

数据挖掘-参数设置

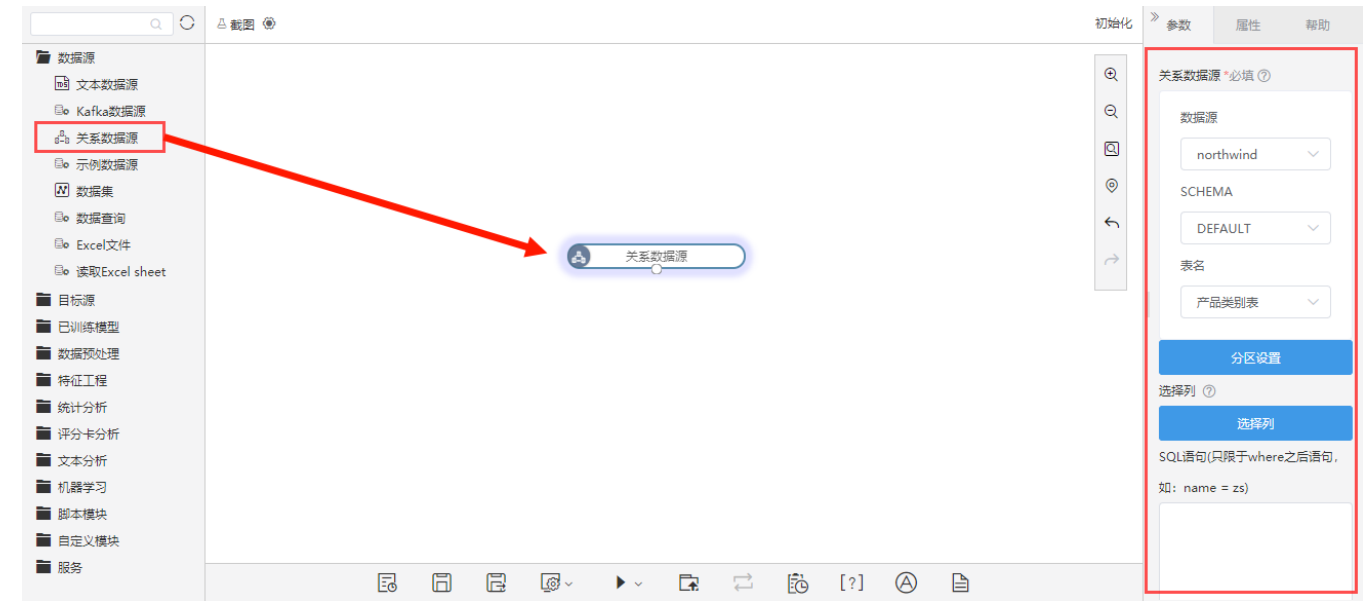
在数据挖掘中，通过参数筛选用户可查询出需要的数据。

- [操作步骤](#)
- [增量抽取场景](#)

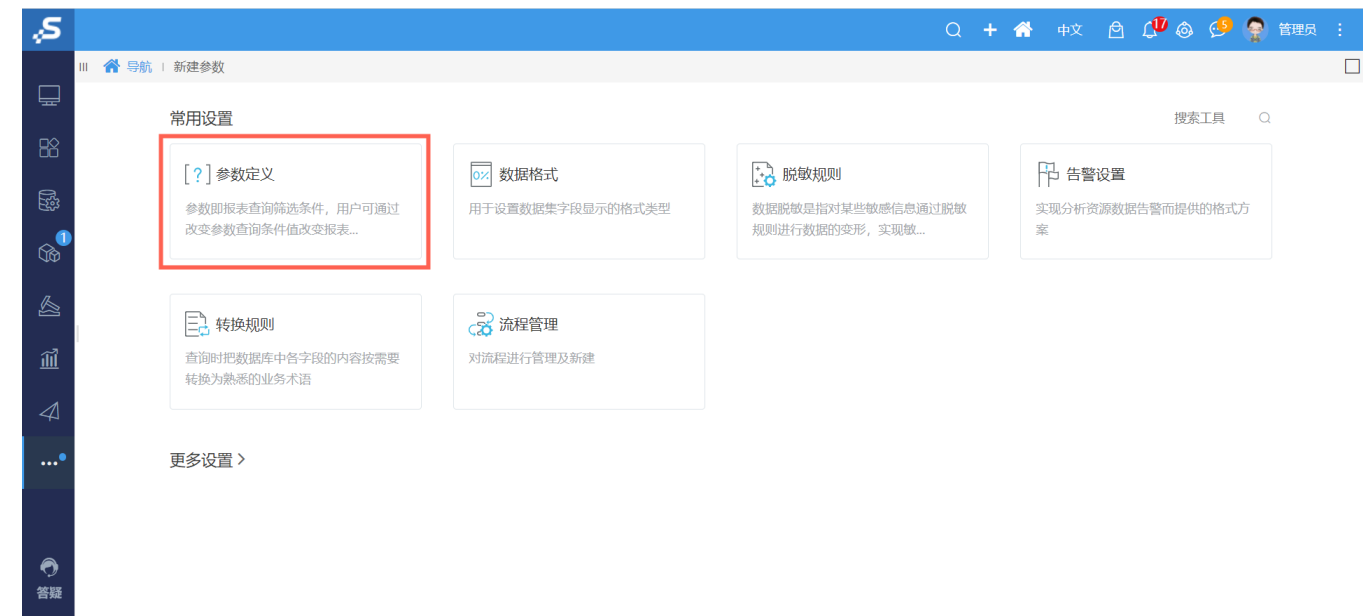
目前只有部分节点支持参数设置功能，如关系数据源、数据查询、关系目标表（追加）、关系目标表（覆盖）、关系目标表（插入或更新）、SQL脚本、源库SQL脚本。

操作步骤

1、新建实验，拖入关系数据源节点，并选择数据源、SCHEMA和表。



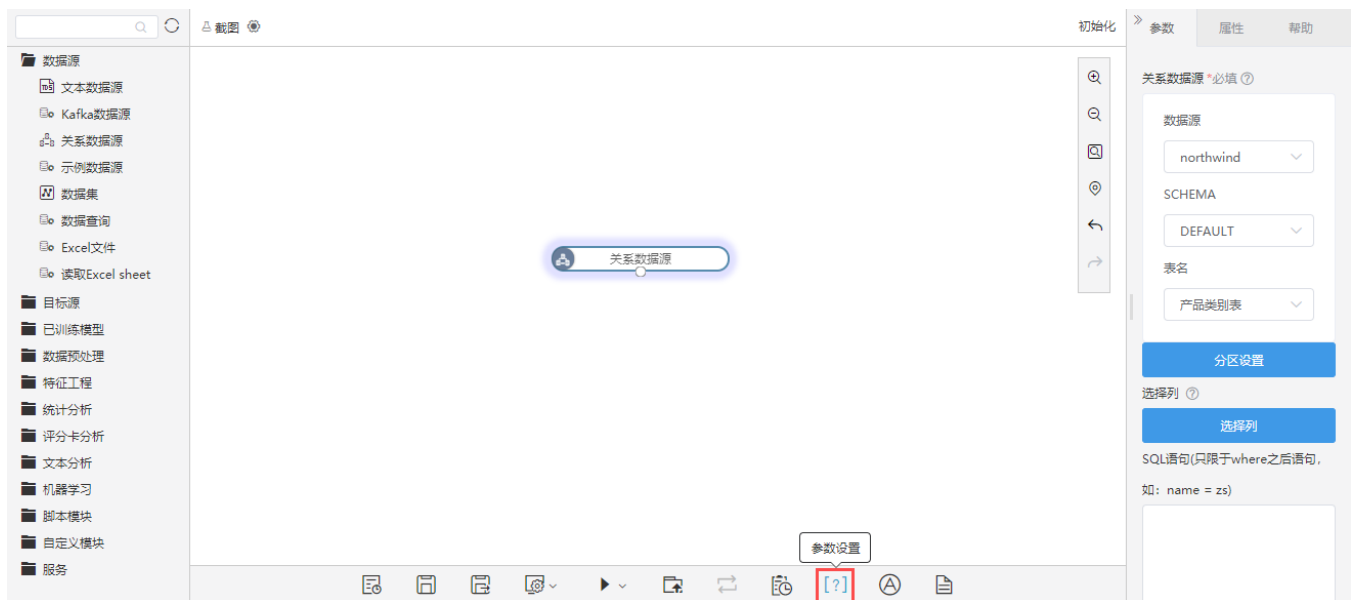
2、在公共空间针对本ETL定义一个参数，具体如何配置，参考[公共设置-参数定义](#)，下图是配置好的参数示例。





3、 点击 **参数设置** 按钮。

注意：在ETL中，需要先保存后，才能对参数按钮进行编辑。原因：涉要设置参数的话，首先需要知道是给哪个资源设置。



4、配置参数，选择参数类型、参数定义、参数值并保存。这里需要在参数定义中选择刚刚设定好的参数。

⌚ 参数设置

*注意：请依次配置参数名及参数值

×

参数名 *

参数类型

参数定义 *

参数值 (可选) ?

参数

请选择

+

param

参数

SMS-20409-sql

饮料

确定

取消

参数设置项说明如下：

设置项	说明	
参数名	参数的名称。	
参数类型	值	参数值：固定参数值
		类型： <ul style="list-style-type: none">int：整数bigint：整数（当整数超过int范围使用）string：字符串double：双精度浮点型
	参数	参数定义：在产品中定义的参数
		参数值（可选）：获取参数定义中的参数值。 不选择参数值为动态参数，选择参数值后为静态参数值。

5、在关系数据源参数的SQL语句输入框中，只需要输入where之后的语句。可输入静态条件，如name=zs，或使用含参数的语句：CategoryName=\${param} （表头真名=\${参数名}）

» 参数 属性 帮助

关系数据源 *必填 ⑦

数据源

SMS_20409_northwii

SCHEMA

DEFAULT

表名

categories

分区设置

选择列 *必填 ⑦

选择列

SQL语句(只限于where之后语句，如：
name = zs)

CategoryName=\${param}

在查看输出中选择表头真名，可查看表头真名。

查看输出

当前显示 4 列 / 总共 4 列， 10 条 / 总共有 10 条数据

列筛选 请选择

表头真名 ☒ 表头别名

#	CategoryID	CategoryName	Description	Picture
1		饮料	软饮料、咖啡、茶、啤酒和淡啤酒	[B@e34780c
2		调味品	香甜可口的果酱、调料、酱汁和调味品	[B@aa13c16
3		点心	甜点、糖和甜面包	[B@6d972f46
4		日用品	奶酪	[B@1a3d7c86
5		谷类/麦片	面包、饼干、生面团和谷物	[B@37e493d5
6		肉/家禽	精制肉	[B@34a658b3
7		特制品	干果和豆乳	[B@609fac7e
8		海鲜	海菜和鱼	[B@11b725a7

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

下载预览数据

6、筛选出产品名称为“饮料”的数据，如图：

查看输出

当前显示 4 列 / 总共 4 列。1 条 / 总共有 1 条数据

列筛选 请选择

表头真名 表头别名

# CategoryID	CategoryName	Description
1	饮料	软饮料、咖啡、茶、啤酒和淡啤酒

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

下载预览数据

增量抽取场景

增量抽取可以显著减少数据传输量和处理时间，比如需要前三天至今天抽取新增的数据，但不希望全量抽取，可以使用参数来实现增量抽取。

1、首先连接好数据源和目标源节点。

ETL增量抽取

运行成功: 1秒

数据源

- 文本数据源
- FTP数据源
- Kafka数据源
- 关系数据源
- 示例数据源
- 数据集
- 数据查询
- Excel文件
- 读取Excel sheet
- 目标源
- 数据预处理
- 脚本模块

关系数据源

关系目标表(追加)

参数

属性

帮助

关系数据源 *必填

数据源northwind

SCHEMADEFAULT

表名订单表

分区设置

选择列

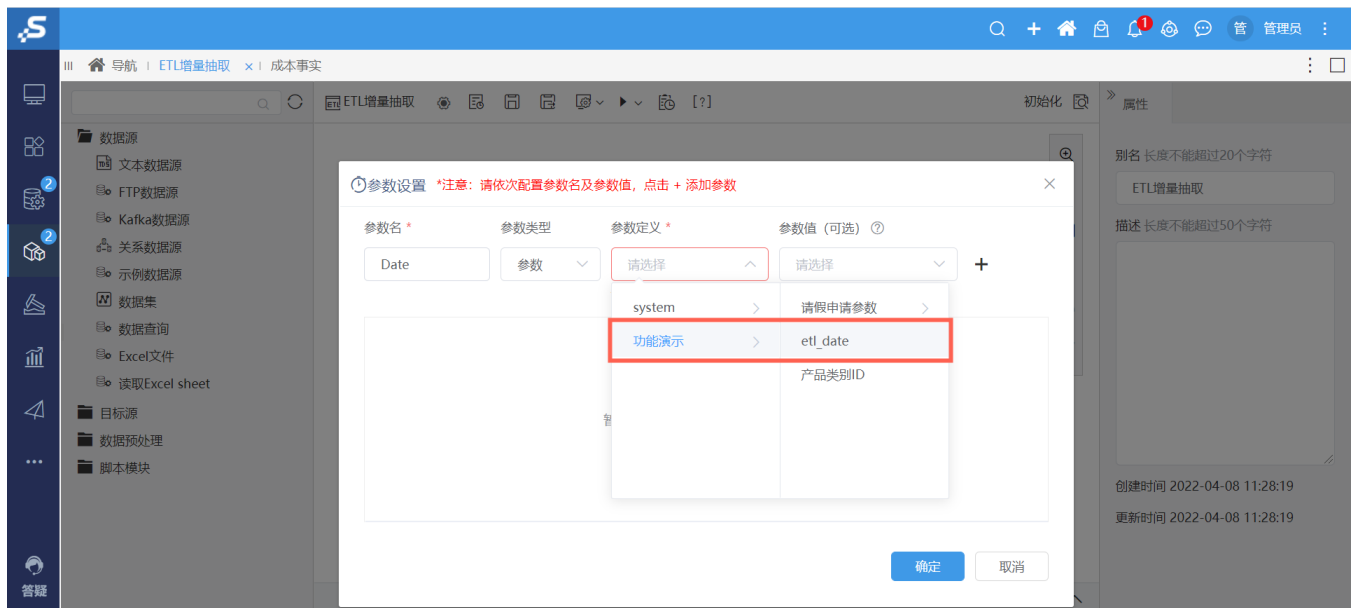
选择列

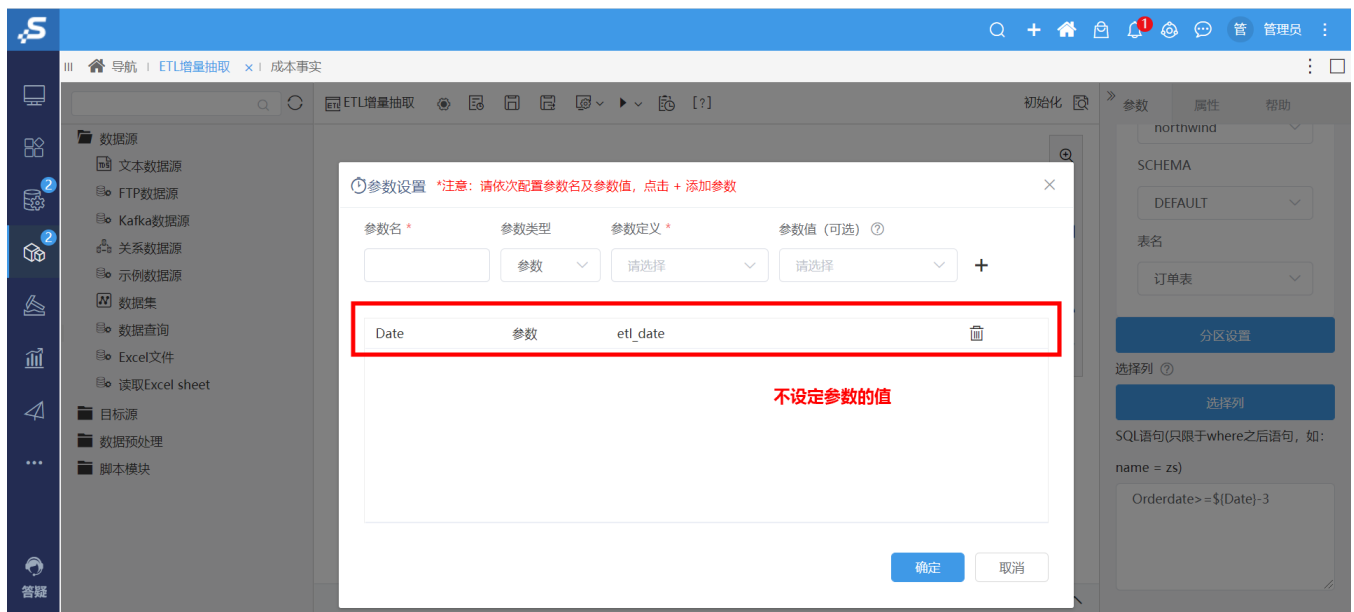
SQL语句(只限于where之后语句, 如: name = zs)

2、在参数定义中编写sql获取当天的日期。



2、返回到ETL界面中，把刚刚新建的参数定义在ETL里。





3、在关系目标源中的SQL语句编辑框中填写如图的内容，即可抽取前三天至今新增的数据。

