

功能示例-跨库数据集

概述说明

自助数据集支持跨库查询，当用户查询数据的范围比较广，并不限于一个数据库时，可以跨多个数据库进行查询。

实现跨库查询的前提条件：保证系统中的“跨库联合数据源”和“高速缓存数据库”有效，且进行跨库查询的数据库已经同步到跨库联合数据源中。

目前支持做跨库的数据源类型包括：高速缓存库、Hadoop_Hive、星环、Vertica、CH、Greenplum、Infobright、Oracle、DB2 V9、MySQL、MS SQL Server、Spark SQL、Teradata_v12、Informix、IMPALA、PostgreSQL。

示例分析

业务需求描述

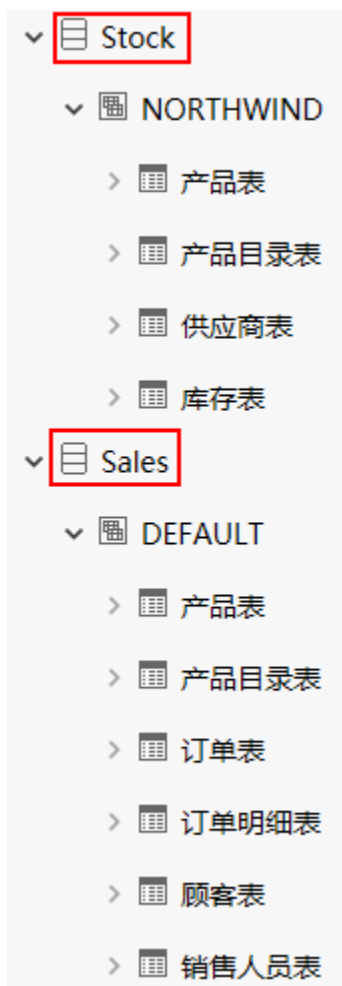
某企业的生产部门和销售部门有各自的数据信息存储库，现在公司上层要求结合库存数据和销售数据进行综合分析。具体要求有：

- 1、各个地区的销售情况，主要通过销售额数据体现，并且能够对区域进行下钻，查看下级省份和城市的销售情况。
- 2、不同类别的产品库存情况。
- 3、不同类别的产品在各个地区的订购量情况。
- 4、不同类别的产品销售情况，主要通过销售额数据体现，通过产品类别选择查看所属产品的销售情况。

需求及数据分析

由于生产部门和销售部门的数据分别存储在不同的数据库：生产部门的数据存储在基于Oracle数据库的“Stock”数据源中，销售部门的数据存储在基于MySQL数据库的“Sales”数据源中。

- 概述说明
- 示例分析
 - 业务需求描述
 - 需求及数据分析
 - 操作步骤
- 注意事项



因此该综合分析需要进行跨库查询，创建一个跨库数据集，并可以用该数据集设计一个自助仪表盘。

操作步骤

本示例以“Stock”数据源和“Sales”数据源来实现。

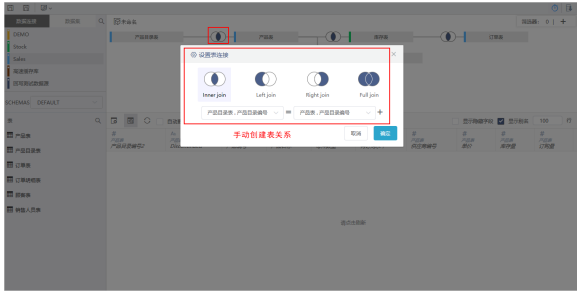
1、新建自助数据集，在左侧数据连接区可以看到能实现跨库查询的数据源。数据源左侧颜色标识条上的空心圆表示该数据源支持跨库查询，如下图所示：



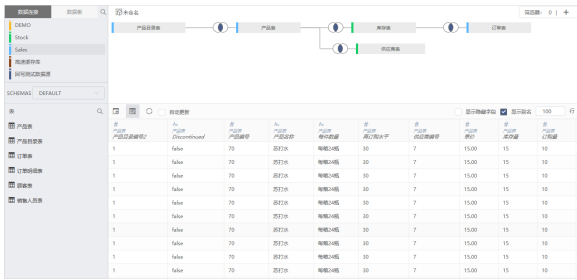
2、双击选择“Sales”数据源下的“产品目录表”、“产品表”和“订单表”，和“Stock”数据源下的“库存表”和“供应商表”到表集合区，表跟表之间会自动建立关联关系。数据表左侧颜色与所属数据源左侧颜色一致，同一个数据源的数据表左侧颜色一致。



若表关系没有自动连接，可以打开设置表连接窗口，手动设置表连接。



3、点击 **刷新** 按钮，可以查看跨库数据集表数据；最后再点击 **保存**，并将该自助数据集命名为“跨库数据集”。



4、新建自助仪表盘，在自助仪表盘左侧的 **数据** 中选择刚刚创建的“跨库数据集”，基于该跨库数据集定义的仪表盘效果如下图所示：



自助仪表盘的定制详情请参见 [自助仪表盘](#) 章节。

注意事项

1. V8.5版本及以上的版本，一旦用户创建了新的关系数据源连接，则自动同步到跨库联合数据源中。
2. V8.5版本之前已存在的关系数据源不会自动同步到跨库联合数据源，需要手动添加。添加入口如下图所示：

