：返回选择的左右表字段中，左表的全部数据和右表中满足关联条件的数据。• 右连接：返回选择的左右表字段中，右表的全部数据和左表中满足关联条件的数据。• 内连接：返回选择的左右表字段中，左表中满足关联条件的数据和右表中满足关联条件的数据。• 全连接：返回选择的左右表字段的所有数据。
关联条件	用于设置查询的字段条件，分别选择关联的左侧表字段和右侧表字段。通过加号按钮可以增加条件。

示例

1、将两个关系数据源的数据进行连接，分别选择左表的字段和右表的字段，连接类型选择“内连接”；关联条件左表和右表选择“顾客编号”。

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 | JOIN x

运行成功: 2秒

参数 属性 帮助

选择左表字段 *必填 ?

选择右表字段 *必填 ?

连接类型

内连接

关联条件

顾客编号 = 顾客编号

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

合并列

合并行

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

JOIN

2、将左右表中顾客编码相等的行连接起来，得到一张新的表。

查看输出

当前显示 10 列/总共 10 列, 100 条/总共有 7488 条数据

列筛选 请选择

表头真名 表头别名

顾客编号	产品大类名称	产品名称	订单编号	数量	销售额	顾客编号2	公司名称	联系人
ALFKI	调味品	酱油	10952.0	16.0	380.0	ALFKI	三川实业有限公司	刘小姐
ALFKI	特制品	烤肉酱	10952.0	2.0	91.2	ALFKI	三川实业有限公司	刘小姐
ANATR	日用品	黑奶酪	10308.0	1.0	28.8	ANATR	东南实业	王先生
ANATR	饮料	苏打水	10308.0	5.0	60.0	ANATR	东南实业	王先生
ANATR	特制品	沙茶	10625.0	3.0	69.75	ANATR	东南实业	王先生
ANATR	调味品	蕃茄酱	10702.0	6.0	60.0	ANATR	东南实业	王先生
ANATR	日用品	白奶酪	10759.0	10.0	320.0	ANATR	东南实业	王先生
ANATR	饮料	苏打水	10835.0	2.0	24.0	ANATR	东南实业	王先生
ANATR	日用品	民众奶酪	10926.0	2.0	42.0	ANATR	东南实业	王先生
ANATR	海鲜	龙虾	10926.0	10.0	80.0	ANATR	东南实业	王先生
ANATR	点心	糖果	10926.0	7.0	64.4	ANATR	东南实业	王先生

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

连接左右表中顾客编号相等的行

下载预览数据