

# SQL数据集

SQL数据集是通过在文本区中输入SQL语句，来定义数据集条件和内容的一种数据集。对于熟悉SQL语句的技术人员来说，这是一个较为快捷创建的方式。



关于SQL数据集的创建入口、资源显示和使用，统一通过V10.5 license控制。如希望在V10.5版本中继续使用SQL数据集，则需更新授权包含SQL数据集的license文件。

- [操作入口](#)
- [功能说明](#)
- [界面介绍](#)
  - [可选资源区](#)
  - [SQL文本区](#)
  - [工具栏](#)
  - [属性区](#)
  - [输出字段](#)
  - [检测输出字段类型](#)

## 纯SQL问题的解答：

SQL数据集可以是纯SQL语句，但是如果在SQL数据集中用到“参数”、“用户属性”、“系统函数”等系统资源，必须把资源拖到SQL语句对应的位置，不能手写。

包含这类资源的SQL数据集就不是纯SQL。

## 嵌套问题的解答：

SQL数据集需要嵌套运行。

如文本语句select distinct Card\_Type\_Cd as c1 from debit\_card，会变成 select T2.c1 as F4 from (select distinct Card\_Type\_Cd as c1 from debit\_card) T2 运行。

这样处理是因为在报表中需要对SQL做处理，如果使用SQL数据集不嵌套一层的话，就无法做到去修改原始的SQL数据集，这样SQL数据集中的预览数据和报表中的过滤、排序等功能就无法实现。

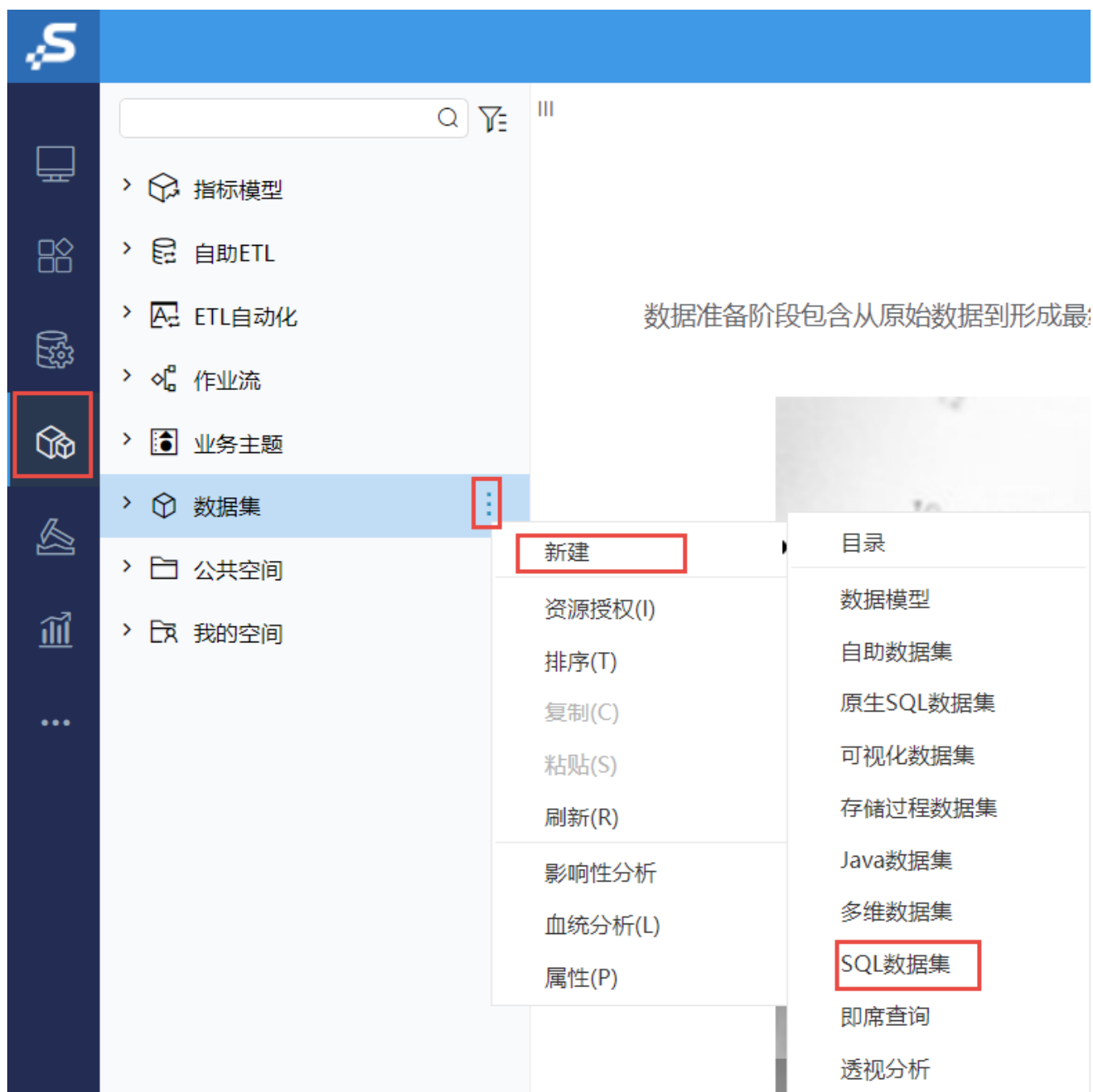
## 嵌套效率问题的解答：

我们测试过SQL Server和MySQL，通过Explain看数据集优化，发现嵌套一层对效率不会有太大的影响。

即上例两个语句执行效率相差并不多，一个执行代价是0.366，嵌套后代价是0.367，相差只有 0.1%。

# 操作入口

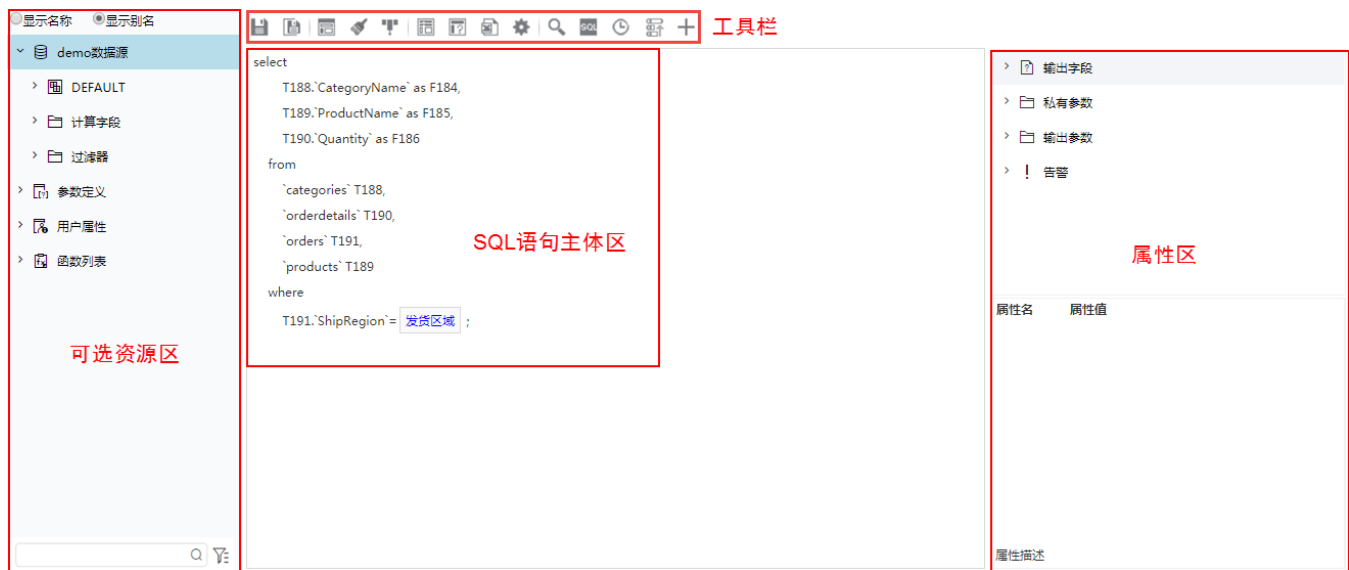
在系统导航栏中选择 **数据准备** ，在数据集的更多  中选择 **新建>SQL数据集** 。



## 功能说明

- 目前SQL数据集的创建只能基于数据源创建。因此若是只使用SQL数据集创建报表，则不需要创建业务主题、语义层定义等操作。
- 技术人员可以先在数据库中写好SQL并执行通过，然后把这个SQL语句复制到SQL数据集中，并从左侧拖拽参数等对象到相应的位置。
- 书写SQL数据集后，一定要进行检测输出字段类型操作。

## 界面介绍



“SQL数据集”界面主要分为以下几个区域：

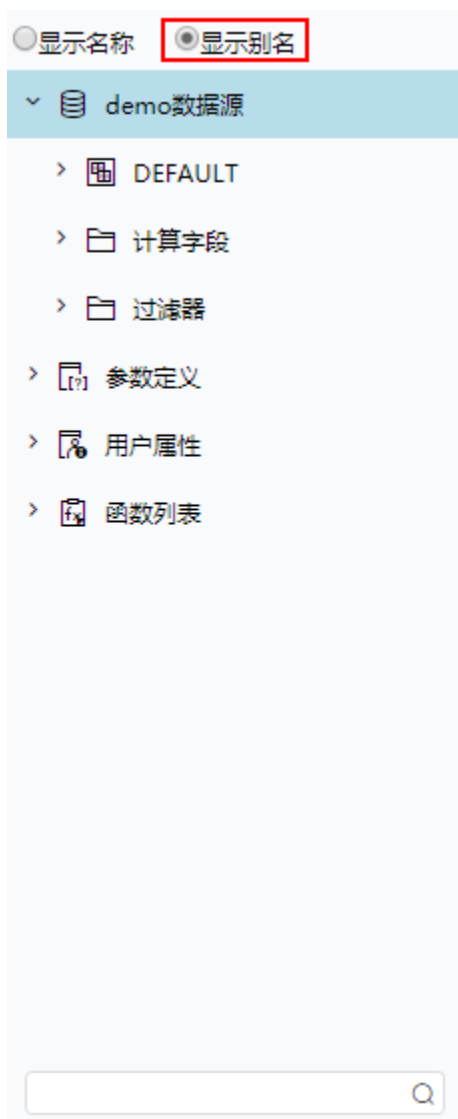
- 可选资源区：该区列出了所有可被SQL数据集使用的资源。
- SQL文本区：该区为用户书写SQL数据集表达式的区域。
- 工具栏：列出了所有可用于编辑SQL数据集的工具按钮。
- 属性区：属性区用于设置数据集相关资源的属性。

## 可选资源区

可选资源区列出了所有可被SQL数据集使用的资源。用户只需将各资源拖到相应的区域，如将参数拖拽到where条件下和字段进行关联。

**可选资源包含：**表字段、计算字段、过滤器、业务视图字段、参数、数据源相关业务主题的业务属性、函数、用户属性、其他数据集的字段等。但是建议SQL语句先写好，只对需要的对象如参数、用户属性、函数等进行拖拽。

**显示别名/名称：**该区针对“关系数据源”、“联合数据源”和“指标数据源”下的各资源：表、字段、视图、存储过程、表副本和视图副本，提供了“显示名称”和“显示别名”的功能。默认以“别名”显示。



**模糊搜索资源：**我们可以该区最下方的搜索引擎对相关资源快速搜索定位。可以在搜索区输入所需字段的真名或是别名进行模糊搜索。



## SQL文本区

SQL文本区使用 [SQL语句编辑器](#) 输入SQL语句及其对象。

**SQL数据集输入的注意事项：**

- SQL数据集可以使用的资源：表/视图、字段、参数、函数、用户属性（使用时请用括号括起来）
- 编写SQL数据集时，支持拖动“参数”、“用户属性”、“系统函数”等对象，其他的建议手写；
- 不支持Select \* from .....的写法；
- 如果有多条语句，用分号和回车隔开；
- 所有的Select字段需要显式给定一个别名，例如：select T1.fieldA FA from table T1。
- 对字段进行重命名时，注意各种数据库使用方式跟平时使用的不同。如：SQL Server不需给别名加单引号，Oracle不需要加as和双引号。
- SQL数据集不能区分输出字段类型，默认设置所有输出字段都是字符型，需要 [检测输出字段类型](#) 。
- 表关系视图的定义对SQL数据集不起效；
- 数据行权限设置对SQL数据集不起效。
- 设置参数缺省值，在SQL数据集中的设置方法为：

```
where {[ T32. CategoryName = CategoryName ]} and {[ T5.OderDate>= 开始日期  
andT5.OderDate<= 结束日期 ]};
```

## 工具栏

工具栏上放置一些常用功能按钮。



工具栏的按钮说明如下表，对应的功能点击超链接跳转到详细说明：

| 按钮名称                      | 说明                                 |
|---------------------------|------------------------------------|
| 保存                        | 保存数据集到默认路径。                        |
| 另存                        | 保存数据集到用户指定路径。                      |
| 视图                        | 用于区域的显示和隐藏。包含显示/隐藏资源区、显示/隐藏属性区。    |
| 格式化SQL                    | 对用户编写的SQL语句按照一定的格式进行展现，便于阅读。       |
| 检查输出字段规范                  | 对输出字段按照规范进行检查，列出不符合规范的字段信息。        |
| <a href="#">定义多重表头</a>    | 对SQL数据集的表头进行修改添加，形成多重表头。           |
| <a href="#">参数布局</a>      | 调整参数在SQL数据集对应的报表中的位置。              |
| <a href="#">Excel模板设置</a> | 把当前数据集的数据结果填充到用户自定义的Excel模板中。      |
| <a href="#">高级设置</a>      | 用于设置当前数据集的缓存策略和分页策略。               |
| 预览数据                      | 对新建或修改的数据集进行预览。                    |
| <a href="#">查看SQL</a>     | 查看该数据集的SQL语句。                      |
| 显示估计的执行计划                 | 对数据库内部的SQL执行过程进行的解释。               |
| <a href="#">数据抽取</a>      | 用于将当前数据集的数据抽取到MPP。                 |
| 创建                        | 可基于当前数据集快速创建透视分析、仪表分析、灵活分析以及自助数据集。 |

## 属性区

属性区用于设置数据集相关资源的属性。它主要包含两个部分：数据集输出区和属性面板。

- 数据集输出区：该区显示数据集输出的相关资源，该区通常与“属性面板”结合起来用于对这些资源的属性进行管理。包含：[输出字段属性](#)、[检测输出字段类型](#)、[私有参数](#)、[输出参数](#)、[告警设置](#)。
- 属性面板：该区需要与“工作区”结合使用，单击工作区中的任一资源，即在属性面板中列出该资源的相关属性信息。用户可以在属性面板中重新设置资源的相关属性。

输出字段

F184

F185

F186

私有参数

输出参数

告警

| 属性名  | 属性值  |
|------|--|
| 标识   | BizViewOutField.l8a8a4c53016c07ef07efe06d0 |
| 名称   | F184                                       |
| 别名   | F184                                       |
| 描述   |  |
| 数据类型 | 字符串  |
| 数据格式 | <字符串-默认值>                                  |
| 排序   | 否  |
| 转换规则 | 否  |
| 脱敏规则 |  |

输出字段

在数据集输出区会列出所有的输出字段。如下图，选中字段，在下面属性区可以对部分属性进行设置。

输出字段

F184

F185

F186

私有参数

输出参数

告警

|      |  |
|------|--|
| 属性名  | 属性值  |
| 标识   | BizViewOutField.l8a8a4c53016c07ef07efe06d0 |
| 名称   | F184                                       |
| 别名   | F184                                       |
| 描述   |  |
| 数据类型 | 字符串  |
| 数据格式 | <字符串-默认值>                                  |
| 排序   | 否  |
| 转换规则 | 否  |
| 脱敏规则 |  |

下面是对字段属性的详细说明：

| 对象   | 设置项  | 说明  |
|------|------|---|
| 输出字段 | 标识   | 字段的全局唯一标识，由字段类型和其ID号共同构成。不允许重新设置。                           |
|      | 名称   | 字段的真名。不允许重新设置。  |
|      | 别名   | 字段的别名，即显示的名字。允许重新设置。  |
|      | 描述   | 对字段的相关描述，被用于其提示信息中。允许重新设置。                                  |
|      | 数据类型 | 字段的数据类型，主要包含：整数、浮点、字符串、日期、时间、日期时间、文本、对象、集合。允许重新设置。          |
|      | 数据格式 | 根据数据类型选择输出字段的数据格式，该格式选项在 <a href="#">数据格式</a> 中进行设置。允许重新设置。 |
|      | 排序   | 选择是否对字段进行排序，字段的排序方式类型有：升序、降序。                               |
|      | 转换规则 | 选择是否对该字段进行转换规则，若设置转换规则，其所选转换规则在 <a href="#">转换规则</a> 中定义。   |
|      | 脱敏规则 | <a href="#">数据脱敏</a> 是指对某些敏感信息通过脱敏规则进行数据的变形，实现敏感隐私数据的可靠保护。  |

检测输出字段类型

SQL数据集中的“输出字段”须通过右键菜单中的 **检测输出字段类型** 功能将当前数据集中输出字段的数据类型从数据库中读取出来。

操作执行原因：

SQL数据集若是不进行检测输出字段类型操作，则默认所有字段类型都是字符串类型。

数据集一旦定义后，其所有输出字段的类型均默认为“字符串”型，当输出字段与数据库中相应字段的数据类型不匹配时，会导致利用该SQL数据集创建报表后进行的分析失误，因此需要用户手工检测输出字段类型。

```
select
  T188.`CategoryName` as F184,
  T189.`ProductName` as F185,
  T190.`Quantity` as F186
from
  `categories` T188,
  `orderdetails` T190,
  `orders` T191,
  `products` T189
where
  T191.`ShipRegion` = 发货区域 ;
```

输出字段

F184

F185

F186

私有参数

输出参数

警告

| 属性名  | 属性值   |
|------|---|
| 标识   | BlizViewOutField.I8a8a4c53016c07ef07efe06d0 |
| 名称   | F186  |
| 别名   | F186  |
| 描述   |   |
| 数据类型 | 字符串   |
| 数据格式 | <字符串-默认值>                                   |
| 排序   | 否   |
| 转换规则 | 否   |
| 脱敏规则 |   |

执行步骤:

选中输出字段节点，右键选择 检测输出字段类型。

输出字段

F184

F185

F186

私有参数

输出参数

警告

检测输出字段类型

修改

删除

如SQL数据集包含参数值，则会弹出“自动检测输出字段”对话框。选择参数值，点击 输出字段检测 按钮。则会进行检测操作。



自动检测输出字段

×

发货区域\* 东北

▼

输出字段检测

取消(C)

**执行结果：**操作成功后，我们可以看到如下字段类型发生变化，为字段本身的类型。

输出字段

F184

F185

F186

私有参数

输出参数

告警

| 属性名  | 属性值  |
|------|--|
| 标识   | BizViewOutField.l8a8a4c53016c07ef07efe06d0 |
| 名称   | F186                                       |
| 别名   | F186                                       |
| 描述   |  |
| 数据类型 | 浮点型  |
| 数据格式 | <浮点型-默认值>                                  |
| 排序   | 否  |
| 转换规则 | 否  |
| 脱敏规则 |  |

浮点型