目录

[1. 节点设置页面 2](#_Toc422493171)

[2. 节点外观页签 4](#_Toc422493172)

[3. 节点基本信息页签 5](#_Toc422493173)

[4. 注意事项 5](#_Toc422493174)

**分布图节点使用说明**

**作者：王方舟**

分布图是Data Studio可视化节点分类中的一个节点，用来展示某一分组列在若干聚合列上分布情况的可视化结果，节点首先将选择的分组属性列的数据按照分组个数进行区间划分，再按照选择的聚合属性计算聚合值（平均值、合计值、计数值等），每一聚合属性使用一种颜色的柱状图作为其可视化展现，以图例说明每一颜色所代表的聚合属性。

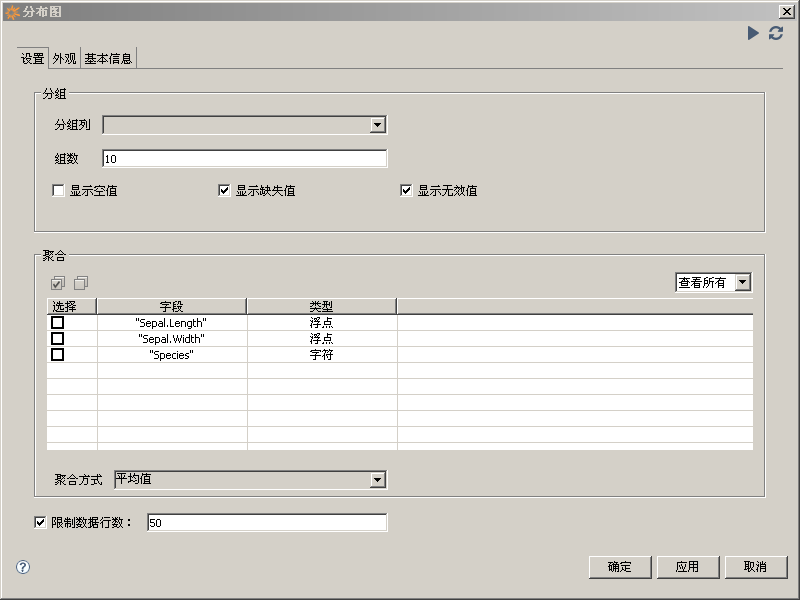
其中节点的分组属性支持数值属性及字符属性，数值属性按照区间个数进行等分计算分组区间，而字符属性则根据其值的个数作为分组的个数。图1为分布图节点样式。

节点icon上方会有节点状态的指示灯，其中红灯代表节点未经设置不能执行，黄灯代表节点已经准备好执行，绿灯代表节点已经执行完成。节点下方的是快捷执行设置按钮，其中点击执行按钮会激活节点的执行行为，设置按钮会弹出如图2所示的设置界面。



图表 1 分布图节点

# 节点设置页面



图表 2 分布图节点浏览界面设置页签

如图2所示，分布图节点设置项大体分为两组，第一组为分组属性，该下拉列表将显示数据节点传过来数据表的数值及字符属性。在下拉列表内选择数值属性后，在组数文本框内键入分组数字；选择字符属性后，组数文本框将置灰，表示不需要输入组数，分组数将按照字符属性值的个数而定。另外选择分组属性后需要选择是否显示空值、显示缺失值和显示无效值，它们具体的意义如下所示：

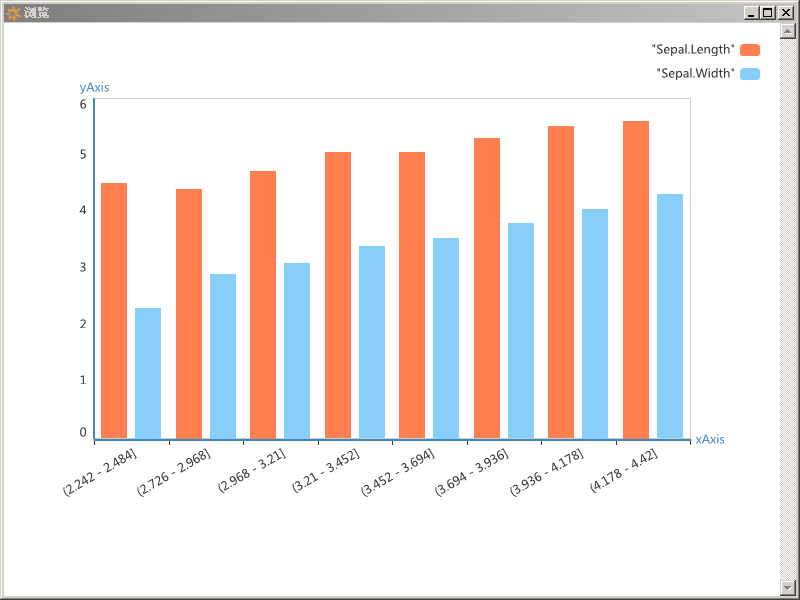
1. 显示空值： 在计算聚合属性在分组属性值区间的分布情况时，会出现某一个区间或某几个区间聚合值为0的情况，如果勾选显示空值则这些区间将在分布图上显示，否则这些区间将不会出现在分布图上；
2. 显示缺失值：缺失值在程序中体现为一个MissingCell，即数据源中该列属性的某一行或几行的数据单元值缺失，若勾选此选项，分组属性中包含缺失值的行将单独作为一列分组生成分布图；
3. 显示无效值：无效值为数值型的Double.NaN（0.0d/0.0）、Double.POSITIVE\_INFINITY（1.0/0.0）、Double.NEGATIVE\_INFINITY（-1.0/0.0），若分组属性包含这种类型的值时，勾选显示无效值将会将这一类的无效值单独作为一列分组生成分布图，否则将忽略这些数据行。

在聚合设置组下，是一个输入数据表模式的table，勾选某几行成为分布图聚合属性并在聚合方式下拉框选择计算的方式（平均值、合计值、计数值、计数值含缺失值），各个聚合方式的计算方式如下：

1. 平均值：合计值除以计数值；
2. 合计值：该区间下聚合属性值的sum值；
3. 计数值：该区间下数据行的个数，忽略MissingCell；
4. 计数值含缺失值：该区间下数据行个数，计算MissingCell。

另外，分布图默认限制展示数据行数为50，若想要展示自定义行数的数据，则在文本框内输入数字，若想将输入数据全部显示，则取消勾选限制行数的复选按钮。

将以上信息都设置完成后，点击右上角的执行按钮，将弹出分布图的对话框，展示计算后的可视化图表如图3所示。



图表 3 分布图节点执行后预览

该可视化展现对话在执行分布图右击节点并选择浏览时将会再次弹出，如图4所示。

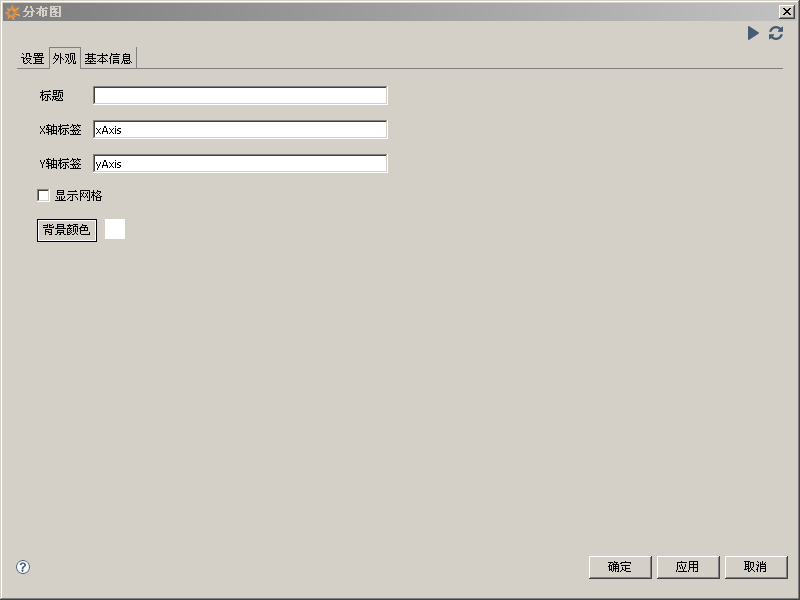


图表 4 可视化节点右键菜单

设置完参数之后点击确定将会保存这次的设置并关闭浏览对话，点击应用只保存设置而不关闭对话，点击取消则本次设置随着对话关闭而丢失。

若在修改设置的过程中想要重置到上次保存的设置结果点击右上角的重置按钮即可实现。

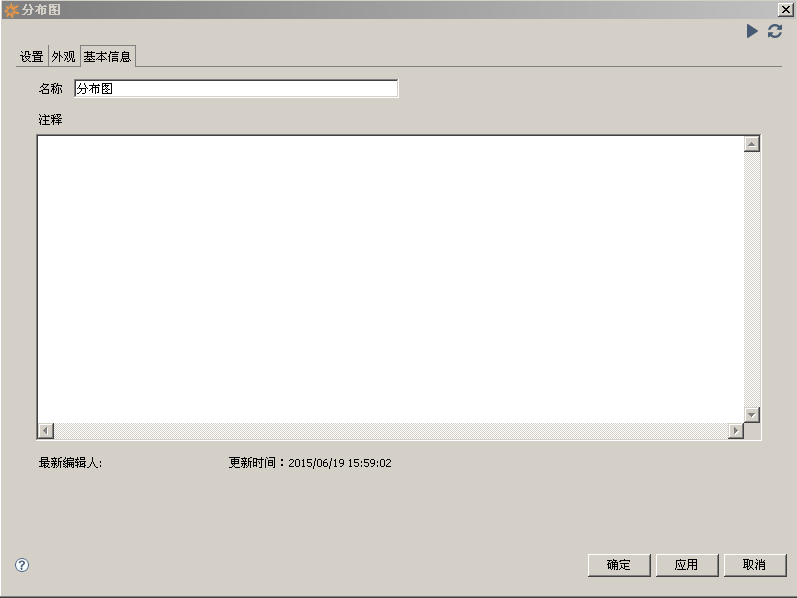
# 节点外观页签



图表 5 分布图节点外观标签

外观标签中的内容主要用来自定义图标的展现，包括图标的标题、X轴、Y轴的标题以及图中网格的显示和背景色的显示。

# 节点基本信息页签



图表 6 分布图节点基本信息页签

分布图的基本信息涵盖一下内容：

1. 节点名称：在文本框内键入新的节点名称将在确定之后修改节点的名字；
2. 注释：节点的备注信息，方便用户将对该节点的用途进行描述和记录，方便下次打开节点时知道节点的自定义信息；
3. 最新编辑人：最后编辑该节点的用户；
4. 更新时间：最后一次对该节点进行编辑的时间。

# 注意事项

1. 在修改分布图节点参数之后不点击右上角执行按钮直接确认时，将弹出如图7的是否确认重置节点的对话；
2. 分布图节点在已设置状态时，右键点击浏览会直接激活节点执行的事件，并展示执行的结果；
3. 由于可视化图表调用Echart，将数据过多时将有可能弹出浏览器警告，因此在用户输入展示数据超过1000行时，将弹出警告；
4. 分布图默认的设置为不显示空值、显示缺失值和显示无效值；
5. 分布图中tooltip提示的格式为：聚合方式 聚合属性1 = 聚合属性值1 …。



图表 7 重置节点警告