

自助数据集-定义流程

自助数据集是一类基于个性化需求的数据集，它面向各阶层用户提供数据查询和抽取服务。

用户可以根据业务要求，通过可视化的方式，使用跨数据库、多表关联、数据转换、复杂逻辑关系计算、数据抽取ETL等功能，将数据提取到自助数据集，这些自助数据集允许被自助仪表盘、透视分析和电子表格进行数据分析应用。

自助数据集的定义流程如下：



- 新建自助数据集
- 选择数据来源
- 定义表集合
- 规范字段
- 保存自助数据集
- 数据抽取

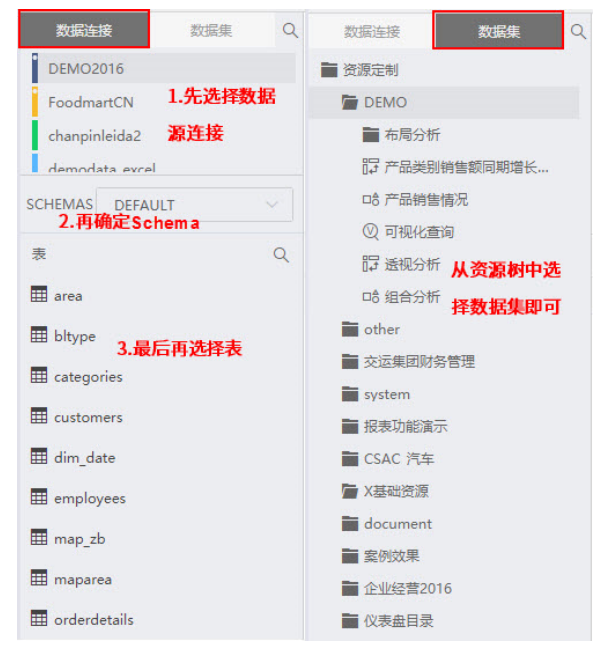
新建自助数据集

在定制管理界面选择快捷菜单 **自助数据集**（入口1），或在“资源目录区”的“资源定制”节点或其子目录节点的右键菜单中选择 **新建查询 > 自助数据集**（入口2）。



选择数据来源

自助数据集的数据来源可以来自数据源，也可以是数据集。



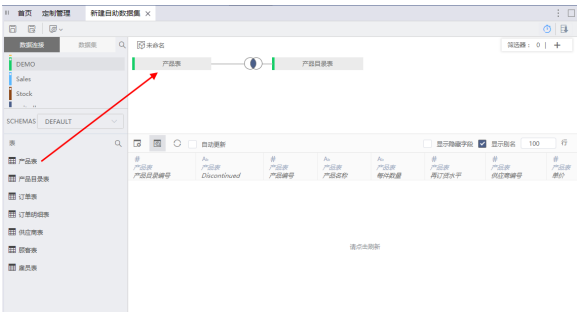
详情请参考 [自助数据集-数据来源](#)

定义表集合

定义表集合包含两部分内容：

- 从数据源中拖拽表或者拖拽数据集到“表集合区”。
- 在“表集合区”定义表关系，系统会根据表间相同别名的字段创建表关系，若无法自动创建需要用户手动创建表关系。

“表集合区”中的表从“数据区”中拖拽而来：

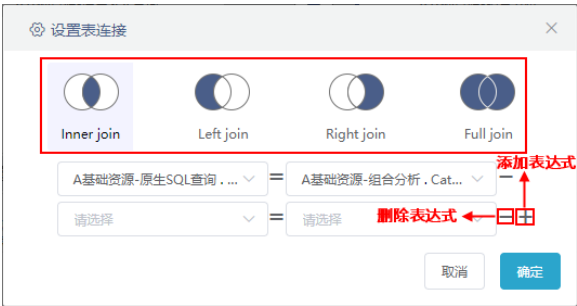


“表集合区”中手动创建表关系的方法：

1、单击表间连接点，如下图所示：



2、弹出“设置表连接”窗口，如下图所示：



- 选择连接方式：内连接、左连接、右连接、全连接。
- 确定表间连接字段。

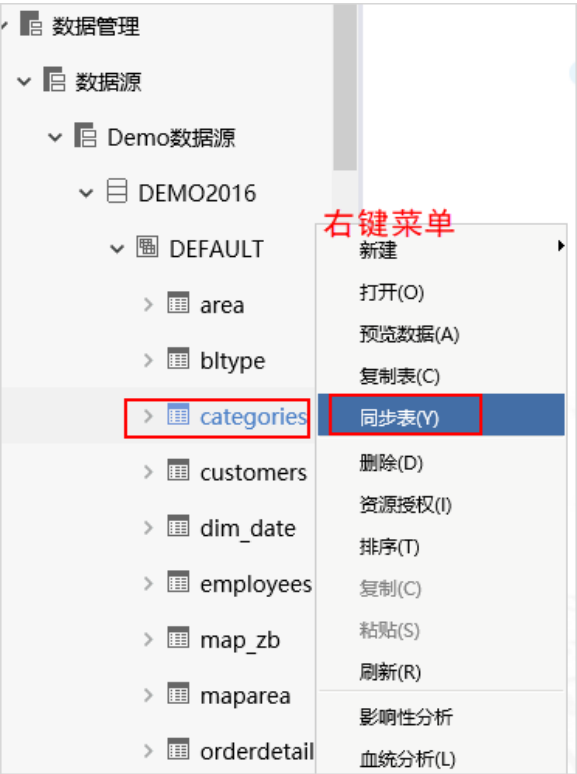
同步：原表或原数据集字段增加或删除时，自助数据集利用同步功能更新。

数据源	操作

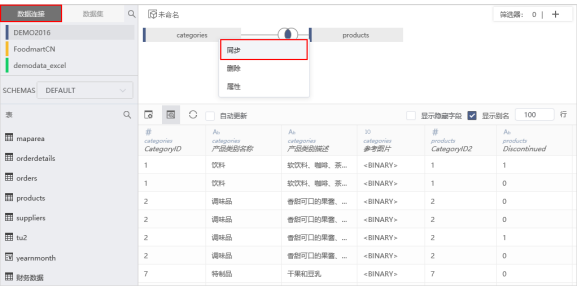
关系数据源/高速缓存

自助数据集引用的表字段增加、删除时，实现同步更新，需要进行如下两个步骤：

1、数据库某表字段增加、删除时，在数据源节点下，选中该表，**右键 > 同步表**，如图：



2、在自助数据集选中该表，**右键 > 同步**，如图：



数据集

选中字段变更的数据集，**右键 > 同步**，如图：



注：目前暂不支持字段类型更改的同步。




使用“同步”功能需保存，再次打开自助数据集才生效。




筛选器：当数据集数据量比较大时，可添加筛选器，对数据集的数据进行筛选过滤，以便得到用户真正所需的数据。




详情请参考 [自助数据集-筛选器](#)

规范字段

自助数据集的“表属性区”对应有维度和度量，系统支持对表和字段进行重新组织、分类、设计等规范化设置操作：[修改别名](#)、[修改数据类型](#)、[设置可见性](#)、[拖拽表和字段](#)、[新建目录](#)、[新建层次](#)、[新建计算字段](#)、[创建时间维](#)、[标记地理维](#)、删除。详细说明如下：

功能	说明	示例效果
修改别名	修改表或字段别名。	<p>手动输入别名，修改“customers”的别名为“顾客”，如图：</p> 
修改数据类型	修改字段的数据类型。	<p>选择字段的数据类型，如图：</p> 
设置可见性	<p>设置表或字段为“可见”或“不可见”。</p> <ul style="list-style-type: none">为了快速查找表或字段，可将自助数据集在特定场景中不需要用到的表或字段设置为“不可见”。将表设置为“不可见”时，表的所有字段都为“不可见”。	<p>点击 可见性的图标切换为“可见”或“不可见”：</p> 

拖拽表和字段	拖拽表或字段移动其位置。	<p>拖拽字段改变字段所属的表，将“产品目录表”的“产品目录名称”拖拽到“产品表”，如图：</p> 
新建目录	用于在“维度”或“度量”中新建目录。	<p>任意区域右键菜单选择“新建目录”，新建一个“产品”目录，如图：</p> 
新建层次	用于新建层次目录，存储层次字段，这些层次字段主要实现自动下钻应用，例如：功能示例-时间层次维下钻、功能示例-地理层次维下钻、功能示例-自定义层次维下钻、交叉表的展开/收缩等。	<p>任意区域右键菜单选择“新建层次”，新建“地区”层次，按层次顺序将字段拖拽到“地区”下。如图：</p> 

新建计算字段	<p>新建一个计算字段。</p> <p>计算字段的数据类型决定了新建的计算字段位于维度或度量。</p> <ul style="list-style-type: none">维度：字符串、日期数据、以ID结尾的字段。度量：数值型数据。 <p>应用详情请参考 自助数据集-计算字段</p>	<p>任意区域或字段右键菜单选择“新建计算字段”，新建“销售额”计算字段，如图：</p> 
创建时间维	<p>用于以维层次的形式标识日期信息字段。</p> <p>入口：日期信息字段的右键菜单选择 创建时间维度。</p> <p>创建时间维度后会自动生成一个“时间维度”的层次目录。</p> <p>应用详情请参考 功能示例-时间层次维下钻</p>	
标记地理维	<p>用于以维层次的形式标识地理信息字段。</p> <p>入口：地理信息字段的右键菜单选择 标记地理维度。</p> <p>已标记地理维度的字段必须存储在层次目录下，才能保证地理维层次字段的下钻效果实现。</p> <p>应用详情请参考 功能示例-地理层次维下钻</p>	
删除	用于删除表。	

保存自助数据集

保存当前自助数据集到指定目录。

数据抽取

将当前自助数据集的数据从原始数据库中抽到到高速缓存库。详情请参见 [数据抽取](#) 章节。