数仓建模介绍

目 录

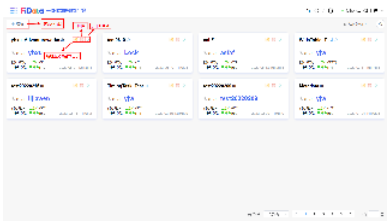
- 1. 概述
- 2. 界面概览
- 3. 数仓建模
  - 3.1 创建业务域
  - 3.2 操作数据仓库
    - 3.2.1 创建维度文件夹
    - 3.2.2 创建维度表
    - 3.2.3 创建日期维度表
    - 3.2.4 维度表字段设置
    - 3.2.5 创建业务过程文件夹
    - 3.2.6 创建事实表
    - 3.2.7 事实表字段设置
  - 3.3 操作分析指标
    - 3.3.1 分析指标维度表
    - 3.3.2 创建原子指标
    - 3.3.3 创建派生指标
    - 3.3.4 创建业务限定
    - 3.3.5 创建宽表

1. 概述

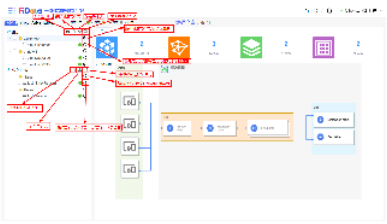
数仓建模功能提供可视化的界面，方便用户通过建模的方法更好的组织、存储数据，以便在性能、成本、效率和数据质量之间找到最佳平衡点

2. 界面概览

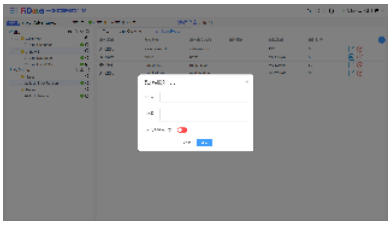
入口



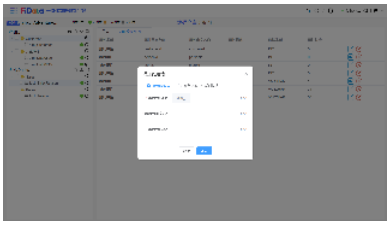
数据仓库



维度文件夹

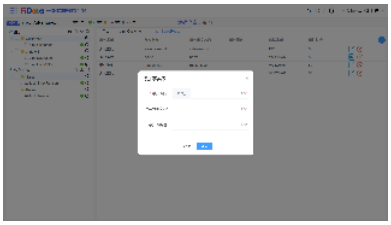


维度表

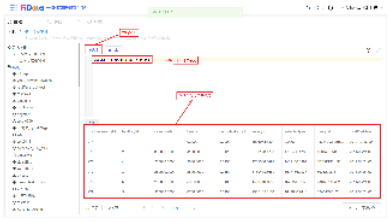


业务过程文件夹

事实表



维度字段设置

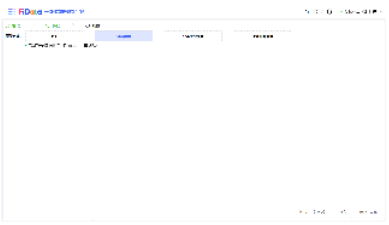


维度字段转换

添加维度键

退化维度&度量字段

覆盖方式



保存&发布

分析指标

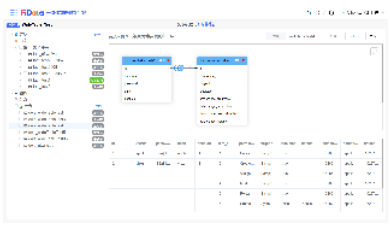
原子指标

派生指标

业务限定

时间函数

宽表



3. 数仓建模

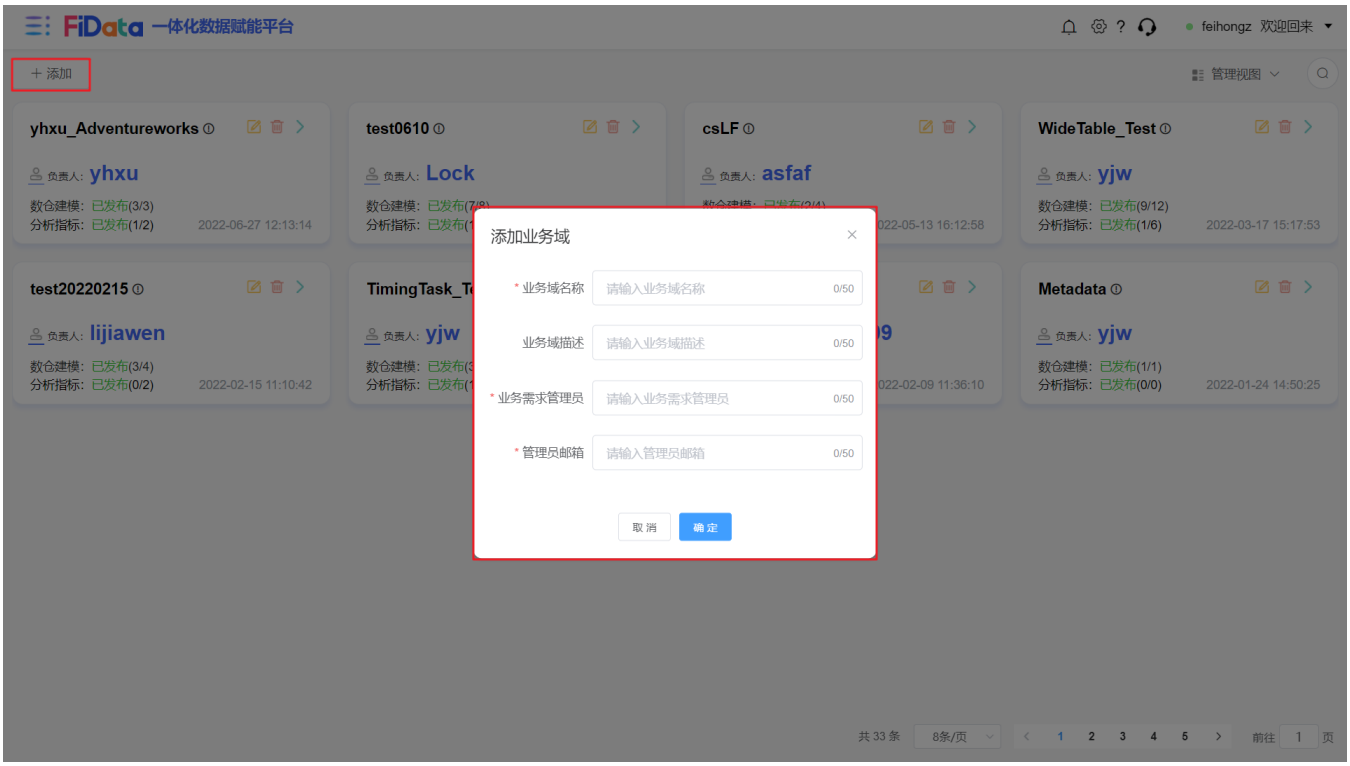
3.1 创建业务域

创建业务域： 点击添加，输入业务域基本信息，即可创建业务域

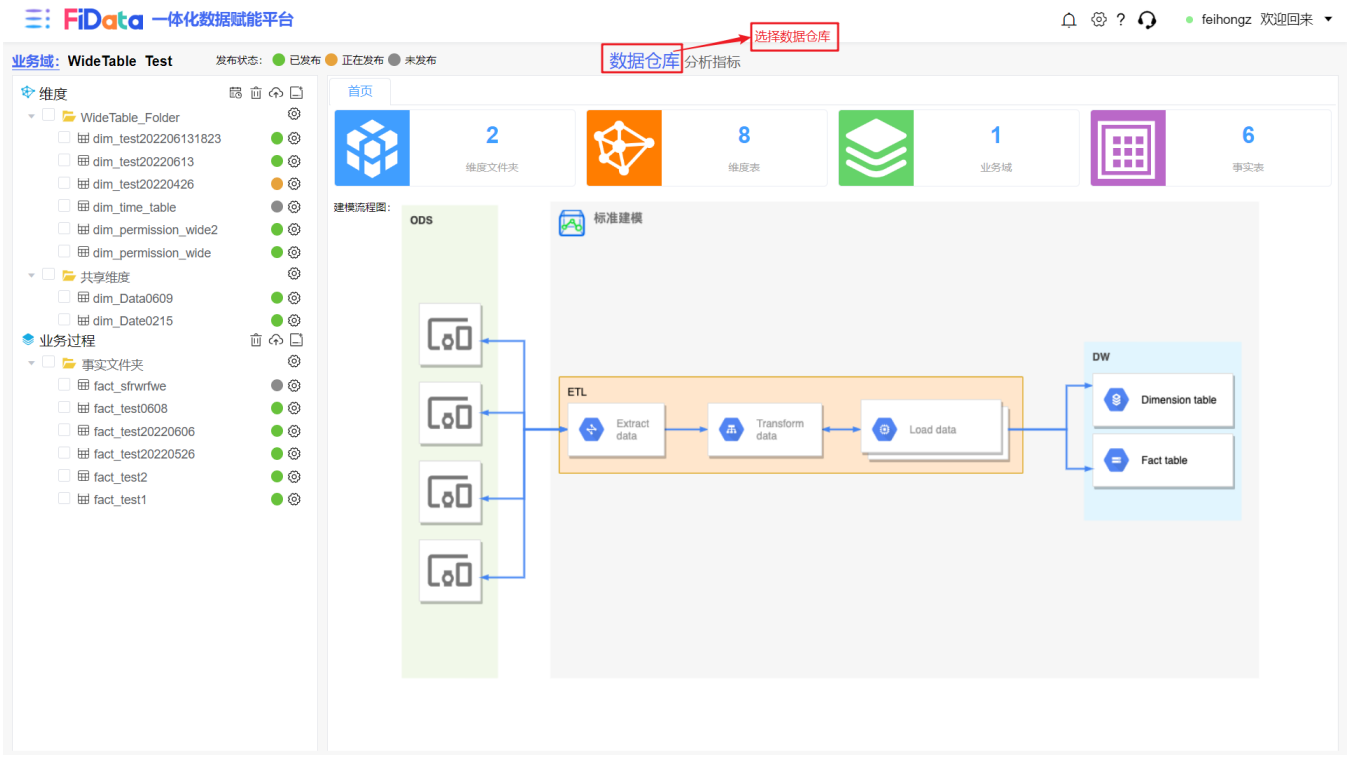
业务域名称：中英文名称皆可

业务需求负责人：该系统的IT负责人

管理员邮箱：该系统的IT负责人邮箱



3.2 操作数据仓库

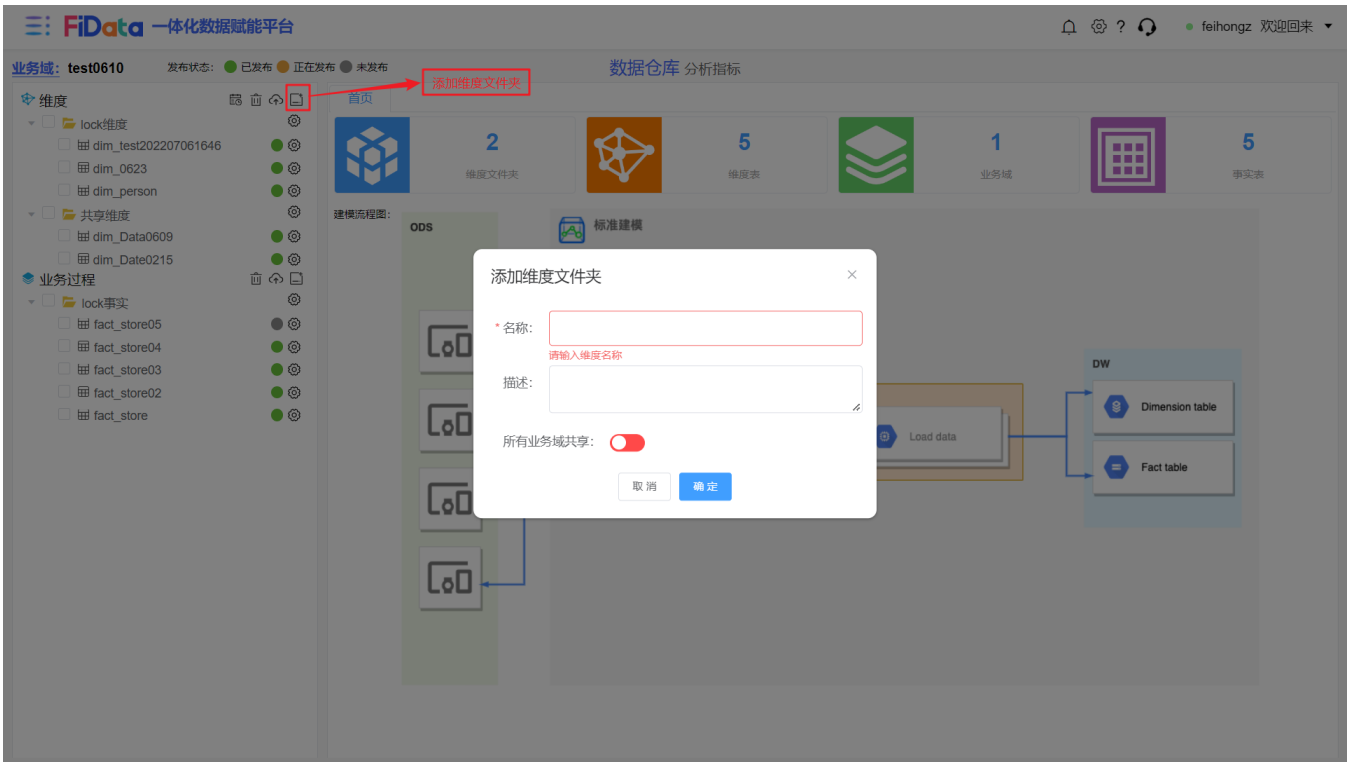


3.2.1 创建维度文件夹

创建维度文件夹： 点击添加，输入基本信息，即可创建维度文件夹

名称：中英文名称皆可

所有业务域共享：开启业务域共享，意味着其他业务域也可以使用共享维度文件夹及其下面的所有维度表。反之，只能当前业务域允许使用

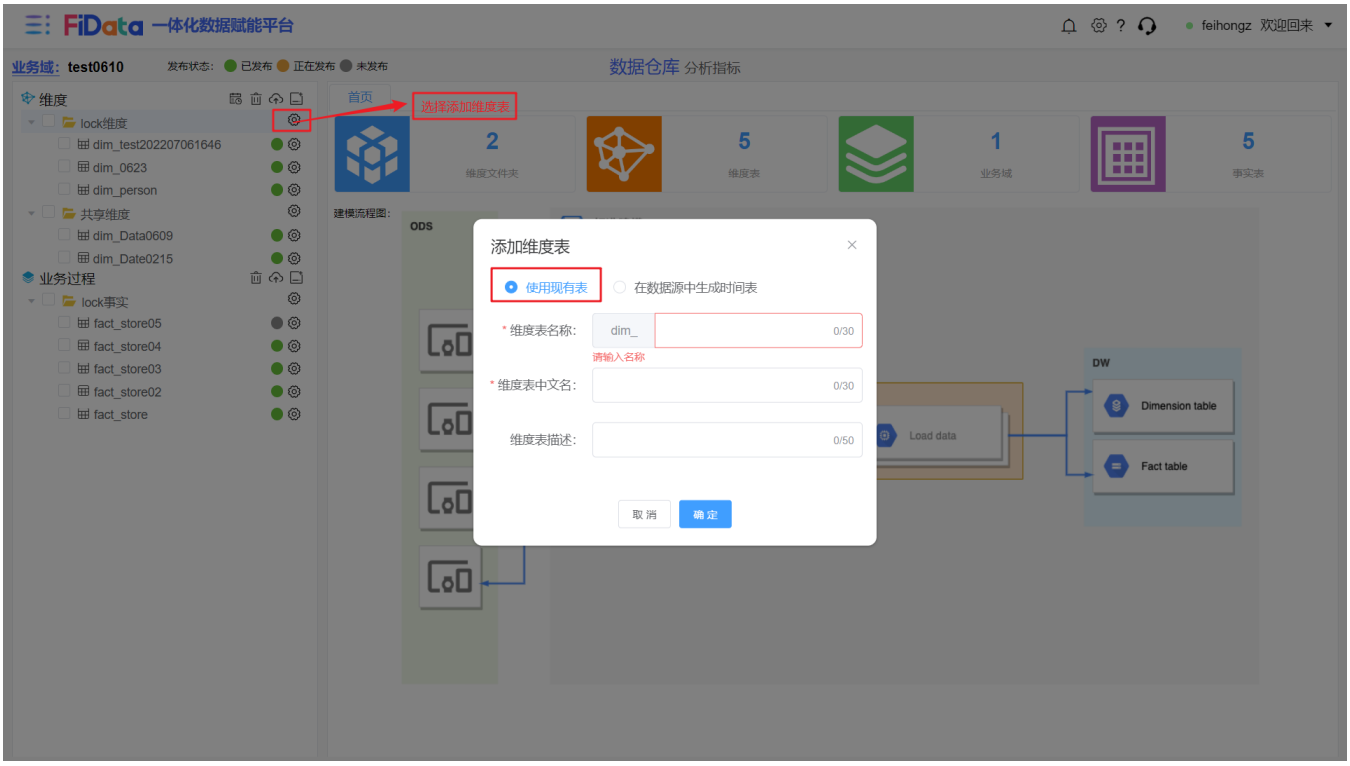


3.2.2 创建维度表

创建维度表：点击添加，选择“使用现有表”，输入维度表基本信息，即可创建维度表

维度表名称：一般是英文名称

维度表中文名：一般是中文名称



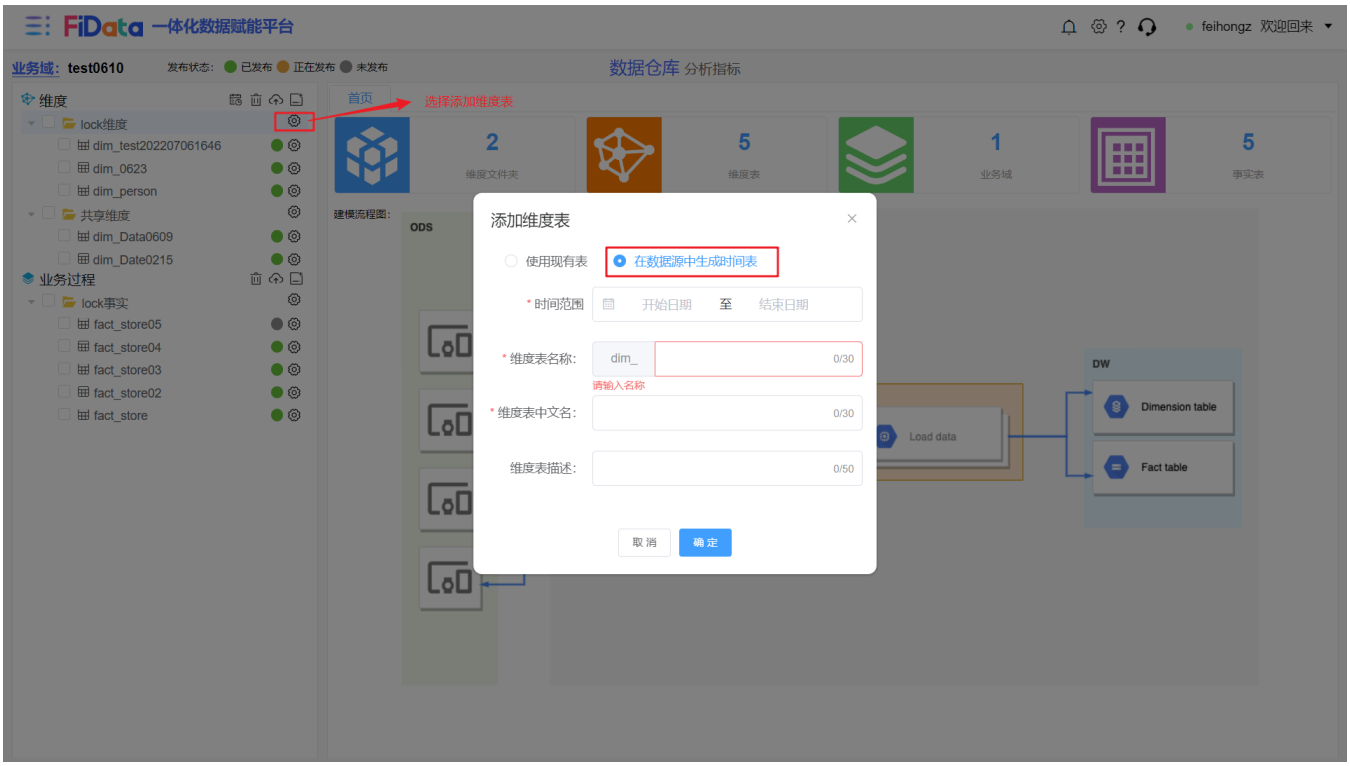
3.2.3 创建日期维度表

创建日期维度表： 点击添加，选择“在数据源中生成时间表”，输入维度表基本信息，即可创建日期维度表

时间范围：用于生成时间范围的维度表

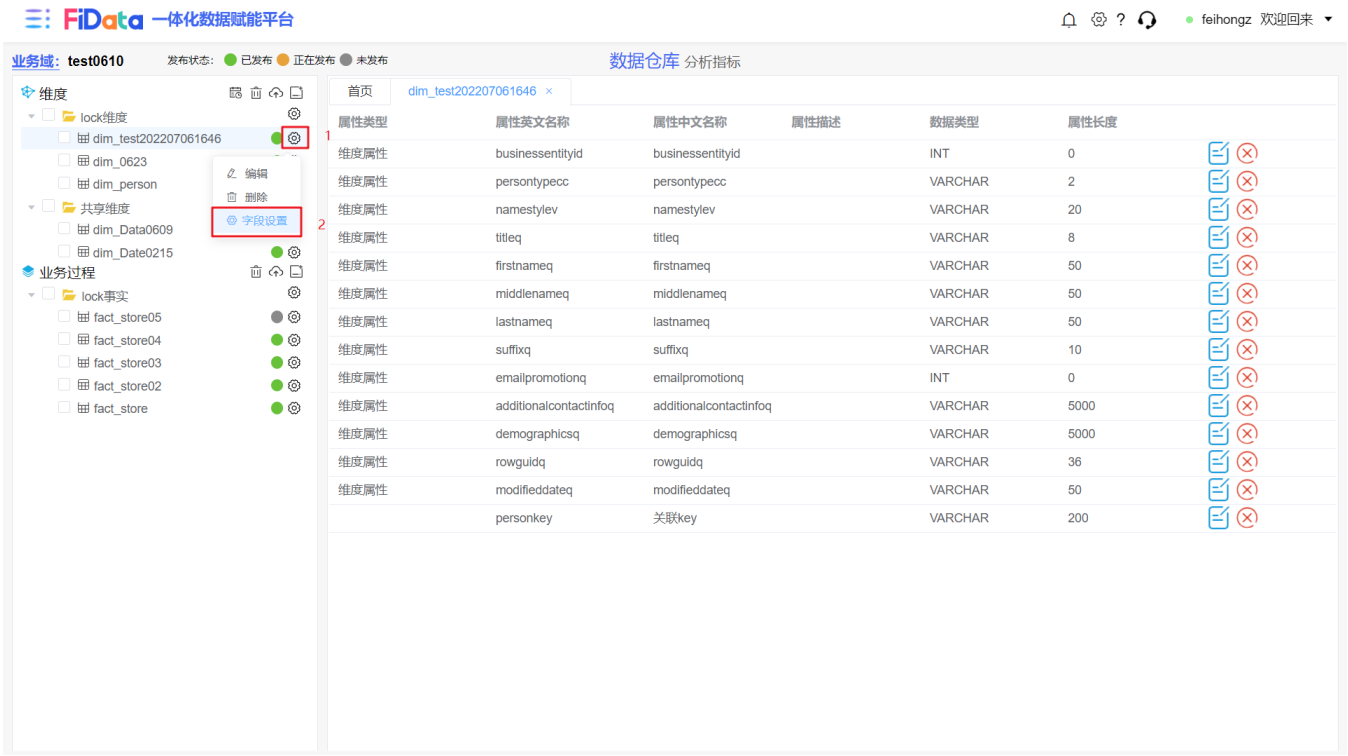
维度表名称：一般是英文名称

维度表中文名：一般是中文名称

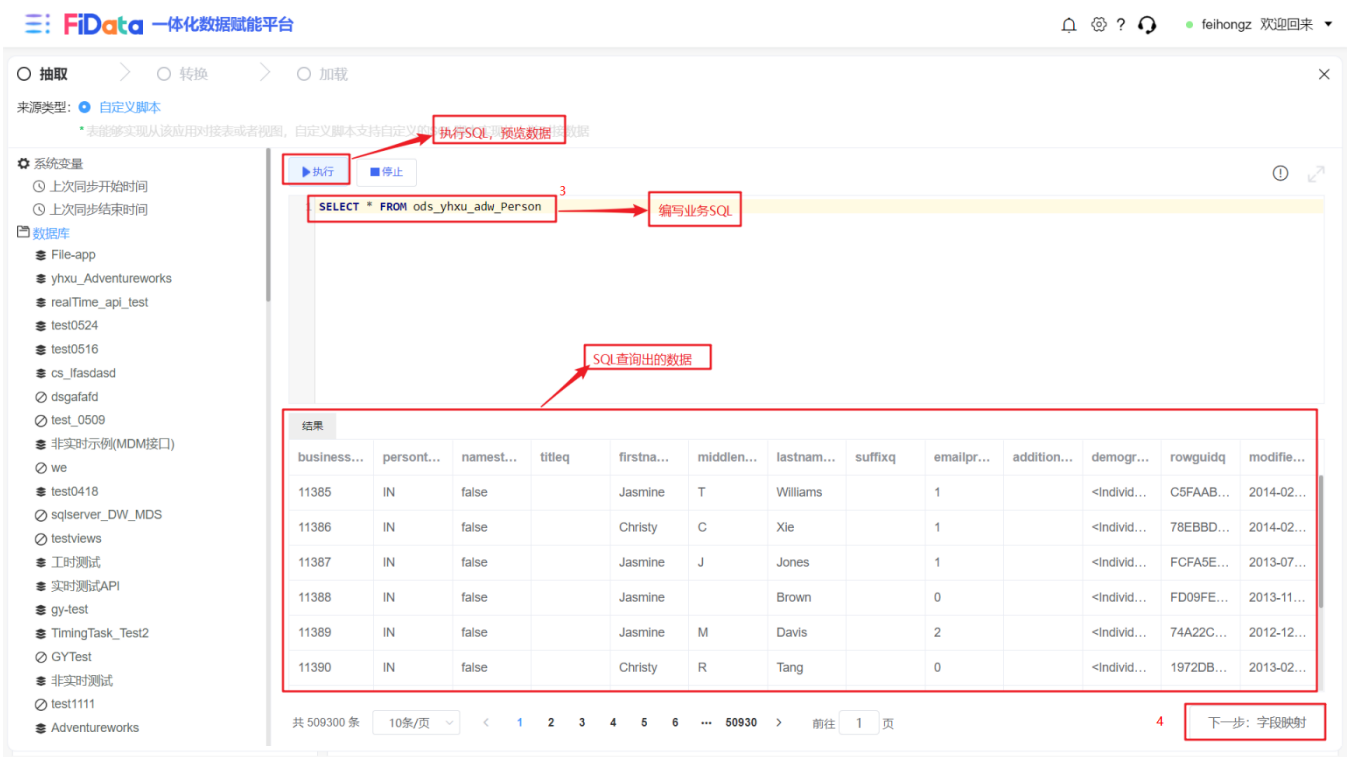


3.2.4 维度表字段设置

1.选择新增的维度表右边的设置logo → 2.选择字段设置；



3.编写业务SQL → 4.进入字段映射

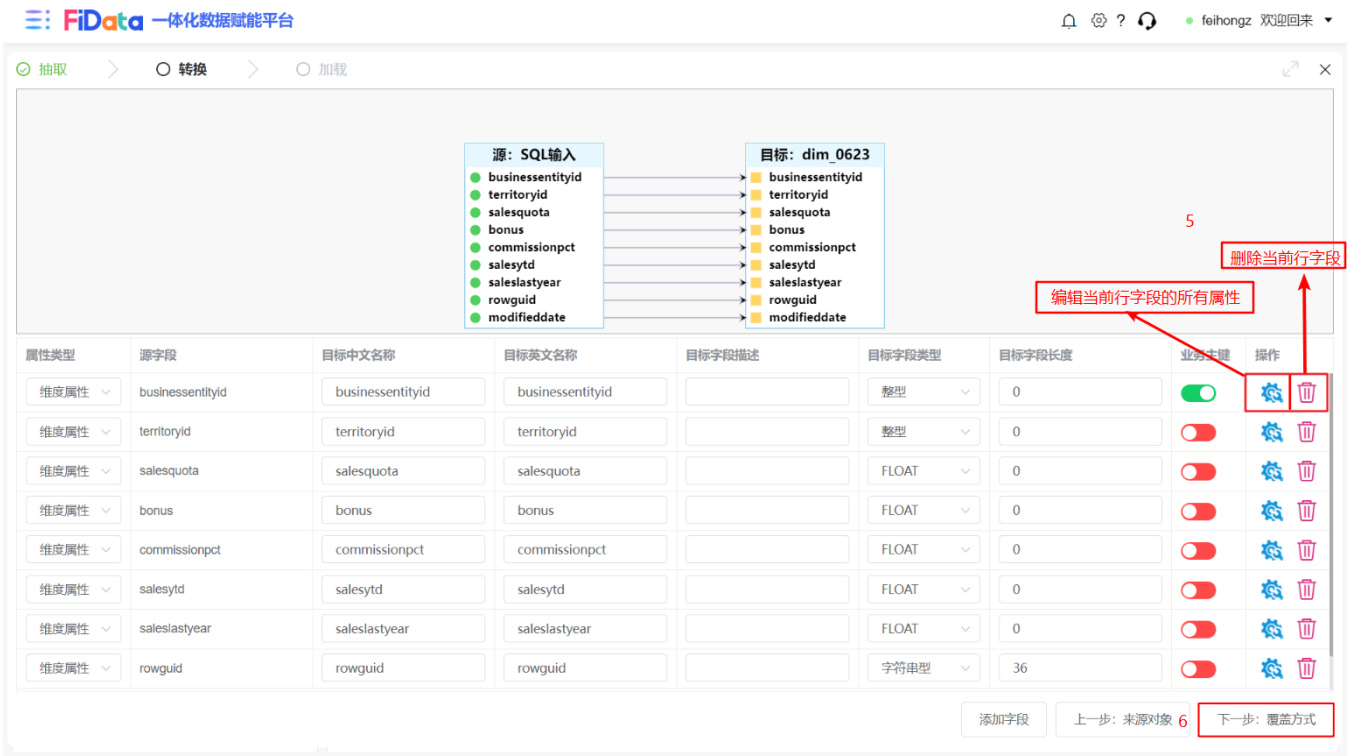


5.操作字段 → 6.进入覆盖方式

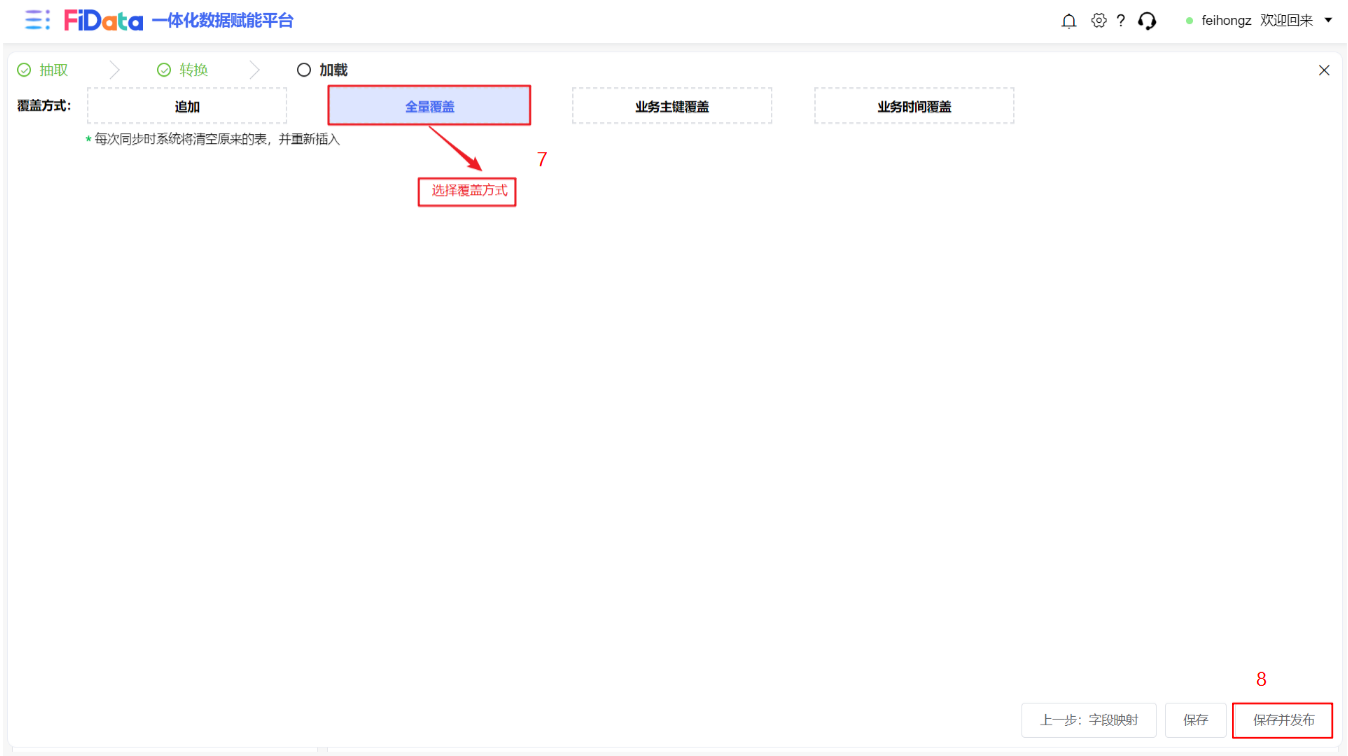
数据类型：

- 整型：
  - 存放数值类型，无小数类型
- 日期类型：
  - 存放日期类型数据。

- 精确数值型：
  - 存放小数类型数据。
  - 设置精度
- 字符串：
  - 存放字符串类型数据。
  - 设置长度



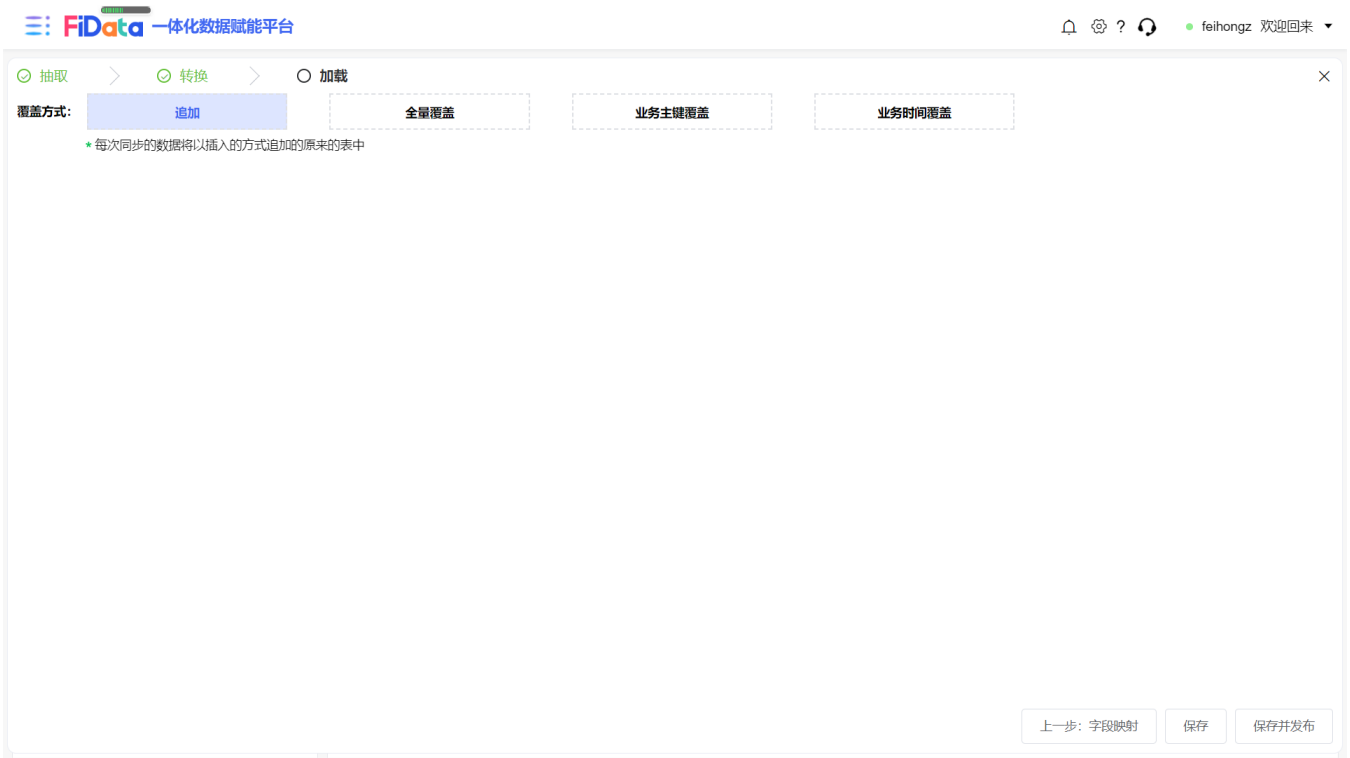
7.选择同步数据的覆盖方式 → 8.保存&发布维度表

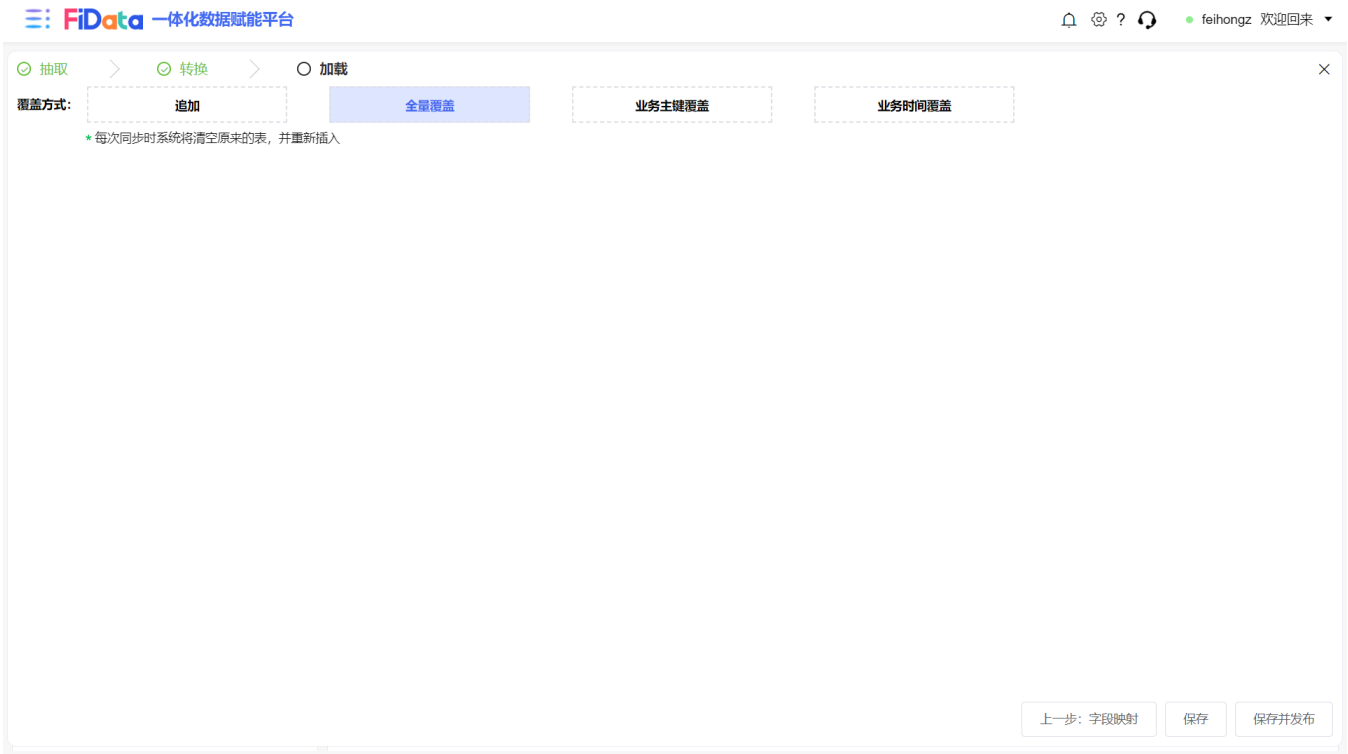


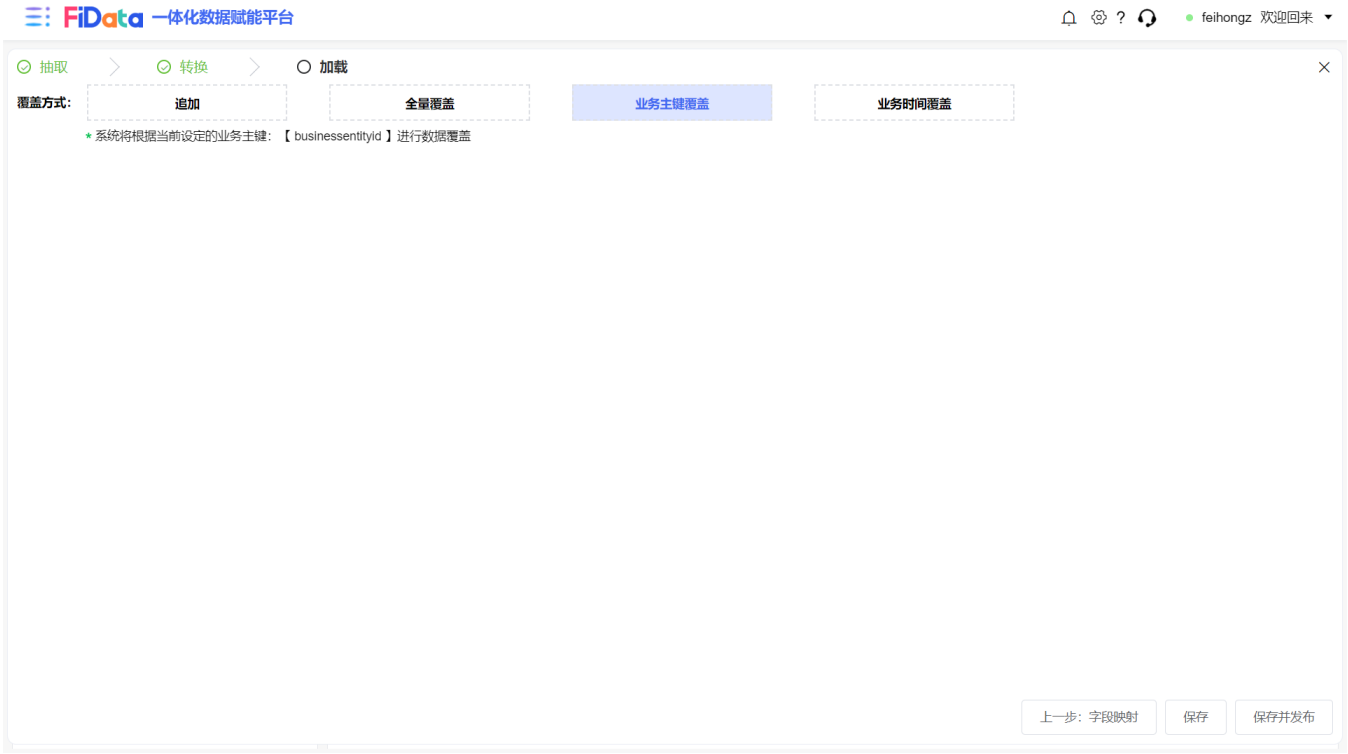
9.填写本次发布描述 → 10.发布

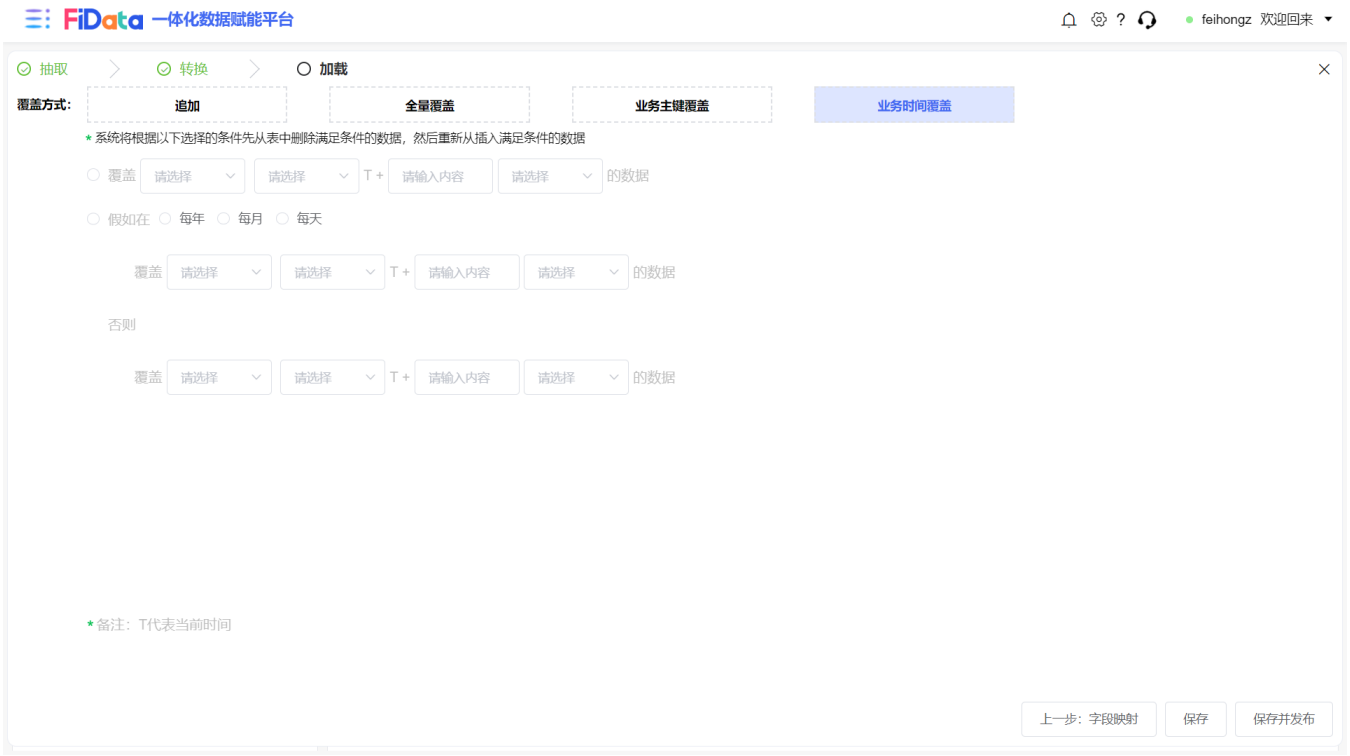
同步方式：

- 追加：每次同步的数据将以插入的方式追加的原来的表中
- 全量：每次同步时系统将清空原来的表，并重新插入
- 业务主键覆盖：系统将根据当前设定的业务主键进行数据覆盖
- 业务时间覆盖：系统将根据以下选择的条件先从表中删除满足条件的数据，然后重新从插入满足条件的数据




FiData 一体化数据赋能平台
feihongz 欢迎回来
抽取
转换
加载
覆盖方式:
追加
全量覆盖
业务主键覆盖
业务时间覆盖
* 每次同步时系统将清空原来的表，并重新插入
上一步：字段映射
保存
保存并发布

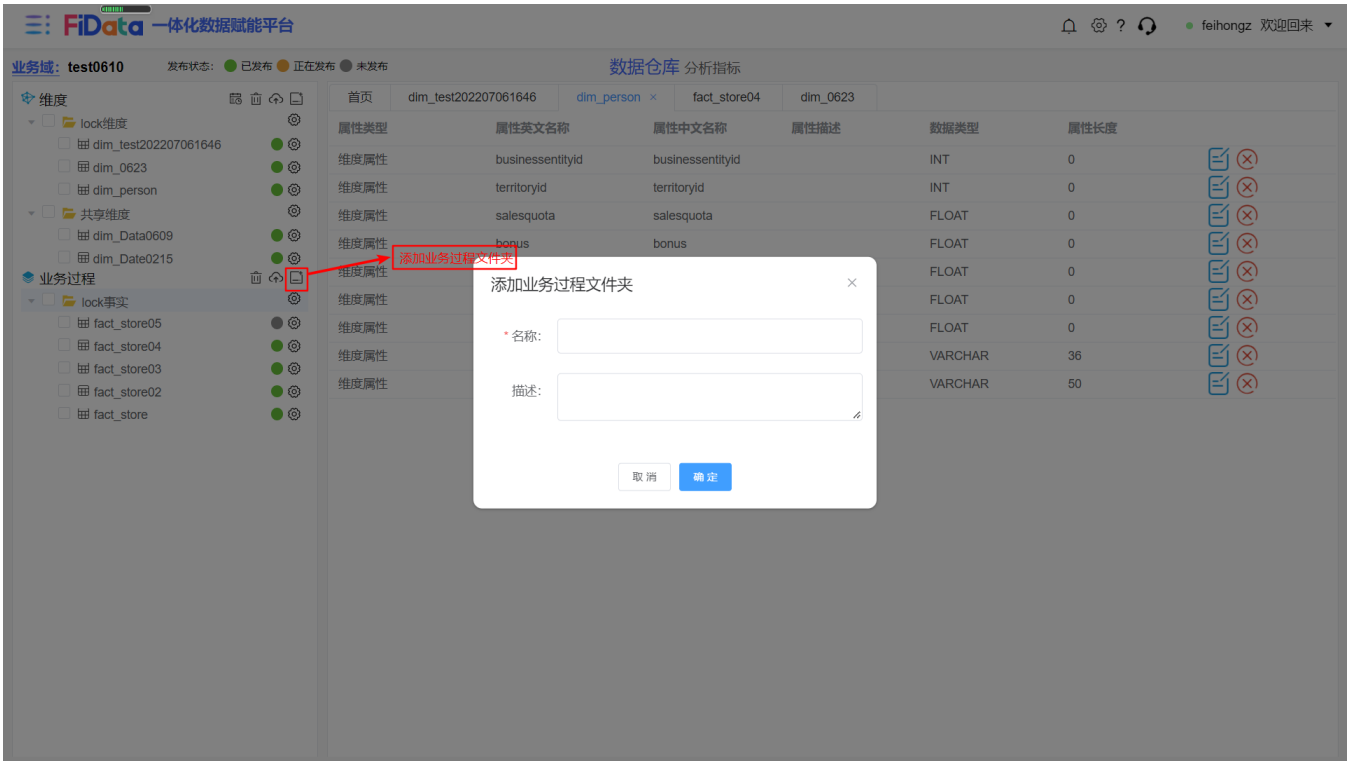

FiData 一体化数据赋能平台
feihongz 欢迎回来
抽取
转换
加载
覆盖方式:
追加
全量覆盖
业务主键覆盖
业务时间覆盖
* 系统将根据当前设定的业务主键：【 businessentityid 】进行数据覆盖
上一步：字段映射
保存
保存并发布



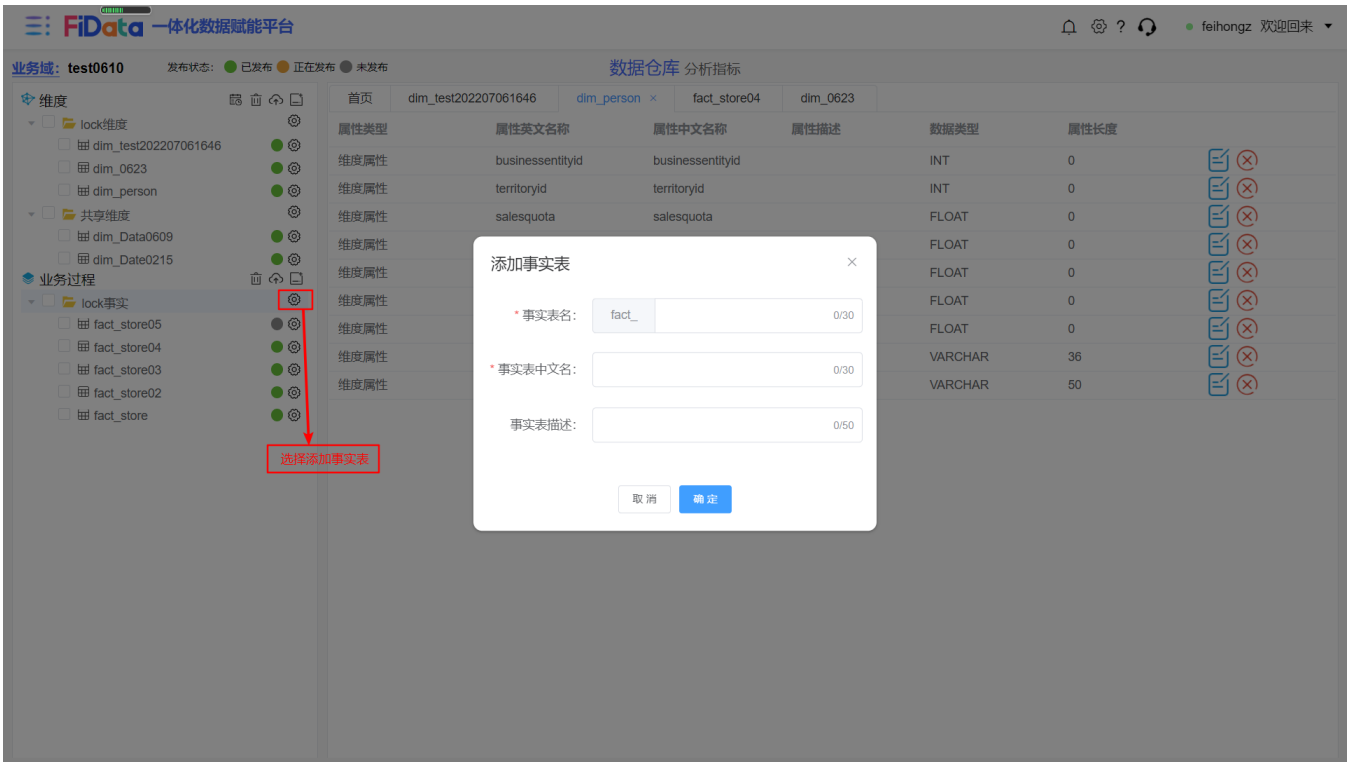
3.2.5 创建业务过程文件夹

点击添加，输入基本信息，即可创建业务过程文件夹

名称：中英文名称皆可



3.2.6 创建事实表



创建事实表： 点击添加，输入事实表基本信息，即可创建事实表

事实表名：一般是英文名称

事实表中文名：一般是中文名称

3.2.7 事实表字段设置

1.选择新增的事实表右边的设置logo → 2.选择字段设置

3.编写业务SQ L→ 4.进入字段映射

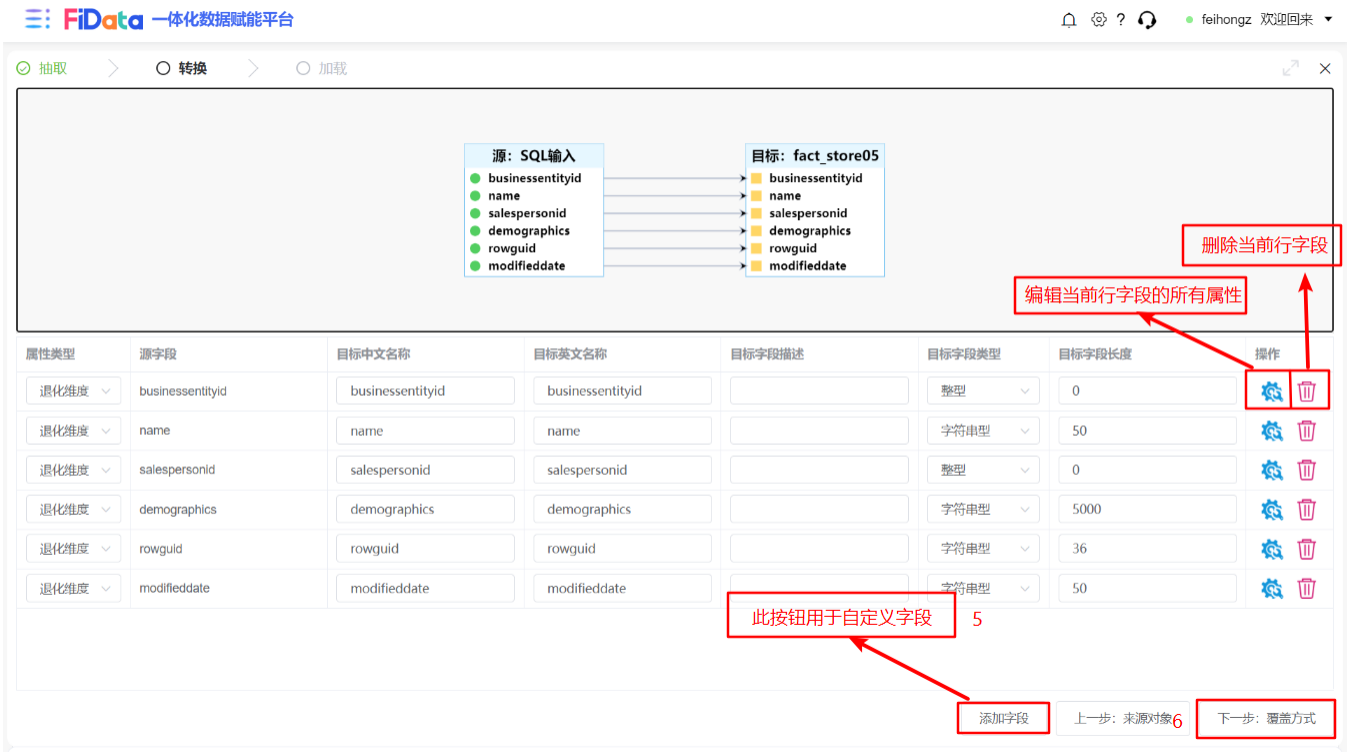
5.添加字段：

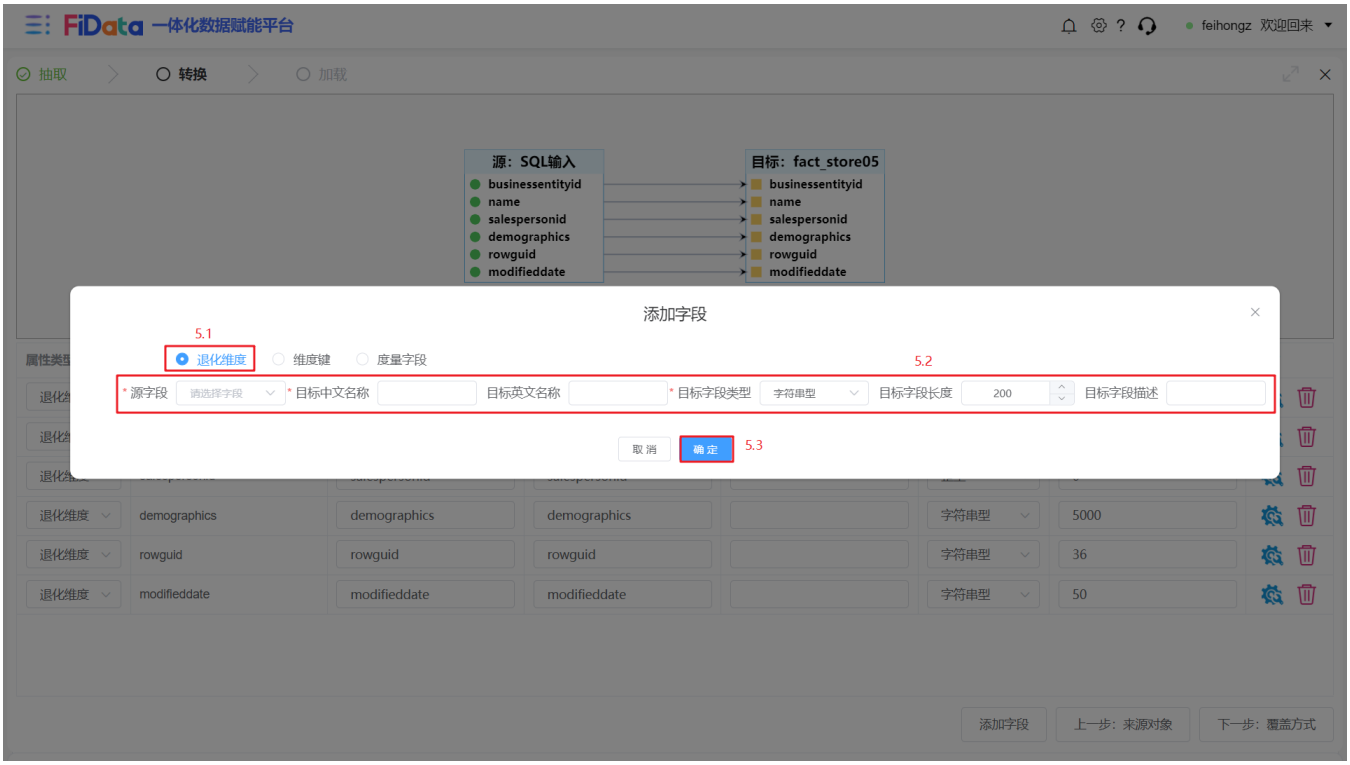
(1)选择字段的属性类型，其中退化维度和度量字段配置一样；

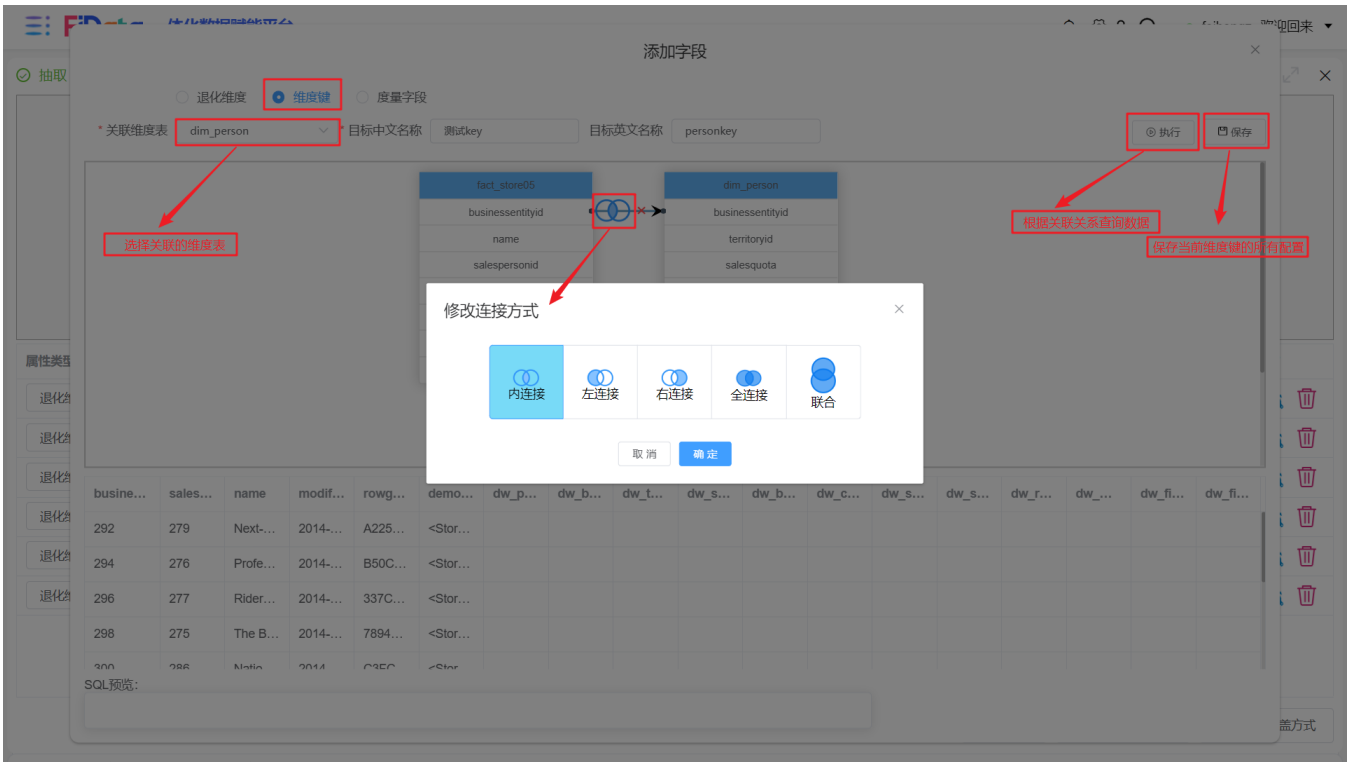
(2)设置字段的其他属性

(3)保存当前字段设置

其他操作同3.4维度表字段设置一致


FiData 一体化数据赋能平台
feihongz 欢迎回来
抽取
转换
加载
源：SQL输入
businessentityid
name
salespersonid
demographics
rowguid
modifieddate
目标：fact_store05
businessentityid
name
salespersonid
demographics
rowguid
modifieddate
删除当前行字段
编辑当前行字段的所有属性
属性类型
源字段
目标中文名称
目标英文名称
目标字段描述
目标字段类型
目标字段长度
操作
退化维度
businessentityid
整型
0
name
字符串型
50
salespersonid
整型
0
demographics
字符串型
5000
rowguid
字符串型
36
modifieddate
字符串型
50
此按钮用于自定义字段
5
添加字段
上一步：来源对象
6
下一步：覆盖方式


FiData 一体化数据赋能平台
feihongz 欢迎回来
抽取
转换
加载
源：SQL输入
businessentityid
name
salespersonid
demographics
rowguid
modifieddate
目标：fact_store05
businessentityid
name
salespersonid
demographics
rowguid
modifieddate
添加字段
5.1
退化维度
维度键
度量字段
5.2
源字段
请选择字段
目标中文名称
目标英文名称
目标字段类型
字符串型
目标字段长度
200
目标字段描述
取 消
确 定
5.3
添加字段
上一步：来源对象
下一步：覆盖方式

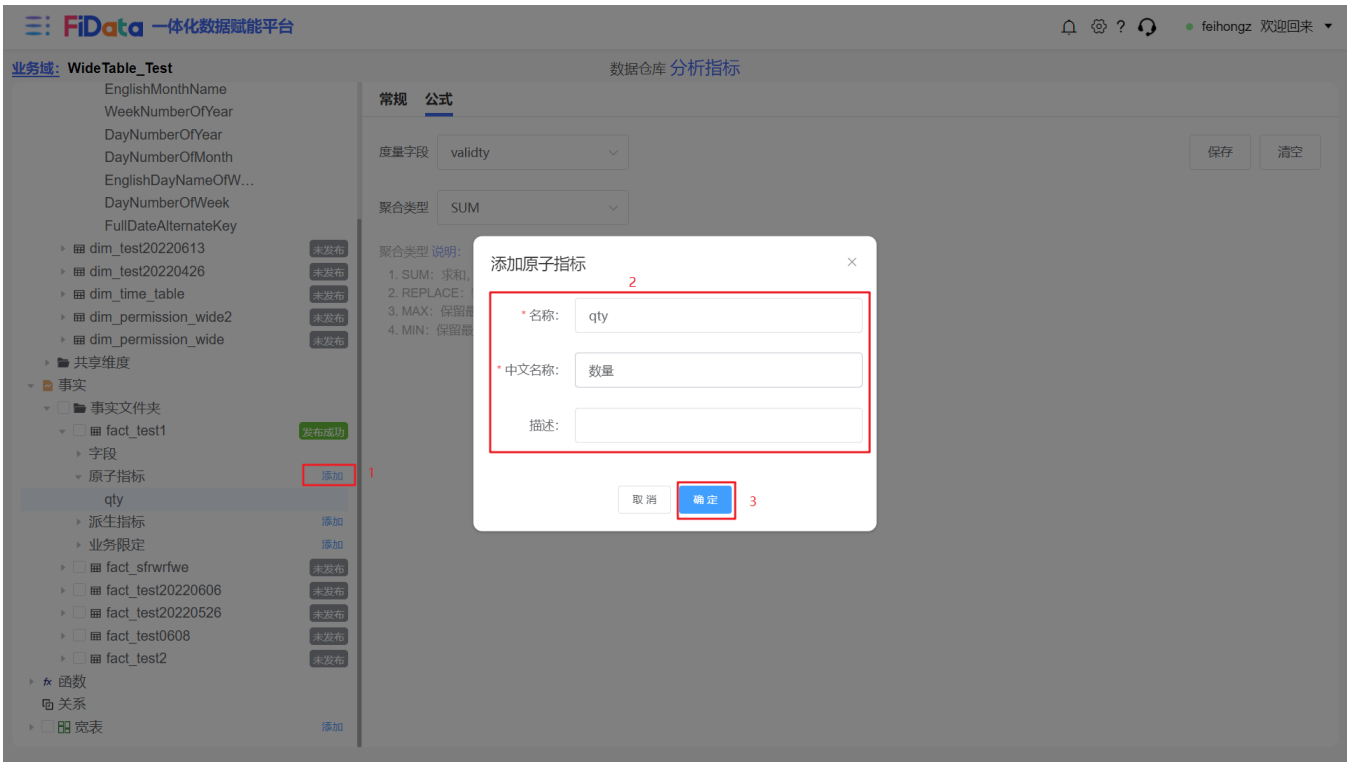


3.3 操作分析指标

3.3.1 分析指标维度表

3.3.2 创建原子指标

1.添加事实表的原子指标 → 2.填写原子指标的基本信息 → 3.保存基本信息



4.选择配置公式模式 → 5.选择具体的度量字段，配置聚合类型 →6.保存所有配置

聚合类型 说明:

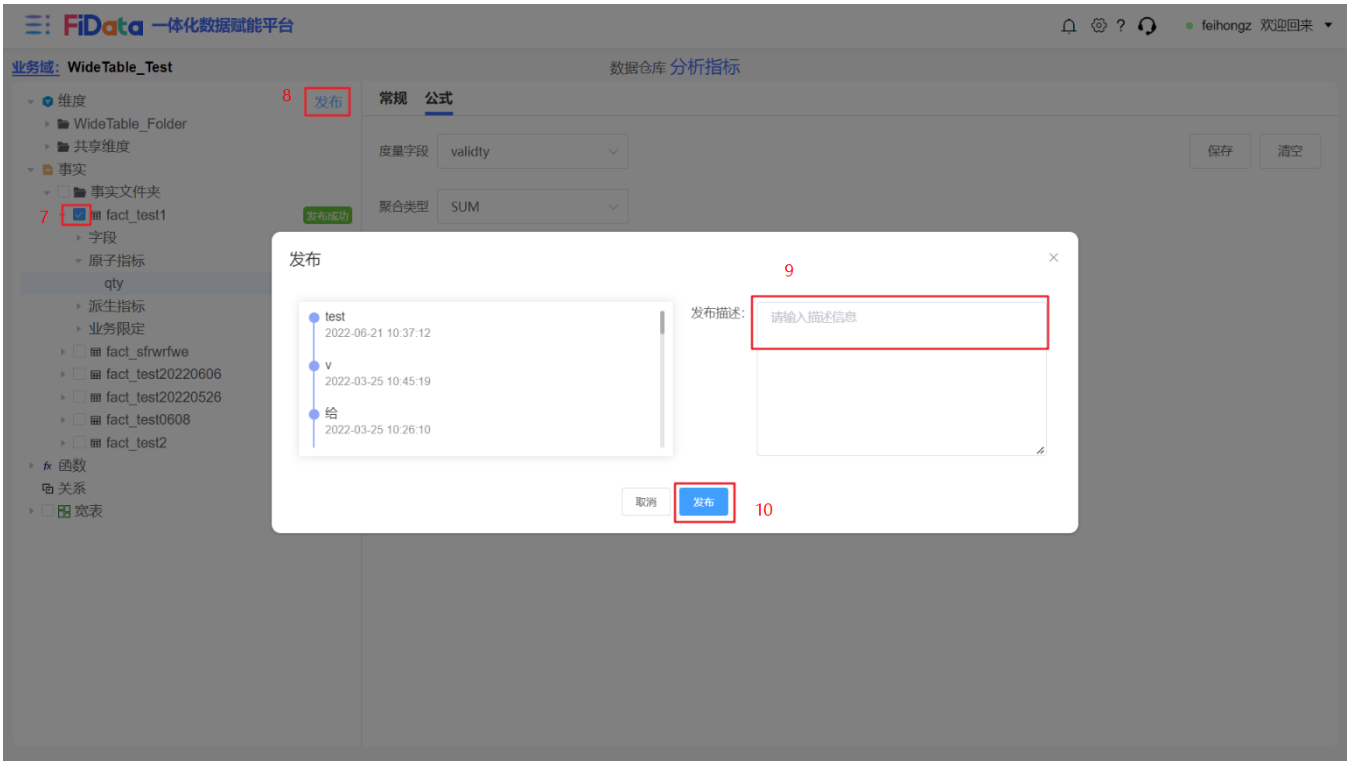
1. SUM：求和，多行的 Value 进行累加

2. REPLACE：替代，下一批数据中的 Value 会替换之前导入过的行中的 Value

3. MAX：保留最大值

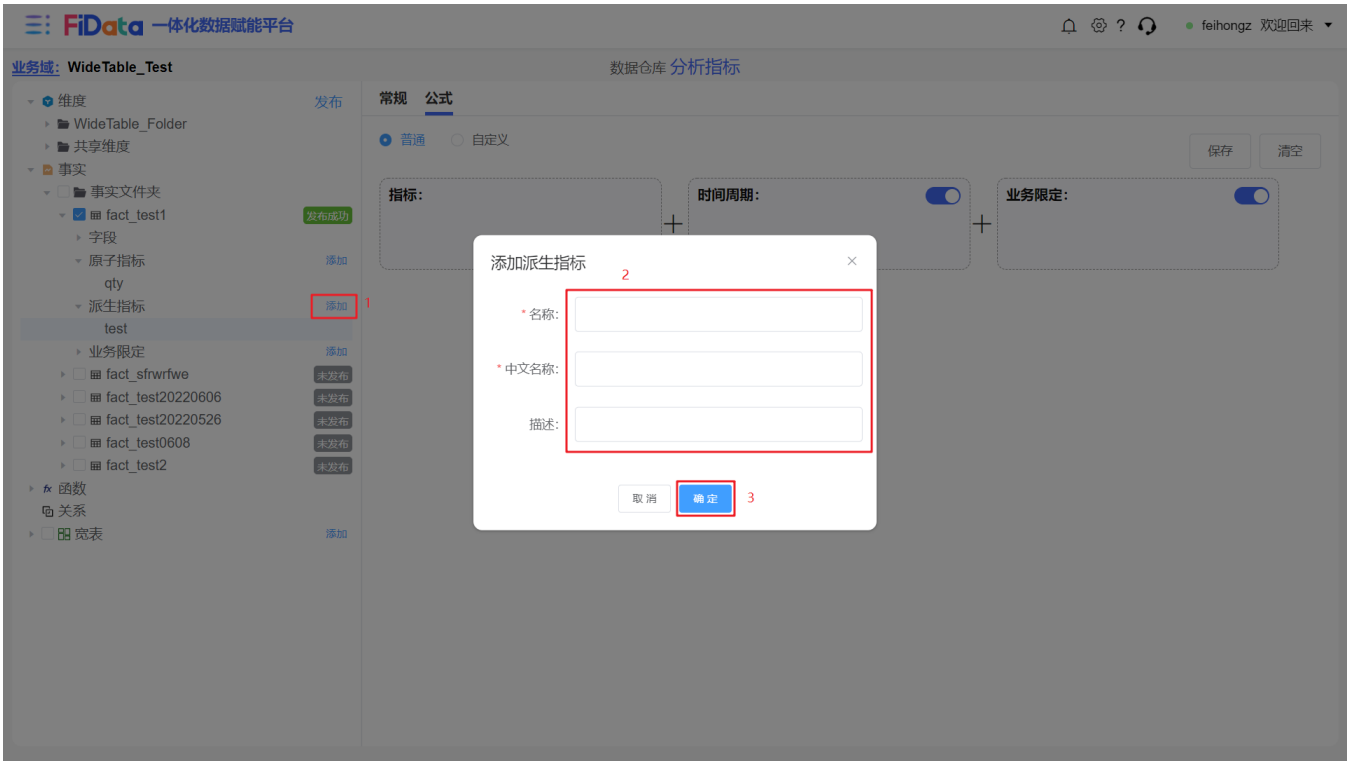
4. MIN：保留最小值

7.选择需要发布的事实表 → 8.选择发布按钮 → 9.填写本次发布的描述 → 10.发布



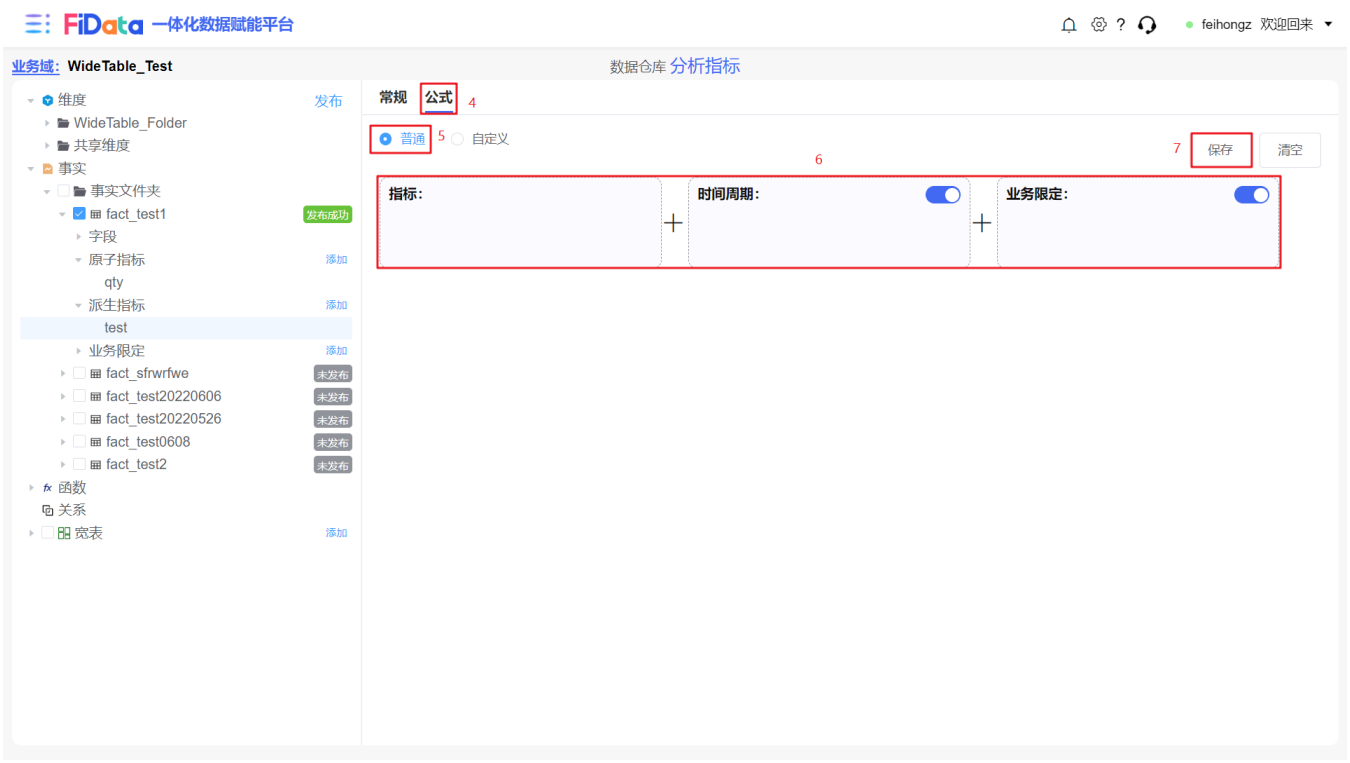
3.3.3 创建派生指标

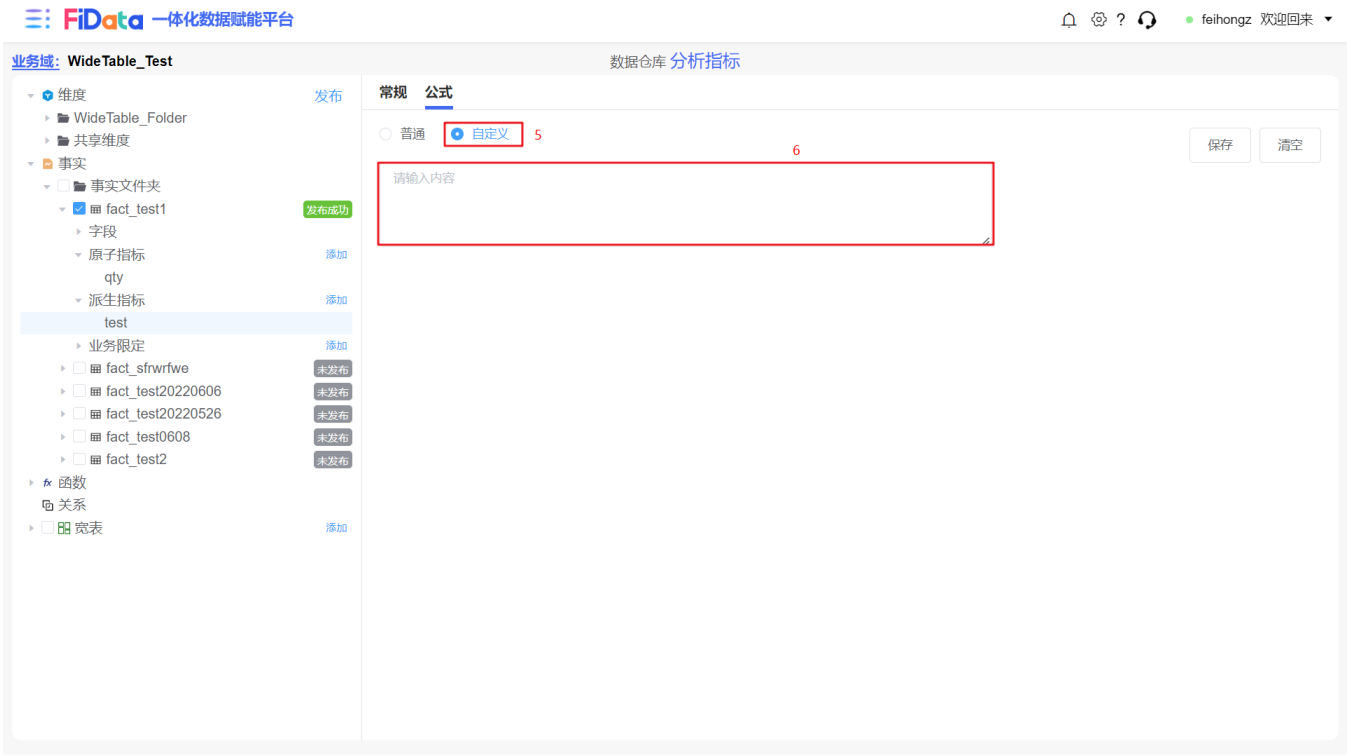
1.添加事实表的派生指标 → 2.填写派生指标的基本信息 → 3.保存基本信息



4.选择配置公式模式 → 5.选择普通公式or自定义公式 → 6.配置公式

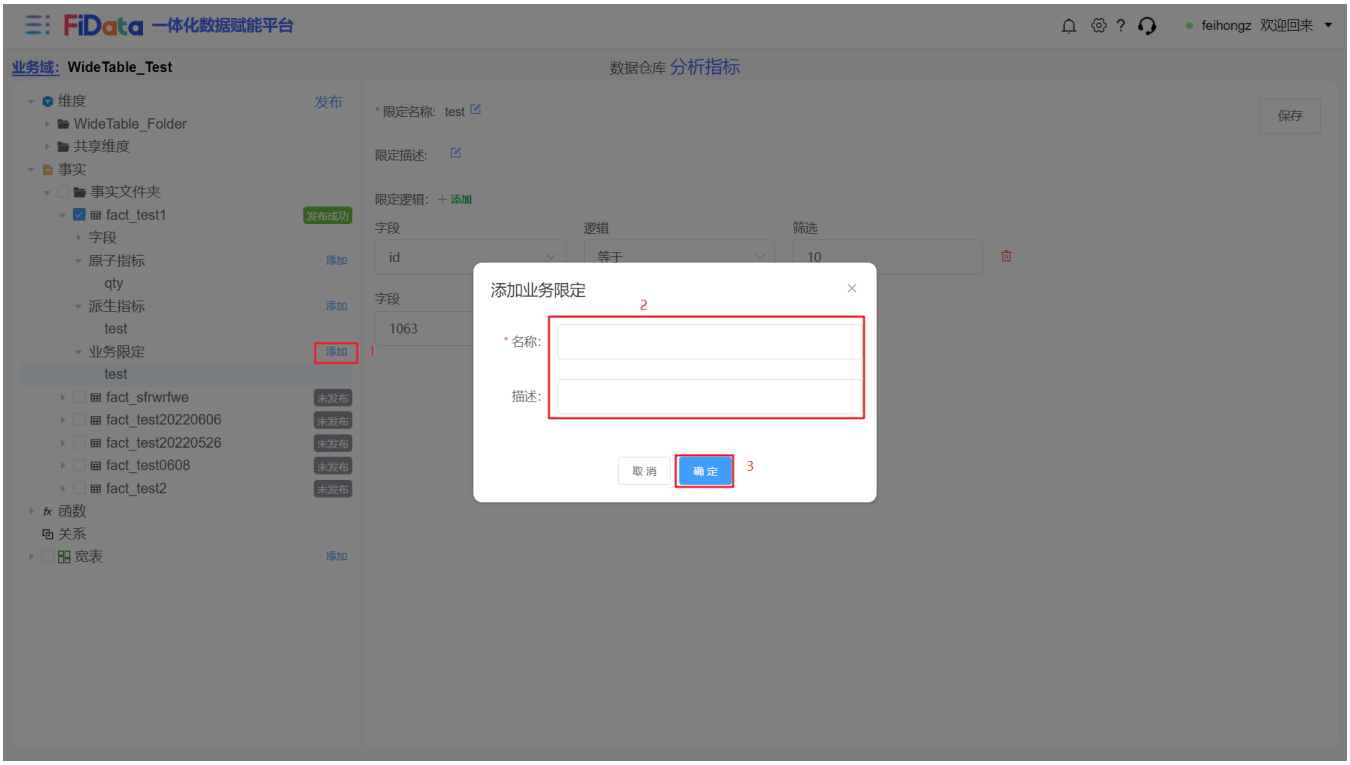
其他发布操作同3.3.2一样





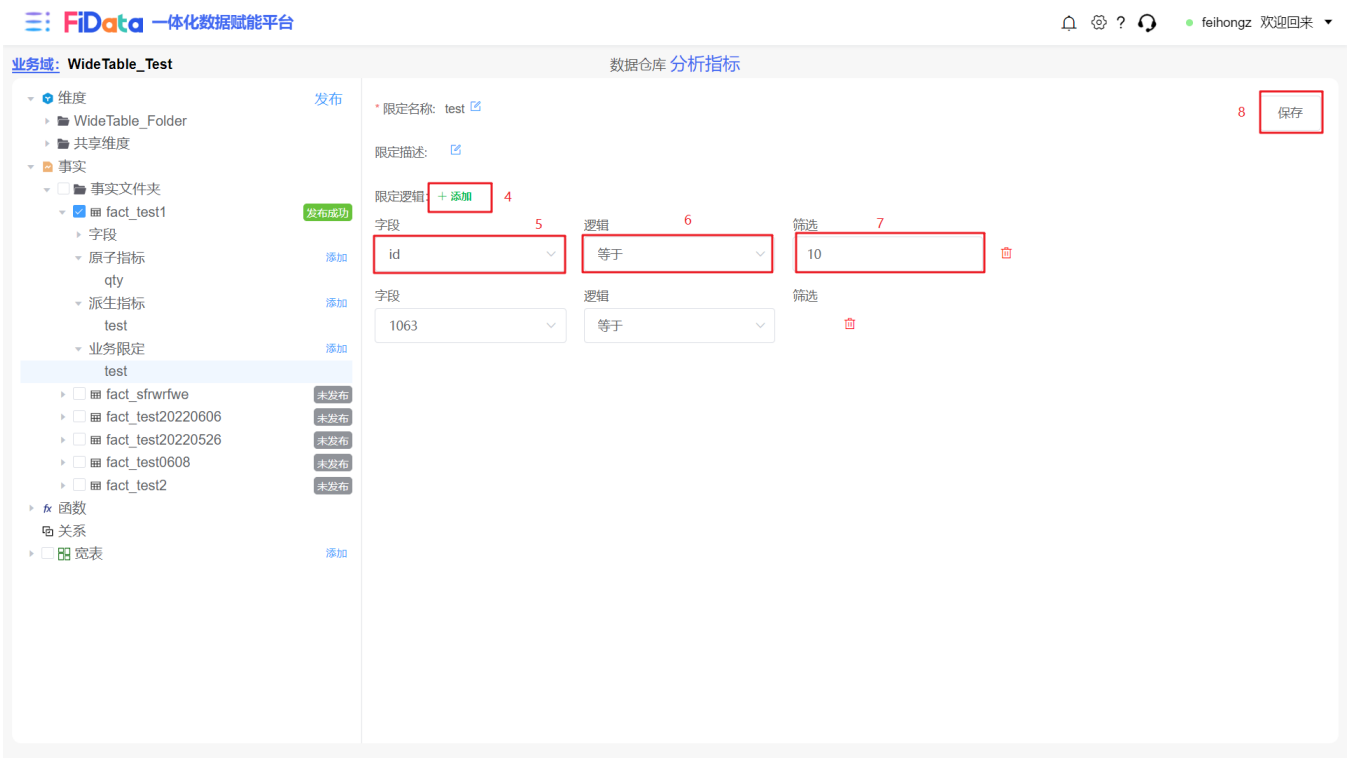
3.3.4 创建业务限定

1.添加事实表的业务限定 → 2.填写业务限定的基本信息 → 3.保存基本信息；



4.添加限定逻辑 → 5.下拉选择需要限定的字段 → 6.选择字段的限定逻辑：等于、大于、小于、包含 → 7.选择字段的筛选条件 → 8.保存所有配置

其他发布操作同3.3.2一样。



3.3.5 创建宽表

1.添加宽表

2.填写宽表名称，一般是英文名称

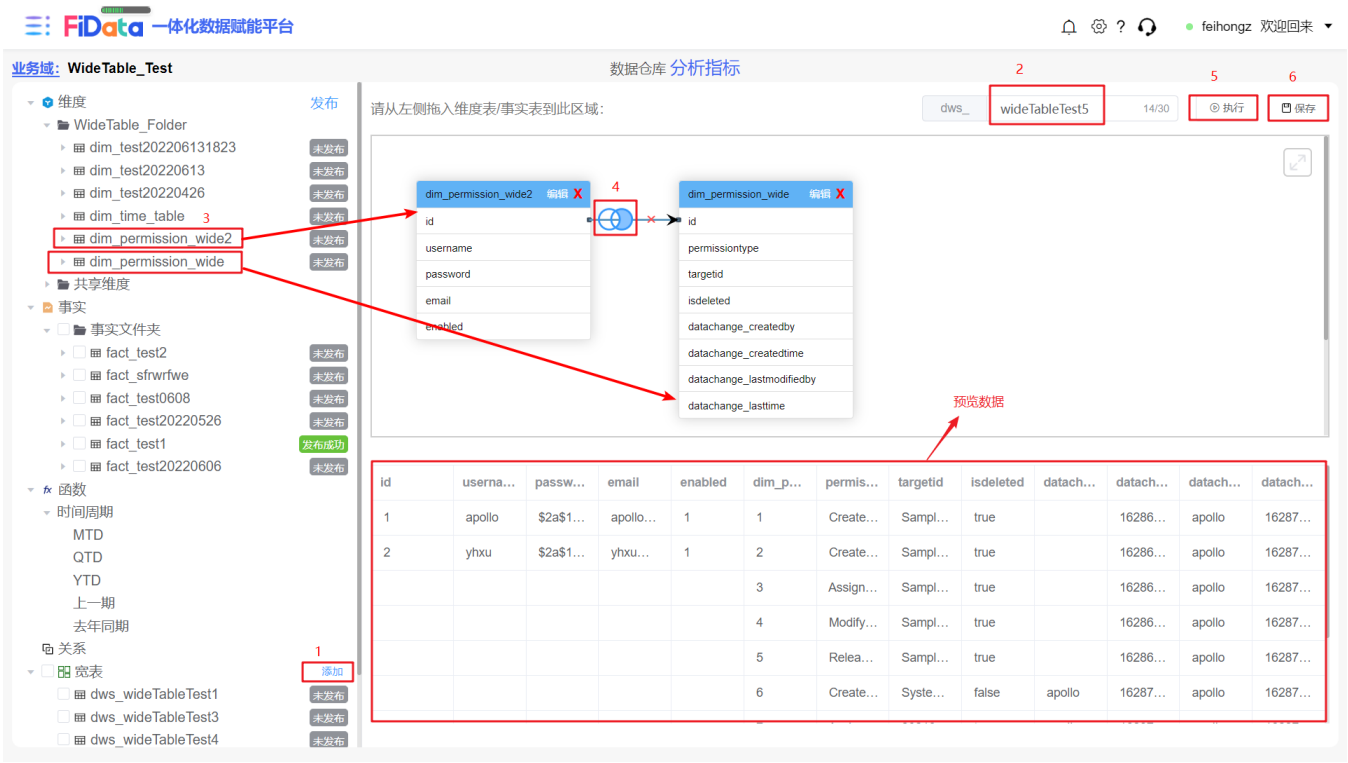
3.从左边维度or事实中拖拽多张表到中间的编辑框

4.设置连接关系：内连接、左连接、右连接、全连接、联合

5.执行多表的连接关系

6.保存所有配置

其他发布操作同3.3.2一样


FiData 一体化数据赋能平台
feihongz 欢迎回来
业务域: WideTable_Test
数据仓库 分析指标
维度
发布
WideTable_Folder
dim_test202206131823
dim_test20220613
dim_test20220426
dim_time_table
dim_permission_wide2
dim_permission_wide
共享维度
事实
事实文件夹
fact_test2
fact_sfrwrfwe
fact_test0608
fact_test20220526
fact_test1
发布成功
fact_test20220606
函数
时间周期
MTD
QTD
YTD
上一期
去年同期
关系
宽表
添加
dws_wideTableTest1
dws_wideTableTest3
dws_wideTableTest4
未发布
1
2
3
4
5
6
请从左侧拖入维度表/事实表到此区域:
dws_
wideTableTest5
14/30
执行
保存
dim_permission_wide2 编辑
id
username
password
email
enabled
dim_permission_wide 编辑
id
permissiontype
targetid
isdeleted
datachange_createdby
datachange_createdtime
datachange_lastmodifiedby
datachange_lasttime
预览数据
id
userna...
passw...
email
enabled
dim_p...
permis...
targetid
isdeleted
datach...
datach...
datach...
datach...
1
apollo
$2a$1...
apollo...
1
1
Create...
Sampl...
true
16286...
apollo
16287...
2
yhxu
$2a$1...
yhxu...
1
2
Create...
Sampl...
true
16286...
apollo
16287...
3
Assign...
Sampl...
true
16286...
apollo
16287...
4
Modify...
Sampl...
true
16286...
apollo
16287...
5
Relea...
Sampl...
true
16286...
apollo
16287...
6
Create...
Syste...
false
apollo
16287...
apollo
16287...