

Smartbi V10.5.8-数据准备

注意：（新特性列表中：+表示新增；^表示增强）

V10.5.8 版本重点详细改进情况如下：

新增	增强
<div>+【指标管理】指标模型支持在线编辑数据</div> <div>+【指标管理】指标模型资源迁移增强</div> <div>+【指标管理】指标模型支持在线编辑数据</div>	<div><【指标管理】屏蔽派生指标和复合指标</div> <div>^【ETL】横版ETL开发</div>
<div>+【数据模型】支持向导创建自定义组</div> <div>+【作业流】支持在作业流中执行数据模型的抽取</div> <div>+【作业流】作业流支持 Foreach 循环器</div> <div>+【作业流】支持更简单“外部ETL”跑完状态</div>	

背景介绍

指标模型维度和指标的数据，当前只能通过自助ETL功能进行数据准备，这个数据加工过程漫长且技术门槛高。

1. 采用模拟数据，直接基于仪表盘制作交互原型，在实施前期与甲方共同确认目标，让前期投入后期直接复用同时，也让报表制作和数据加工过程更简单。

2. 简化产品演示、demo演示、功能试用等数据准备过程，不需要通过自助ETL进行数据准备，仅通过Excel导入或手工编辑即可完成，增强产品易用性。

解决方案

在Smartbi V10.5.8 版本中，指标模型的维度和指标数据支持在线编辑表数据的功能，支持 **手工录入、复制粘贴** 等操作，快速生成模拟数据即可完成数据准备过程。

指标模型的维度和事实表数据预览界面，点击 **在线编辑** 。



⚠ 在线编辑的是指标存储库中的数据，此操作不会影响源业务系统的数据。

参考文档

在线编辑数据操作详情请参见：[在线编辑数据](#)。

+【指标管理】指标模型资源迁移增强

背景介绍

指标模型在持续迭代的过程中，往往需要先在测试环境进行验证，数据准确后再迁移至生产环境，且需保证生产环境中指标模型的定义信息和指标存储库中的数据信息保持一致。如何提高运维人员在做资源迁移时的效率，快速在新环境中恢复资源，成为亟待解决的问题。

解决方案

在Smartbi V10.5.8 中，为检测指标模型的中定义信息和数据信息是否匹配，提供 **检查表结构** 功能，可快速识别不一致的项并提供 **全量覆盖建表** 功能和 **增量建表** 功能，最终保证定义信息和数据信息的一致性。

- 指标模型右键菜单增加 **检查表结构**，找出当前模型结构定义和指标存储库中物理表结构不一致的维表和事实表，并支持对比查看表中有哪些字段不一致。

- 提供 **全量覆盖建表**，根据当前指标模型中定义的所有维度和事实表，全部覆盖指标存储库中的同名表。

- 提供 **增量建表**，根据当前指标模型中定义的维度和事实表，只批量创建指标存储库中不存在的表。



参考文档

检查表结构操作详情请参见：[检查表结构](#)。

+ 【指标管理】事实表关联维度/指标自动填充字段

背景介绍

企业在持续发展的过程中，维度会根据实际的业务情况而发生变化，维度的变化将会影响到所有相关的指标，。

功能简介

在Smartbi V10.5.8 中，事实表关联维度和指标时，能自动填入对应字段，极大地提高ETL工程师批量修改关系时的工作效率。

事实表中，如果字段名和字段别名没有被修改过，则：

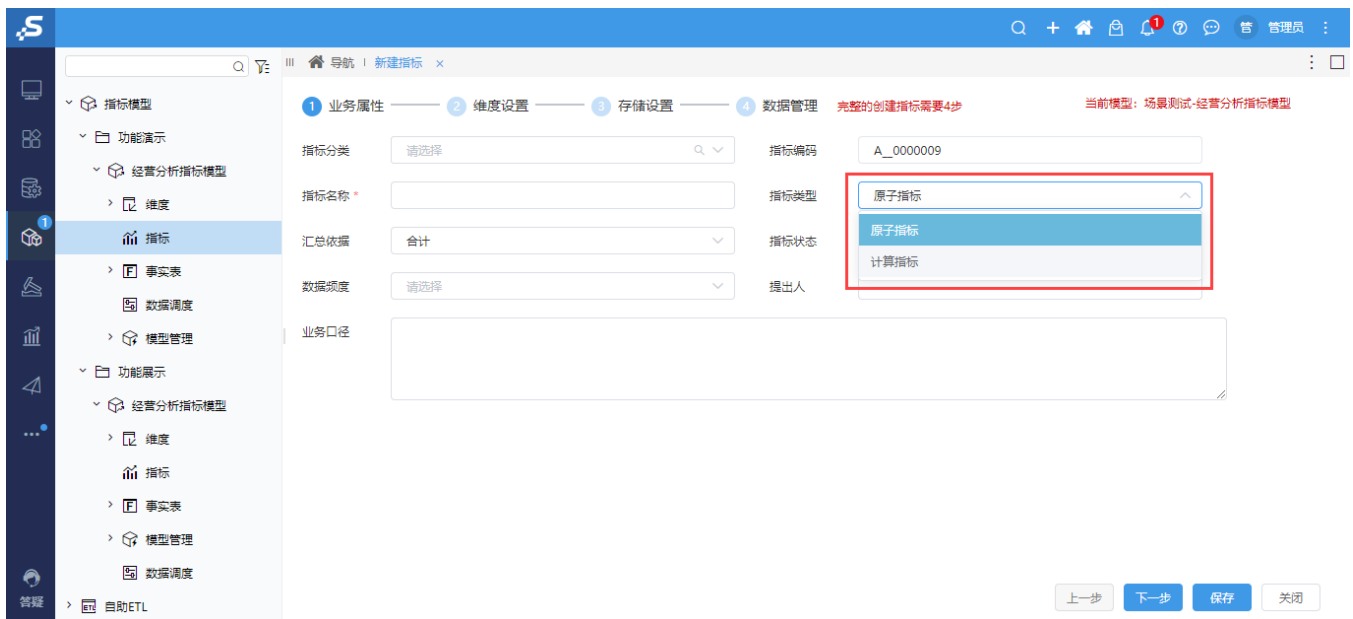
- 关联维度，自动填充对应维度的主键字段名和别名作为字段名和字段别名。
- 关联指标，自动填充该指标的指标编码和指标名称作为字段名和字段别名。



◀【指标管理】屏蔽派生指标和复合指标

功能简介

- 指标类型中，屏蔽派生指标和复合指标。
- 原 虚拟指标 更名为 计算指标。



派生指标和复合指标的功能目前还不完善，所以在Smartbi V10.5.8 中暂时先屏蔽派生指标和复合指标，等后续功能优化完后再开放使用。

+【数据模型】支持创建自定义分组字段

背景介绍

分组字段是指对原始数据根据条件将相同的数据先合并到一组，然后按照分组后的数据进行汇总计算。

在V10.5 beta版本中，如果要新增一个分组字段，有2种方法：

- 1) **写MDX语句**：但是大部分人不会写
- 2) **写SQL语句**：不懂技术的业务人员不会写SQL，只能依靠技术人员

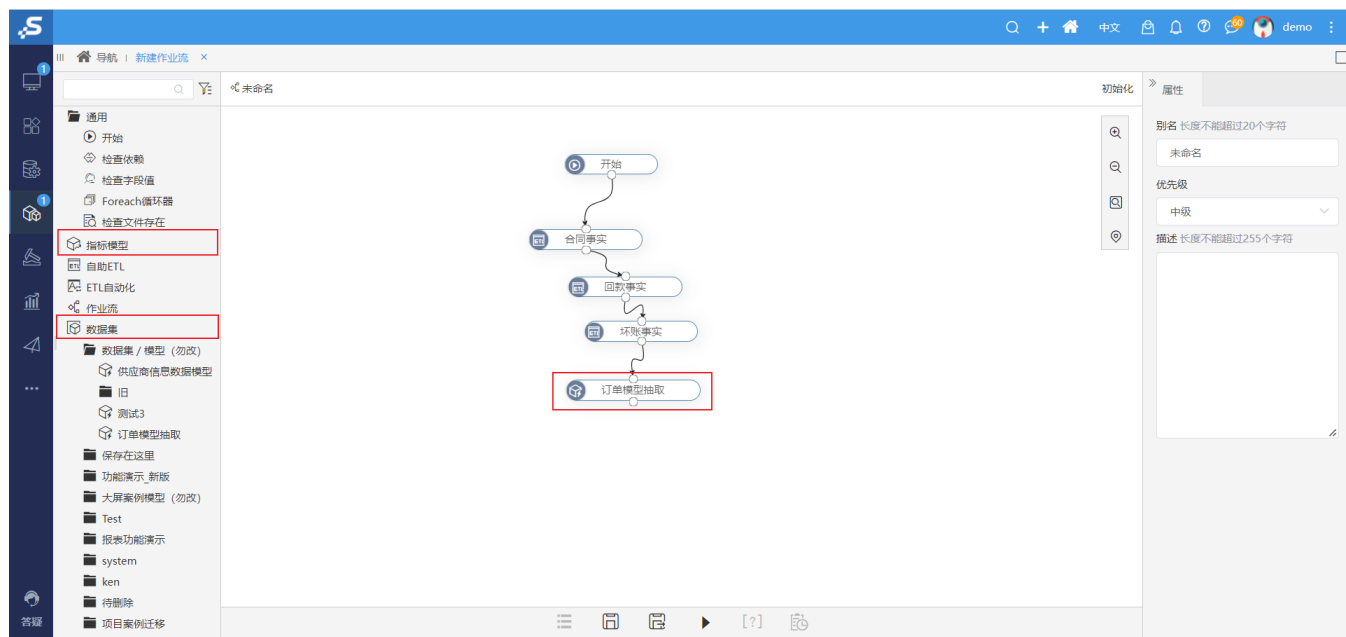
所以在V10.5.8版本中，数据模型以及交互式仪表盘等均可通过界面可视化自定义分组字段，让业务人员轻松创建分组字段，减少对技术人员的依赖，提高了功能的易用性。

功能介绍

1) 在数据模型支持自定义分组，支持对字符串、数值、日期进行枚举分组以及范围分组；

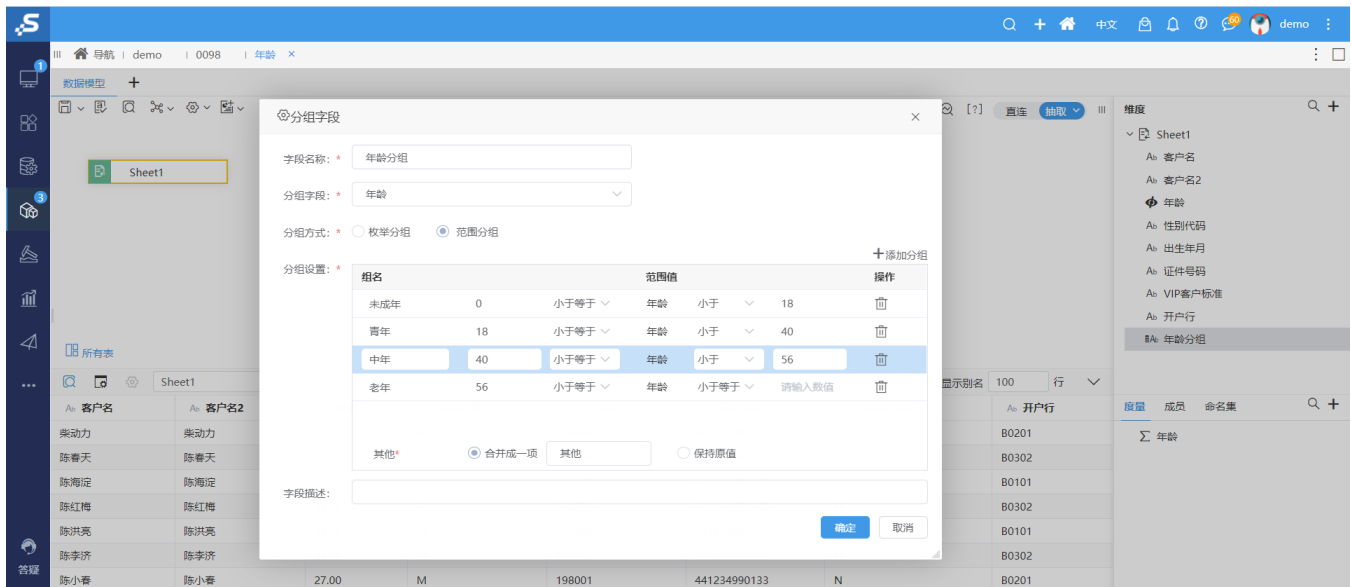
例如，设置城市分组，划分出几个大区如下：

- 一线城市：北京、上海、深圳
- 二线城市：南京、南昌、厦门、大连、常州、天津
- 三线城市：张家口、昆明、成都、济南、烟台



例如，对年龄分组。例如，客户年龄字段分组，划分出几个大区如下。

- 0岁≤未成年<18岁
- 18岁≤青年<40岁
- 40岁≤中年<56岁
- 老年≥56岁



2) 交互式仪表盘、即席查询也支持创建报表资源级别的自定义分组字段即只对该报表生效，不会保存到数据模型。

参考文档

详情请参考 [在数据模型中如何自定义分组](#)。

+【作业流】支持在作业流中执行数据模型的抽取

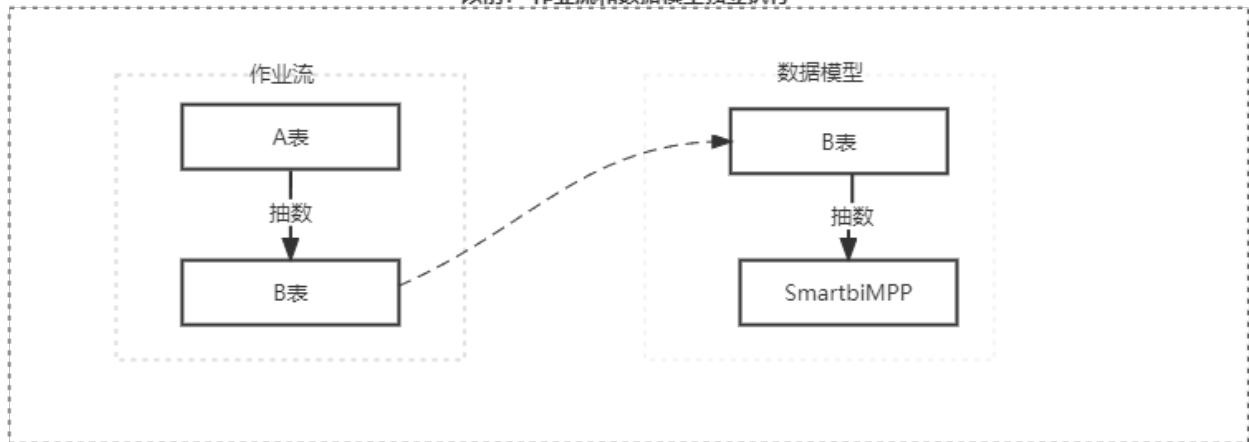
背景介绍

在V10.5 beta版本：

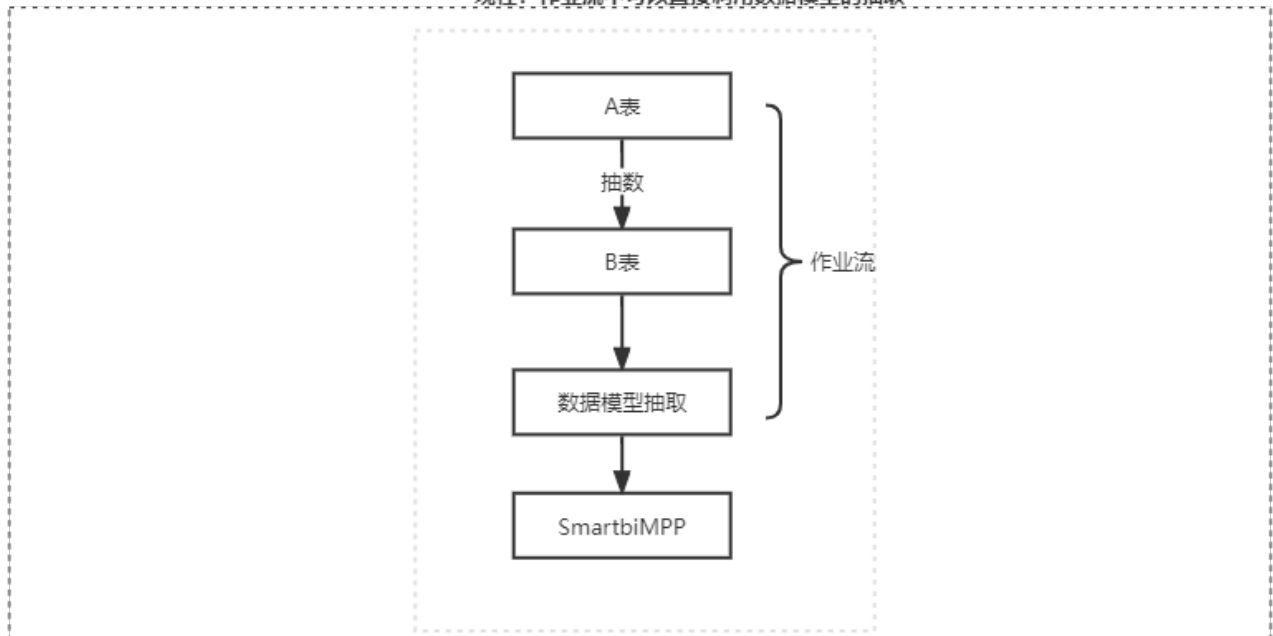
- 1) 作业流和数据模型的抽取是**独立分开**进行的，但是数据模型的数据依赖于作业流即作业流先抽完，数据模型才能开始抽数，否则会出现数据不正确；
- 2) 作业流执行完了之后数据模型再抽取数据，需要自定义任务(写代码)才能实现，而且**非常复杂，耗时耗力**；

而在V10.5.8，作业流可以通过拖、拉、拽轻松设置数据模型与ETL执行的先后顺序，无需写代码、轻轻松松确保生产安全！

以前：作业流和数据模型独立执行

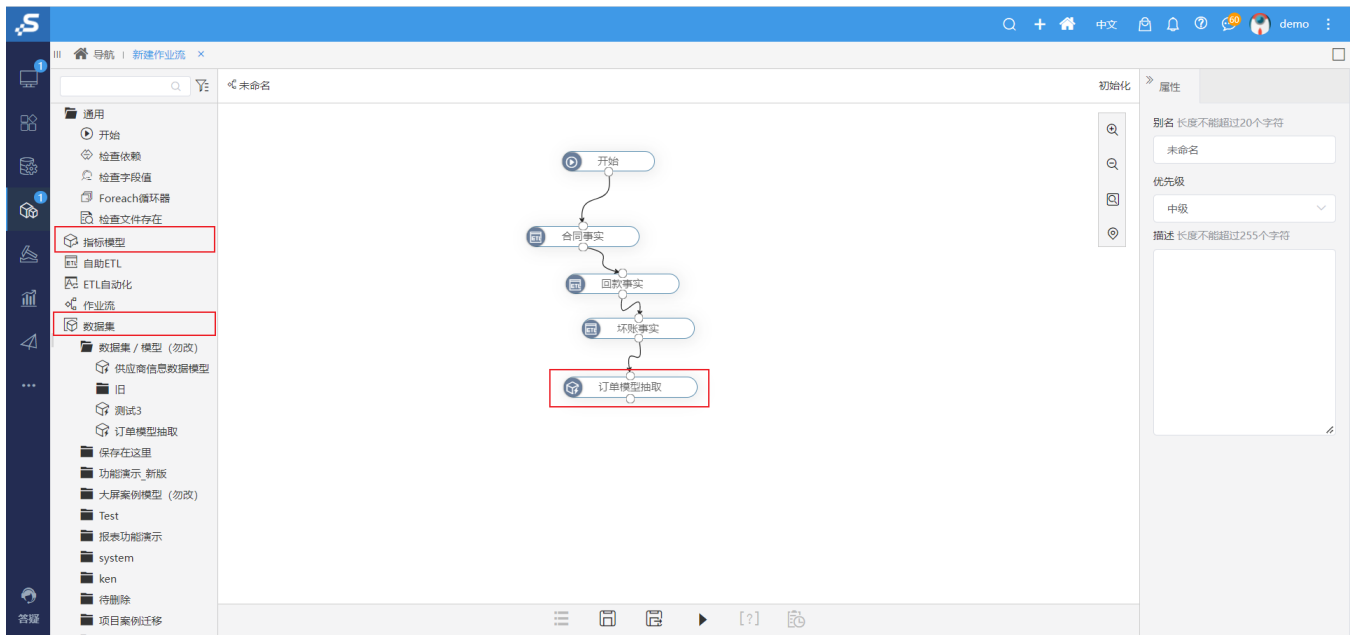


现在：作业流中可以直接调用数据模型的抽取



功能介绍

在作业流左侧增加了”指标模型“、“数据集”目录(目录下仅展示抽取的数据模型并支持把数据模型当作节点拖入到画布中，并且与其他节点建立执行的先后顺序，如下图：



参考文档

详情请参考 [如何在作业流中执行抽取的数据模型。](#)

+【作业流】作业流支持 Foreach 循环器

背景介绍

在V10.5 beta版本中，读取文件数据主要有以下方式：

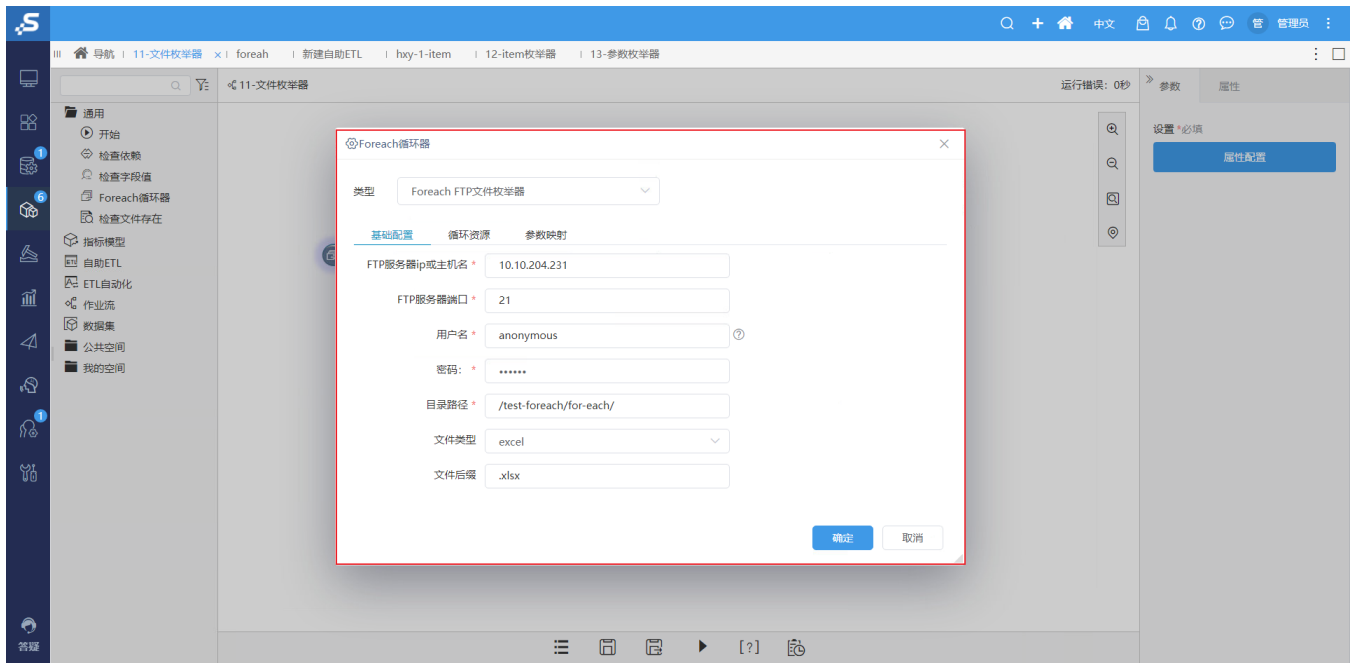
- 1) **使用Excel导入模型**：要根据模板导入，并且一次只能导入一个文件
- 2) **数据源/ 加载文件数据**：一次只能导入一个文件
- 3) **ETL的”FTP数据源”节点**：一次只能读取一个文件

而在V10.5.8版本中，支持一次性批量读取多个excel文件数据并且合并成大宽表存储到指定表中，业务人员无需重复导入，提高了其工作效率。

功能介绍

Foreach 循环器包含：Foreach FTP 文件枚举器、Foreach Item 枚举器、Foreach 参数枚举器；其中 Foreach FTP文件枚举器

通过获取ftp文件下的文件名(FileName)+Sheet名称(SheetName)，再配合”循环资源”循环读取数据合并到一张大宽表，如下图：



参考文档

详情请参考 [如何使用“Foreach 循环容器”节点。](#)

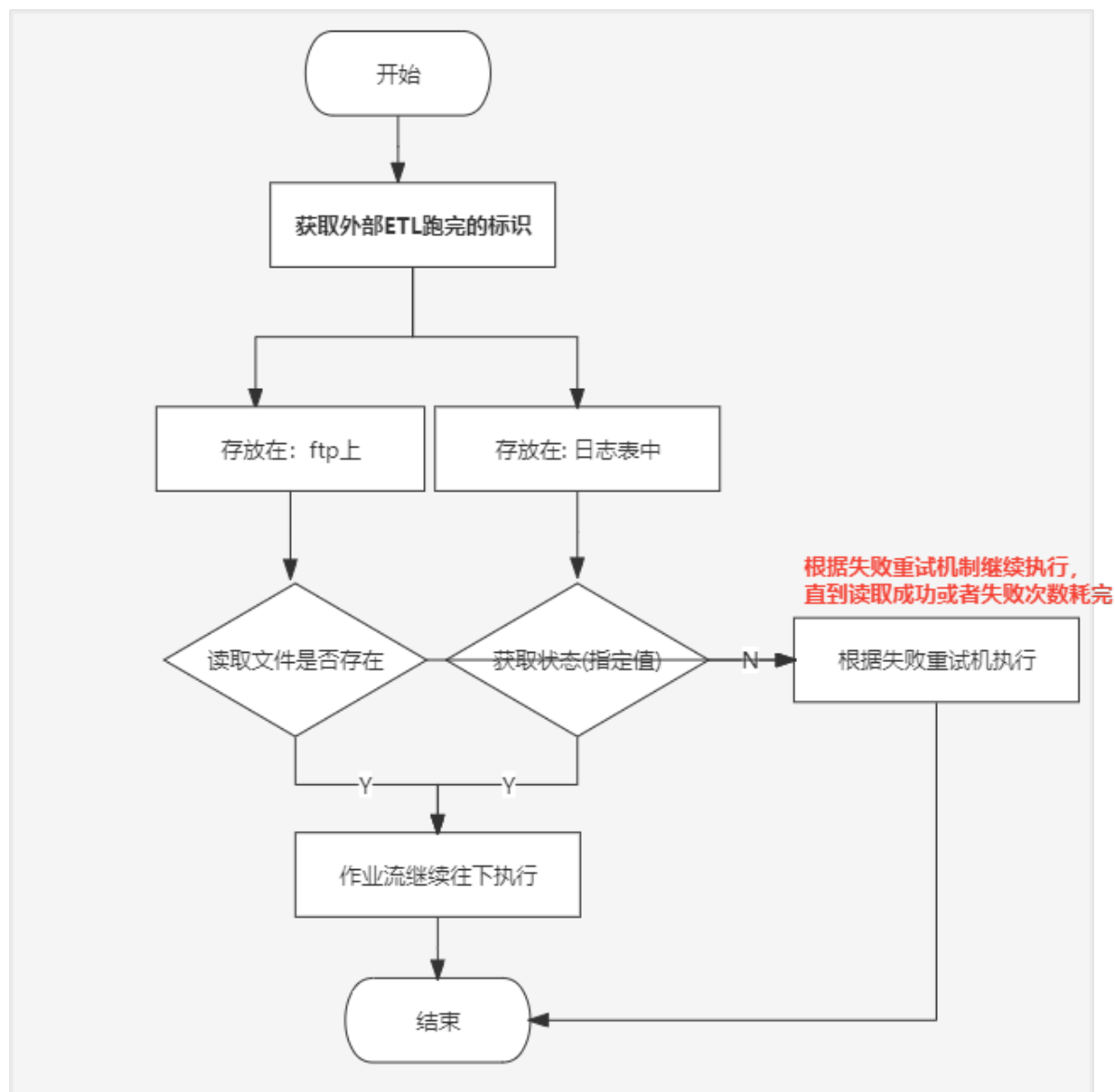
+【作业流】支持识别“外部ETL”跑完状态

背景介绍

在V10.5beta版本中，识别“外部ETL”跑完状态有以下方法：

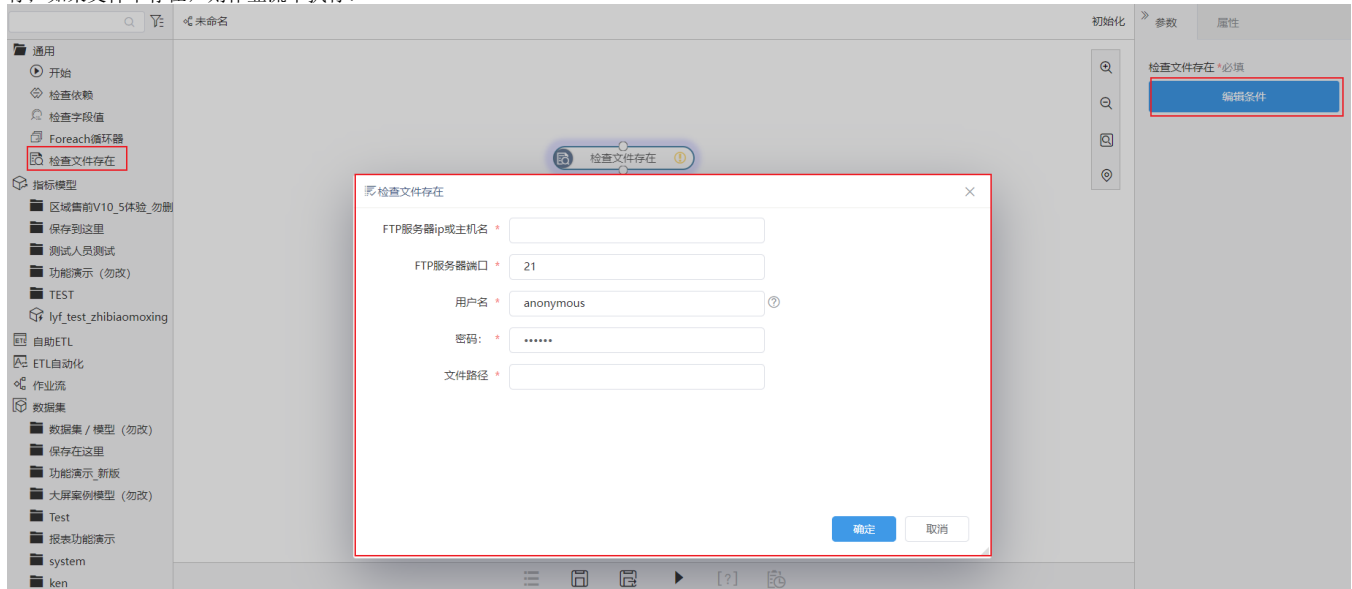
1) “**源库脚本**”节点：需要写SQL语句并且该节点无法与其他节点关联，需要单独建一个ETL资源，不但操作麻烦而且增加了维护成本。

而V10.5.8，支持”标识“存放在ftp上，也支持”标识“存放在具体的表中，如下图，扩展了识别方法，操作也更为直观简单，节省了实施成本，确保了数据正确。

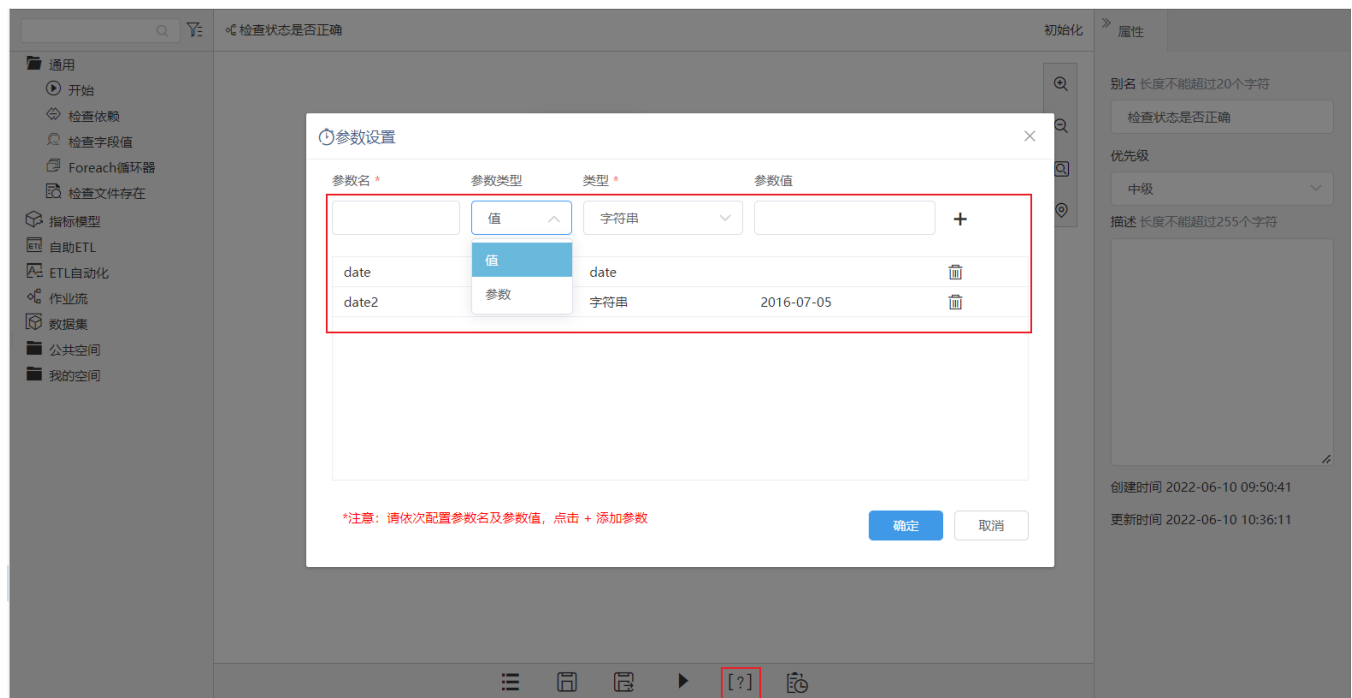


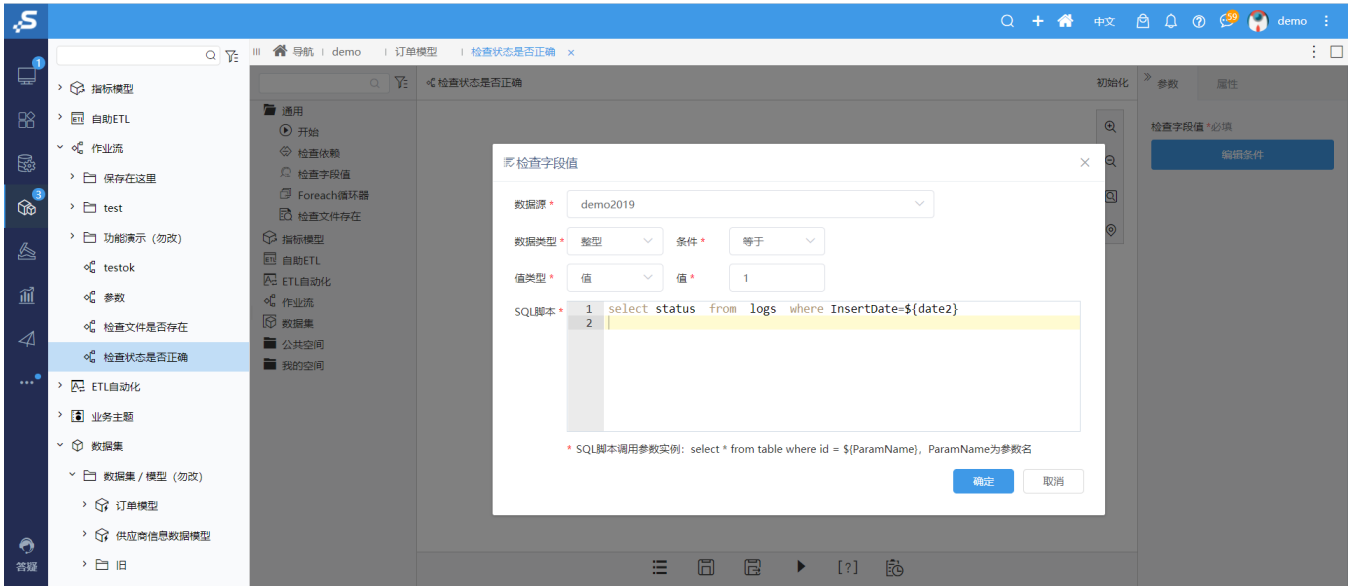
功能介绍

1)通过“ftp方式”检测文件是否存在：在作业流左侧增加了”检查文件存在”的节点，可以通过ftp读取指定文件，如果文件存在，作业流则往下执行；如果文件不存在，则作业流不执行：



2. 获取每天ETL执行状态：作业流增加了参数，用以动态判断记录状态





参考文档

详情请参考 [作业流如何识别“外部ETL”跑完状态。](#)

^【ETL】横版ETL开发

背景介绍

由于数据准备处理需要经常查看数据，改为横版ETL后，留出画布下方空间放节点配置面板和查看数据面板。用户点击执行按钮后能马上从数据面板上看到效果，节省了多次点击的时间。

功能简介

ETL改为横版DAG图，配置面板和数据面板调整至画布下方。选择节点时，配置面板上增加“执行该节点”按钮。左侧的节点树目录结构也做了调整。



参考文档

详情请参考 [自助ETL - 界面介绍。](#)