**功能：**
与分组小提琴图类似，把两组合并为一个小提琴图展示，便于直观比较两组之间的数据分布状态，适合展示多因素数据分布。

**输入：**
1. 文件必须是带有表头的以tab分隔的.txt格式文件，表格名称以字母、数字、下划线组成，不可以有后缀名称。
2. 文件第一行“a/b/c/d”为X轴参数，可以为不同样本、不同基因、不同cluster、不同时间梯度等；第一列为Y轴参数属性，可以为基因ID、细胞barcode、样本等，对应内容（数值）可以为基因表达量、细胞UMI、肿瘤体积等。
3. 两组文件中，每组文件的**列数**和**对应列的名称**需保持一致。

示例文件如下：

Group1 Group2

 

**参数：**

1. 输入文件，依次输入两个文件，文件1对应分组1，文件2对应分组2，可自定义分组名称及颜色，如未选择则为系统默认名称及颜色。

2. X轴标题：自定义

3. Y轴标题：自定义

4. 图形标题：自定义

5. 散点：展示或不展示

**输出：**
1. 图：程序将输出PNG及PDF格式的豆荚图。

1. 表：程序将输出包含每组每样本平均值、最小值、最大值、上下四分位和中位数信息的.xls格式的表格。

**结果说明：**

1. 图形解读如下：


①图中不同颜色代表不同组分，每个点代表一个数据；

②黑色矩形中的白点代表中位数Q2，组间白点相对位置越高，说明该组数据整体分布越靠上，对应生物学意义即为基因的表达量越高/检测到的细胞UMI数越多/肿瘤体积越大等；
③黑色矩形长度代表这组数据中上下四分位数差值（IQR），矩形越长说明数据越分散，越短说明越集中；矩形上边缘为上四分位数Q3；下边缘为下四分位数Q1；
④豆荚图的轮廓代表数据分布的主要区间，轮廓以外的数据为离散数据；轮廓的纵轴两端分别代表最小值和最大值，纵轴方向长度越长代表数据弥散程度越大；横轴方向长度代表数据分布密度，越长说明分布在对应纵坐标位置的数据越多。

1. 表格解读如下：



表头属性：第一列对应豆荚图X轴内容；第二列为分组信息；第三列往后依次为平均值（mean）、最小值（min）、下四分位数（Q1）、中位数（Q2）、上四分位数（Q3）、最大值（max）、上下四分位间距（IRQ）。