

目录

目录	1
同环比计算详解	5
使用复杂报表模块制作自由报表	5
复杂报表的数据准备	6
1. 新建数据连接，同步数据表到复杂报表	6
2. 在复杂报表中添加数据集	6
制作基础的报表模板	8
1. 先绘制出报表基本框架	8
2. 将数据集中的字段拖入对应的单元格	8
3. 单元格插入图片	9
将基础报告扩展为自由报表	9
1. 扩展	9
2. 预览	10
理解表计算的分区与寻址	10
1- 如何实现快速排名（取前10）	10
2- 表计算细节	11
散点图使用场景	12
离散和连续字段的简单概念	12
什么是维度？什么是度量？	12
如何添加用户？	12
如何才能更好使用数据分析工具分析数据？	13
商业智能分析模块简介	13
什么是BI？	14
如何对接restful api	14
概要讲解：	14
解析restful api主要步骤：1. 启动一个json的网络请求， 2. 将该请求放入系统中解析	14
具体demo如下：	14
数据库迁移、数据连接信息变化，如何操作才能正常使用	14
报表跳转/跳转报表如何使用	14
普通图表的跳转	14
跳转报告页	16
EXCEL文件在数据连接中表头显示异常	16
如何替换数据模型里的数据连接	17
在数据抽取中，客户的数据会抽取到什么地方？	17
fixed 和include 的异同？各筛选功能的优先级？	17
自定义sql没问题，为什么保存没有反应？	18
数据连接不上有哪些原因？	18
如何知道一个模型被几个报表给引用？	18
先后抽取的时候，都提示抽取成功，但是数据记录对不上	18
参数控制器的经典实用场景	18
平台的数据模型如何查数据？	19
在指标卡中 对比时间截止到当前时刻的逻辑	20
平台如何控制列权限	20
平台抽取优先级的作用	20
设置抽取优先级的意义：	21
依赖抽取两个表同时更新才做触发的逻辑	21
底层字段变化，数据模型和报表层该做什么？	21
excel 类型的数据模型报object异常	21
抽取为什么被禁止了及如何恢复抽取	22
如何改变抽取创建人	22

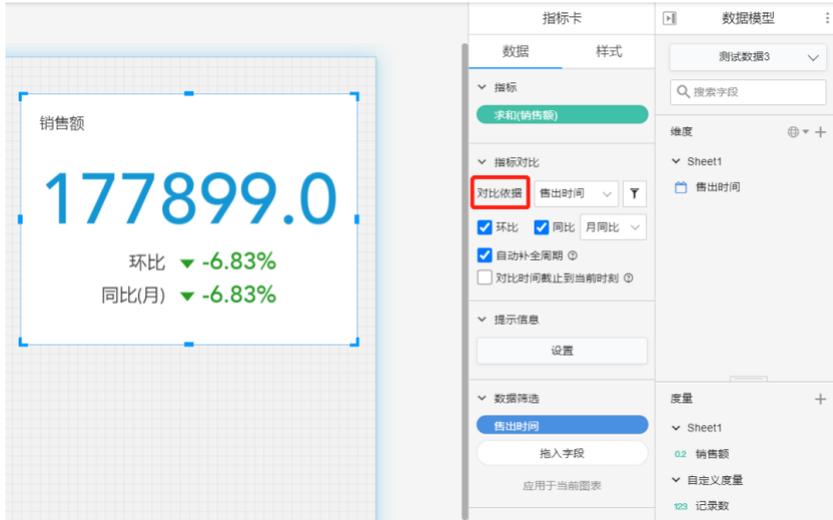
增量抽取的重跑机制	22
增量抽取的增量键设置？增量抽取选不到增量键？	23
如何正确的删除一个抽取？	23
误操作：自定义sql经常重命名	24
在数据模型中点击宽表为什么会处于加载中？	25
如何创建跨数据连接的模型	25
跨数据连接关联表	25
如何对大型的excel文件进行报表展示？大Excel数据量如何抽取？	25
盘点AxisBI的所有抽取模式及答疑	25
1. 目前AxisBI抽取有几种模式？	25
2. 在介绍抽取的时候介绍几个概念：	26
3. 下面我们介绍着三种抽取模式的特点。	26
4. 每种抽取模式的原理：	26
5. 抽取的答疑	26
数据模型变动了, 报告这块我们该做什么？	26
数据模型切换	27
为什么在数据模型中数据会加倍？	27
AxisBI自定义sql模块执行原理	27
数据抽取总是在加载中？	27
LOD算法其一：fixed的含义及其用法	27
使用LOD函数解决impala去重计数报错问题	28
处理聚合和非聚合字段的混合计算	29
1. 聚合和非聚合字段是无法进行混合计算的	29
2. 将聚合字段转为非聚合字段	29
取数模块和报告模块的区别	30
1. 场景介绍：	30
2. 报告和取数模块使用：	30
使用步骤是：	30
优缺点：	30
了解IF ELSE 解决场景中的一些问题	30
如何建立平台dataset和echart组件的映射？	31
如何彻底删除报表？	32
如何对一个数字开方？	32
AxisBI打开报告很慢？	32
简单理解可视化分析的分区和寻址？	32
度量字段和维度字段在AxisBI的含义	33
报告分享和集成的区别	34
如何排查报表数据问题？	34
批量选择字段，批量拖拽字段的功能	35
跳转报告的时候如何设置浮窗大小使其适应窗口？	35
如何将拥有共同特征的字符串都提取出来	35
excel导出问题排查方法	36
标题时间宏的用法	36
时间宏	36
同一个表格上快速切换度量/维度？动态维度如何使用	36
显示双轴，双轴的坐标不一致，如何解决？	36
查看有数据的最近一天	37
将表格的null值显示为0的几种方法	38
AxisBI数轴的几种设置	38
跨数据模型错误	39
日期类型和日期显示格式的区别	40
如何在图表上的标签里展示最大值最小值和头尾数据显示	40

如何快速查看报表关联的筛选器	41
数据转型导致的数据库SQL执行错误	41
为什么透视表展示相同模块的内容会出现两次?	41
计算字段里的粒度不一致问题	42
Excel导出后, 打开提示Excel文件损坏	42
计算字段的日期比较、日期常量、日期加减	43
SQL执行报错?	43
数据库SQL执行错误, 如何排查问题?	43
如何解除报表和数据模型的关系	43
如何制作多个度量的折线图? 双轴图?	44
如何排查DC错误或数据库错误?	44
如何添加参考线? 如何描绘一条用于比较的标准值? 如何设定一个目标值/目标线?	44
参考线	45
报告/指标卡查询出来没有数据或者数据不准确? 查询数据为空?	45
AxisBI报表上, 怎么查看某块内容对应的抽取任务是什么	47
如何替换报表中的数据模型	49
如何隐藏透视图的度量名称	50
如何设置图例的内容?	50
报表迁移中遇到的问题事项?	50
如何给二维表做自定义排序?	51
热力地图的介绍	52
双y轴如何设置刻度?	53
如何设置上下文筛选器	53
字段重命名和数据字典的区别	54
属性聚合方式是什么?	54
并列柱形图如何制作?	54
度量值如何做饼图?	55
如何对度量值设置百分比?	55
如何调表格的宽度?	56
筛选器不起作用常见的原因有哪些?	56
大屏如何排查数据问题	56
大屏数字翻盘器为什么会动	57
AxisBI内数据行级权限的交并集关系解释	57
角色权限设置数据连接的注意点	58
对hive类型的行级权限设置字段和设置值慢, 实现的新策略	58
行级权限实质	59
项目管理员权限的注意点	59
定时邮件发送的方式	59
定时邮件是否支持邮件组?	59
当碰到定时邮件发送失败, 失败原因是访问数据超时, 怎样处理?	59
怎么样提问? 怎么样快速得到有效的答案?	59
域主注册后怎么登陆AxisBI?	60
报告已经删除了, 再去删除模型的时候提示: 有报表或大屏依赖, 不能删除数据模型	60
数据填报如何把表落地到自己数据库	60
如何删除域中用户	60
回收站功能	61
登录系统后, 提示用户不存在?	61
个人用户搭建一套AxisBI试用学习环境, 以MySQL为例	61
第一步	61
第二步	62
第三步	62
将AxisBI的报告嵌入网页后, 为什么报告中的图表无法加载到内容?	62

通过血缘分析来查询资源和数据的依赖情况	63
报告的导出导入，报告的迁移？	64
报告的导出导入	64
如果一个用户被删除了，那么如何删除他私人文件夹下的报告	64
Impala/Hive数据源：创建自定义sql报错	65
数据刷新管理的三种模式	65
浏览器的支持情况	65
用户删除了，私人文件下面的报告会不会消失？	65
推荐：如何提交工单？	65
数据导出的极限是多少？	66
平台报503错误是什么原因引起的？	66

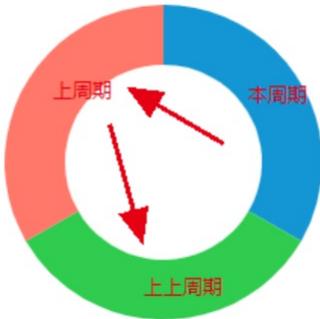
同环比计算详解

环比增长率=(本期数-上期数)/上期数×100% 同比增长率=(本期数-同期数)/同期数×100% 那么，我们以指标卡为例，介绍一下AxisBI里同比和环比的各类场景是如何计算的。



在指标卡这里，计算同比和环比最为关键的部分是**对比依据**，也就是我们选择的**周期**。

首先我们将**环比**，环比是本周期对比上周期，假如我们对**对比依据**里选的是**今日**，那么上周期指的是什么呢？答案是**昨日**；假如选的是**本周**，那么上周期即是**上周**；假如选的是**本月**，那么上周期即是**上月**，同理可得其他。可以得知，所谓**环比**，也就是将本周期等距的移动一个范围，并进行对比的过程。



环比概念图

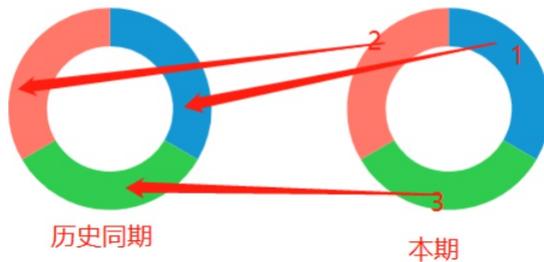
再说**同比**，指标卡有三种**同比**方式，**周同比**、**月同比**、**年同比**，顾名思义，是指**同比**的历史同期是在**周级别**、**月级别**和**年级别**。

比如我们对**对比依据**选择**今日**选择**周同比**，那么也就是今日（假设今日为周四）和上周的周四比较；假如我们选择**月同比**，那么也就是今日（假设今日为1号）和上月1号比较；同理，假如我们选择**年同比**，那么也就是和去年此日比较。

所以我们也知道，你所选的**对比依据**不能超过历史同期的级别，也就是说如果你选的**对比依据**是**本月**，那你就没有**周同比**之说了；如果你选的**本季**，也自然没有**月同比**了。

PS. 假如现在是2019. 3. 22日，**对比依据**选择**本月**，那么请问**环比**和**月同比**是否一致呢？

答案：不一致，因为**环比**是平移这段周期，2019年3月合计31天，2019年2月合计28天，因此**环比**计算的是【2019. 3. 1至2019. 3. 31】和【2019. 1. 29至2019. 2. 28】这两段时间的比较。而**月同比**，则是计



同比概念图

算3月自然月和2月自然月的比较。

在指标卡中，新增加的一个功能：**对比时间截止到当前时刻**

如果用户勾选了该选项，当前时间的选择中，当前时刻取到分钟，{00:00-当前时分}=今日时间；对比时间中，对比今日的时间截取到为{与今天的相对日 00:00-与今天的相对日当前时分} 如果用户勾选了该项，在筛选中，筛选里的当前时间并不包含今日，此时将对当前时间段和对比时间段无影响。

使用复杂报表模块制作自由报表

传统的列表式报表，有时并不能满足用户的需求，这种情况下可制作自由格式的报表，通过卡片式报表，把数据以卡片的形式展现，让用户轻松的了解信息的相关性。接下来，我们通过制作一个三国武将的卡片式报表，来了解自由报表，效果如下图：

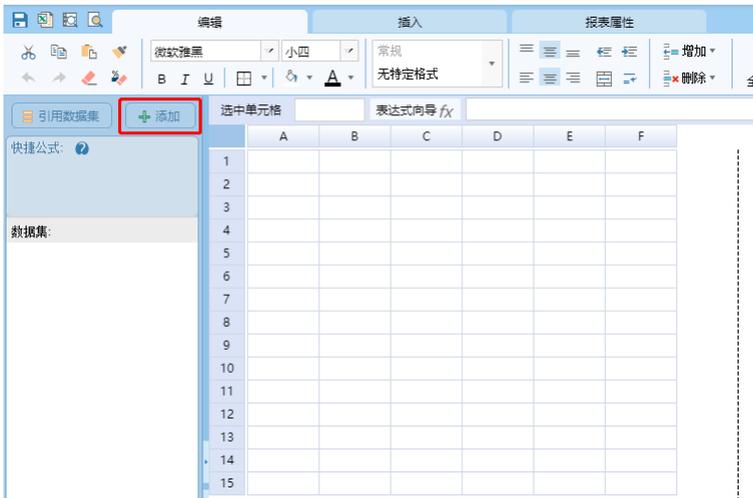
复杂报表的数据准备

1. 新建数据连接，同步数据表到复杂报表

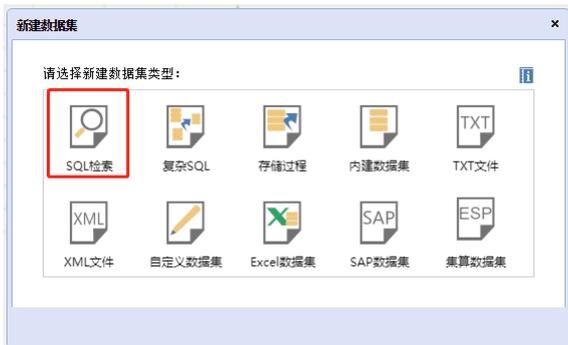
示例数据连接为mysql，打开数据连接中的“同步至复杂报表”开关，在“登记表”中，登记要使用的数据表，用于在复杂报告中建立数据集

2. 在复杂报表中添加数据集

进入复杂报表，点击“添加”，新建复杂报表的数据集



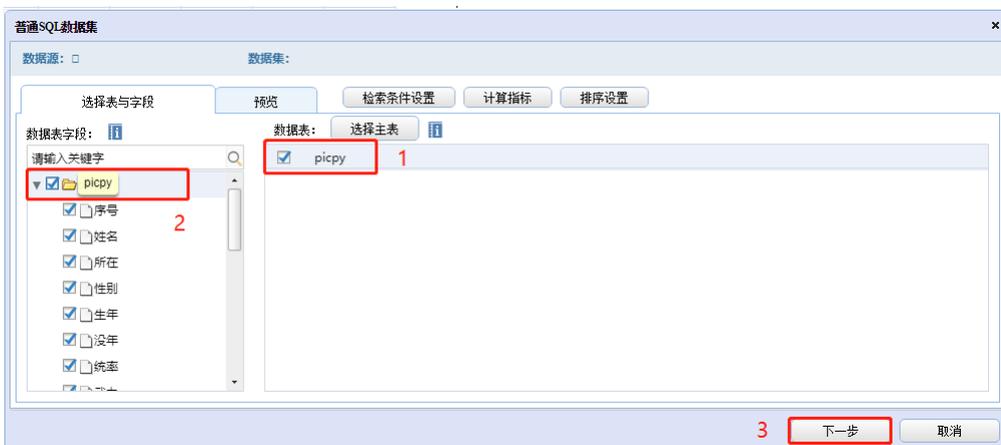
选择“SQL检索”



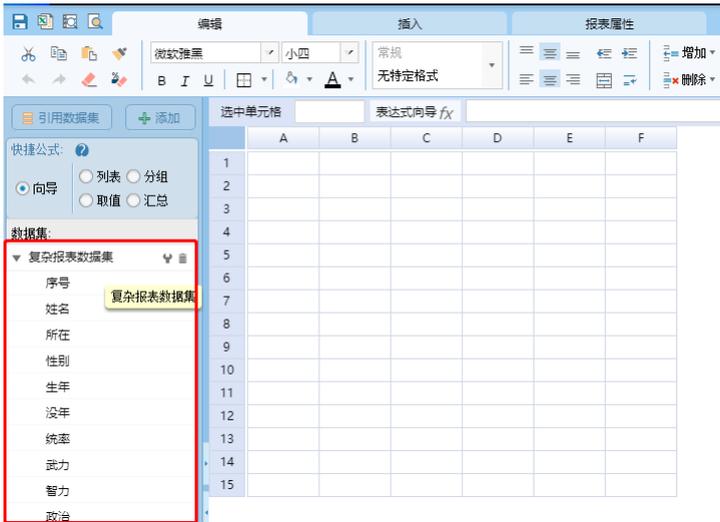
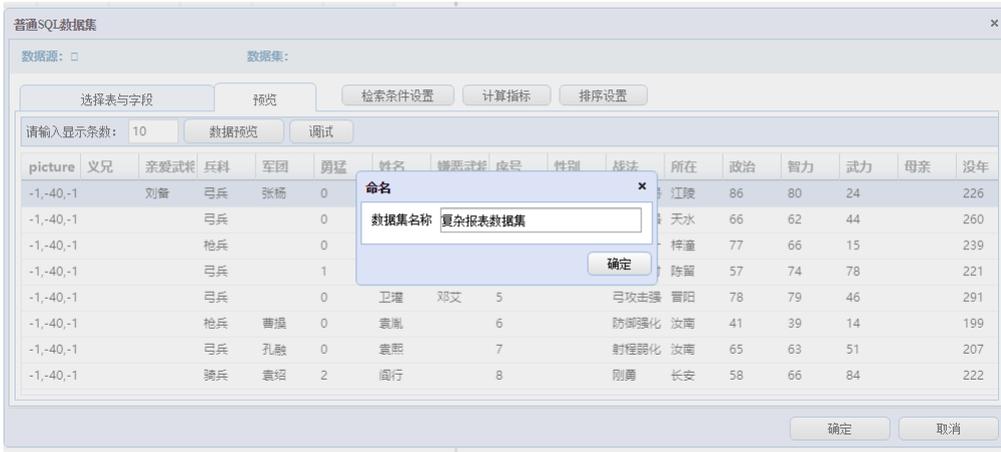
选择在数据连接中登记过的数据表



按选中数据表后，选择需要的字段，示例中选择了所有字段



点击右下角“确定”后，命名该数据集为“复杂报表数据集”，数据集会自动生成在最左侧的数据集栏



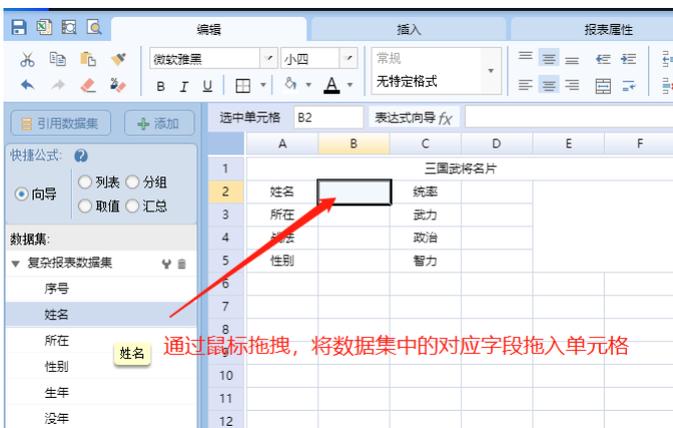
制作基础的报表模板

1. 先绘制出报表基本框架



2. 将数据集中的字段拖入对应的单元格

此时会自动弹出单元格向导，根据向导的提示选择相应的公式和属性



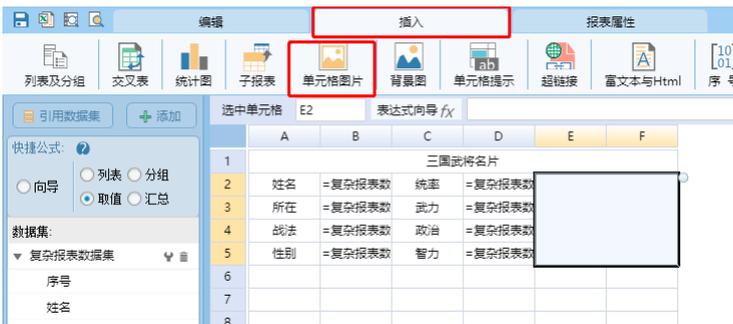
通过鼠标拖拽，将数据集中的对应字段拖入单元格

“姓名”字段选择“列表”，其余字段选择为“取值”，也可以在快捷公式中直接选择“取值”，此时拖拽到单元格的字段默认属性为“取值”



3. 单元格插入图片

选中单元格，在最上方选择“插入”-“单元格图片”



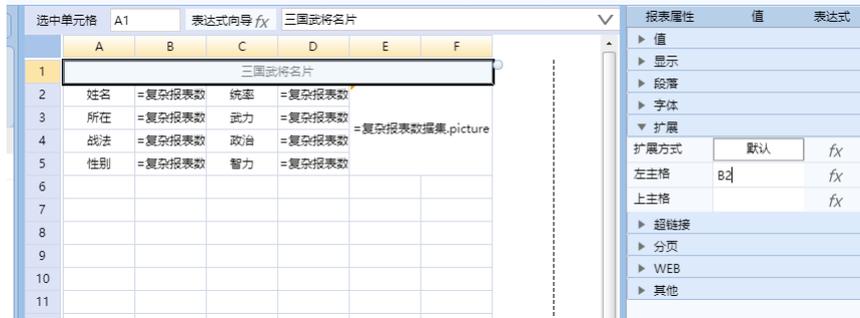
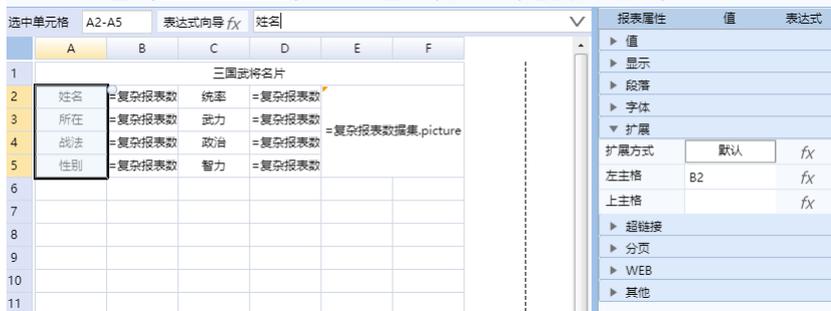
在弹框中，选择“单元格图片”-“图片字段”，选择数据集，然后手动输入图片字段，在示例的数据集中，picture是blob类型字段，存储了三国武将的头像图片。



将基础报告扩展为自由报表

1. 扩展

将纵向的A2到A5,以及横向的A1到F1分别选中后，在右侧“扩展”中分别设置以B2单元格为扩展左主格



2. 预览

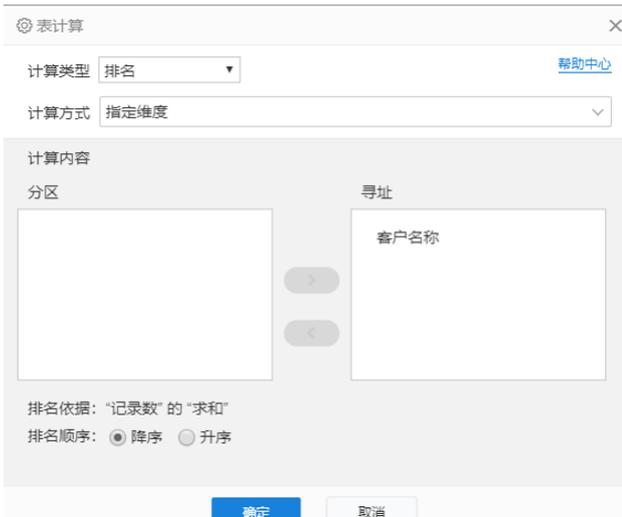
保存后我们预览报告，就可以得到一个基础的自由报告了

三国武将名片				
姓名	伊籍	统率	29	
所在	江陵	武力	24	
战法	破坏力强化	政治	80	
性别		智力	86	
三国武将名片				
姓名	尹赏	统率	51	
所在	天水	武力	44	
战法	弓攻击强化	政治	62	
性别		智力	66	
三国武将名片				
姓名	尹默	统率	26	
所在	梓潼	武力	15	
战法	智力上升	政治	66	
性别		智力	77	
三国武将名片				
姓名	于禁	统率	83	
所在	陈留	武力	78	
战法	弓箭强射	政治	74	
性别		智力	57	
三国武将名片				
姓名	卫瑾	统率	69	
所在	晋阳	武力	46	
战法	弓攻击强化	政治	79	
性别		智力	78	
三国武将名片				
姓名	襄胤	统率	32	
所在	汝南	武力	14	
战法	防御强化	政治	39	
性别		智力	41	

理解表计算的分区与寻址

1- 如何实现快速排名（取前10）

首先是将度量值拉入数据筛选，然后选择范围1-10，然后选择表计算，然后选择排名，计算方式指定维度，寻址拉入指定的维度，根据选择按照升序或者降序排名。



2- 表计算细节

表计算中的计算方式：单元格，表，区

在表计算中有两种类型的字段：分区和寻址

分区字段：将您的数据分组到不同的桶中，这些分区中的每一个然后通过计算来执行。

寻址字段：决定您的计算方向

计算方式：

① 单元格

理解为单独的一个单元格数据，分区为‘日期/季度/地区’，寻址为空，如下

日期	季度	华东	华北
2017	第一季度	1	1
	第二季度	1	1
2018	第一季度	1	1
	第二季度	1	1

② 区

区有‘边界的概念’，这个边界取决于另外的维度，比如以‘日期/地区’为分区，区的方向是先横穿，再向下，如下

日期	季度	华东	华北
2017	第一季度	1	2
	第二季度	3	4
2018	第一季度	1	2
	第二季度	3	4

以‘日期/地区’为区，区的方向是先向下，再横穿，如下

日期	季度	华东	华北
2017	第一季度	1	3
	第二季度	2	4
2018	第一季度	1	3
	第二季度	2	4

以‘季度/地区’为区，先横穿，再向下，如下

日期	季度	华东	华北
2017	第一季度	1	2
	第二季度	1	2
2018	第一季度	1	2
	第二季度	1	2

表则是指全表，默认方向为向下

日期	季度	华东	华北
2017	第一季度	1	1
	第二季度	2	2
2018	第一季度	3	3
	第二季度	4	4

还有一种为‘先向下 再横穿’（先横穿，在向下的方向不再赘述）

日期	季度	华东	华北
2017	第一季度	1	5
	第二季度	2	6
2018	第一季度	3	7
	第二季度	4	8

散点图使用场景

散点图是指在回归分析中，数据点在直角坐标系平面上的分布图，散点图表示因变量随自变量而变化的大致趋势，据此可以选择合适的函数对数据点进行拟合。

所以散点图适用于两组数据相互关系或变化趋势的分析问题。

散点图的使用文档如下：

<https://docs.ksyun.com/documents/38156>

离散和连续字段的简单概念

离散型数据一般是指该变量值可以一一枚举，它可以是数字或者字符或者日期时间。而连续型数据则是指相邻两个值之间是连续不断的，可以无限分割的。在实际运用中，离散和连续不是绝对的，比如日期，可以认为一天一天之间是离散的，但是也可认为是连续不断的。

离散和连续取决于数据希望的表现方式，比如散点图和折线图的区别，我们希望散点图可以表现数据离散的特征，而折线图则表现数据连续的特征。

在AxisBI里，一组数字通常被认为是连续的，比如销量、销售额等等。一组字符通常被认为是离散的，一组日期通常被认为是离散的。

什么是维度？什么是度量？

维度：说明数据，维度是指可指定不同值的对象的描述性属性或特征。例如，地理位置的维度可以包括“纬度”、“经度”或“城市名称”。“城市名称”维度的值可以为“旧金山”、“柏林”或“新加坡”。

度量：度量就是基于某个属性对其属性的指标进行衡量。

如：中国男人平均身高是170cm。

维度就是：中国男人

度量就是：身高

所以度量值参与聚合计算，算平均身高就是把所有中国男人的身高做汇总除以个数算出平均值。

从sql上简单理解：

```
select a,b ,count(c) from table group by a,b
```

维度字段a, b参与分组，度量字段c 参与统计。

注意：维度和度量没有绝对的概念。可以相互转化。在AxisBI中日期字段只能是维度。

如何添加用户？

1. 如果您是域管理员，首先是点击[企业域管理]，进入[人员信息]，点击[新增用户]，这样平台会给对应的用户发送一封激活邮件，被添加的用户激活后即呈现已激活的状态，此时您可以将其加入项目中。



2. 选择对应的项目，点击项目中心，进入用户管理，即可添加用户，添加用户之后，进行对应的角色权限管理即可。



如何才能更好使用数据分析工具分析数据？

如何才能更好使用数据分析工具分析数据？

那么首先需要理解什么是数据分析？

数据分析是指用适当的统计分析方法对收集来的大量数据进行分析，提取有用信息和形成结论而对数据加以详细研究和概括总结的过程。

那么第一个问题，如果采集过来的大量数据不适合当前的分析主题，那么数据分析是不是毫无意义呢？

现实中很多人随意给出一组数据，把数据填充到分析工具得到一个可视化图表。认为这就是数据分析，但却忽略这个图表体现出来的含义实际意义。

还有一些人随意给出一组数据，把数据填充到分析工具认为一定会出来一个什么样的图表，结果图表没出来。那时是因为分析工具支持的不够强大呢？

答案显然不是。

其实在数据分析的时候我们需要提前做这么一件事情。

保证数据的质量。

数据质量是数据分析结论有效性和准确性的基础也是最重要的前提和保障。

如何控制数据质量？

1. 明确自己的分析主题

如你拥有来自不同部门的交易数据，那么此时你想做一件什么事情，变得尤为重要，你是想分析部门的业绩情况，还是想看哪个客户是优质客户？

这时我们要明确分析主题。

2. 合理的数据建模

明确主题后，我们要业务建模，业务模型继而确定技术建模。简单的说，如：本次课题，看哪些是优质客户？

那么要确定优质客户的标准是什么，如收入问题，有联系方式，交易额超过10万。那我们此次建模只需要采集这些拥有的信息字段就好。

合理的数据建模可以给分析减负。

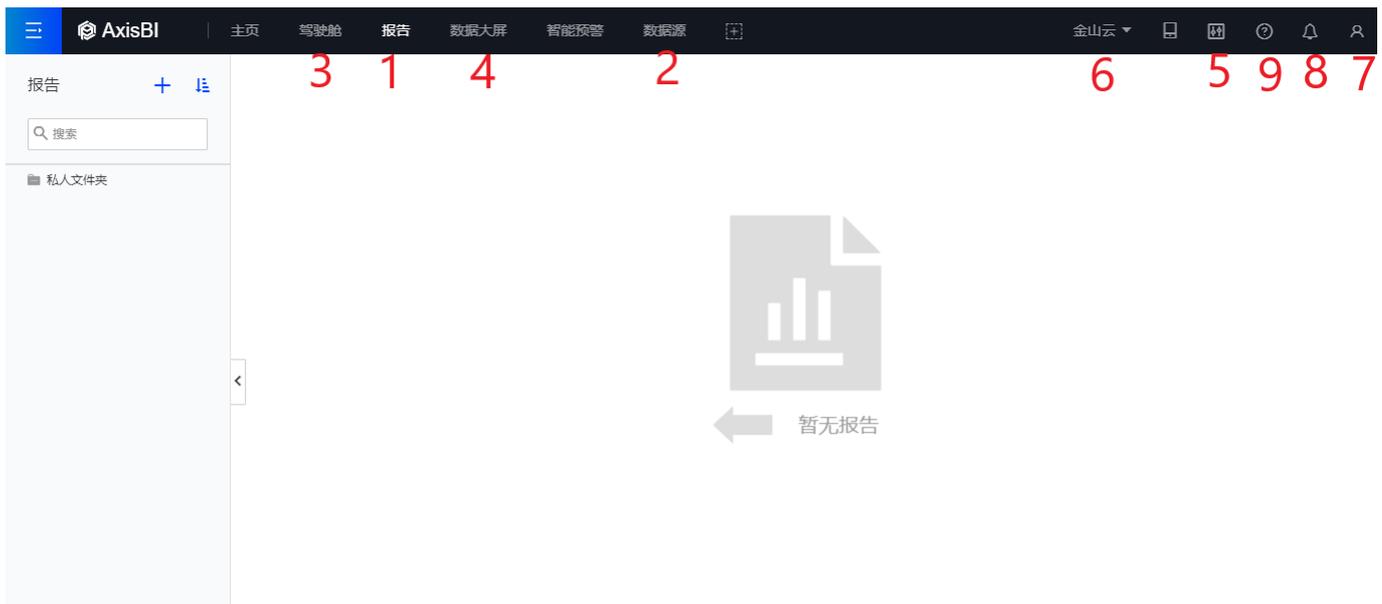
3. 去除脏数据

脏数据可以理解为异常数据，如电话号码中存在汉字，性别中存在其他字符。那这部分我们需要清洗，严格要求。

如果提前做好这些准备工作，我相信数据可视化分析会更加如虎添翼。

商业智能分析模块简介

首次进入商业智能分析，会进入如下页面，在介绍如何使用商业智能分析前，先向大家简单介绍下各个模块的功能。



1. 报告：最重要的模块之一，我们可以像制作PPT一样，简单地完成数据分析报告的制作，对数据进行分析跟呈现，并快捷地与他人分享我们的分析结果。

2. 数据源：在建立数据分析报告前，我们需要在“数据源”模块建立数据连接，将我们的数据（可以是Excel文件数据，也可以是Mysql、Oracle等数据库，抑或是Hadoop、Spark等大数据平台）接入商业智能分析。同时我们需要在“数据源”模块，根据分析的目的，选择一个或多个数据表，进行多表关联跟数据处理，建立数据模型，之后需要基于数据模型来建立报告。

3. 驾驶舱：对于一些重要的报告，若希望相关的同事或领导可以快速看到，可以将这些报告发布“驾驶舱”给指定的用户，则对应的用户登录系统后，会直接访问“驾驶舱”中的报告。在“驾驶舱”，我们可以及时了解企业关键的KPI指标，就像查看飞机的驾驶舱一样，因此我们将这个模块命名为“驾驶舱”。

4. 大屏：“大屏”是独立于报告的另一项工具，不同于侧重数据分析的“报告”，“大屏”模块主要偏向于数据呈现，提供由专业视觉设计师设计的各类行业主题模版，用户可以借助“大屏”模块制作出精美的数据大屏投射至展厅的电视屏幕上。
5. 项目中心：项目管理员可以在该模块进行资源的管理跟权限的分配。
6. 项目列表：通常企业里不同的业务线，或者不同的业务部门都有数据分析需求，这时可以为不同的业务部门建立不同的项目，不同项目之前的数据跟资源是相互隔离的，我们可以通过“项目列表”在不同的项目间切换。
7. 用户中心：用户可以在“用户中心”查看或修改自己的个人信息，并可进行密码设置。还可以在此查看自己拥有哪些项目的权限，以及申请加入项目或创建新的项目。
8. 消息中心：系统的消息提示会显示在这里。
9. 帮助中心：通过“帮助中心”，可以访问商业智能分析提供的帮助文档跟教学视频，并可以向开发团队提交意见反馈。

什么是BI？

BI (Business Intelligence) 即商务智能，它是一套完整的解决方案，用来将企业中现有的数据进行有效的整合，快速准确的提供报表并提出决策依据，帮助企业做出明智的业务经营决策。

商业智能的概念最早在1996年提出。当时将商业智能定义为一类由数据仓库（或数据集市）、查询报表、数据分析、数据挖掘、数据备份和恢复等部分组成的、以帮助企业决策为目的技术及其应用。而这些数据可能来自企业的CRM、SCM等业务系统。

商业智能能够辅助的业务经营决策，既可以是操作层的，也可以是战术层和战略层的决策。为了将数据转化为知识，需要利用数据仓库、联机分析处理（OLAP）工具和数据挖掘等技术。因此，从技术层面上讲，商业智能不是什么新技术，它只是数据仓库、OLAP和数据挖掘等技术的综合运用。

把商业智能看成一种解决方案应该比较恰当。商业智能的关键是从许多来自不同的企业运作系统的数据中提取出有用的数据并进行清理，以保证数据的正确性，然后经过抽取（Extraction）、转换（Transformation）和装载（Load），即ETL过程，合并到一个企业级的数据仓库里，从而得到企业数据的一个全局视图，在此基础上利用合适的查询和分析工具、数据挖掘工具、OLAP工具等对其进行分析和处理（这时信息变为辅助决策的知识），最后将知识呈现给管理者，为管理者的决策过程提供数据支持。商业智能产品及解决方案大致可分为数据仓库产品、数据抽取产品、OLAP产品、展示产品、和集成以上几种产品的针对某个应用的整体解决方案等。

如何对接restful api

概要讲解：

解析restful api主要步骤：1. 启动一个json的网络请求， 2. 将该请求放入系统中解析

具体demo如下：

1. 首先起一个http服务，或用tomcat，wamp等工具作为容器。

以tomcat为例：把json文件放入www文件夹，我的文件为 demo1.txt

json文件内容如下：

```
``` [{"地区": "华东", "省份": "浙江省", "销售额": "13", "利润": "12"}, {"地区": "东北", "省份": "黑龙江", "销售额": "17", "利润": "14"} ]
```

正常情况下可以在浏览器上访问到：

![image.png](http://fe-frame.ks3-cn-beijing.ksyun.com/project/cms/f0689c70af40d11384d440cc490cb1ae)

\*如果tomcat 部署在其他服务器上，需要保证和部署商业智能分析的服务器互通。

2. 在商业智能分析上新建一个restful api的数据连接，将该请求放入商业智能分析中解析：

![image.png](http://fe-frame.ks3-cn-beijing.ksyun.com/project/cms/9e1423a968e352e470e18999388cf4bb)

\*最简单的一个demo就完成了。上面的json格式是商业智能分析通用的json格式。

## 数据库迁移、数据连接信息变化，如何操作才能正常使用

1. 首先确认数据库迁移以后表结构是否发生变更了，若没有变更，则直接修改数据连接内的IP/PORT等信息使之可以连上新的数据库。若发生了变更，请参考第3步。

2. 需要确认新的数据库是否配置了正确的白名单信息，白名单信息如下：

106.2.127.56

106.2.124.188

59.111.29.246

59.111.165.166

59.111.165.165

59.111.29.247

请注意，原来如果做了堡垒机、端口转发的，也需要联系您的运维或者DBA再进行配置。

3. 如果新库有发生了表结构的变更，那么需要尽量不要修改原有的数据连接信息，而是选择再新建一个数据连接信息，然后在数据模型端做替换数据连接的操作，如何替换数据连接可以参考FAQ：<https://docs.ksyun.com/documents/38307?preview=1>

## 报表跳转/跳转报表如何使用

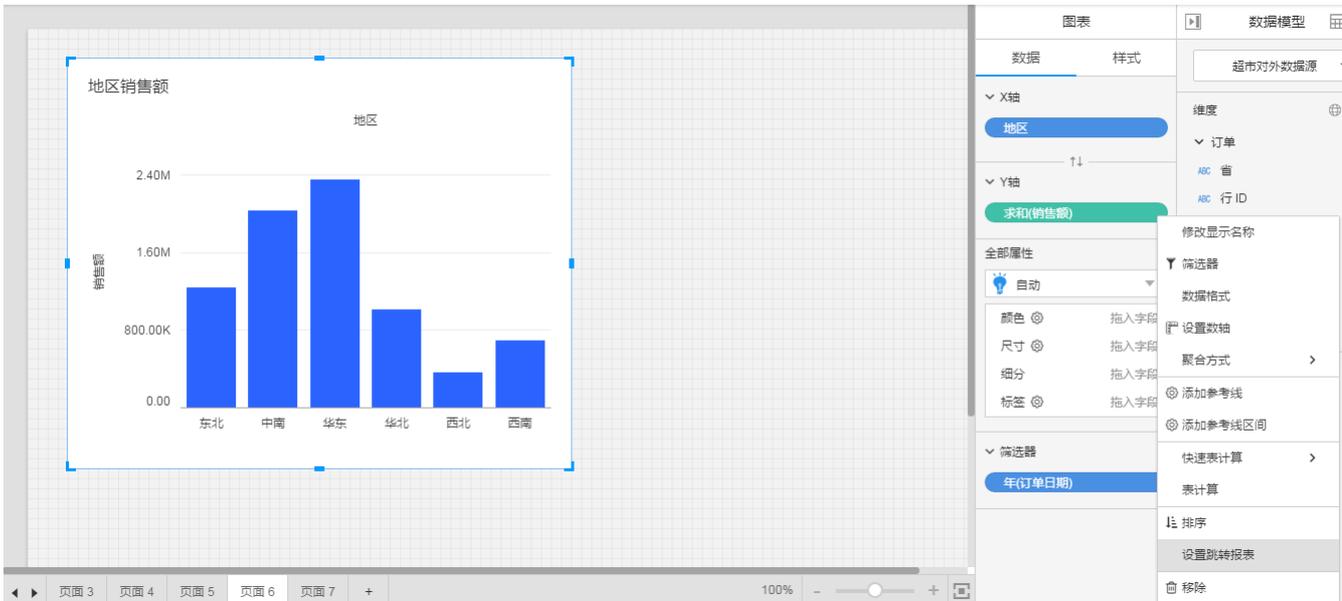
功能说明：通过一张报表，点击该报表的图形，会跳转到另外一张报表，这个另一张报表可以是不同报告，也可以是同一报告的不同页面。

### 普通图表的跳转

比如：地区销售额的图表，点击图表的柱子，可以查看某个地区的销售明细。本例中，只取了2012年和2013年的数据，应用了一个筛选器。

举例说明设置跳转报表的步骤：

1. 鼠标移动到Y轴的度量“销售额”，点击“销售额”出现下拉菜单，选择“设置跳转报表”



2. 在设置跳转窗口，选择需要跳转的报表。点击确定完成了跳转报表的设置

设置跳转报表 ✕

本报告  其他报告

页面	报表
页面 1	地区销售明细
页面 2	
页面 3	
页面 4	
页面 5	
页面 6	
页面 7	

[高级设置](#)

确定
取消

3. 高级设置的说明

设置跳转报表 ✕

页面	报表
页面 1	地区销售明细
页面 2	
页面 3	
页面 4	
页面 5	
页面 6	
页面 7	

[高级设置](#)

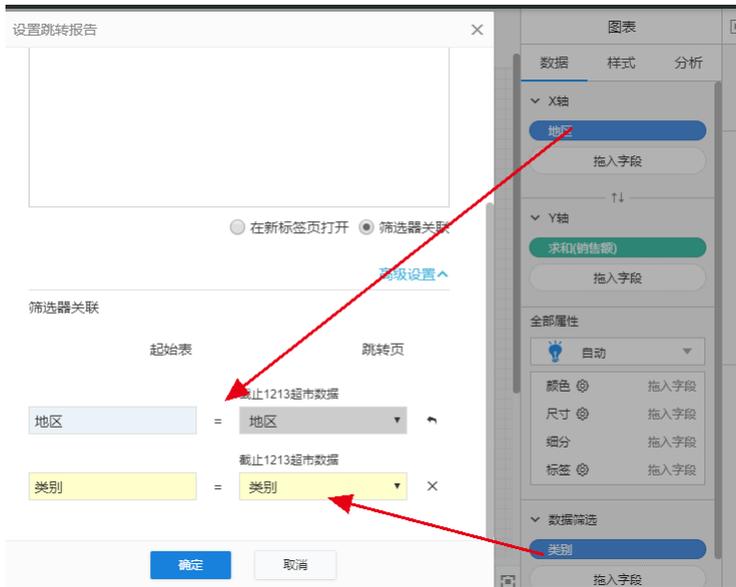
筛选器关联

	起始表		跳转表
地区	=	地区	▼
订单日期	=	订单日期	▼

确定
取消

蓝色区域的筛选：是起始报表传递的筛选条件

黄色区域的筛选：是起始报表拥有的筛选器，如下

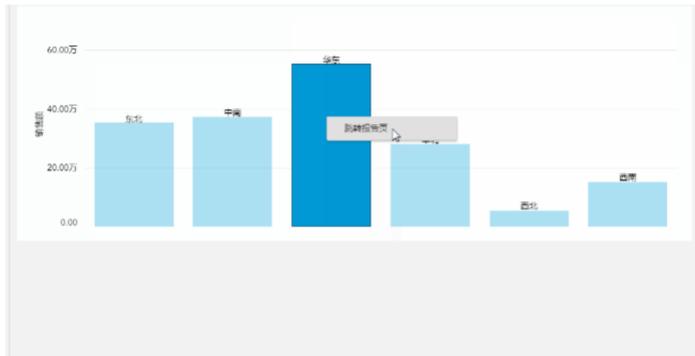


这样我们就设置好了。

### 跳转报告页

平台还支持跳转至报告页，如下：

如下方动图所示，我们在柱形图中选中关心的数据（这里我们选择销售额最高的“华东”地区），右键，在弹出的工具栏中选择“跳转报告”，我们将会跳转至更详细的报告页，同时会对该报告进行筛选，只查看“华东”地区的数据。



设置方式和跳转报告类似，就不再赘述，请大家自行尝试。

有一点是特别要划重点的，那就是跳转带的条件



其实每一次跳转都会带上这些条件，如果你不想这么做就点击“x”按钮将其条件取消，否则跳转过去的报表会带上这些条件。

进一步解释一下，所谓“带条件”，如果用伪代码分析一下：

报表a，目前展示的是按地区划分的销售额，且仅展示【类别=家具】的数据，在平台里是一个非常简单的柱形图，它的X轴是地区，Y轴是求和【销售额】，数据筛选是类别=家具。

用sql表达：

```
select sum(销售额),地区 from table where 类别="家具" group by 地区
```

当我们选择地区的一个条件进行跳转到B报表的时候，会根据高级设置的条件来进行数据的匹配，比如此时我们点击了东北的柱形。

也就是

```
select * from (B报表模型的宽表) B where B.地区="东北" and B.类别="家具"
```

然后得到一个过滤后的宽表，通过这个报表再去形成B报表。

通过这一点我们也知道了，优先级的关系，即跳转的条件先生效还是B报表本身的条件先生效，答案自然是跳转的条件先生效。但是这个跳转的条件其优先级又低于上下文筛选，筛选优先级的关系，大家可以在这篇文章进行详细了解：

<https://docs.ksyun.com/documents/38309?preview=1>

此处不再赘述。

如果我们只是在维度上设了跳转，那么跳转代入的维度的条件参数，那么如果在度量上设置呢会怎么样？

答案是会代入此度量对应的所有维度条件。

## EXCEL文件在数据连接中表头显示异常

EXCEL文件做数据连接时，表头显示异常（表头显示为F0, F1等，而不是本来的Excel首行标题），解决：

1. 清除excel 的排版格式，可以用格式刷清除掉。
2. 标题行的前后不能有空格，否则无法识别；

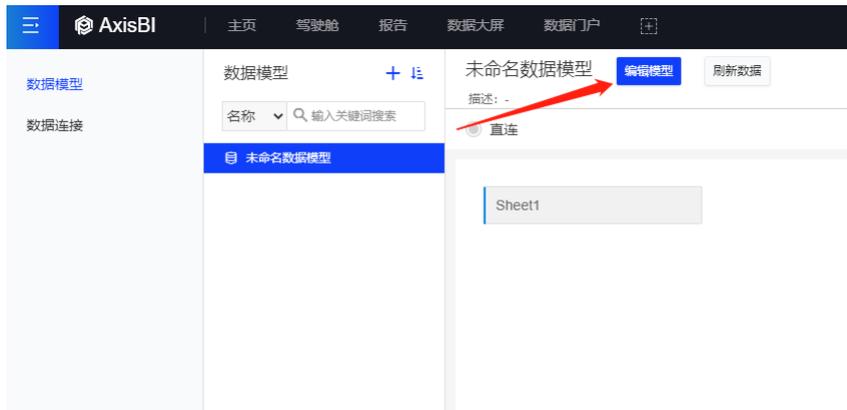
## 如何替换数据模型里的数据连接

如何替换数据模型里的数据连接？

条件：两个库必须具有相同的表结构。 字段名称，类型都必须完全匹配。

操作步骤如下：

1.选择要修改的数据模型，编辑模型



2.编辑模型界面，选择切换数据连接按钮，进入数据连接切换界面，选择目标数据源，点击确定即可切换。



## 在数据抽取中，客户的数据会抽取到什么地方？

会抽取到我们的分布式版本的postgresl。

## fixed 和include 的异同？各筛选功能的优先级？

相同点：两者都能根据多个维度去聚合一个度量。

不同点：（1）.筛选器作用:fixed 和included的时候，优先级不一样。一般的控制器组件对fixed不受影响。一般都是先执行fixed再执行筛选。如果想先执行筛选就必须勾选优先筛选才行。而include一般都是先执行筛选在执行include的函数。（执行优先级请看下图）

优先级  
(数字越大优先级越高)



注：图表内的维度筛选，默认优先级：列表筛选 = 文本筛选 > top筛选 = 条件筛选

(2).fixed 不受视图上面（图表）维度的影响。

Include会受视图上面（图表）维度的影响。

## 自定义sql没问题，为什么保存没有反应？

引起的原因是自定义查询太耗性能了。一般超过一分钟没有结果返回。AxisBI这边就会处于没有反应的状态。

打开浏览器的开发者模式（F12），可以看到控制台有504的错误。

遇到这种情况，建议您稍后重试，或者选择数据库资源充足的时候重试。

如果您是修改自定义SQL，要注意，此自定义SQL如果在抽取中，可能也会导致失败，修改前请先取消抽取，再进行修改。

## 数据连接不上有哪些原因？

白名单没有设置好。咨询AxisBI工作人员，拿取ip白名单，并设置好对AxisBI服务器访问通道。（客户的服务器那边设置）

防火墙有拦截

数据库对外的权限没有打开。

可能数据版本不兼容。这种情况可能性很小。

## 如何知道一个模型被几个报表给引用？

在数据连接里面点开，相关内容。



## 先后抽取的时候，都提示抽取成功，但是数据记录对不上

现象：定时抽取数据不对，但抽取提示是成功的。于是手动调抽取，数据就能和源头数据对上了。数据正常。

这种情况是定时抽取之前数据还没有完全准备好。当你手工抽取的时候，数据源数据已经更新完毕了。

建议合理调整抽取时间，以避免此种情况。或者设置2个定时任务，比如早晚各一次，这样可以避免疏漏。

## 参数控制器的经典实用场景

参数控制器的经典实用场景：

场景：当点击字符A，我以字段A来计算，当点击字段B我用字段B度量来计算。

第一步：先建立一个参数列表

名称:

类型:

列表项:

- A
- B
- C

默认值:

第二步：建立一个参数控制器。把该参数拉进去。



第三步：用计算字段和参数关联。

名称:

式:

```
IF [参数].[参数1]="A" THEN [usetimes] ELSE
IF [参数].[参数1]="B" THEN [zaijuid]
ELSE
[zaijuusetimes]
```

#### 4. 用计算字段拉入度量

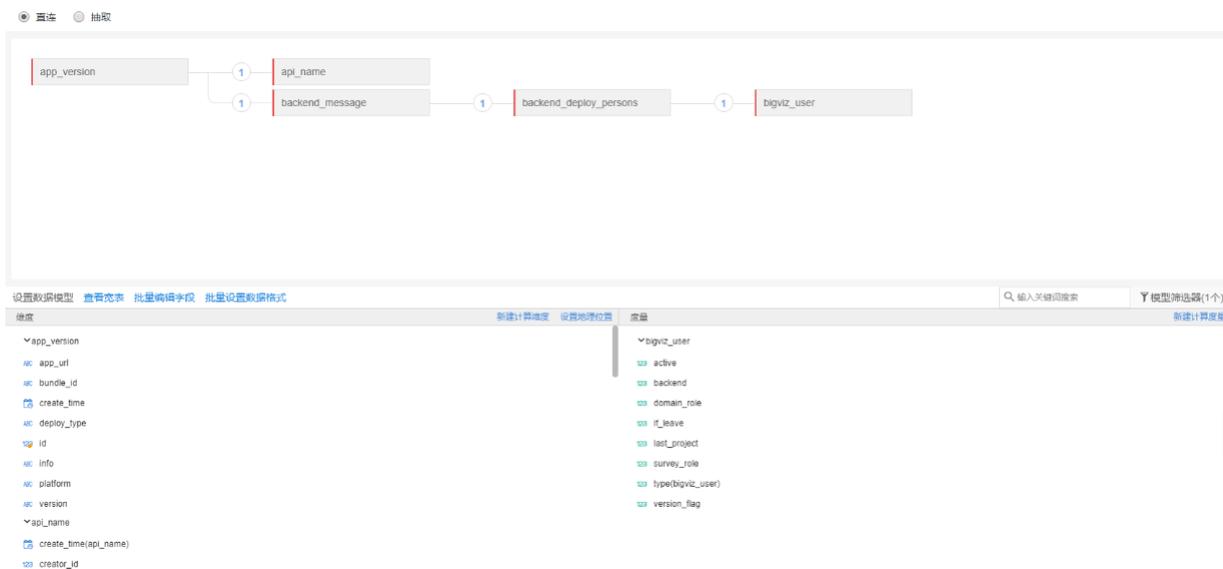


这就是参数控制器的经典应用。

有兴趣的同学可以模仿一下。这个报告是什么样的哦。

## 平台的数据模型如何查数据？

如图的数据建模，平台如何执行sql？



我们从后台截取sql如下：

```
SELECT * /*查询字段已经删除，主要看下面的join逻辑*/
FROM `dashboard`.`app_version` AS `t1`
INNER JOIN `dashboard`.`api_name` AS `t2` ON (`t1`.`id` <=> `t2`.`id`)
INNER JOIN `dashboard`.`backend_message` AS `t3` ON (`t1`.`id` <=> `t3`.`id`)
INNER JOIN `dashboard`.`backend_deploy_persons` AS `t4` ON (`t3`.`creator_id` <=> `t4`.`creator_id`)
INNER JOIN `dashboard`.`bigviz_user` AS `t5` ON (`t4`.`id` <=> `t5`.`id`)
```

WHERE (t5.id IN (8082))

从sql可以看出，平台是以节点对应关系来执行sql的。

拓展：

(1) 如果在列表筛选器针对其中一个表的字段设置了筛选，sql 如何执行？

sql是先把模型中表与表的join关系理出来。然后基于整个结果做的筛选字段筛选。

这样的缺陷是：如上图，如果t1和t5在不做筛选的时候会出来多条结果，对性能不是特别好。

优化：可以把t5创建一个自定义sql， 内容如： select \* from t5 where 条件=xxx 。 然后把这个sql视图代替上图的t5单表。

(2) 这样的数据模型，在报告中，如果只用了某张表的一个字段，会如何？ 平台会把数据模型的sql 当成一个子sql，在该sql上做查询

## 在指标卡中 对比时间截止到当前时刻的逻辑



在指标卡中，增加一个功能， 对比时间截止到当前时刻

如果用户勾选了该选项，

当前时间的选择中，当前时刻取到分钟，(00:00-当前时分)=今日时间；

对比时间中，对比今日的时间截取到为(与今天的相对日 00:00-与今天的相对日当前时分)

比如 在筛选中，对发货日期勾选了今天，勾选了环比，勾选了对比时间截止到当前时刻：

当前时间为3-18-14:20

当前时间段为，filter (发货日期) =today ∩ (00:00-当前时分)=3月18日 00:00-14:20

对比时间段为，filter (发货日期) =(today-1) ∩ (00:00-当前时分)=3月17日 00:00-14:20

如果用户勾选了该项，但是筛选时间如果勾选本月等快捷时间，则不起作用。

## 平台如何控制列权限

场景：有一张表，表中有5个字段分别是Co11, Co12, Co13, Co14, Co15，如何让不同的用户看到不同的字段？如A用户 看到字段是Co11, Co12, Co13, Co14, B用户看到字段是Co12, Co13, Co14, Co15

方案1：建立多个数据模型，引用相同的表，对不同字段进行隐藏。在权限管理里面，至给数据模型查看权限，不给编辑权限。



方案2：建立多个数据模型，在原来的表的基础表上套自定义sql。用该sql 来建立数据模型。对于不同的用户对应各自的模型（即各自的自定义sql）

## 平台抽取优先级的作用

如何设置抽取优先级：点击：“项目中心”->“抽取管理”进入以下界面。

全部连接 0
抽取完成 0
抽取中 0
抽取失败(全部) 0
禁用 0

发送日报
立即抽取
禁用
启用
设置禁用规则
自定义列
搜索

<input type="checkbox"/>	表名	数据连接	类型	状态	调度信息	占用空间	成功率	创建人	优先级	操作
--------------------------	----	------	----	----	------	------	-----	-----	-----	----

**!** 您没有可浏览的数据。若有疑问，请联系该项目管理员

### 设置抽取优先级的意义：

设置优先级是为了保证平台一些重要抽取的数据产出。为重要抽取提供一个绿色通道。  
 抽取目前分为默认队列和高优先级队列。  
 举例说明：如果某个时间点，用户抽取的并发数是300个。但是平台默认并发数100个。  
 这样会导致200个任务堵塞无法进行，如果200个中有重要任务就必须等待。  
 因此抽取提供了一种优化策略，将默认队列的并发数设置成90个，高优先级队列并发数为10个。  
 当遇到上面情况，平台90个任务数走默认队列抽取，10个任务数预留给高优先级队列。

## 依赖抽取两个表同时更新才做触发的逻辑

当两个表都触发才抽取的逻辑如下：

demo：如果我有一个数据模型，宽表构成A和B表组合而成。

假设A和B都是每天一更新

日期	基础表	更新时间点	备注
2019-07-17 A		1时	
2019-07-17 B		3时	数据更新完毕已经满足同时触发需求，所以被触发
2019-07-18 A		源头故障，数据没有更新	
2019-07-18 B		3时	今天没有满足，条件不触发
2019-07-19 A		1时	数据更新完毕已经满足同时触发需求（昨天有一次B），所以被触发
2019-07-19 B		3时	

从上述可以看出，我们的规则是没有周期限制的。

所以一旦出现源头故障。数据会持续错乱。

解决措施：

手工去触发打断之前的逻辑。这样后期数据正常，不会出现问题。

方案优化：

设置依赖的时候，在自定义sql 识别的时候，选择一张产数据最晚的表。做依赖触发条件。

如上图就把B表设置依赖触发表，不要把A、B同时选择。

## 底层字段变化，数据模型和报表层该做什么？

### 1. 数据模型层：

在数据模型预览状态下点击刷新数据。（这样就可以保证底层数据元信息一致）



刷新数据完成后，如果是抽取，请找到对应的抽取，重新执行一下抽取任务。

### 2. 在报告层：在数据模型编辑状态，重新点击保存，然后覆盖应用到报表。

具体可见：

<https://docs.ksyun.com/documents/38342>

## excel 类型的数据模型报object异常

自定义sql字段没有变，为什么会报错。

在excel 中，一个excel文件就相当于关系型数据的一个库。每个sheet页就相当于表。

报该错误 object异常，说明excel的sheet页名字变化了。



## 抽取为什么被禁止了及如何恢复抽取

目前出去禁止的规则如下：可能符合一些条件，系统会自动禁止。



如果抽取修复完毕，如何打开抽取？

选中改抽取，点击启用。就解除禁止了。提前：需要项目管理员操作



## 如何改变抽取创建人

如何变更抽取创建人？

点击“项目中心”->“抽取管理”会弹出以下界面：

点击每个抽取的“更多”，会弹出一个菜单，选择变更创建人，输入交接人的名称。



## 增量抽取的重跑机制

抽取是将用户建立的模型的数据从源库转移到AxisBI的高性能MPP内，形式上分为全量覆盖抽取，全量追加抽取和增量抽取。

全量覆盖：每次会删除前一个副本，只保留一个副本，并且每次抽取都是全量的

全量追加：不会删除副本，保留所有副本，每次抽取都是全量的

增量抽取：每次抽取都是增量的，需要一个增量键作为增量的标志

这次我们主要来说一下增量抽取的重跑机制：

功能路径：项目中心->抽取管理->对应增量抽取->更多->重跑



可以看到，重跑分为两种，一种是任务时间，一种是数据范围。为了方便解释，我们设定一种增量抽取场景，我们抽取一张表T，增量键为每天都不同的DAY字段。今天是2019.3.22，我们发现2019.3.20的数据有了更新，我们需要更新这部分数据，因此需要用到我们的重跑功能，也就是将2019.3.20的数据重新从数据库抽取到mpp。使用数据范围和任务时间的原理略有不同，用伪代码解释一下。

数据范围：

①删除MPP内的老数据，`delete from MPP.T where day between`

`"2019/03/20 00:00:00" and "2019/03/20 24:00:00"`。

②从源库查找数据，`select * from 源库.T where day between "2019/03/20 00:00:00" and "2019/03/20 24:00:00"`

③将查询所得数据插回mpp，即完成数据的重跑更新。

这个过程的关键是利用了增量键【day】直接来判断数据的范围。因此，重跑若要成功，需要保证第②步的时候能够成功的查询到数据。

任务时间：顾名思义，也就是任务执行的时间，实际上mpp将数据抽取入库的时候会为数据打上一个时间戳，也就是任务时间time。如果我们选择了任务时间，mpp的动作不变，但是区别在于把增量键的条件改成了time的条件，伪代码如下

①查询MPP库，得到数据的边界 `select min(day), max(day) from MPP.T where time between "2019/03/20 00:00:00" and "2019/03/20 24:00:00"`

②做一个删除的动作，`delete from MPP.T where time between "2019/03/20 00:00:00" and "2019/03/20 24:00:00"`

③查询源库，`select * from 源库.T where day between min(day) and max(day)`

④将查询所得的数据插回mpp

这种场景适用于一天内做了多次增量抽取的情况

这里再介绍一种特殊情况，增量抽取抽的不是一张表而是一条SQL的时候，该如何分析是否可以重跑呢？很简单，把上面所述的{T}表，用你的SQL代替即可，一条自定义SQL实际上是一个视图，视图也可以当成表来看待。

比如，我的SQL写的是(`select SYSDATE AS DAY, * from T`)，然后设置day作为增量键，也就是总是以当前时间作为增量。那么这种情况重跑会如何？

答案：很遗憾，这种写法重跑是跑不到数据的，我们可以分析一下

①mpp删除数据

`delete from MPP.(select SYSDATE AS DAY, * from T) a where a.day between`

`"2019/03/20 00:00:00" and "2019/03/20 24:00:00"`

这一步删除了0条数据，因为 day总是等于当前时间，但是执行上是不会报错的，因此系统会继续执行第二步。

②从源库查找数据，`select from 源库.(select SYSDATE AS DAY, from T) a where a.day between "2019/03/20 00:00:00" and "2019/03/20 24:00:00"`

这一步，检索到0条数据，但是执行上是不会报错的，因此系统会继续执行第三步。

③将查询所得的数据插回mpp

通过分析可以发现，实际上这个重跑相当于啥也没做。

## 增量抽取的增量键设置？增量抽取选不到增量键？

增量抽取要求使用一个类型为int或者date的字段作为增量字段，且必须数据库内的该字段类型为int或者date。

（注意事项：如果在宽表中做了日期类型转化或int转化，对于抽取选增量字段还是会选不到，因为增量字段会按原始表或者自定义sql的原始字段类型判断）

这个增量键会在增量抽取的生成查询的sql的where条件进行使用。

对于HIVE的分区表，我们建议增量键和分区键一致，因为一般生产库有设置hive.limit.query.max.table.partition 参数，随着时间增长，分区增多，如果抽取生成的查询上没有用到分区键，可能会导致全表扫描，从而触发报错。

注意点：hive分区字段是string 类型，而这个字段想做增量抽取，建议大家可以写个自定义视图。

如一个表有n个字段。分区字段是day。

`select cast(day as date) day1, * from A`

day1字段一定不要和day字段同名。

上面是最快捷的办法，但是最好不要写\*。

## 如何正确的删除一个抽取？

在有些时候，有一些已经过时的报告和抽取仍然占用着资源，这个时候我们需要去一一删除掉这些报告、模型、抽取来释放资源。

操作方法如下：

1. 步骤如下，“项目中心”->“抽取管理” 找到要删除的抽取。点击“更多”



2. 删除抽取。



3. 如果抽取不能删除, 会提示抽取有依赖关系无法删除。此时该如何删除?

我们可以通过查看依赖项。看这个抽取有哪些依赖。



点击显示如下, 我们可以到该抽取上面有以下的数据库模型和报表依赖



我们需要删除该抽取上的所有数据库模型和报表, 抽取才可以顺利被删除。

删除顺序: 删除报表->删除数据模型->删除抽取 (情况1: 删除前请注意数据模型和报表是不是真的废弃了, 必要时可以做报告导出备份 情况2: 一张报表里面可能会关联有多个数据模型, 我们只要删除抽取对应的那个数据模型和报表即可, 不一定整个报表都删除)

PS. 只想删除抽取, 不想删除报告和模型, 怎么做?

可以将抽取依赖的数据模型改为直连, 选择覆盖到报告以后, 就可以直接删除抽取了 (删除抽取以后还可以将数据模型改回抽取模式使得抽取重建)。这种情况适用于发生了一些抽取上的BUG, 需要删除原有抽取并且重建抽取的时候使用, 日常不建议使用此方法。

### 误操作: 自定义sql经常重命名

场景: 当前的数据模型满足不了我业务需求的时候, 我们可能会基于老的数据模型上做修改。

那么有一种情况就是变更抽取模式下面的自定义SQL?

误操作: 修改sql, 然后重命名, 然后把重命名的sql 加到数据模型中。这种会出现: 以后点击宽表的时候, 数据模型在加载中。

原因: 这里有别名操作, 导致数据模型和抽取无法匹配。

如何解决:

1. 将数据模型切换到直连模式, 保存覆盖应用到所有报表。
2. 到'项目管理' -> '抽取管理' 找到之前的抽取名, 然后删除。删除的时候记住当时的定时任务配置和抽取方式, 方便后面抽取重新配置。
3. 返回数据模型, 进入编辑状态, 把新的sql 拖拽进来。然后切换到抽取模式, 按照之前的抽取设置重新配置一遍, 保存即可。

## 在数据模型中点击宽表为什么会处于加载中?

可能是如下两个原因导致的?

1. 自定义sql里面有计算字段, 但是没有做别名。

解决方案: 加个别名, 然后保存。

2. 在抽取模式的时候, 自定义sql 被重命名了。

解决方案: 参考文档

如下:

[点击跳转到解决方案](#)

## 如何创建跨数据连接的模型

### 跨数据连接关联表

平台支持将不同数据连接中的表进行关联, 比如一张数据表来自MySQL数据库, 一张数据表来自Excel文件, 要将两张数据表关联成一张宽表后分析。此时需要将不同数据连接的表抽取至平台提供的MPP数据库中。



## 如何对大型的excel文件进行报表展示? 大Excel数据量如何抽取?

如果你有几百万行的Excel数据。这时候可能你的excel可能无法上传至平台, 或者无法在平台上面分析, 此时就必须使用平台的抽取功能, 将Excel数据转化为关系型数据库数据。

可以建议使用以下方法:

首先观察excel 有几个sheet 页。Excel的每一个sheet页在平台里会映射为独立的一张表。

1. 单个Excel过于巨大的时候, 需要拆分为多个Excel (一般不建议单个Excel超过50m)。

如一个excel 有20个sheet 页, 可以把一个excel 文件分成多个excel文件, 每个文件放几个sheet 页。

然后新建多个数据连接去连接这些Excel。

之后在数据模型里引用这些数据连接的表, 并设置为抽取模式。

2. 还有一种情况, 一个sheet页里面数据量太大。这时建议把一个sheet页分为多份Excel (保证分出来的Excel仍然拥有同样的表头)。

之后在数据模型里引用这些数据连接的表, 并设置为抽取模式, 同时因为这些表都具有相同的表头, 在模型层使用union的方法将数据再合并在一起。

3. 根据方法2延伸出来的场景, 如果是多个Excel, 也可以分开建立数据连接, 最后在模型层进行union或者join处理。

拓展内容:

跨数据连接建模怎么做?

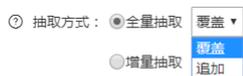
见文章: <https://docs.ksyun.com/documents/38351>

## 盘点AxisBI的所有抽取模式及答疑

### 1. 目前AxisBI抽取有几种模式?

大类分: 全量抽取和增量抽取。而全量抽取又分覆盖和追加。

所以细分有3种抽取模式, 如下图



## 2. 在介绍抽取的时候介绍几个概念：

源头库：客户存原始数据的库。（数据连接连的那个库）

目标库：AxisBI的数据库（mpp）。

抽取的过程：就是将客户的数据存到AxisBI的数据库中。

## 3. 下面我们介绍着三种抽取模式的特点。

抽取模式	优点	缺点	适用场景
全量抽取-覆盖	简单，数据质量好	大数据量的时候，占用资源	小数据量表
全量抽取-追加	可以帮源头库存历史数据	抽取成功部分，不支持更新	源头库只存增量数据，不存历史数据
增量抽取	对数据增量部分处理，占资源小	抽取成功部分，不支持更新	大数据量表

## 4. 每种抽取模式的原理：

- 1. 全量抽取-覆盖：

举个例子：第一天源头库有100条数据，第二天源头库新增10条，源头库即110条。那么第一天抽取：目标库100条，第二天抽取会删除昨天抽取的100条，然后重新抽取，此时目标库也是110条。

- 2. 全量抽取-追加：

举个例子：第一天源头库有100条数据，第二天源头库新增10条，源头库即110条。那么第一天抽取：目标库100条，第二天会抽取源头库的110条数据，但是不会删除昨天的数据。所以第二天目标库的数据是210条。

这种抽取适用场景很小，源头库只存增量数据，不存历史数据。

场景：源头库存的是当天的交易数据。那么到了明天，该表就会把今天交易数据删掉，存明天的数据，该表只保留一天的数据。

此时我用追加抽取，在目标库我可以到每个时间节点的数据。

- 3. 增量抽取：

增量抽取的前提必须有增量抽取的字段（增量字段）。一般都是时间类型。

举个例子：今天是2018年8月9日，源头库里面存了2018年1月到8月8号的数据。目标库存了2018年1月到2018年8月4号的数据。（这些日期不是系统日期，是增量字段体现出来的日期）

那么增量抽取如何进行，首先系统会拿增量字段去源头库找到最大日期，找到8月8号是最大时间。那然后去目标库找，找到最大日期是8月4日。那么系统就判别：目标库缺少源头库数据5号到8号的数据。

## 5. 抽取的答疑

1. 增量抽取是按系统时间来的？

不是的，是按照增量字段的来的。

2. 为什么我不能点击增量抽取

因为缺少增量抽取字段，增量抽取字段必须是数值类型和日期类型，但是按照数据要求，正常增量字段能反应数据增量变化的，一般都是时间类型。

3. 抽取中因为有部分数据错误，导致抽取失败，那么有效部分的数据会不会加载到目标数据库中。

只有抽取成功，我们才会把这次抽取的所有数据加载到目标库中。

4. 由数据原因引起增量抽取失败，该怎么办？

修复好数据，直接在抽取管理里面找到该抽取，点击立即抽取即可。

5. 全量抽取覆盖模式，定时抽取完，发现目标库和源头库数据不一致，什么原因？

一般是源头数据库的数据，在定时抽取之前还没有完全更新好。当你手工抽取的时候，数据源数据已经更新完毕了。导致这种原因是一般源头库数据量增多或者库的性能下降，本来数据源2小时可以更新完所有数据，但是现在需要更新3小时。这时候抽取就会不准。建议联系dba 优化表性能，或者合理调晚抽取时间。

6. 建好模型并进行抽取后，只有抽取成功和抽取失败的提示，但是不能确定数据是否被成功抽取并插入到宽表中，我能找到的验证方法是点击模型一直往下翻看数据，如果宽表中历史数据较多的话，很难找到刚跑出来的记录，该怎么办？

可以在数据模型里面做一些筛选，然后看宽表。只要数据模型不保存，那些筛选就不会被添加。

7. 在增量抽取下，重新抽取是什么意思？

重新抽取会选择时间段，抽取时间和增量字段无关。重新抽取会重跑在选择的时间段下面该数据模型的所有抽取任务。

如果点击立即抽取会有会怎么样的操作呢？首先系统会识别目前源头库最大的时间，8月13号，然后目标库目前也有13号数据，所以点击立即抽取，只会重抽最新的一天数据。

8. 抽取是以数据模型为单位的吗？

不是的。数据模型里面有多个物理表，那么该模型就有多个抽取任务。同时自定义SQL中有多个物理表，那么抽取是怎么进行的，我们只会把这个自定义sql的结果存到目标库中，不会对其中的物理表抽取。

所以抽取的基本单位是物理表和自定义SQL，而不是数据模型。这也解释了增量抽取中，在数据模型那里将字符型转化成日期型做增量抽取不行，但是在自定义sql 和物理表做转化，该字段才会在增量抽取中添加成功。

## 数据模型变动了, 报告这块我们该做什么？

数据模型变化，比如说增加一个字段，接着我们该做什么呢？

我们要在数据模型中点击保存，然后会弹出下面信息，



我们要选择覆盖，接着会告诉你有多少报表受到影响，点击应用。

未命名报告(5)报告将受到影响	<a href="#">查看详情</a>
地图测试手机报告将受到影响	<a href="#">查看详情</a>
地图报告将受到影响	<a href="#">查看详情</a>
未命名报告(43)报告将受到影响	<a href="#">查看详情</a>
自定义地图报告将受到影响	<a href="#">查看详情</a>
未命名报告(3)报告将受到影响	<a href="#">查看详情</a>

应用
取消

为什么要这么做？

因为在AxisBI中数据模型和报表的关系是这样的，

我们创建报表会引用数据模型，但是这个引用的数据模型是原始的数据模型的一个副本。如原始的数据模型有变化的时候而且有些字段被报表引用了的话，不点击覆盖应用到报表，那么报表这边不会看到问题，但是运行报表会出错。如果点击模型，应用到报表，当前的副本模型会更新到原始模型的最新版本。这时报表有问题我们可以很快看出，那时候我们只要修改报表就可以了。

这样做的好处是防止长时间不运行的报表，突然运行，报错，确找不到原因。

## 数据模型切换

数据模型切换目前可以跨数据源切换数据模型，不过要数据模型的结构完全一致。

但是目前不能批量切换。只能一个个模型切换。

## 为什么在数据模型中数据会加倍？

出现这种情况，该数据模型是一个join 类型的数据模型。

如果join，那么主表有一条，通过连接条件，但是您的副表多条符合条件的数据。那么会导致数据加倍。

解决方案，重构数据模型，重新确定连接条件。或者看看副表2条数据其中是不是脏数据。

这种问题很常见，主要是业务模型的原因。

## AxisBI自定义sql模块执行原理

目前我们AxisBI自定义sql 的语法是完全和数据源的数据库有关的。

如：连oracle，我们的语法就必须写成oracle的语法。

同理连mysql，我们就必须写成mysql 的语法。

我们会拿着自定义的sql直接去数据库解析该sql。其实就是一个查询返回的结果集。

## 数据抽取总是在加载中？

- (1) 抽取的数据本身就非常庞大。（建议瘦身数据）
  - (2) 在同一个数据源，其他用户有大批量数据在抽取，数据耗性能的sql太多的，数据库负载太大。
- 大数据量建议改变抽取频率，适当调整抽取时间段，尽量找资源空闲的时间抽取，防止资源总是被占用。

## LOD算法其一：fixed的含义及其用法

LOD算法的全称是详细级别表达式（level of detail expressions）。它主要是为了克服一些表达式之间的算子粒度不一致的问题。例如，我们在回答一个简单的问题的时候：我想知道每个人的身高与平均身高之间的差别。通常，为了得到结果，我们需要进行两步查询，第一步，查出所有人的平均身高：SELECT AVG(height) FROM humans，假设该值为H； 第二步，将每个人的身高减去这个平均值：SELECT height - H FROM humans；

这是一个两步的过程，如果想使用一个表达式完成这个查询该怎么做呢？它的过程就会比较复杂：

```
SELECT all.height - avg.H
FROM humans as all
JOIN (SELECT AVG(height) as H FROM humans) as avg
```

而LOD就是相当于实现了后者的过程，以便用户在一个简单的表达式内就简单的可以实现这个需求： [身高]- (FIXED : AVG([身高]))

详细级别表达式可以解答非常多这类的问题，比如单日销售额和月总计销售额的占比如何，总售出超过100万元的产品是来自哪些订单……凡是明细粒度和聚合粒度不一致，但是却又需要比较的场景，我们都可以通过LOD简单的实现。

再举一个例子，AxisBI的表格如下

浙江	杭州	100
浙江	宁波	150
江苏	南京	200
江苏	苏州	150

如上图，假设我们需要求省级别的总和，我们只需要做成如下

浙江	250
江苏	350

那如果我们需要做成

浙江	杭州	100	本省之和 (250)
浙江	宁波	150	本省之和 (250)
江苏	南京	200	本省之和 (350)
江苏	苏州	150	本省之和 (350)

该如何实现这个本省之和呢？

这就引入fixed做跨粒度计算，本省之和= {FIXED [省份]:SUM(额度)}

这样就相当于在省份的粒度上对额度求和了。这样也更容易计算各城市额度占省份之比等各类指标。

扩展阅读1:

假设我在省份的面前加上一个日期维度:

2019-8-1	浙江	杭州	100	{FIXED [省份]: sum (额度) } =450
2019-8-1	浙江	宁波	150	{FIXED [省份]: sum (额度) } =450
2019-8-2	浙江	杭州	200	{FIXED [省份]: sum (额度) } =450
2019-8-1	江苏	南京	200	{FIXED [省份]: sum (额度) } =200

可以看到这里的维度其实是不对了，因为我们要求杭州在8-1号的省占比，其实应该是100/(100+150)而不是100/(100+150+200)。所以此时需要将计算字段改为{FIXED [日期],[省份]:SUM(额度)};

扩展阅读2:

由于fixed语法其实是利用了一个子查询，如

```
SELECT all.height - avg.H
FROM humans as all
JOIN (SELECT AVG(height) as H FROM humans) as avg
```

那么如果我在界面上添加了一个过滤器比如 性别筛选器=男，那么这个查询会变成什么样呢？如下:

```
SQL1:
SELECT all.height - avg.H
FROM humans where SEX="男" as all
JOIN (SELECT AVG(height) as H FROM humans) as avg
```

但其实我们需要男性平均身高差其实应该是

```
SQL2:
SELECT all.height - avg.H
FROM humans where SEX="男" as all
JOIN (SELECT AVG(height) as H FROM humans where SEX="男") as avg
```

这个时候就需要将我们的性别筛选器的优先级提升，勾选其“优先筛选”才可以做到SQL2的效果。



FIXED的优先级可以查看此文:

<https://docs.ksyun.com/documents/38309>

## 使用LOD函数解决impala去重计数报错问题

问题说明:

impala不支持在同一条select语句中对多字段去重统计

具体说明,可以参考impala官方文档:

[https://impala.apache.org/docs/build3x/html/topics/impala\\_count.html](https://impala.apache.org/docs/build3x/html/topics/impala_count.html)

解决方法有两个:

1) 修改APPX\_COUNT\_DISTINCT参数

APPX\_COUNT\_DISTINCT=true

开启参数后，相当于使用了ndv(函数)，会得到不那么精确的结果

2) 使用LOD函数通过子查询绕过impala的限制

示例:

将度量【去重计数有效会员数】改为ATTR({INCLUDE : COUNTD([有效会员数])})

下面是用LOD函数实现的示例:

使用LOD函数通过子查询绕过impala的限制:

每新增一个去重统计的计算字段，在新增的计算字段里增加一重include的嵌套，如下

流失会员数: ATTR({INCLUDE: COUNTD([流失会员数])})

新增会员数: ATTR({INCLUDE:ATTR({INCLUDE: COUNTD([新增会员数])})})

有效会员数: ATTR({INCLUDE:ATTR({INCLUDE:ATTR({INCLUDE : COUNTD([有效会员数])})})})

## 处理聚合和非聚合字段的混合计算

### 1. 聚合和非聚合字段是无法进行混合计算的

当使用非聚合度量计算时，计算使用的是最基础的明细数据，一次处理一行

当使用聚合度量时，是将多行数据合并为一行，所以二者不能混合计算。

具体可参考: <https://docs.ksyun.com/documents/38361>

### 2. 将聚合字段转为非聚合字段

通过【跨视图粒度计算】的fixed函数，将聚合度量转化为非聚合度量。

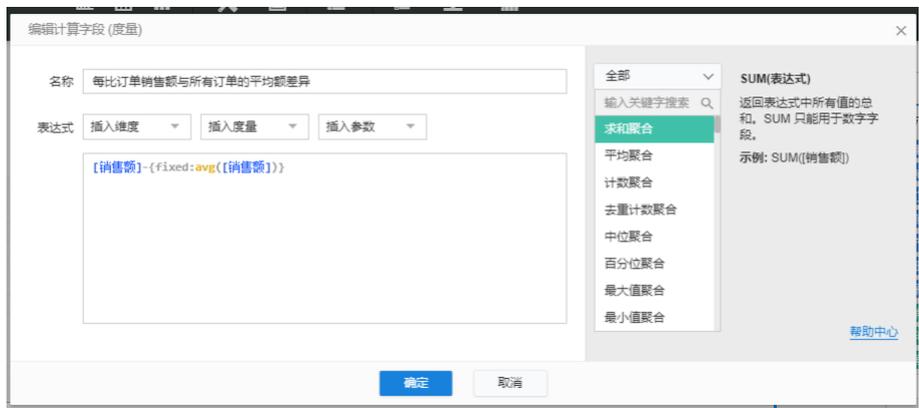
【示例】:

场景: 比较每笔订单的销售额和所有订单平均销售额

1) 所有订单平均销售额: {fixed:avg([销售额])}



2) 每比订单销售额与所有订单的平均销售额差异: 销售额-所有订单平均销售额差异



3) 得到结果:

发货日期	订单_ID	子类别	销售额	所有订单平均销售额	每比订单销售额与所有订单的平均差异
2018-09-01	CN-2980-3497234	收纳具	1129.80	1611.75	-481.95
2018-09-01	CN-2984-2547654	用具	662.76	1611.75	-948.99
2018-09-01	US-2982-1357144	用品	129.70	1611.75	-1482.06
2018-09-01	CN-2979-3930358	电话	1241.52	1611.75	-370.23
2018-09-01	CN-2984-2547654	标签	164.22	1611.75	-1447.53
2018-09-02	US-2980-5728538	收纳具	810.18	1611.75	-801.57
2018-09-02	US-2979-1069751	器具	4167.07	1611.75	2555.32
2018-09-01	CN-2978-5334809	复印机	1346.52	1611.75	-265.23
2018-09-02	US-2979-1069751	纸张	248.08	1611.75	-1363.67

关于【跨视图粒度计算】，详见用户手册:

<https://docs.ksyun.com/documents/38092>

## 取数模块和报告模块的区别

### 1. 场景介绍：

假设现在有交易，地区，客户三张表

#### 1. 交易表结构

客户id	地区id	销售额
------	------	-----

#### 2. 地区表结构

地区id	地区名
------	-----

#### 3. 客户表结构

客户id	客户名称
------	------

在AxisBI上建模如下：（关联方式是内连接）



### 2. 报告和取数模块使用：

使用步骤是：

（1）报告模块使用方式：

数据连接->数据模型（固化）->报告

（2）取数模块使用方式：

数据连接->数据模型（动态）->取数

从使用步骤可以知道，差异点是在是否是动态。

假设现在有两个需求：

需求1：按客户维度去分析销售额。

需求2：按地区维度去分析销售额。

需求1：

（1）报告侧使用：

在表格中拖入客户名称和销售额。形成的sql如下：

```
select 客户名称,
sum(销售额) from (select * from 订单表 t1, 地区表 t2, 客户表 t3 where t1.客户id=t3.客户id and t1.地区id=t2.地区id) group by 客户名称
```

（2）取数侧使用：

在表格中拖入客户名称和销售额。形成的sql如下：

```
select 客户名称, sum(销售额) from 订单表 t1, 客户表 t2 where t1.客户id=t2.客户id group by 客户名称
```

需求2：

（1）报告侧使用：

在表格中拖入地区名和销售额。形成的sql如下：

```
select 地区名, sum(销售额) from (select * from 订单表 t1, 地区表 t2, 客户表 t3 where t1.客户id=t3.客户id and t1.地区id=t2.地区id) group by 地区名
```

（2）取数侧使用：

在表格中拖入地区名称和销售额。形成的sql如下：

```
select 地区名, sum(销售额) from 订单表 t1, 地区表 t2 where t1.地区=t2.地区id group by 地区名
```

结论：

（1）报告侧，一旦模型确定，报告侧都是基于模型固化的sql做的分析。模型的sql：

```
select * from 订单表 t1, 地区表 t2, 客户表 t3 where t1.客户id=t3.客户id and t1.地区id=t2.地区id
```

（2）取数侧，表格中使用了哪些字段，就把这些使用字段的表进行关联。

### 优缺点：

1. 取数是动态建模，可以适用于星型模式（一张事实表，多个维度表关联）

当数据量大的时候，少关联表可以大幅度提升性能。

2. 取数是动态建模，是一种探索式分析。

3. 报告是固化建模，有一定的使用场景。两者不可替代。

## 了解IF ELSE 解决场景中的一些问题

该表达式语法如下：

```
IF 条件 THEN 返回值1 ELSE 返回值2
当条件满足时，返回返回值1，否则返回返回值2。
```

使用场景

场景1: 假设有份超市数据, 有个地区字段。需求: 想把东北和华北的数据统一划分“北方”, 其他地区继续使用原来字段。

解决方案:

方法一: 利用AxisBI本身的创建组功能。

方法二: 利用if else 。

```
if [地区] in("东北","华北") then "北方" else [地区]
```

场景二: 假设有份全国人口数据, 有个年龄段字段。需求: 想统计年龄段的人数分布。

方法一: 如果年龄段, 是等距的。如0-10岁, 10-20岁, 20岁至30等分布。可以使用数据桶功能。

方法二: 利用if else, 好处可以实现不等距的分布。

```
if [年龄]>=0 and [年龄]<10 THEN "0-10岁" else if [年龄]>=10 and [年龄]<20 THEN "10-20岁" else if 依次类推
```

场景三: 假设有份超市数据, 有地区和销售额字段。需求: 想知道全国销售额和东北区地销售额。

方法: 创建if else 字段, 名称为: 东北销售额 。

```
if [地区]="东北" then [销售额]
```

然后把销售额和东北销售额拉入表格组件。即可显示全国销售额和东北销售额同时展示。

场景四: 假设有超市数据, 有地区字段, 字段内容001, 002。需求: 想把001, 002变成东北和华北。

方法一: 利用数据字典实现别名。

方法为: 利用if else 。

```
if [地区]="001" then "东北" else if [地区]="002" then "华北" else [地区]
```

场景五: 假设有超市数据, 有地区字段, 内容有脏数据, 如null, ""等。需求: 分析时需要把这些字段剔除。

方法: 利用if else, 新建计算字段, 然后把该字段放入图表筛选或者筛选器。设置排除“Y”值的。

```
if isnull([地区]) or [地区]=" " then "Y" ELSE "N"
```

### 如何建立平台dataset和echart组件的映射?

AxisBI开发了js组件。这意味着AxisBI具有使用echart组件的能力。

AxisBI使用echart, 会用自己的dataset数据集来填充echart的数据部分。

问题: AxisBI的dataset如何存储?



从上图的左边部分可知AxisBI对dataset是列数存储。从右图可知dataset实际存储情况。dataset实际上是一个二维数组。所以我们要取“地区”的数据我们只要遍历dataset的第一列数据, 取“折扣”即遍历第二列数据。依次类推。

现在我们做一个最简单的echart 图形:

下面代码为echart官网的折线图实例:

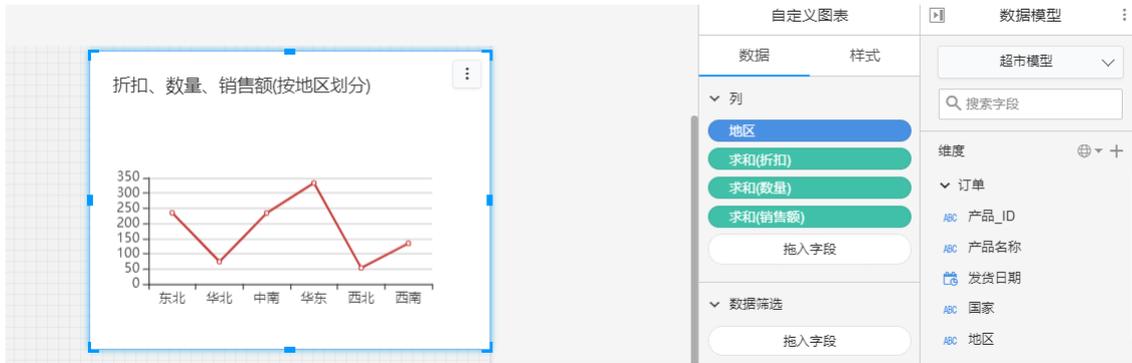
```
option = {
 xAxis: {
 type: 'category',
 data: ['Mon', 'Tue', 'Wed', 'Thu', 'Fri', 'Sat', 'Sun']
 },
 yAxis: {
 type: 'value'
 },
 series: [{
 data: [820, 932, 901, 934, 1290, 1330, 1320],
 type: 'line'
 }];
}
```

下图为填充AxisBI的dataset的折线图实例:

```

// 向右侧数据面板中拖入字段, 可在代码中引用dataSet字段获取二维数组数据
//对dataset进行遍历
dataSet = dataSet.map((data, i) => dataSet.map(d => d[i]));
var 地区 = dataSet[0];
var 折扣 = dataSet[1];
var 数量 = dataSet[2];
var 销售额 = dataSet[3];
option = {
 xAxis: {
 type: 'category',
 data: 地区
 },
 yAxis: {
 type: 'value'
 },
 series: [{
 data: 折扣,
 type: 'line'
 }];
}
```

如图效果:



## 如何彻底删除报表？

AxisBI具备回收站功能。该功能为了防止用户错删报表，可以通过回收站还原。

如何彻底删除报表？

用户在公共区域删除报表，并不直接删除资源。而是把删除报表放入回收站。

如彻底删除需要在回收站在一次删除。（删除操作由系统管理员定时清理）

## 如何对一个数字开方？

计算字段内支持power函数，但是没有sqrt函数，那么如何开方呢？

POWER(底数, 指数)

返回底数的指数次方。

示例:  $POWER(2, 3) = 2 \times 2 \times 2 = 8$

我们知道开方的定义即为数字的1/2次方，因此，我们可以使用

$power(2, 0.5)$  即 根号2 d

## AxisBI打开报告很慢？

方案1: 打开电脑的任务管理器看看有什么耗内存的软件，关掉对应的应用程序。如: acture软件。

方案2: 浏览器不要多开，浏览器每个页面本身也是很耗内存，会导致AxisBI变慢。

这个和方案1是相同的原因。

方案3: 网络不稳定，请检查网络。

方案4: 报表是直连，数据源对sql支持的不是很好。

方案5: 报表上存在很多组件，很多分页，很多表计算其中的一种或几种。

对AxisBI报表尽量做多分页，或者分报表做。

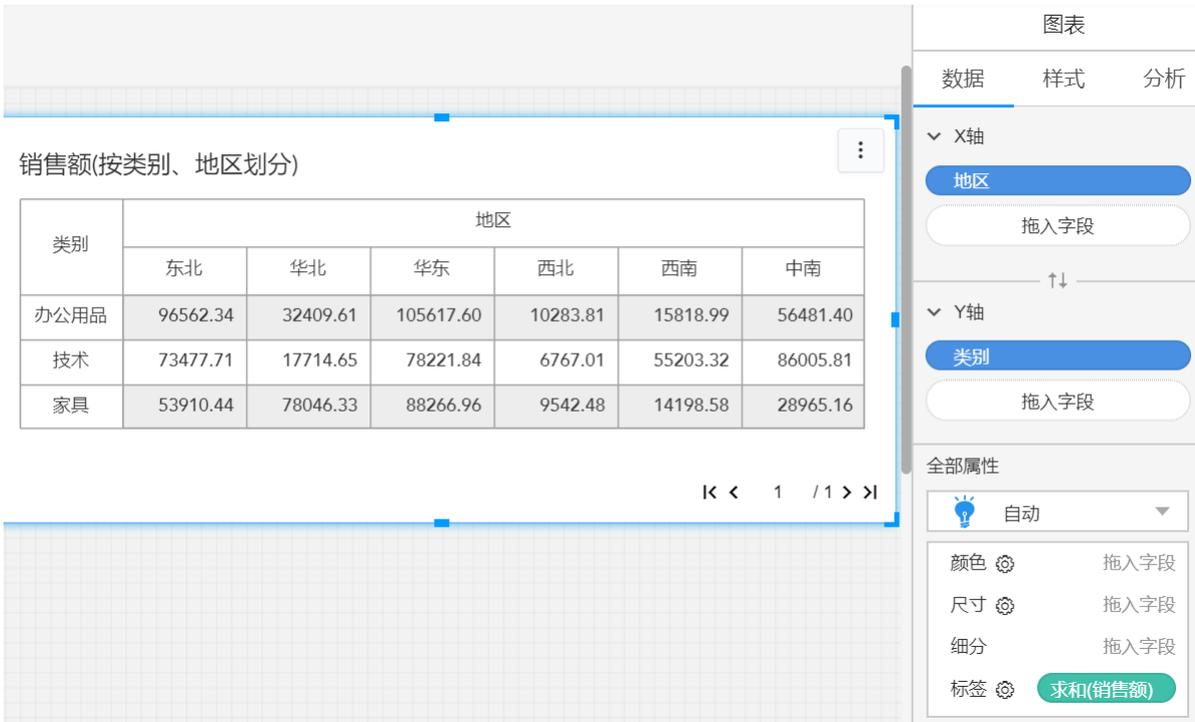
方案6: 排除以上5种方案请联系AxisBI技术人员。

## 简单理解可视化分析的分区和寻址？

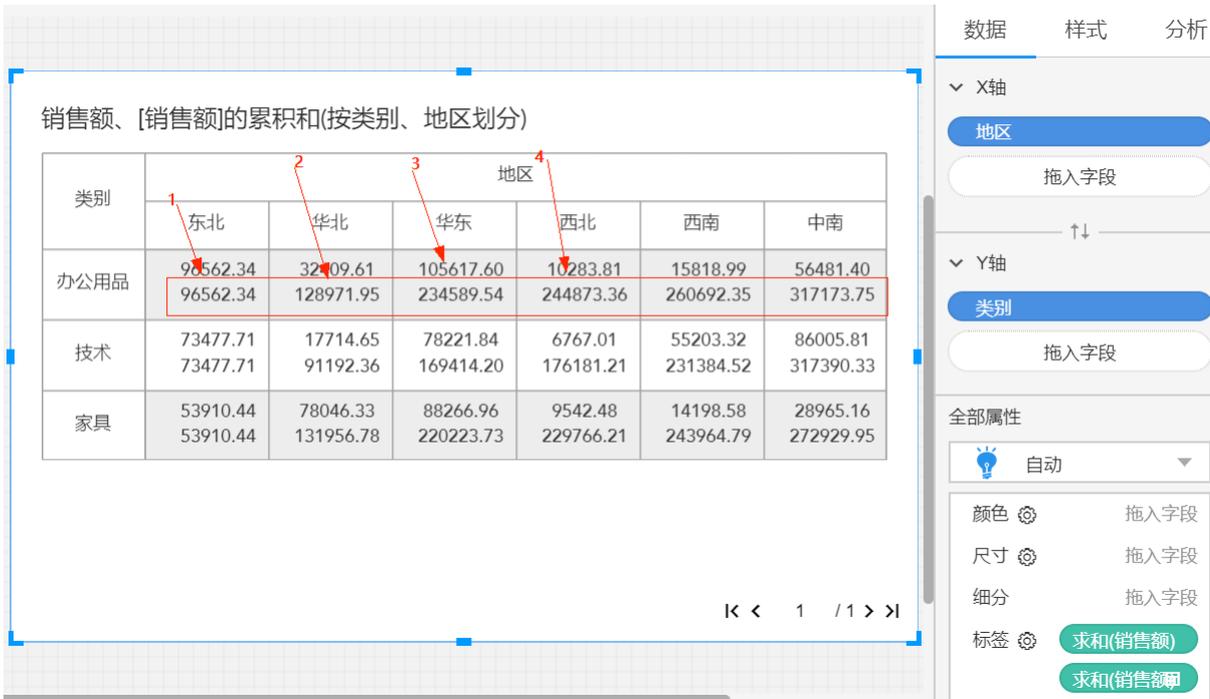
何为分区: 定义用于计算的字段如何分组;

何为寻址: 定义用于计算的字段在该视图中的该计算的依据和方向。

demo: 如图数据:



如果对销售额按照类别分区，按照地区寻址，做累加求和。会得到什么样的数据？



结果如上：

在办公用品这类：

第一排是每个地区的销售额。

我们观察第二排数据：，1号位销售额是他本身，2号位销售额等于东北加华北销售额，

3号位是：东北加华北加华东的销售额。

同理在技术这类：结果也是如此计算。

所以分区是将数据分成一组，寻址按照原有的数据的顺序，按照一定的规则计算。

## 度量字段和维度字段在AxisBI的含义

1. 首先需了解度量和维度的基本定义：

<https://docs.ksyun.com/documents/38295>

2. 对于AxisBI：维度和度量

在下图中最右侧系统会默认会把你的表分成维度和度量字段（一般规则数值类型为度量，其他字符，日期等都为维度字段）。在图中中间，图表部分，蓝色的数据胶囊为维度字段，绿的数据胶囊为度量字段。

在一个图表中数据展现，以图表配置数据区为准。

图一



3. 如何理解：维度参与分组，度量参与聚合？

如图一：目前维度字段是：地区和类别，度量字段：数量和销售额。

所以AxisBI形成的sql 如下：

```
select 地区, 类型, sum(数量), sum(销售额) from 超市表
group by 地区, 类型
```

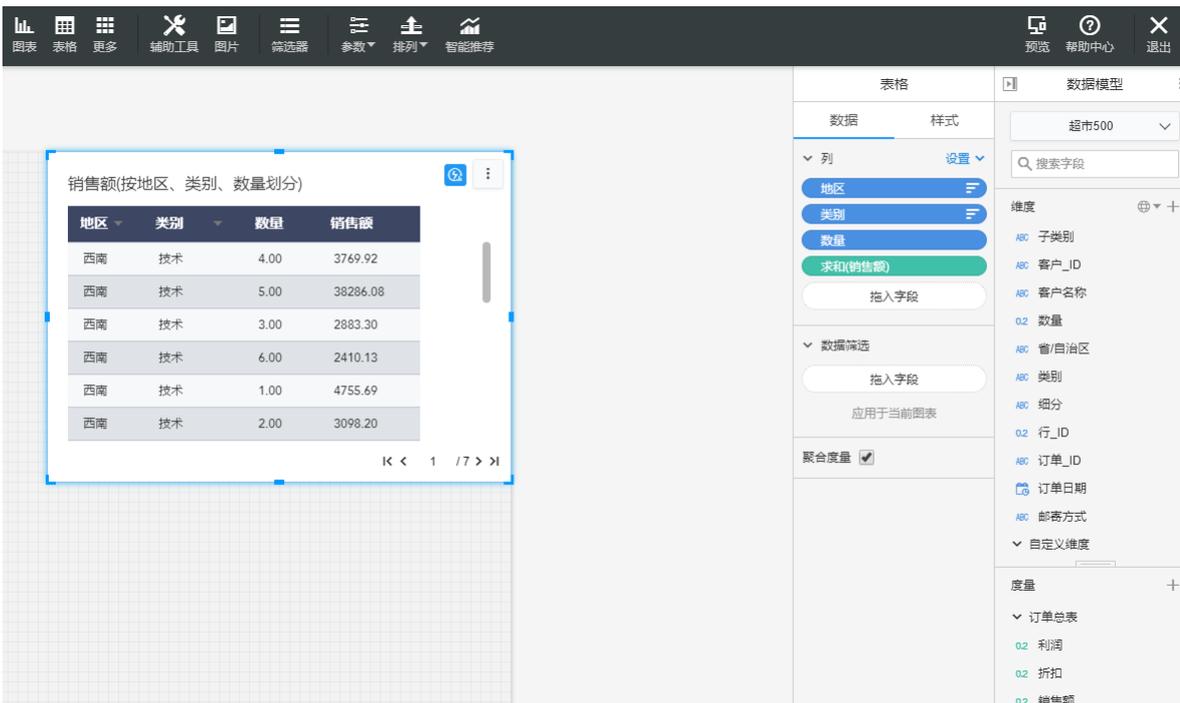
如图二（下图）：目前维度字段是：地区，类别和数量，度量字段：销售额。

所以AxisBI形成的sql 如下：

```
select 地区, 类型, 数量, sum(销售额) from 超市表
group by 地区, 类型, 数量
```

从数据量上看数据明显增多（由图一的1页数据变成了下图7页数据），同时：相同地区和类别的数据出现多行，显然数量这个字段也参与分组。

图二



### 报告分享和集成的区别

报告提供分享功能，就意味着通过分享链接看报表不需要用户登录权限。

相关知识点：

报表分享的页面是可以做集成的。

报表分享是最简单的一种集成，不能控制权限，不能设置页面大小等一些参数。

只有私有化部署的报表，才提供灵活的集成功能。

### 如何排查报表数据问题？

关于数据准确性问题，一定是基于参考数据做对比（首先要保证参考数据的准确性）。

排查方法：

首先要确定是数据哪个环节出了问题？

对于AxisBI报表，数据的传递环节有：

1. 数据连接->数据模型->报表

排查思路：从报表到数据模型到数据连接。层层比较。

环节点	排查点
报表	筛选器，计算字段
数据模型	数据模型类型，筛选器，数据关联
数据连接	数据连接配置

说明：

(1) 报表端常见的问题有：

报表的筛选：有报表页面筛选，筛选器组件，图表配置区的筛选。

多关联了筛选器导致数据变少或者变没

展示的字段是计算字段，结果逻辑写的有问题

计算字段里面用了fixed函数，但报表中有筛选器。

(2) 数据模型常见的问题有：

数据模型是抽取类型，数据变化只在源头，却没有执行抽取调度

数据模型上有筛选器

数据模型有多表关联，因为连接条件有问题导致数据变多倍，或者数据减少。

(3) 数据连接常见的问题有：

参考数据是生产环境，数据连接连的是测试环境。

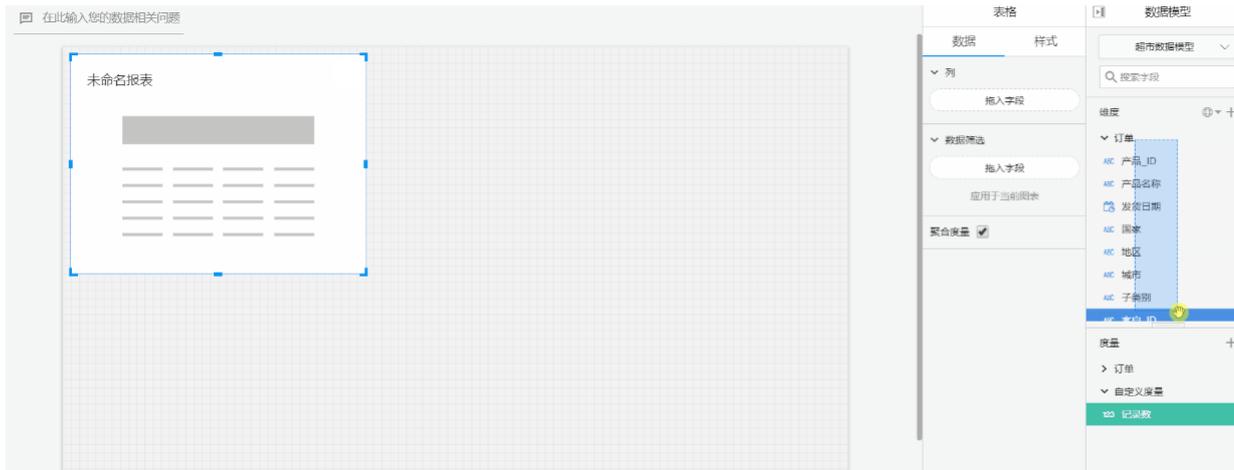
温馨提示：

如果数据有大规模的问题，可以关注一条数据，以点代面排查，层层排查。

## 批量选择字段，批量拖拽字段的功能

一般框选在自定义维度和度量或者表名的空白处，比较容易选中，使用光标划过。或者按照shift键进行多选。

图片如下：



## 跳转报告的时候如何设置浮窗大小使其适应窗口？

一般来说为了更好的跳转浏览体验，当被跳转的报告过大的时候可以：

- ①被跳转的报告其页面属性设置为自适应窗口
- ②不使用跳转浮窗，使用跳转至新开页面

## 如何将拥有共同特征的字符串都提取出来

比如有一个字段customer\_name，它是客户的名称，但是由于是人工录入可能命名并不规范，比如金山云，录入的时候可能叫做“金山云公司”，“北京金山云网络公司”等等，我们需要将其统一称呼为“金山云网络技术有限公司”，该如何处理呢？

可以使用计算字段实现，有两个函数可以实现：

①使用contains函数

**CONTAINS(字符串, 子串)**

如果给定字符串包含指定子字符串，则返回 true。如果给定字符串不包含指定子字符串，则返回 false。

示例: CONTAINS("Calculation", "alcu") = true,  
CONTAINS("Calculation", "beu") = false

IF CONTAINS([customer\_name], "金山云") then "金山云网络技术有限公司" else [customer\_name]

②使用find函数

**FIND(字符串, 子串, [起始位置])**

返回子字符串在字符串中的位置。如果未找到子字符串，则返回0。起始位置为可选参数，如果给定此参数则从此位置开始查找，字符串中的第一个字符为位置1

示例: FIND("Calculation", "alcu") = 2

IF FIND([customer\_name], "金山云") > 0 then "金山云网络技术有限公司" else [customer\_name]

## excel 导出问题排查方法

excel 导出出现如下问题:



可能原因:

1. 报表生成的SQL有异常。

思路: 首先查看报告本身展示是否有异常, 若无异常, 可以尝试打开下载下来的这个失败文件, 文件内有时候会有对应提示信息, 例如报表重名导致的导出失败或者是网络因素导致的导出失败。

2. 底层数据模型变更。

思路: 数据模型发生变更以后, 如果保存模型的时候未选择覆盖到报告, 在导出报告的时候可能会导致导出异常 (因为导出数据的时候必须实时查询底层数据)。若因此失败, 可以尝试刷新报告/刷新模型/编辑模型并且覆盖到对应报告, 这样的话可以让错误更加明显的暴露出来, 方便排查。

如何覆盖到对应报告: <https://docs.ksyun.com/documents/38342>

3. SQL运行超时。

sql 查询超时, 目前导出的查询时间是1分钟, 如果sql 查询时间大于1分钟, 会出现类似的问题。

建议性能调优。

4. AxisBI系统故障

思路: 若导出任何报告的Excel都会失败, 则考虑是AxisBI本身系统故障。若您有运维能力, 可以尝试进入schedule模块, 查看服务是否正常运行。若您没有运维能力, 则需要联系运维或者技术支持人员。

## 标题时间宏的用法

### 时间宏

AxisBI中有些地方 (邮件名称、图表标题、文本框、注释框) 支持插入时间宏, 我们可以插入一些时间关键词, AxisBI解析后, 会显示成关键词对应的具体时间, 比如输入 "{{today}}", AxisBI会显示成今天的具体时间, 假设今天是2019年3月14日, 则AxisBI会显示成 "2019-03-14"。

时间宏的输入格式为三个中括号+关键词, 比如 {{{today}}}, {{{today - 2weeks}}}。

AxisBI支持如下关键词:

关键词	含义
today	今天
yesterday	昨天
tomorrow	明天
first day of this month	本月第一天。其中, "first day" 可以替换成 "last day"。"this" 可以替换为 "last"、"next"。"month" 可以替换为 "week"、"quarter"、"year"
monday last week	上周一。其中, "monday" 可以替换为 "tuesday"、"wednesday"、"thursday"、"friday"、"saturday"、"sunday"。"last week" 可以替换为 "this week" 或 "next week"。

您还可以在关键词的基础上进行加减法, 更灵活地表示时间, 比如输入 "today - 4days", AxisBI会解析显示成4天前的具体时间。其中, "days" 可以替换为 "weeks"、"months"、"quarters"、"years"。数字可以任意替换, "-" 可以替换为 "+"。

时间宏一般会显示成标准的日期格式, 比如 "{{today}}" 显示成 "2019-03-14", 也支持显示成自定义的格式, 定义方式为 {{{today % 格式串}}}。

格式串示例如下:

格式串	显示格式
YYYY	2019
YYYY年	2019年
YY年	19年
YYYY-M	2019-3
YYYY/M	2019/3
YYYY年M月	2019年3月
YYYY-M-D	2019-3-14
YYYY/M/D	2019/3/14
YYYY年M月D日	2019年3月14日
M-D	3-14
M月D日	3月14日
M月	3月
M	3
YYYY年第Q季度	2019年第1季度

更多用法可以继续参考 <http://moment.js.cn/docs/#/displaying/format/>

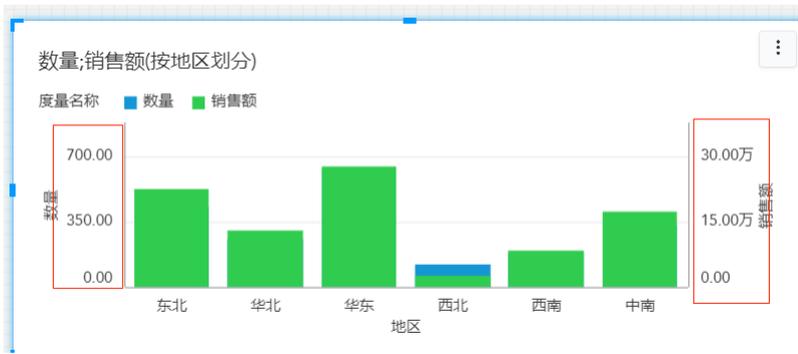
## 同一个表格上快速切换度量/维度? 动态维度如何使用

有一种需求, 需要按照【地区—销售额】来展现, 但是有时候又希望按照【类别—销售额】来展现, 亦或者希望按照【地区—利润额】等等, 一般情况下我们需要制作这三种图表, 方便查看。

但是AxisBI的【动态维度】、【动态度量】可以帮助你实现在一个图表上自由切换。

## 显示双轴, 双轴的坐标不一致, 如何解决?

如图所示两边的数轴不一致如何解决?



点击图表，样式，同步轴。

图表

数据   样式   分析

> 标题

> 图例

> 绘图区

同步轴

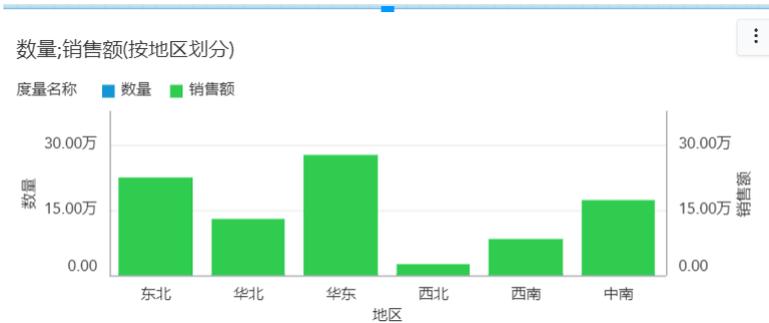
> 显示设置

缩略轴

背景

边框  实线

设置完效果：



## 查看有数据的最近一天

有时候BI工具使用的数据源会延迟，比如在跑批的情况下，按理说今天查看的是昨天的数据，但是有时候跑批任务延迟了，导致昨天的数据还没出来。此时如果报告选择了昨天，那么会显示报告查询数据为空。



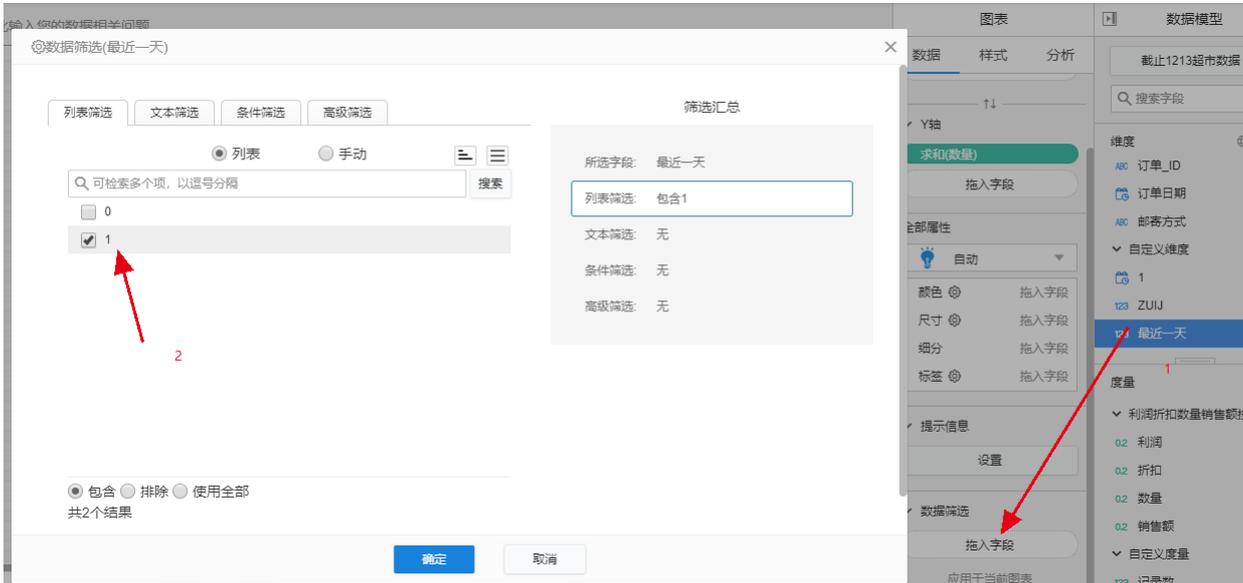
那这个时候我们希望展示出来的可以是前天的数据（或者说有数据的最近一天），该怎么做呢？

以上图报告为例，我们首先写一个计算字段：`IF datediff("DAY",[发货日期],[FIXED : MAX([发货日期])))=0 THEN 1 ELSE 0`

这个字段的目的在于获取有数据的最大一天。



然后将该计算字段拖入数据筛选，列表选择1，即可得到最近一日的数据



### 将表格的null值显示为0的几种方法

137	null	null
1590	58394	37983
12	202	0
1290	32185	24969
102	null	null

有时候数据会显示为null，想将其改为0，方法如下：

#### 1. 使用计算字段ifnull

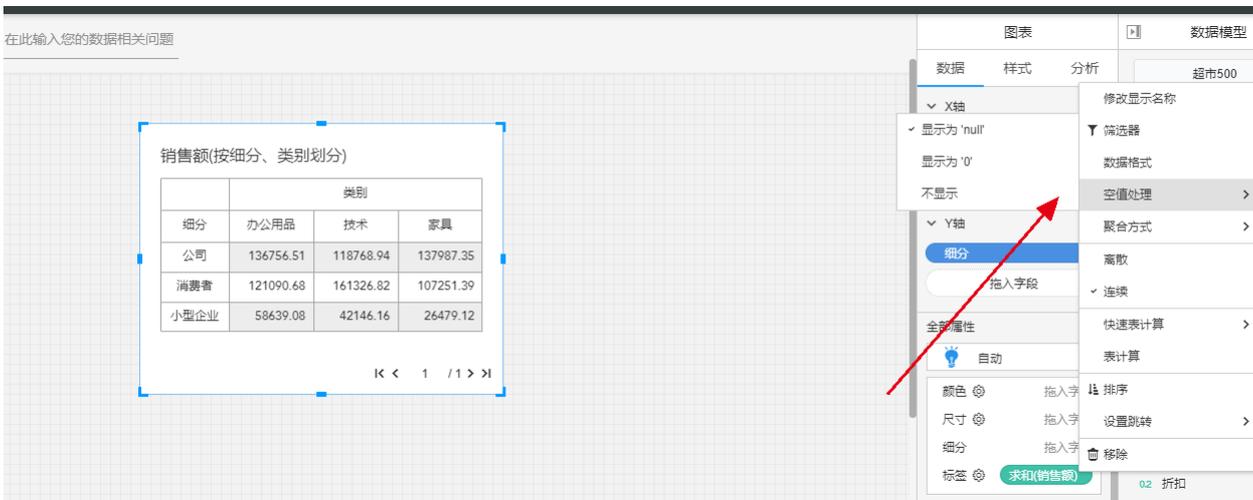
IFNULL(表达式1, 表达式2)

如果表达式1不为null，则返回表达式1，否则返回表达式2

例如字段A含有空值，我们需要将字段A里的空值显示为0，那么新建一个计算字段：IFNULL([字段A], 0)，该函数会判断，当字段A为空的时候，将其显示为0。

同理也可以，IFNULL([字段A], "你好")，当字段A为空的时候，将其显示为"你好"。

#### 2. 使用AxisBI的度量空值设置

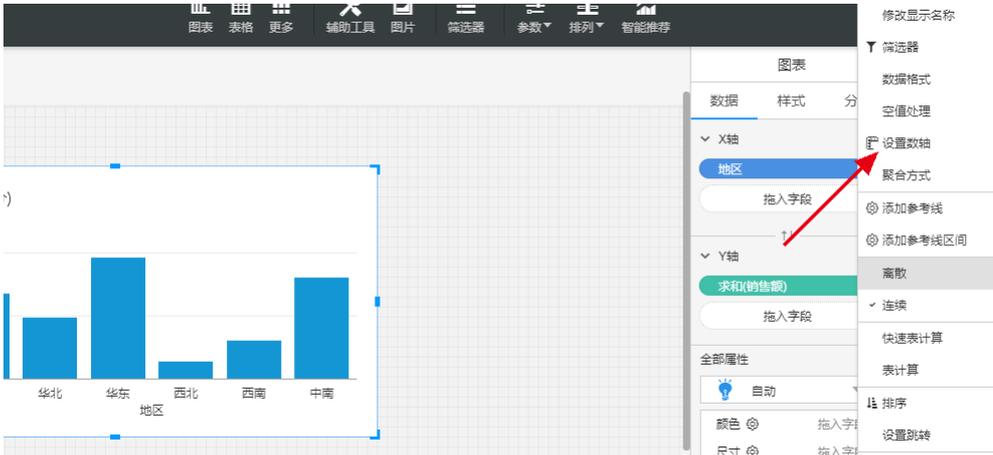


### AxisBI数轴的几种设置



如图所示，

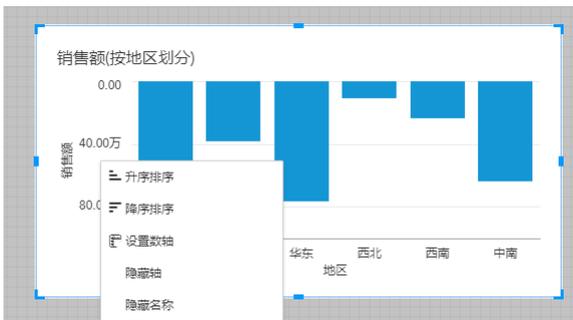
(1)、将数轴倒序



如此可以将数轴倒置。

(2)、隐藏数轴、隐藏名称

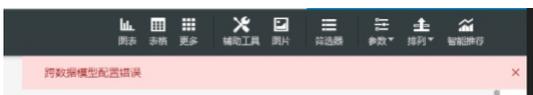
首先要双击选中报表，然后点击数轴，选择隐藏数轴或者隐藏数轴名称



(3)、排序

同(2)操作，可以对数轴进行降序或者升序排序

## 跨数据模型错误



报表点开出现跨数据模型配置错误。

问题原因：存在筛选器在做跨数据模型关联的时候出现问题。之所有出现这个原因是因为数据模型在后期做关联的时候，关联的两个字段类型不一致，导致的。

排查方法：查找所有筛选器，看哪个筛选器在跨模型关联的地方出现红色的框框，重新点击并关联匹配。

快速查看关联筛选器：<https://docs.ksyun.com/documents/38383>

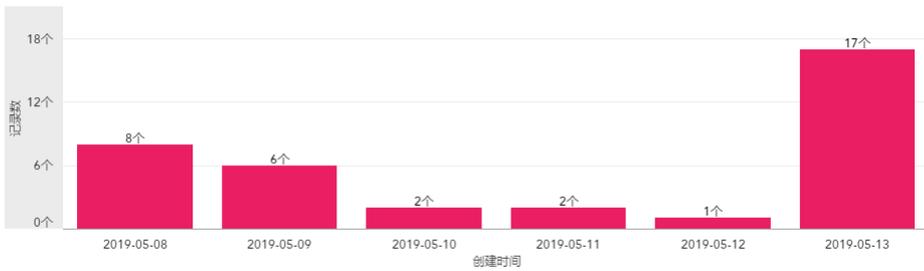
## 日期类型和日期显示格式的区别

AxisBI图表上有这两个选项，他们之间的区别是什么？



区别是：日期类型会以选中格式，做维度聚合计算；日期显示格式仅仅是在格式上做显示的处理。

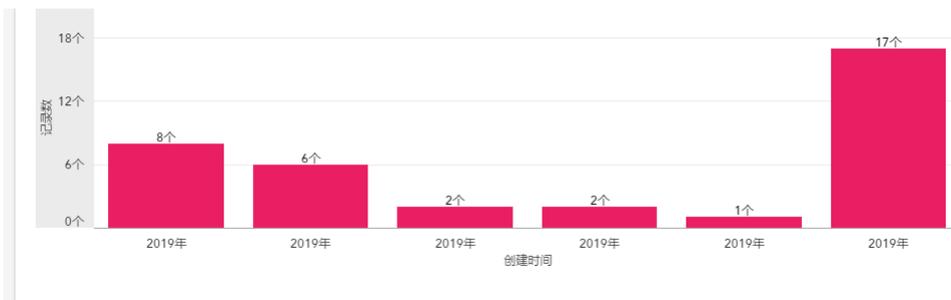
以下图我们用显示年展示。



日期类型的结果如下：



日期显示格式



## 如何在图表上的标签里展示最大值最小值和头尾数据显示



目前AxisBI有4种显示方式，但是头尾显示和最大值最小值显示，不能同时出现。

目前有策略，复制一张一模一样的图表，一张设置最大值和最小值，一张设置头和尾，把其中一张图表背景和标题设置透明即可。

### 如何快速查看报表关联的筛选器

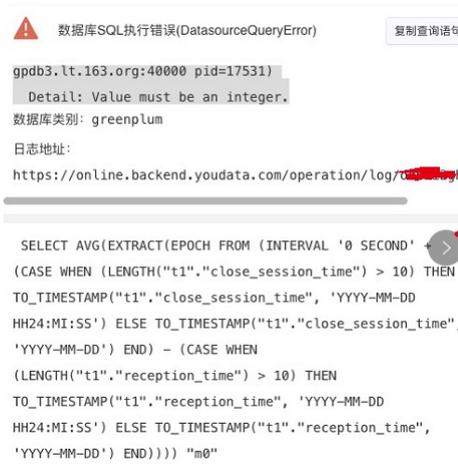
因为AxisBI采用了浮动式布局，经常有用户会将筛选器置于图层底层，导致不显示，亦或者使用了tab标签，将筛选器覆盖了，导致报表查看的时候会出现一些“神奇”的情况，比如数据莫名对不上等。针对这种情况，AxisBI新增了图表关联筛选器功能，鼠标点击图表信息处，可以看到：



可以看到此时有2个筛选器对筛选数据产生了影响。

### 数据转型导致的数据库SQL执行错误

这类错误的现象常常如下：



仔细查阅错误，一般会有更多的提示内容，比如

ERROR: invalid value "pres" for "YYYY" (seg6 slice3 gpdb3.lt.163.org:40000 pid=1\*\*\*)

Detail: Value must be an integer.

从这个错误提示其实大致能知道，应该是某个字段转型为日期的时候，其字段含有脏数据。比如这个错误内容里就提示，脏数据是字符串“pres”

在报错的SQL里，我们进一步可以查出来哪些字段参与了日期转型，在这个例子里可以看到reception\_time参与了日期转型，那么我们可以这么做：

我们做一个表格，然后把reception\_time还原为原始的格式（原始格式一般是字符串，就是从字符串转成日期的这个过程有问题），将reception\_time拉到表格上。然后使用数据筛选器，在筛选器的列表里我们可以看到，大部分reception\_time都是正常的“2019-04-01”等格式，但是随后发现AxisBI十个字段值里内容是“presale”，也就是说“presale”是无法转型成日期的，就是这些内容导致了报错，所以要解决这个报错，需要在源表里对这些数据进行处理。

如果不适合在报表层做这些操作，可以尝试进入模型，利用模型筛选器做。

关闭宽表 最近更新时间: 2019-05-09 14:51:58 最多显示1000行

dt	id	reception_time
2019-04-27	6	presale
2019-04-28	1	presale

### 为什么透视表展示相同模块的内容会出现两次？

透视表如果在抽取中相同的维度会出现2次。

场景：

在2个维度或者2个维度以上的维度做透视表的时候。

如省份，城市做展示的时候聚合销售额。

正常情况一般相同省份的城市会聚合到一起做展示。

省/自治区	城市	销售额
北京	北京	5801.60
吉林	白城	696.78
	蛟河	15693.16
	临江	18035.36
	双阳	3465.00
	四平	726.46
内蒙古	洮南	295.34
	包头	3578.79
	呼和浩特	689.44
	济宁	19304.46
	满洲里	224.95
上海	上海	30726.08
	成都	5148.36

但是如果抽取情况下，可能相同的省份，不同的城市会分两条展示。如上图就会出现两个吉林。

这是因为我们抽取是分布式的数据库，结果是分批返回的，这个时候会出现这个问题。

如何解决这个问题。

可以对y轴的数据胶囊做一个排序。

### 计算字段里的粒度不一致问题



大家写计算字段的时候可能会遇到这类“聚合粒度错误”，那么这个到底是什么意思呢？

其实所谓粒度是指能够区分每一行数据的字段称之为该数据的最细粒度。

通常，我们认为一份明细数据里，每一行都是一个最低粒度。与最细粒度对应的就是聚合粒度，比如SUM，聚合是把多行数据汇总的结果，也就是跨越了多个最细粒度。

上述错误其实是因为[人均奖励领取次数]这个字段本身是一个已经聚合过的计算字段，这个字段的写法其实是SUM(奖励次数)/COUNT(ID)，我们把上述错误简化一下，其实相当于：IF A>B THEN SUM(C) ELSE SUM(D)。

细心的用户可能发现了，A>B是最细粒度的比较，而SUM(C)、SUM(D)则是聚合粒度的结果。因此出现了跨粒度的错误。

修改的方法其实也非常简单，有两种：

①SUM(IF A>B THEN C ELSE D) ②IF SUM(A)>SUM(B) THEN SUM(C) ELSE SUM(D)

看到这里应该就明白了跨粒度是怎么一回事了。当然，①和②是完全不同的两种写法，其结果也大不相同，具体采用何种写法应该取决于你的业务目的。

PS，进一步分析①和②的不同，比如有如下明细数据

A B C D  
 1 2 1 0  
 1 2 1 0  
 3 0 1 0

那么显然①得到的结果是(D1=0)+(D2=0)+(C3=1)=1 ②公式得到的 sum(C)=3

A B C D  
 1 2 1 0  
 1 2 1 0  
 3 0 1 0

让我们继续展开说一个例子，在数据的聚合里，还有一个非常经典的例子，那就当我们需要得到一个倍率的时候，比如成功率，我们会用成功次数/总次数，其实应当表达为SUM(成功次数)/SUM(次数)=成功率

但是可能在报告制作过程中，有人会做成SUM(成功次数/次数)，这二者区别我们用一段数据表示

次数	成功数
5	2
4	1

这里的成功率=(2+1)/(5+4)，理解为全部实验的成功率

如果变成(2/5) + (1/4)那是完全不同的结果，意义变成了单次实验成功率之和了

因此，我们需要深入理解聚合与粒度的关系，才能得到我们希望的结果

### Excel导出后，打开提示Excel文件损坏

问题如图所示：



解决方法：

造成问题的原因主要是有重名报表的存在，可能是因为复制报表或者其他原因导致报表重名。导出Excel以后，每个报表都会形成一个sheet，同名的报告会变成同名的sheet导致Excel报错。

因此按照以下步骤查找同名的报告：

先删除一个报告页查看导出是否异常，如无异常继续删除。如有异常，则说明该tab有重名报告。检查每个报告的标题（注意即使是隐藏报告标题也会引发错误）

## 计算字段的日期比较、日期常量、日期加减

日期比较：

日期的大小比较可以使用<(小于)、=(等于)、>(大于)、!=(不等于)、>=(大于等于)、<=(小于等于)

日期加法：

日期增加或者减少数天的语法可以使用DATEADD

DATEADD(日期部分, 整数, 日期型表达式)

将日期型表达式的日期部分增加整数，可使用的日期部分包括“year”，“quarter”，“month”，“week”，“day”，“hour”，“minute”，“second”；函数返回一个日期。

示例：DATEADD(“MONTH”，1, [订单时间]) 表示将 [订单时间] 增加一个月

日期相减：

DATEDIFF(日期部分, 开始日期, 结束日期)

计算开始日期和结束日期的日期部分的差值，可使用的日期部分包括“year”，“quarter”，“month”，“week”，“day”，“hour”，“minute”，“second”；函数返回一个整数。

示例：DATEDIFF(“DAY”，[发货时间]，[收货时间]) 表示从发货到收货所用的天数

日期常量：

比如直接写“2019-01-01”会被系统识别为字符串，应该使用DATE(“2019-01-01”)将其格式化为一个日期常量

其中now() 函数表示此时的系统时间

today() 函数表示此时的系统日期

## SQL执行报错？

有一种常见的情况，一直用的好好的报告突然提示sql执行报错了：

因为AxisBI对于字段是允许进行类型变更的，但是类型变更的时候不会进行检查（AxisBI是读时检查），当你将一个字段phone（比如电话号码）转为数字类型的时候，AxisBI可以完成转换。如果是字符串类型，在AxisBI内生成的SQL类似：select \* from table where phone='137123456789'

如果你做了转型，那么查询会变成 select \* from table where to\_number(phone)=137123456789；若此时你的phone字段内有一个号码是这样的：0574-1234567；显而易见，这个号码不能被成功转型数字，因为有一个“-”符号。这样就会导致SQL查询错误。

因此，我们需要检查哪些做了类型转换的字段，原始的数据里是否有了不能被转型的脏数据

## 数据库SQL执行错误，如何排查问题？

查看报告的时候，会遇到数据库SQL执行错误，这种情况一般有几个方面的原因：

数据源问题，这种情况比较普遍，数据库发生了一些问题，导致sql执行失败。

在AxisBI上做了不合适的数据类型的转型导致，常见的有字符串转日期失败，字符串转数字失败。

AxisBI服务器问题，许可证过期等因素。

为了排查问题，首先需要获取原因，一般发生此错误的时候，在编辑页面上会出现“查看详情”按钮，



点击查看详情，可以将其中的内容复制出来，查看一下具体错误

还有一些语法兼容问题，比如impala不支持两次count(distinct)

## 如何解除报表和数据模型的关系

我们需要理解：

1. 在一张报表中我们要理清以下关系：

(1) 报表->数据模型 1对N

(2) 数据模型->图表组件 1对N

(3) 图表组件到数据模型是 1对1

2. 在多张报表中我们要理清以下关系：

一个数据模型可以被多个报表引用。

一个报表里面可以有多个数据模型。

我们制作报表的步骤是：

新建报表->添加数据模型->添加图表组件->用数据模型来填充图表的数据项

所以要解除报表和数据模型的关系。

1. 看一下这个数据模型被多少张报表。解除相互关系。用下面的方法。

2. 一个报表里面如何删除数据模型？

根据这个步骤：新建报表->添加数据模型->添加图表组件->用数据模型来填充图表的数据项

(1) 我们需要删除该报表中所有引用该数据模型的图表组件。（这里还没有结束哦）

(2) 删除完以后，我们可以看到这个数据模型有个“垃圾桶”图标这表明该数据模型只是在报表中添加，没有被任何组件所引用。

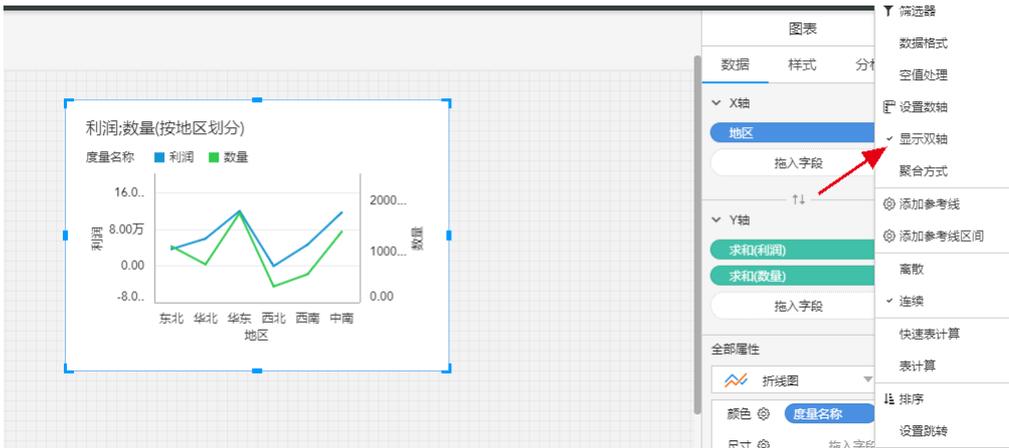


(3) 这时点击垃圾桶，移除该数据模型。

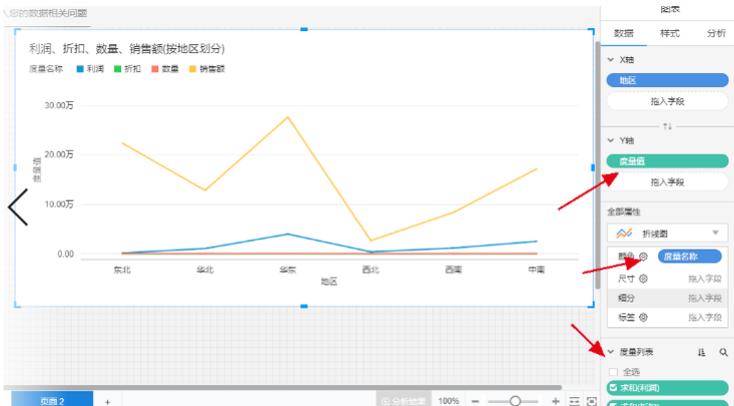
此时数据模型和报表的关系才算完全解除。

### 如何制作多个度量的折线图？双轴图？

一般来说，如果有两个度量需要双轴显示的时候，只需要在第二个度量的数据胶囊上点击“双轴显示”即可，如下图：



如果有三个及以上的度量又该如何操作呢？如下图：



巧妙地使用度量值就可以做到了。

### 如何排查DC错误或数据库错误？

如果第一次创建图表正常，后面报表出现DC错误或数据库错误。（排除升级导致）

首先定位引起问题的字段：

思路：控制变量

1. 首先选中问题报表。

复制一份。（养成备份习惯）

2. 在复制的报表上把其他的正常组件都删除。如果报表有多处有问题。只保留一个错误组件其他都删除。保存一下看看错误是够还存在。（减少其他干扰项）

3. 拉一个表格组件，

把问题图表的上的字段都拖到表格。

运行一下，看看。

看看表格是否报错。

如果表格报错，每次拖走一个字段，然后看看是哪个字段导致错误。

然后检查这个字段是不是计算字段。或者有没有做类型修改。

(1) 如果是类型修改了。还原成原始字段类型观察是不是有不符合要求的数据。

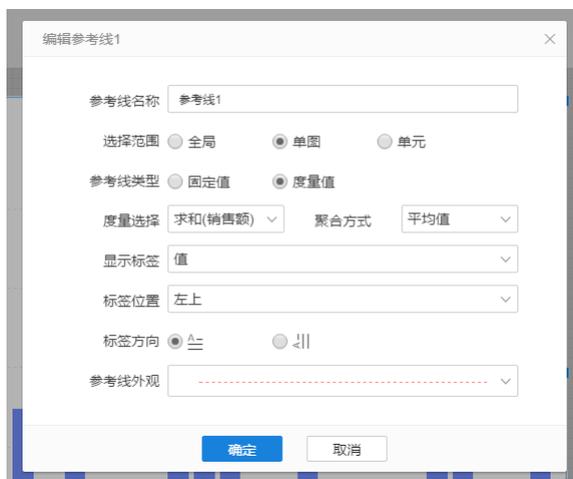
(2) 如果是计算字段，查看一下是否有正则表达式（比如mysql数据库不支持正则表达式），是否有不正确的跨粒度计算。

(3) 使用计算字段后，对字段的一些排序会引起一些错误，尝试取消字段排序

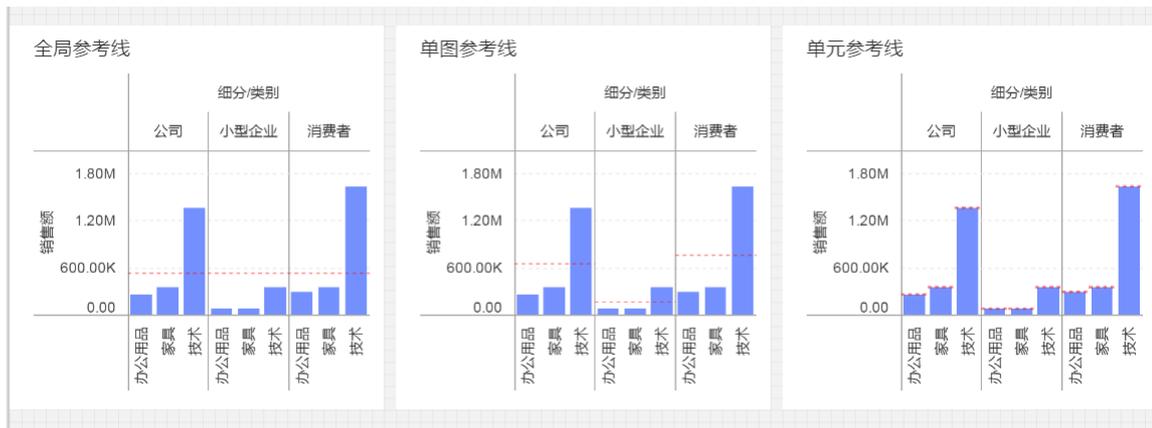
### 如何添加参考线？如何描绘一条用于比较的标准值？如何设定一个目标值/目标线？

参考线

为了辅助分析工作，网易有数中引用了参考线与参考区间的功能，这一功能的入口在X轴与Y轴的度量数据胶囊的下拉菜单中，如图所示：



参考线的范围会根据数据透视表的范围来画参考线，范围有三种，分别为：全局、单图与单元，它们的意义如下图所示：



参考线既可以是常数，也可以根据一个度量计算的值，比如这里求和(销售额)的平均值。

参考区间的编辑与参考线的编辑类似，只是需要设置两条参考线。

## 报告/指标卡查询出来没有数据或者数据不准确？查询数据为空？

问题现象：报告页显示查询数据为空或者指标卡显示0或者数据不准确



请按以下思路进行排查：

首先如果您的版本>=5.4（SaaS版本肯定是高于此版本，私有部署则需要咨询一下）。

则可以在刷新数据后，直接查看关联筛选器来判断：



通过排查其中的筛选器条件来判断数据是否有异常。

若您的版本低于5.4则：

①首先刷新报告数据：



②是否有筛选器导致数据被过滤；若您使用的AxisBI版本支持“查看SQL”功能，则可以点击查看sql，在sql内我们可以快速定位使用了哪些筛选。如下：



```
最近一次SQL语句 SELECT SUM("t1"."利润") "m0", (CASE WHEN "t1"."发货日期"::TIMESTAMP IS NULL THEN CAST("t1"."发货日期" AS VARCHAR(255)) ELSE TO_CHAR("t1"."发货日期"::TIMESTAMP, 'YYYY-MM-DD HH24:MI:SS') END) "d0", "t1"."国家" "d1", "t1"."城市" "d2" FROM "3246"."订单总表58c32e4fd1bb2cd7"."t1" WHERE (((("t1"."订单日期" >= CAST('2018-12-00 00:00:00' AS TIMESTAMP)) AND ("t1"."订单日期" < CAST('2018-12-04 00:00:00' AS TIMESTAMP))) AND ("t1"."类别" IN ('办公用品', '技术', '家具'))) GROUP BY (CASE WHEN "t1"."发货日期"::TIMESTAMP IS NULL THEN CAST("t1"."发货日期" AS VARCHAR(255)) ELSE TO_CHAR("t1"."发货日期"::TIMESTAMP, 'YYYY-MM-DD HH24:MI:SS') END), "t1"."国家", "t1"."城市" LIMIT 20
```

从SQL里我们可以很直观地看到这里有两个筛选，一个是[订单日期]的筛选，一个是[类别]的筛选。若您使用的AxisBI版本不支持查看SQL，则您可以在表格查看是否有筛选，或者页面是否筛选器表格的维度筛选：



报告页面筛选:



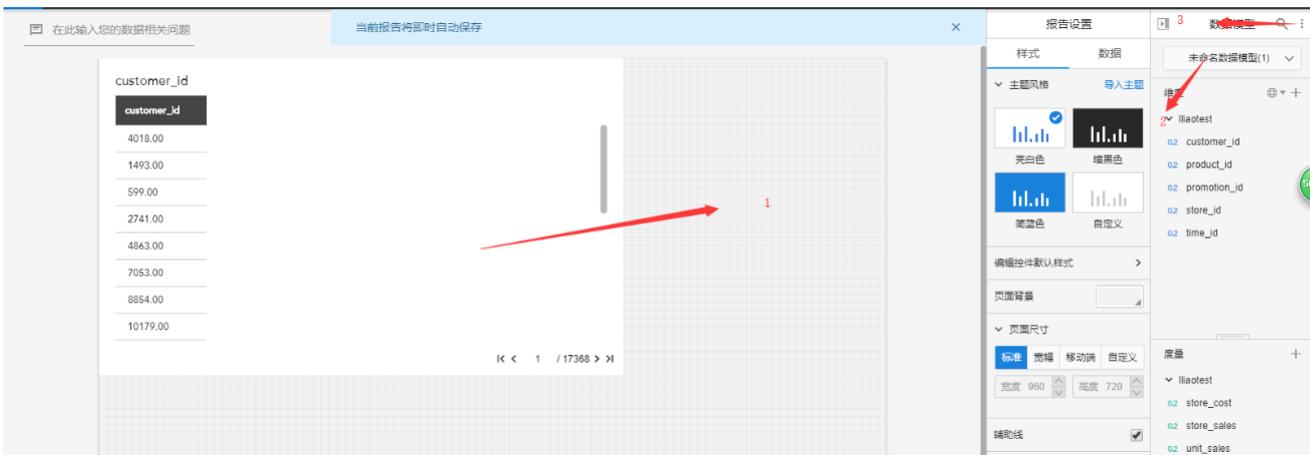
③数据模型层的排查: 编辑对应的数据模型, 数据模型分为直连和抽取, 直连情况下: 您使用的是源表还是自定义sql, 若为源表且查看无数据, 则为数据源问题, 请联系您的数据库管理员; 若为自定义sql, 您可以将该sql复制至源数据库执行, 若数据源也查无数据则请联系数据库管理员, 若数据源没问题, 则请联系技术支持人员;

若为抽取模式: 查看抽取记录, 是否有抽取失败, 若有失败请根据具体原因联系技术支持人员; 若抽取正常, 仍无数据, 请联系数据库管理员 (如果是全量抽取, 则可以尝试立即抽取, 有可能数据源的数据在你设置的抽取任务时间之后才入库, 那么抽取任务执行时是没有抽到数据的)

请注意, 请务必确认, AxisBI使用的数据连接信息(包括数据连接使用的用户、队列)和您源数据保持一致, 否则可能因为权限、环境等原因导致您无法得到准确的信息。

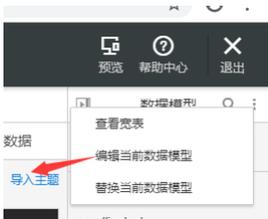
## AxisBI报表上, 怎么查看某块内容对应的抽取任务是什么

首先找到报表。进入编辑状态。



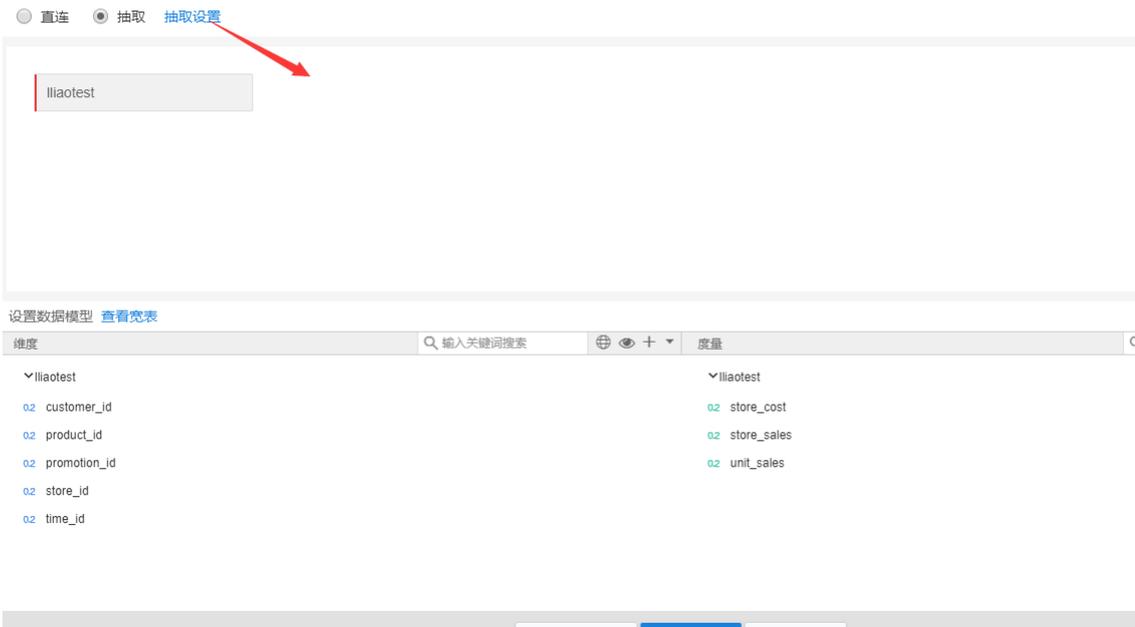
1. 选中报表的1部分。报表的右边会自动做映射。右边的2是该图表的数据模型。

2. 然后在3部分有 “三个点的图标”处点击, 找到“编辑当前数据模型”



3. 此时会跳转到对应的数据模型部分。查看抽取设置。

未命名数据模型(1)



4. 会弹出如下界面， 箭头所标就是抽取名称。

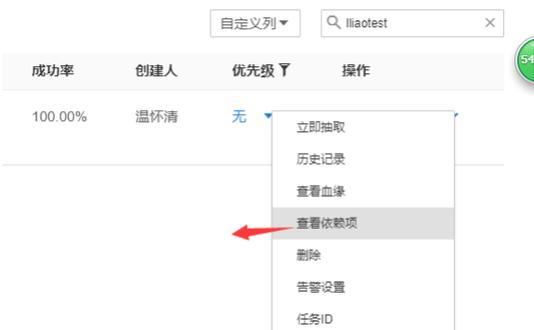


5. 可以在项目管理的抽取管理中查看。输入抽取名称，进行查找。



6. 小窍门: 如果知道抽取我们可以也快速知道数据模型和报表。

在上图“更多”的地方点击。会弹出如下界面。然后点击“查看依赖项”



7. 如下图所示：我们可以知道该抽取建了什么数据模型和报表。

数据模型名称/报告名称	创建者	创建时间
未命名数据模型(1)	王	2018-11-14 10:54
未命名报告(1)	王	2018-11-14 11:08

## 如何替换报表中的数据模型

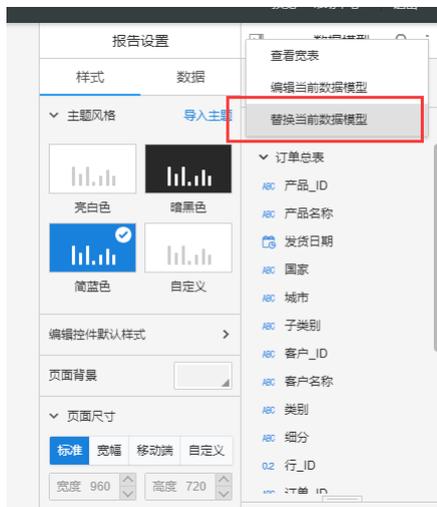
如何把报表中某个图表的数据模型替换？

最好的办法是替换数据模型，而不是直接去修改原始的模型。

替换数据模型的好处是：可以保留报表中的计算字段，并且原报告最大可能不出错。

1. 双击选中需要修改数据模型的图表

2. 点击更多，选择替换数据模型，选择目标模型，目标模型字段个数不得少于现有模型，且需要字段类型完全一致，最后确认后即可完成模型替换。匹配度低于100%进行替换的情况，仍然有可能会发生少量飘红的情况。





新版本不需要100%匹配。

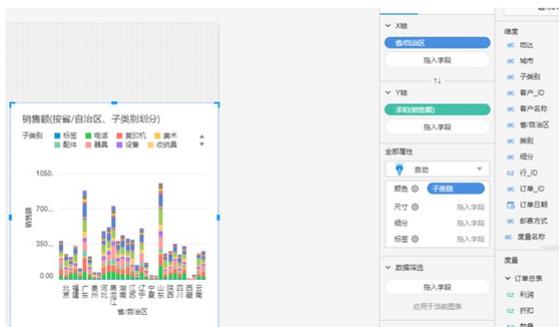
### 如何隐藏透视表的度量名称



选中后，右键会有一个隐藏名称的菜单。



### 如何设置图例的内容？



如何修改图例里面的内容？

图例也就是控制颜色的字段，每个图例都是字段内的一个值，我们只需要对字段的数据字典做一下调整，就可以改变图例的文字。

如果点击【颜色】边上的齿轮，则可以自定义图例的颜色。

### 报表迁移中遇到的问题事项？

在AxisBI同版本报表迁移中，步骤是报表从旧项目导出，然后在新项目导入。

如果报表的数据模型是抽取模型，在导入的时候，会要求切换数据模型。此时该怎么做?应该在新的项目里面新建一个和旧项目里面一样的数据模型（这里的一样指的是模型的字段一致即可）。然后在报表里面替换。接着导入。

报告的迁移方法，可以参考这篇文章：

<https://docs.ksyun.com/documents/38402?preview=1>

## 如何给二维表做自定义排序？

如何实现二维表的自定义排序呢？

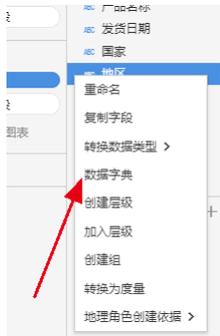
方法有两种：

①建立数据字典：比如下面这份数据，自然排序的结果是东北，西南，华东，西北，中南，华北这样的顺序，那么我现在想自定义排序，顺序是华北，西北，东北，西南，华东，中南(任意你要的顺序)，那么只需要以[地区]这个维度，建立数据字典

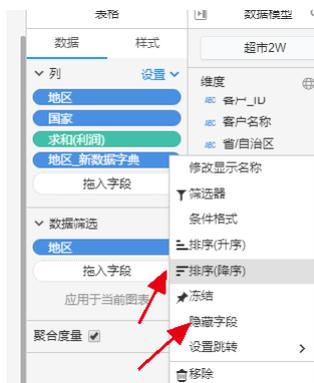
原始数据：

地区	国家	利润
东北	中国	1495439.69
西南	中国	585820.37
华东	中国	3643991.48
西北	中国	592819.42
中南	中国	4034966.84
华北	中国	2586319.30

编辑数据字典，形成新的字段



原名	别名
中南	6
西南	4
西北	2
华东	5
华北	1
东北	3



然后将该字段拖入二维表，按此排序，再隐藏该字段即可。

先建立一个计算字段，在该字段设置排序的规则。

然后拖入二维表中，选择升序和降序。

这时如果觉得这个计算字段不想显示在二维表中，可以选择隐藏。



由方法1, 我们还可以引申得到方法二:

建立一个计算字段, 其值为

```

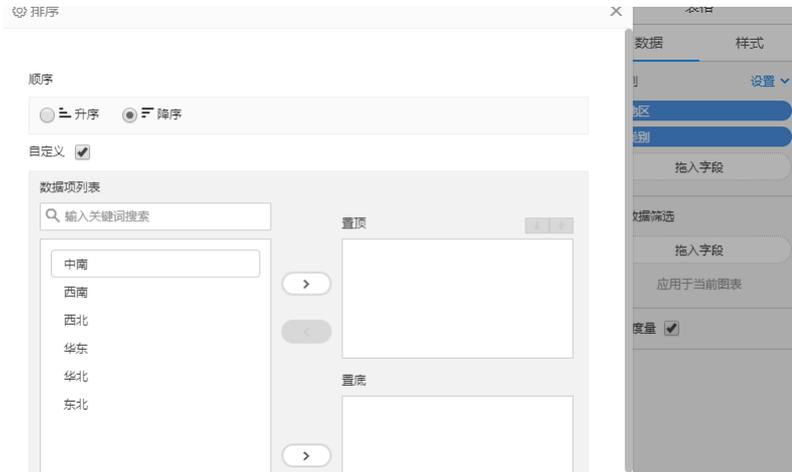
你 计算字段方法
式 插入维度 插入度量 插入参数
CASE [地区] WHEN "华北" THEN 1 WHEN "西北" THEN 2
ELSE "6"

```

然后同样的, 利用该字段排序并且隐藏该字段即可。

AxisBI已经开始支持自定义排序, 如下:

点击对应的维度, 选择排序, 勾选自定义



## 热力地图的介绍

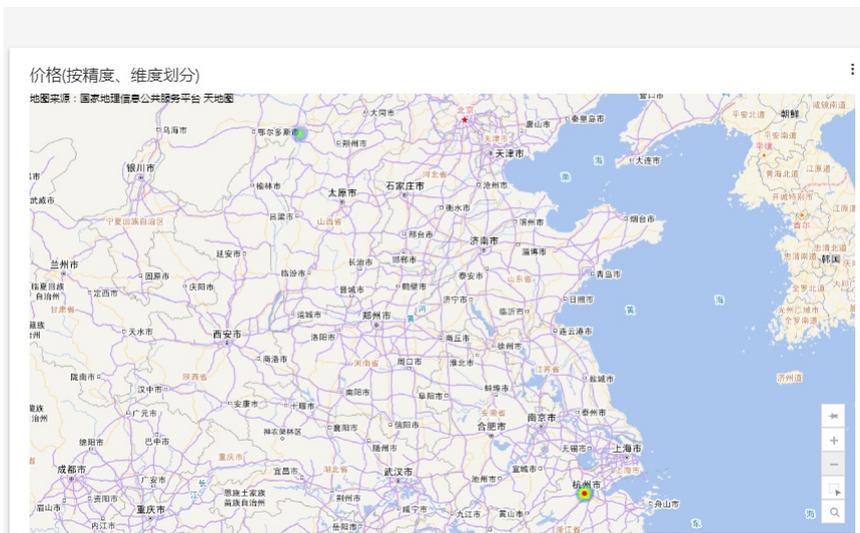
首先绘制地图就是先放一张很大很全的地理信息图片, 利用真实的地理经纬度和图片的像素位置产生关联。

热力地图的实现就是:

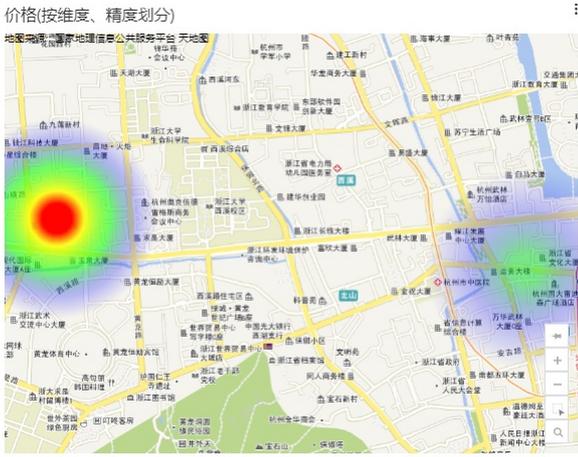
通过图表塞入经纬数据, 然后去调用地图的api, api会去建立这张地理信息图片像素位置和你提供的经纬度数据的关联, 然后把这像素点显示出来。

解析:

如果塞入一个蒙古市的一个小区经纬度, 然后在塞入浙江省的一个小区的经纬度, 你就发现图片是这样的



如果塞入两个相同城市的不同小区的经纬度，你就发现图片是这样的，几乎可以看到每个街道。



问题：两个图片的内容相差很大。他们两个是同一张图片吗？

答案是同一张图。

在预览地图的时候，图的右下角有个加号符号。你可以点击如果无限放大图一你也可以看到图二的信息。

这是因为屏幕大小是固定的，为了显示展示的数据，只是对图片进行了缩放。

## 双y轴如何设置刻度？

我们可以右击y轴中的字段，选择轴设置，然后自定义数值。（其中范围：勾选 包括零 。决定数值最小值是不是从0开始。）

设置数值(销售额) ✕

数值名称:

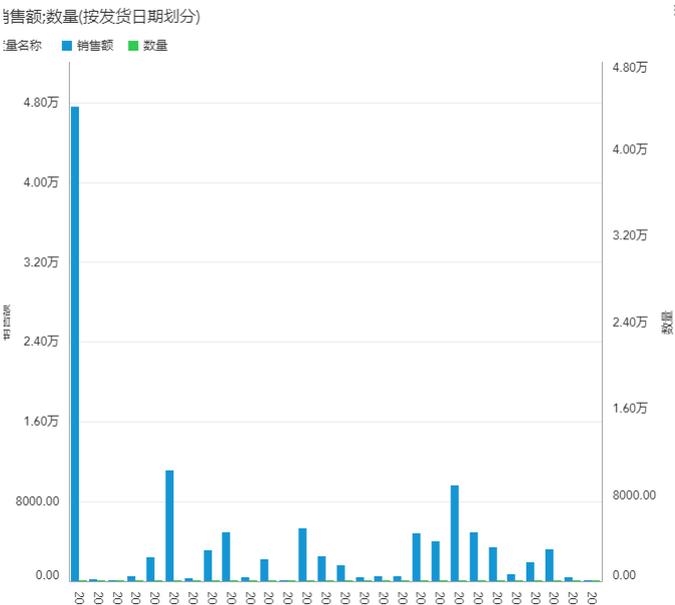
数值显示范围:  自动  统一轴范围  独立轴范围  自定义

-  (输入框留空则在数值上显示默认数值)

标签显示方向:  自动  水平  垂直

范围:  包括零

顺序:  倒序



数值设置，一定要两边都要自定义。否则效果就会如此图。刻度一样，但是两个值高度不一致。这个原因是：左边的数值是自动的，右边的数值是设置过的。自定的数值看似4.80万，但实际上不一样是真实值。

## 如何设置上下文筛选器

只有维度字段才可以设置上下文筛选。



上下文筛选器的作用是提高筛选器本身的优先级，各个筛选器优先级如下：

**优先级**  
(数字越大优先级越高)



注：图表内的维度筛选，默认优先级：列表筛选 = 文本筛选 > top筛选 = 条件筛选

### 字段重命名和数据字典的区别

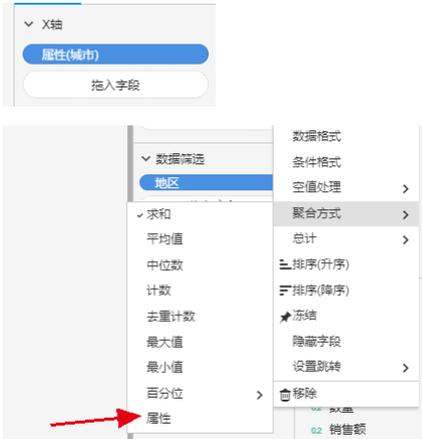
重命名：一般就是修改显示名称，针对的是字段名。

数据字典：是让字段里面内容重新映射。所以数据字典会新建字段。

所以最大的一个区别是：一个是在原有字段上的操作，一个会新建字段。

### 属性聚合方式是什么？

聚合方式是属性，只有所有值相同的时候才会返回当前的值，否则就返回null。



就像上面的城市，如果10条数据都是北京，那么就会返回北京，如果10条数据有一条不是北京，那么返回就是null。

PS: 属性聚合的算法是比较该列内的最大值与最小值，若最大值=最小值，则返回该值(差异最大的2个值都一样了那说明所有列的值都一样)，若最大值 != 最小值 则返回null

### 并列柱形图如何制作？

并列柱形图如何制作？

我做一个例子：查看各个地区，不同类别的销售额。

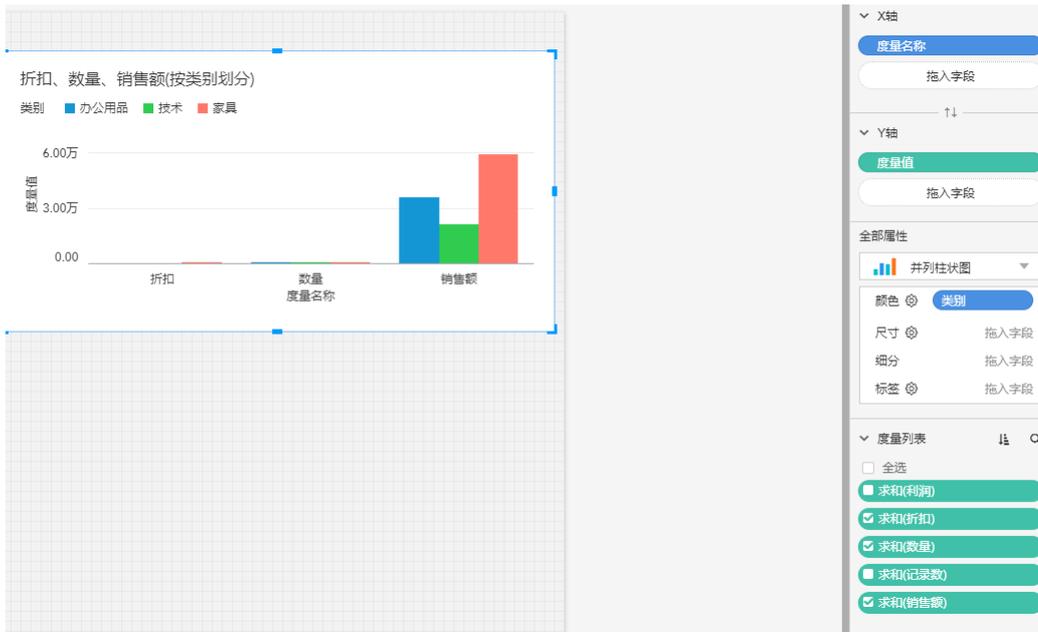
我们可以把地区放入X轴，销售额放入Y轴，颜色放入类别。然后选择并列柱状图就好了。



多度量怎么做并列柱形图?

首先x轴放度量名称, y轴放度量值, 颜色选类别。

我做一个例子, 在不同分类下面求销售额, 求利润。



从上图我们知道一般度量值和度量名称是成对使用的。

还有一种方法, 用双轴显示成并列的。这种方法只限于2个度量。(不推荐)

拓展:

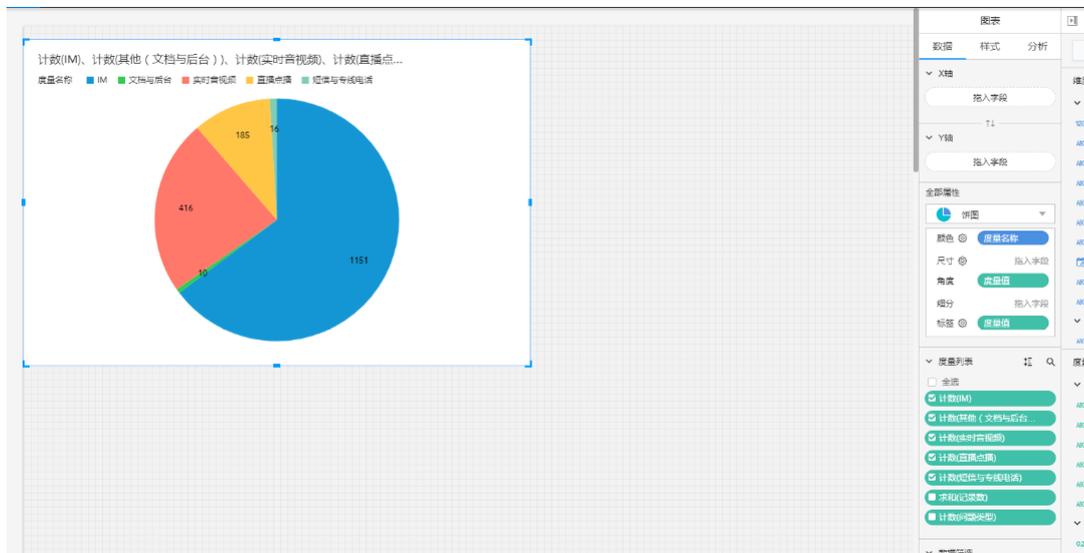
一般可以生成并列柱状图, 都可以生成堆叠柱状图。

### 度量值如何做饼图?

普通饼图实现就不做介绍了。

现在我们做个用度量值做饼图的方法。

以度量名称做颜色, 以度量值做尺寸, 标签拖入度量名称和度量值。但是度量值不能进行表计算, 所以不能以百分比显示。



如何对度量值设置百分比?

可以先创建度量字段计算百分比后，再作图。

具体实现：假设有ABCD四个度量：

需要创建 4个计算字段，

计算字段A: sum (A) /SUM(A+B+C+D)

计算字段B: sum (B) /SUM(A+B+C+D)

计算字段C: sum (C) /SUM(A+B+C+D)

计算字段D: sum (D) /SUM(A+B+C+D)

然后配置饼图，颜色放入：度量名称，角度放入：度量值。度量值只勾选对应的计算字段。

### 如何调表格的宽度？

第一种：选中表格后双击，会出现辅助线，拖动辅助线可以调整列宽。

第二种：也可以考虑在样式面板勾选“自动换行”，内容过长时会换行。

### 筛选器不起作用常见的原因有哪些？

一般而言，并不是筛选器“不起作用”，目前大致有这么几种情况：

- (1) 存在报告的页面筛选器，页面筛选器对全局产生影响，同时比较隐蔽。
- (2) 有筛选器被置于图层底层没有发现，但实际上已经发生了作用。
- (3) 筛选器没有勾选对应的报表关联。
- (4) 在做LOD的时候，没有将筛选器勾选“优先筛选”或者“上下文筛选”。

### 大屏如何排查数据问题

大屏出现组件异常：可能是数据原因或者服务器问题

下面介绍：数据问题排查方法（通用）

先复制一份大屏，在复制的大屏上操作。（非常重要，养成备份的好习惯）

步骤1：

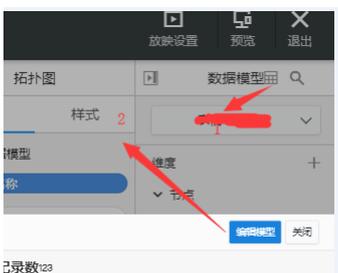
编辑大屏，看看大屏的组件是不是在编辑状态还存在。组件丢失有可能就是服务器问题。

步骤2：如果组件存在，选中出问题的组件。在选中状态的时候，在右边中间的组件配置区切换到”数据”，看看配置中是不是有红色字段，然后看看有没有筛选器。（红色字段主要是怕数据模型那边变化了引起了字段缺失，筛选器会影响数据）



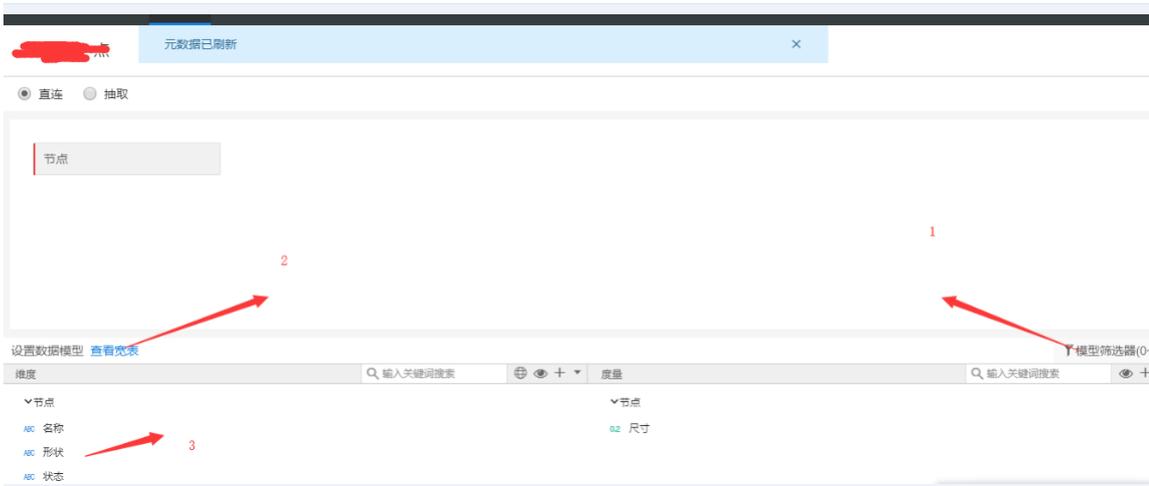
步骤3：查看数据模型，正常组件在选中状态。右上角的数据模型也会自动切换到对应的组件的数据模型。

此时点击下图中数据模型的“预览宽表”（1部分）会弹出宽表。如下图所示的2，点击“编辑模型”。



步骤4: 此时会进入数据模型的编辑界面。

- (1) 如下图所示: 先查看1部分, 有没有筛选器。(主要是排查自己的数据没有被过滤掉)
- (2) 然后查看2部分, 宽表是不是有数据。(主要是确保宽表是正常运行的)
- (3) 然后查看3部分, 对比一下大屏字段类型和数据模型是不是一致。(主要是看看大屏那边有没有做过字段的类型修改, 因为有时候类型修改会引入脏数据, 如原始字段是字符型, 在大屏变成了时间型。如果数据本身不符合时间规范在大屏就会报错。)



步骤5: 如果都没有问题, 在数据模型的非编辑状态下面。刷新一下数据。主要是更新源头最新的表结构。然后进入编辑状态。重复步骤4的检查。



步骤6: 如果都没有问题, 点击保存。选择覆盖。覆盖的作用是为了让数据模型和报表引用数据模型同步。然后进去大屏, 重复2步骤检查一下大屏。

步骤7: 如果在数据模型这部分有问题就追溯数据连接。如果都没有问题。至少数据部分是没有问题的。

该方法试排查数据问题的通用方法。同样适用于报表。

## 大屏数字翻盘器为什么会动

在大屏中有个刷新频率的概念。就是每隔几秒去数据库读数据。

而翻牌器出来的数据是一个汇总数据, 如果我刷新频率是5秒, 第一次数据库统计出来数据是50, 5秒后读取数据库统计库统计出来的数据是100。那么数据的变化是50逐渐变化到100的过程。到100就停止。

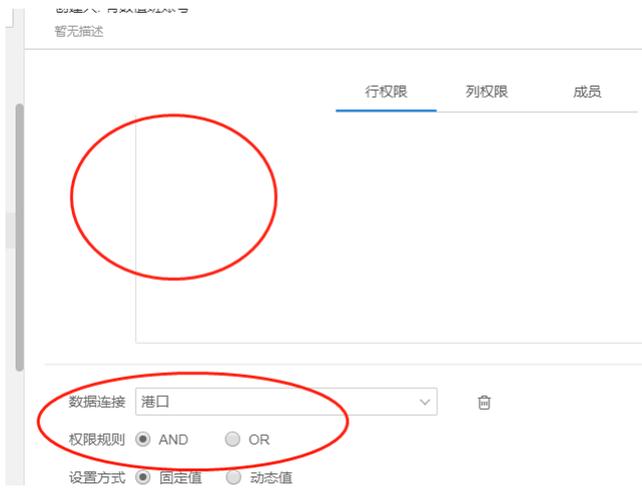
## AxisBI内数据行级权限的交并集关系解释

如下图所示, 行级与行级之间是并集关系 (OR)

而行权限字段与权限字段之间则是由权限规则控制 (默认选中AND)



另外一个行级权限内部, 权限块与权限块之间是 (OR) 的关系

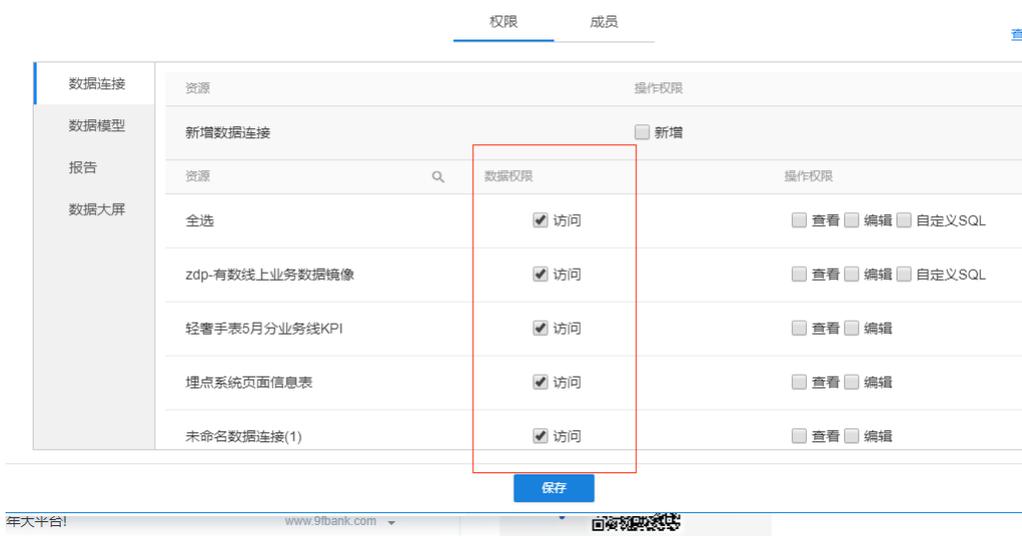


## 角色权限设置数据连接的注意点

数据连接的权限有4中：1. 访问；2查看；3编辑；4. 自定义sql

1. 访问权限 系统默认是勾选的，不建议去掉
2. 查看：让授权用户能不能看到这个数据连接
3. 编辑：可以对数据连接做编辑修改。
4. 自定义sql：开启后，可以在创建数据模型的时候，创建自定义sql。

我们在设置角色权限的时候，数据权限的“访问”，默认是勾选的，不要去掉。 如果不想用户看到特定的数据连接建议不授予查看权限。 如果去掉勾选，会提示无数据访问权限类型的错误。



## 对hive类型的行级权限设置字段和设置值慢，实现的新策略

之前说过行级权限的实质是在sql 上加条件（where）。具体可以参考：

<https://docs.ksyun.com/document/s/38418>

在实际中对hive设置行级权限，因为hive 本身对查询就比较慢，所以在行级权限设置字段和设置值的时候，可能几乎读不出来。（读字段会去查询表结构，列表读值会去读表中内容，都涉及到sql 查询）

面对这种现象我们该如何处理？

建议添加一个维度表。

如之前讲的例子超市表现在是存在hive 中。此时设置行级权限响应就很慢。如果我们相对不同的人设置不同的地区权限。

我们需要创建一个维度表，这个表可以是excel 或者关系型数据库。

现在维度是地区，那么我们只要创建地区表就好了。我们只要对该维度表设置权限即可。

然后在数据模型中，原始模型和该表进行内连接。

1. 最开始设置行级权限：

```
select * from 超市表
where 地区=xxx
```

2. 现在加了纬度表，行级权限设置就成了：

```
select * from 超市表 , 地区表
where 超市表.地区=地区表.地区
```

```
and 地区表.地区=XXX
```

因为地区表是excel 或关系型数据库，查询很快，所以设置行级权限比较好。

还有一种方法：创建一个地区的维表设置行级权限。（和上面相同）

然后做报表层把行级权限设置的表放入列表筛选器，最后利用筛选器的跨模型关联功能，和其他数据模型设置跨视图关联。

但是使用这个方法需要几个前提：①这个筛选器需要默认选择全部来影响 ②需要隐藏这个筛选器，以免通过清空选项来导致行级过滤无效

## 行级权限实质

行级权限实质是往sql中加一个where 条件。

行级权限的设置操作如下：

该链接是针对超市表（包含地区字段）为例子。

<https://docs.ksyun.com/documents/38055>

在上述行级权限设置的案例中，我们针对超市表。

1. 如果我们在报表看全部数据。

sql 如下：select \* from 超市表

2. 如果特定的人看特定数据，需要用行级权限。

sql 如下：select \* from 超市表 WHERE 地区=xxx

在使用中我们需要注意：

1. 行级权限设置只能针对表（数据模型左侧中库出现的表）和自定义sql。

2. 如果有1个用户分别授予两个行级权限的逻辑是or的关系。

如果一个行级权限可以看东北数据，一个行级权限是华北数据，那么两个行级权限同时作用，可以看到的是合集，这个用户可以看到东北和华北的数据。

## 项目管理员权限的注意点

在项目里面，项目管理员（角色）的权限是最大的。

1. 项目管理员可以查看编辑任何报表。

只有私人文件夹下的报表，需要私人提供报表路径。

此时管理员只有查看的权限，并不能编辑。

2. 管理员目前会受行级权限影响。

3. 超管不受行级权限影响。

## 定时邮件发送的方式

定时邮件目前是以图片形式发送的。

收到邮件，我们会以超连接的形式跳转到报表。

跳转时需要用户权限验证。

## 定时邮件是否支持邮件组？

定时邮件是否支持邮件组？

目前不支持。

## 当碰到定时邮件发送失败，失败原因是访问数据超时，怎样处理？

该失败原因是：报告数据加载太慢，超时了。可能是数据量太大，可能是数据库性能慢，也可能是其他原因。

解决方案1：该定时邮件失败是偶发，可以进入报告，等待报告绘制成功，手动重新发送邮件

解决方案2：如果失败是常见。可以查看是否数据量太大，导致加载慢。如果可以，可以通过筛选等方式减少数据。如果是数据库性能较差，可以使用AxisBI的抽取模式。通过AxisBI的MPP集群提升性能。

解决方案3：如果失败是常见。比如定时邮件设置在9:00发送。可以对该报告设置刷新计划，设置为8:00。注意在这种情况下，定时邮件需要选择缓存数据。另外如果定时邮件设置在8:00/9:00等热门时间点，建议改一下，比如九点整的改为八点五十五或者九点零四，越冷僻的时间点越不容易和其他调度任务重叠、竞争

解决方案4：如果底层字段发生变化。通过报告找到对应的数据模型。点击覆盖应用到所有报表。

问题举例：

失败原因：访问数据超时，定时邮件相关报表未绘制成功

解决方案：请进入报告，等待报告绘制成功后，手动重新发送邮件

## 怎么样提问？怎么样快速得到有效的答案？

问问题？谁不会呀？这看似是一个简单的事情。

举一个例子：今天有一个用户提了一个问题：

发现模块一直报未知异常是什么情况啊

我看到这个问题的第一印象是：我无法回答。

任何的问题都需要足够的信息才能诊断，医生看病也需要望闻问切，只是打电话跟医生说：我肚子疼。

医生能直接开药吗？当然不能，必须要了解到具体的病情才可以下诊断。

不懂得如何提问，浪费的不仅是技术支持人员的时间，更重要的是浪费自己的时间。

要你将想要做到一件什么事描述清楚该怎么做：

1.1 问题标题应涵盖主要关键字

1.2 需求类的问题描述：现在情况是什么样子的，你想使用什么方法，最后达成什么结果。

故障类的问题描述：你做了什么操作，导致了什么样的结果，错误的情况如何。

1.3 仅文字描述不足以清晰说明问题或者你认为有必要时，最好配上相应的截图。

1.4 用词准确，切勿前后矛盾

1.5 表意清晰，别自己都看不懂自己在说啥，想象自己是解决问题的人，能否看懂自己的问题

1.6 有错误代码的、错误信息、日志、公式的请尽量贴出来

仍然以文章开头的问题作为例子，至少要描述清楚：什么模块，报了什么错误，做了什么导致这个错误，最后加上截图。

提问是一种智慧，如果提问前能够三思而后行，那能大大节约彼此的时间，最终受益的还是提问者自己。

## 域主注册后怎么登录AxisBI?

域主注册成功后。

直接点击直接登录域，然后填写登录账号和密码。（密码为注册时填写的密码）

## 报告已经删除了，再去删除模型的时候提示：有报表或大屏依赖，不能删除数据模型

这个是因为删除报告以后，默认报告会进入回收站。此报告物理上仍然存在，只是只在回收站内显示。

所以此时仍然无法删除数据模型，会提示依赖关系。

进入项目中心—回收站以后，再次清空此报告，此时报告就真正删除了，数据模型的依赖关系也就解除了。



## 数据填报如何把表落地到自己数据库

数据填报如果默认落地到AxisBI，会存入AxisBI的系统库中，所有的填报只有两张表，

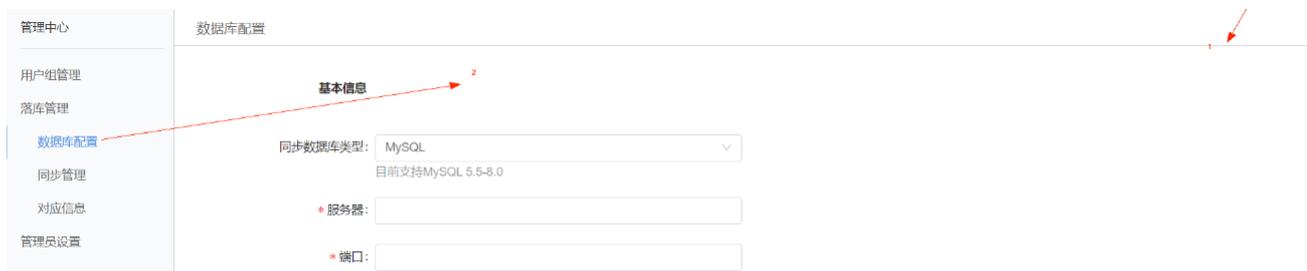
一张表用来存填报元信息，一张表存填报数据。

如果对数据填报数据做ETL设置。

我们可以把数据填报的表落地到自己数据库中（目前支持mysql）。

首先进入填报界面，点击右上角的齿轮图标，“设置”

会弹出如下界面：



填写完成后，所有的数据填报，会存入指定的mysql数据库中。

此时一个表单就对应着指定数据库的一个表。

注意点：

填报名不等于表名。但是我们可以根据填报的id找到对应的表。

填报字段名不等于表字段名，因为在一个填报中可能会有同名字段出现，所以我们会把填报名存入数据库的备注中。

## 如何删除域中用户

删除用户的条件首先你是域主或者域管理员。

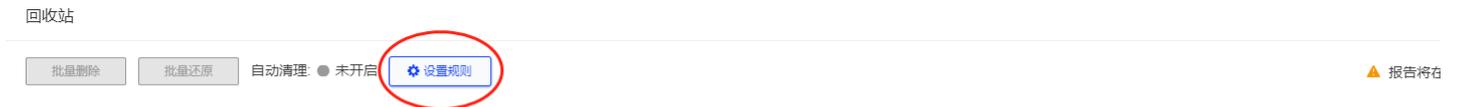
如何你是以上身份点击“平台管理”->“人员信息”会弹出如下界面，可以勾选用户批量删除也可以单个删除。



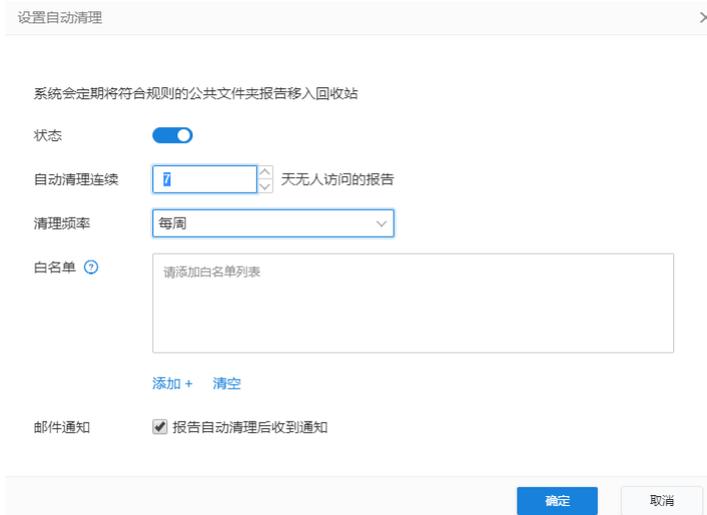
## 回收站功能

回收站功能支持自定义规则自动回收无用报告。

点击 项目中心--回收站



可以看到回收站保留时长是90天，90天之后将永久删除



通过设置自动清理，我们可以将长期不访问的报告自动转入回收站。

该新功能若需要使用，必须在首次打开回收站—设置自动清理，点击“确定”按钮才会生效。

对于一些低频访问的报告，我们可以将其设置白名单以免“误删”。

## 登录系统后，提示用户不存在？

错误的原因有可能：

1. 尚未加入该域，请联系域管理员，添加进入域，添加方法参考

<https://docs.ksyun.com/documents/38296>

2. 还有一种是加过域了，但是登录仍然如此提示，这种情况往往是域名输错了。

## 个人用户搭建一套AxisBI试用学习环境，以MySQL为例

个人用户搭建一套AxisBI试用学习环境，以MySQL为例

硬件需求：windowPC一台

网络：连接公网

软件需求：MySQL；花生壳；

重要申明：以下教程仅供个人学习使用，商业使用或者公司使用请咨询贵司的DBA或者信息技术人员指导下，进行相关的数据库安全/网络安全设置，本教程不承担因个人设置不当造成的数据泄露责任。

### 第一步

1. 安装MySQL，安装过程不再赘述，官网下载下来，一路默认安装即可。

2. MySQL进入root用户，对root用户进行远程授权访问

```
mysql> use mysql;
```

```
Database changed
```

```
mysql> update user set host = '%' where user = 'root';
```

```
Query OK, 1 row affected
```

```
mysql> commit;
```

```
Query OK, 0 rows affected
```

务必注意，商业使用的情况下，一般不会开放root，需要进行单独的白名单设置。这里只是为了举例方便，商业使用一定请咨询DBA或者系统管理员进行正确的设置。

## 第二步

1. 下载花生壳软件
2. 设置内网穿透

### 内网映射

- 1、用于**微信推广**的映射，请使用已经备案的顶级域名（推荐：**花生壳铂金版**，送顶级域名（含）
- 2、若顶级域名需使用内网穿透功能，**NS管理**必须选择“Oray DNS解析”，48小时后生效即可使用

应用名称:

选择域名:

映射类型: 应用类

外网端口: 动态端口号

内网主机:

内网端口:

一般来说，固定端口比动态端口要好（但是要贵）。映射后生成的域名是

MySQL 编辑 诊断

外网访问: :18974

内网主机: 127.0.0.1:3306

映射带宽: 1M 已用293.92KB流量

域名: xx.imwork.net

端口: 18974

## 第三步

1. 申请一个试用账号。
2. 登录AxisBI—数据源—数据连接—新建连接，选择数据源类型MySQL，依次输入我们的设置参数：数据源名称由你自己定义，服务器填写的是花生壳的域名xx.imwork.net，端口填18974，用户填写root，填入root用户密码，点击保存，即可使用。

MySQL(类型: MySQL)

基本信息	表信息	相关内容	操作记录
<p>基本设置 <input type="button" value="导入数据源"/></p> <p>*数据源名称: <input type="text" value="MySQL"/></p> <p>*服务器: <input type="text" value=""/></p> <p>*端口: <input type="text" value="18974"/></p> <p>*用户名: <input type="text" value="root"/></p> <p>*密码: <input type="password" value=""/></p> <p><input type="button" value="连接"/></p> <p>缓存有效期: 1天 <input type="button" value="▼"/> 缓存失效点: 00 时 00 分</p> <p><input type="button" value="手动刷新数据"/> <span style="color: orange;">⚠ 点击刷新数据连接下所有报告的数据</span></p>			

将AxisBI的报告嵌入网页后，为什么报告中的图表无法加载到内容？

请根据按以下步骤确认问题：

1. 确认是否只有在chrome浏览器中，您的网页无法加载到内嵌的AxisBI的图表内容。

备注：从Chrome 51开始，浏览器的Cookie新增了一个SameSite属性，用来防止CSRF攻击和用户追踪。如果iframe的地址和父级的地址不同源，则可能无法传递cookie

2. 如果只是chrome无法打开，请参考以下步骤：

在打开chrome后，输入 chrome://flags/#same-site-by-default-cookies，

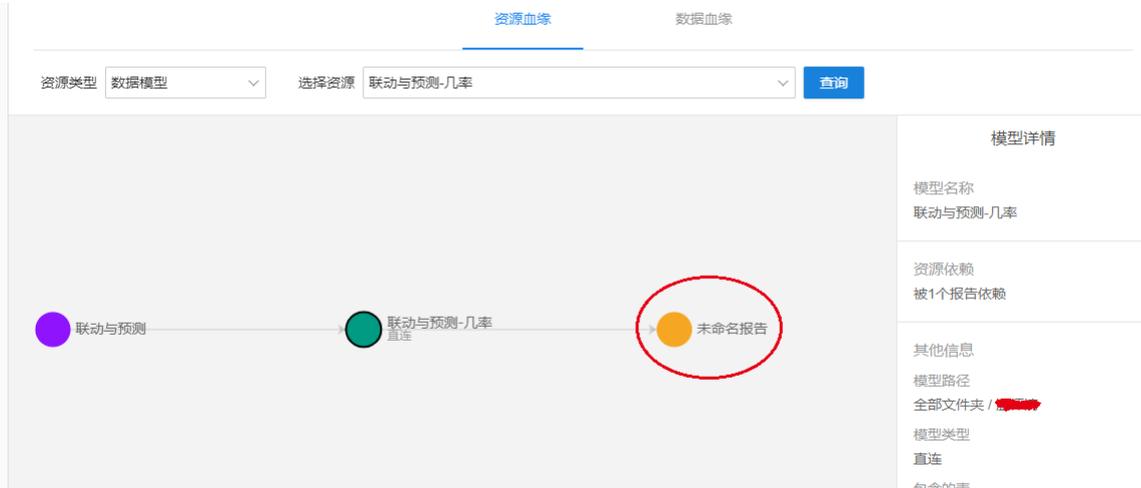
将【SameSite by default cookies】的default选项改为disable后，清除浏览器缓存，再打开您的网页。

3. 如果所有浏览器都无法加载嵌入的AxisBI报告内容，请在工单系统联系技术支持人员。

### 通过血缘分析来查询资源和数据的依赖情况

场景1. 已知数据模型，需要删除改模型，但是提示有资源依赖无法删除：

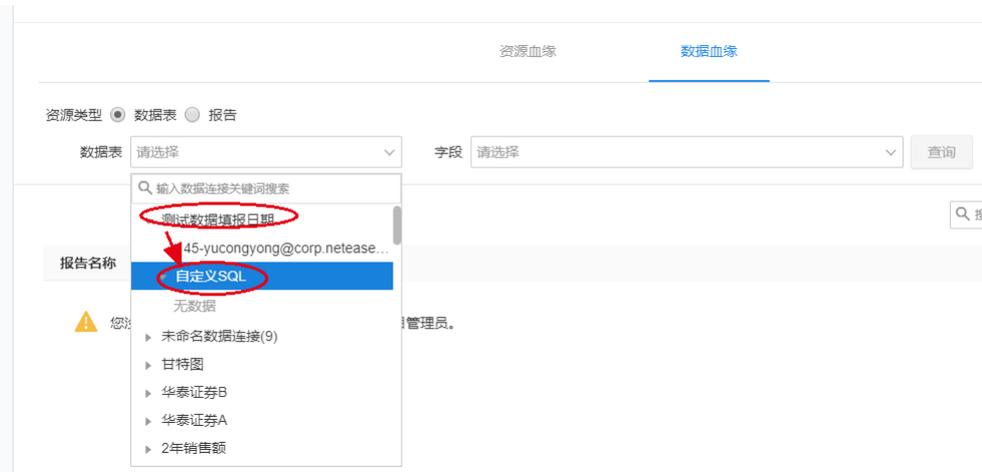
依次进入项目中心--血缘管理--资源血缘--选择数据模型，输入模型名称进行查找，可以看到有许多的报告依赖此模型，我们需要删除对应的报告或者[解除对应的报告-模型关系](#)。



同理，也可以依此法删除数据连接和抽取

场景2. 已知某个自定义SQL，需要删除此SQL，但是提示有资源依赖无法删除：

①依次进入项目中心--血缘管理--数据血缘--选择数据表--选择你要查找的数据连接--自定义SQL--输入自定义sql名称，点击查询。可以看到有许多的报告依赖



②自定义SQL的依赖关系还可以在数据源--数据连接--表信息--自定义表内看到其依赖的抽取，根据抽取表的依赖关系可以进一步推断其依赖关系。

查看抽取



查看抽取表的依赖



## 报告的导出导入，报告的迁移？

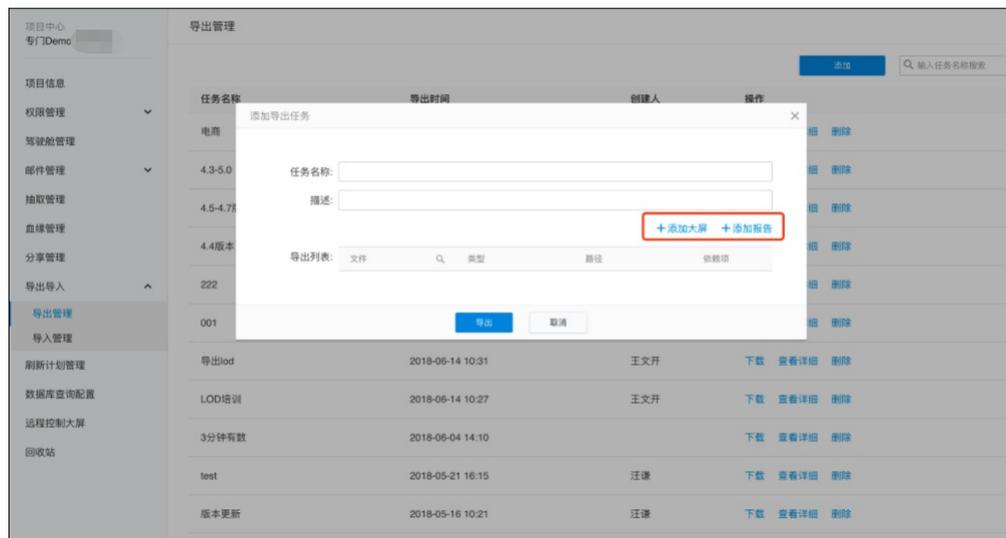
### 报告的导出导入

当需要跨项目复制/移动报告，或者需要将测试环境中做好的报告迁移到生产环境中时，我们可以将需要的报告导出成离线文件，然后将该离线文件导入至目标项目。

第一步：进入项目中心的“导出导入”模块



第二步：创建一个导出任务，勾选需要导出的报告，导出成离线文件



第三步：在需要导入报告的项目中，导入之前得到的离线文件，完成报告导入

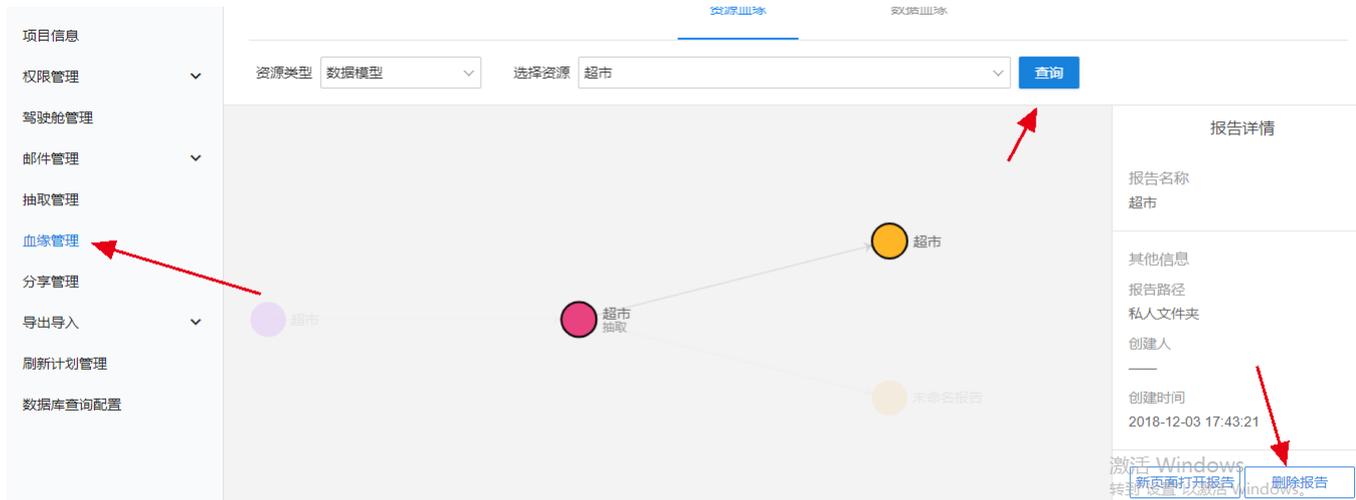


需要注意的是，导出报告时，会同时导出报告所依赖的数据模型跟数据连接。导入时，我们可以选择是否要导入报告依赖的数据模型跟数据连接，如果不导入，则需要替报告指定在新项目中所要连接的数据模型。

## 如果一个用户被删除了，那么如何删除他私人文件夹下的报告

删除一个用户后，我们可能忘了让该用户删除其私人文件夹，这些文件夹下的报告可能会占用一部分空间，同时也会影响抽取的依赖关系。

那么此时项目管理员可以使用项目中心—血缘管理，通过这个可以删除私人文件夹下的报告



## Impala/Hive数据源：创建自定义sql报错

impala数据连接下，创建自定义sql报错，提示类似：

```
org.apache.hive.service.cli.HiveSQLException: Create execute request failed:
```

一般来说自定义sql在数据源那边已经验证过语法，但是在AxisBI自定义sql的时候仍然报了错误，这个问题原因有如下几种可能性：

1. 字段别名重复
2. sql末尾含有注释(目前AxisBI识别不了最末尾的注释语句，请去掉或者写在SQL中间)
3. 如果SQL内有UNION关键字，则union的上下文的字段名称必须完全一致

## 数据刷新管理的三种模式

刷新计划管理分出了3种方式。

分为是数据连接，数据模型和报表。

1. 预加载是数据连接，我们会加载这个数据连接上面用到的所有报表。并非连接下的所有表。
2. 预加载是数据模型，我们会加载这个数据模型上面用到的所有的报表。
3. 预加载是报表则是制定报表进行预加载。

所以刷新计划管理最后的对象报表。

还有的刷新优先级怎么理解？



如果同一时间有很多并发刷新，目前AxisBI最大并发数是8个。

设置了高优先级会先刷新。

## 浏览器的支持情况

AxisBI在谷歌浏览器展示最佳，也可以使用搜狗浏览器等。在其他的浏览器上，可能存在样式偏差等情况（这是因为浏览器采取了不同的标准，可能使得前端渲染存在几个px的误差）

## 用户删除了，私人文件下面的报告会不会消失？

用户删除了，无法还原。需要管理员重新添加。

但是私人文件下面的报告不会消失，4.8以后的版本可以在项目中心—血缘管理内找到私人文件夹下的报告并删除之。

操作方法：<https://docs.ksyun.com/documents/38433?preview=1>

## 推荐：如何提交工单？

登录金山云控制台：<https://passport.ksyun.com/>。

点击页面最下方账户管理—提交工单，即可提交工单。

工单页面如下：



选择对应产品，填写相应问题后，即可创建并提交工单。

## 数据导出的极限是多少？

SaaS版本数据导出的极限是10万条，这个条数受到数据字段数量的影响，也就是说你的字段如果越少，导出的行数越接近10w，其余条数规律约等于200万/N列，也就是说一般20列以下，导出数量约等于10w，超过20列逐渐增加。

不过这是粗略的算法，便于理解。

如果是私有部署版本，可以根据自身服务器的性能和使用频率上调该极限。

那如果要导出超过10万的数据怎么做呢？

可以采用分批多次导出，也就是说可以设置筛选器，比如分为A/B两部分，先筛A，导出一部分，再筛B，再导出和之前合并即可。

强调一下，AxisBI作为BI工具，不是数据导出工具，所以不会提供无限导出的功能。如果需要做数据分享，可以用导出截图、AxisBI分享链接都可以。

不同的部署环境，导出的极限设置也不一样。有可能是2万条。域主的设置有关系。平台最大设置是10万条。

## 平台报503错误是什么原因引起的？

这种错误一般都是系统服务引起的。我们需要检查一下docker 里面所有服务是不是开启状态。

一般都是web服务端口挂掉所致。