如何将数据集数据通过计划任务抽取到临时表中

有这样两类需求,可能需要用到临时表。

1) 主要的查询数据在Oracle数据库中,而在查询时需要关联的另外一个表却在SQL Server数据库中, Smartbi中查询数据时是没法直接跨库关联的,因此 我们需要将SQL Server库中的表抽取到Oracle的一临时表中,便于关联查询。

2)在数据库中有一个超大的上亿条记录的数据表,查询时还需关联多张表,查询效率极低,可能需要几分钟,甚至更长时间。一个可能的解决方案就 是,把复杂查询的数据,先定时抽取到一个临时表中,减少数据量,在此基础上再去查询,即可极大地改善查询效率。

上述两种情况,在Smartbi中都可以通过计划任务,将数据集数据先定时抽取到联合数据源的临时表中,然后再进行查询。完整的操作步骤描述如下:

1、点击 **数据准备 > 数据集**,进入"数据集"界面选择 新建数据集 > 原生SQL数据集。输入SQL语句后,在输出区的"输出字段"右键选择 检测输出字段。

select <u>ProductID</u> , <u>ProductName</u> , <u>CategoryID</u> , <u>QuantityPerUnit</u> , <u>UnitPrice</u> from products	< 図 輸出文段		
1			
I ? 修改			
> 日 私有参数 删除			
> 臼 输出参数			
居性名			

2、接着点击工具栏上的 **数据抽取** 按钮,进行相关抽取设置。

li li 🖬 🚿 🖬 🗊 🖻 🌞 🔍 🔤 🕒 🚟

select ProductID, ProductName, CategoryID, QuantityPerUnit, UnitPrice from products	> ☆ 諭出字段	
	ProductID	
	ProductName	
	CategoryID	
	QuantityPerUnit	
	UnitPrice	
	> 🗅 私有参数	
	属性名 属性值	
	属性描述	

数据抽取设置	×
 ○ 实时 ● 全量抽取 ✓ 清空数据 	
增量抽数按时间戳 抽取表设置	
抽取表名 tmp_ds_products_1	
高级设置: 异常处理	
抽取出错时 💿 回滚 💿 继续	
 资源创建者 管理员 	
特定用户 日立即抽取 1	
确定(O)	取消(C)

上图的各设置项详情请参见 数据抽取 章节。

计划修改		×
计划基本信息		
计划名称: *	抽取到临时表	
计划别名:	抽取到临时表	
计划描述:	抽取到临时表	
待执行任务:*	抽取到临时表 选择任务	
触发类型:*	时间	
间隔类型:*	毎天 ∨ 図是否启用 □是否指定生效范围 状态: ○	
生效范围:	从 🗸 到 🗸	
	触发时间:* 12:00 (例如13:30)	
运行设置:*	间隔时间:* 每 1 天	
	起始日期:* 2019-11-21 ~	
执行设置		
失败重试机制 :	重试次数 2 次 重试间隔:3 时 く	
	手动执行(M) 确定(O)	取消(C)

上图定时任务设置项详情请参见 计划任务 章节。

4、抽取成功后,在"高速缓存库"节点下可以看到该原生SQL数据集的临时表,以及在"跨库联合数据源"下该原生SQL数据集的库表信息:

^{3、}在抽取设置中单击 **设置定时抽取** 项,进行计划任务的抽取任务设置:



5、等所设置的任务"触发时间"过了之后,直接查询跨库联合数据源中的表,确认其中是否已正确插入数据。

III 谷 导航 I 预览数据[tmp_ds_products_1] ×						: 🗆
表名称: tmp_ds_products_1			表别名:	tmp_ds_products_1		
ProductID	ProductName	CategoryID		QuantityPerUnit	UnitPrice	
1	苹果汁	1		每箱24瓶	18.00	
2	牛奶	1		每箱24瓶	19.00	
3	蕃茄酱	2		每箱12瓶	10.00	
4	盐	2		每箱12瓶	22.00	
5	麻油	2		每箱12瓶	21.35	
6	酱油	2		每箱12瓶	25.00	
7	海鮮粉	7		每箱30盒	30.00	
8	胡椒粉	2		每箱30盒	40.00	
9	鸡	6		每袋500克	97.00	
10	璧	8		每袋500克	31.00	
11	民众奶酪	4		每袋6包	21.00	
12	德国奶酪	4		每箱12瓶	38.00	
13	龙虾	8		每袋500克	6.00	
14	沙茶	7		每箱12瓶	23.25	
15	味精	2		每箱30盒	15.50	
						关闭(C)