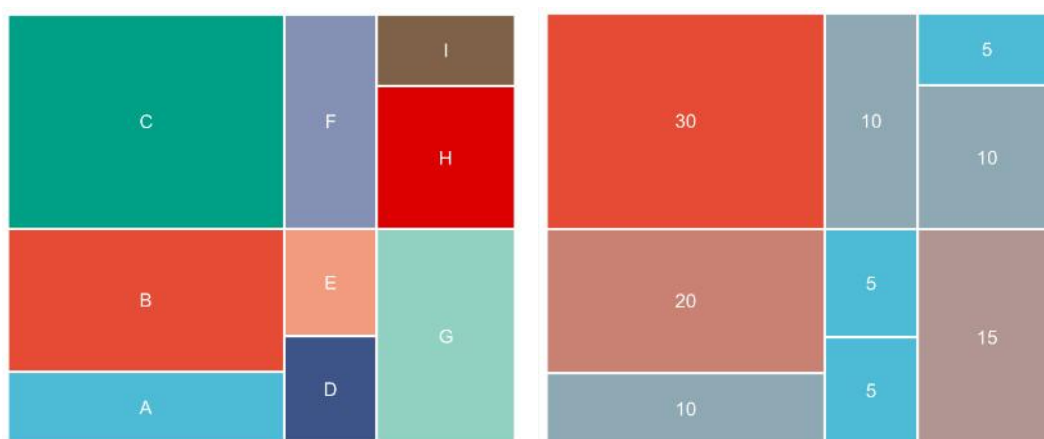


基础绘图 - [局部整体] - 堆砌方块图



网址: <https://www.xiantao love>



更新时间: 2023.07.14

目录

基本概念	3
应用场景	3
分析过程	3
结果解读	5
数据格式	6
参数说明	7
映射	7
方块	8
标注	10
分面	12
坐标轴	13
标题文本	14
图注 (Legend)	15
风格	16
图片	17
结果说明	18
主要结果	18
方法学	19
如何引用	20
常见问题	21

基本概念

- 堆砌方块图：又名矩形树状图，是一种用于可视化层次数据的图表，通过矩形区域的大小和颜色来表示不同层次的数据关系。按照数据总体为 1 去计算每个变量的百分比，并将结果以方块图的形式展现来描述数据的大小即占比情况，变量占总体数据的百分比越高，方块面积越大。

应用场景

- 查看数据大小关系和整体分布情况
- 不同组之间数据的比较
- 其他…

分析过程

上传数据 → 数据处理(清洗) → treemapify 计算方块的排布位置

→ 可视化

- 数据格式：（具体数据格式要求可以看后面过程的“数据格式”部分）

- 数据第 1 列必须为**分类类型**，对应映射部分的内容（颜色映射等）
- 数据第 2 列必须为**数值类型**，对应方块的大小
- 数据第 3 列及以后必须均是**分类类型**
- 必须提供第一列和第二列（至少两列）；至少三行，至多 260 行

	A	B	C	D
1	id	value	subgroup	group
2	A	10	Group1	Dep
3	B	20	Group1	Dep
4	C	30	Group2	Dep
5	D	5	Group2	Dep
6	E	5	Group2	Dep
7	F	10	Group2	Adv
8	G	15	Group3	Adv
9	H	10	Group3	Adv
10	I	5	Group3	Adv
11				

- 数据处理：对每一列数值类型的数据及其他列数据进行相应处理
 - 数值类型数据只能是纯数值类型数据，不能包含 0，负数、非数值与不规则的值
 -
- 可视化：将清洗后的数据使用 treemapify 包计算每个变量在总体数据的占比（即计算出 xmin, xmax, ymin, ymax），使用 ggplot2 包进行可视化

结果解读



- 堆砌方块图方块的颜色表示不同变量/样本，对应数据第 1 列；当选择其他列作为颜色映射时，会根据所选列的类型给方块上色
- 堆砌方块图方块的大小代表分类变量占总体数据的百分比的大小，对应数据第 1、2 列
- 可以直观比较不同组的数据情况和差异情况

数据格式

堆砌方块图

	A	B	C	D
1	id	value	subgroup	group
2	A	10	Group1	Dep
3	B	20	Group1	Dep
4	C	30	Group2	Dep
5	D	5	Group2	Dep
6	E	5	Group2	Dep
7	F	10	Group2	Adv
8	G	15	Group3	Adv
9	H	10	Group3	Adv
10	I	5	Group3	Adv
11				

数据要求：

- 第一列是用于绘图的变量（一个变量代表一个方块），**数据至少需要 2 列，3 行**，第一列均需要是字符类型，每二列均需要是数值类型。
- 方块的排布顺序是根据选择可视化中的“方块 - 排序”（从小到大/从大到小/不排序），若选择“不排序”，则排布顺序默认和上传数据的顺序一致。
- 数据**最多支持 260 行，10 列**，若验证数据时返回报错，需要在上传数据内进行相应的调整，然后再上传数据。
 - 数值类型数据只能是纯数值类型数据，不能包含 0、负数、非数值与不规则的值
- 数据每一列列名不能重复，不能有空值
- 列名中不能以 ymax, ymin, xmin, xmax, label_x, label_y 命名

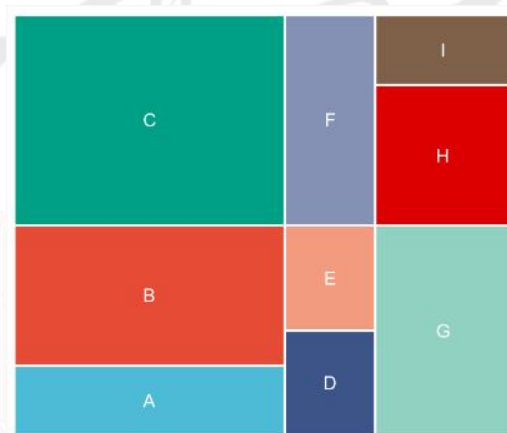
参数说明

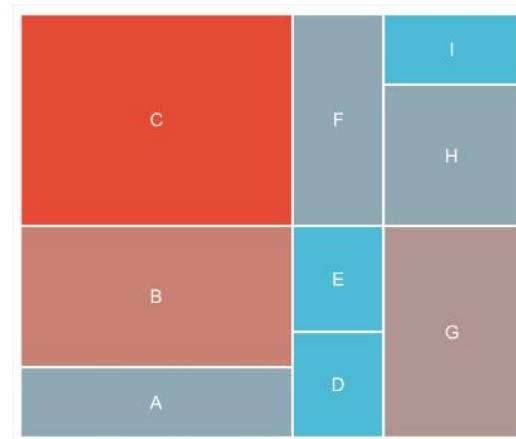
(说明：标注了颜色的为常用参数。)

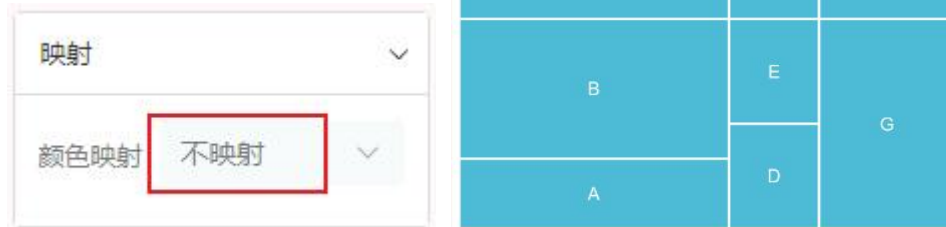
映射



- 颜色映射：可选择“id”、“value”或者“不映射”，默认为上传数据的第一列，即“id”（分类变量）。可选项的内容是上传数据的列名，不同数据的映射内容是有差别的。当选择“不映射”时，即为不用颜色映射，使用默认颜色。





方块



- 起始点：指定方块布局的起点，默认从左下角开始；还可以选择左下、右上和右下。

- 排序：按第二列的数值大小顺序放置方块，默认为“不排序”。还可以选择从小
到大排序或者从大到小排序。
- 填充色：方块的填充色颜色选项，有多少个分类变量（数据第一列的分类变量的数量）就会提取多少个颜色，最多支持修改 10 个颜色，超出会使用随机颜色。受配色方案全局性修改。
- 描边色：方块的描边色颜色选项，默认为白色。
- 描边粗细：方块描边的粗细，默认为 0.75pt。
- 不透明度：方块的透明度。0 为完全透明，1 为完全不透明。



标注

标注

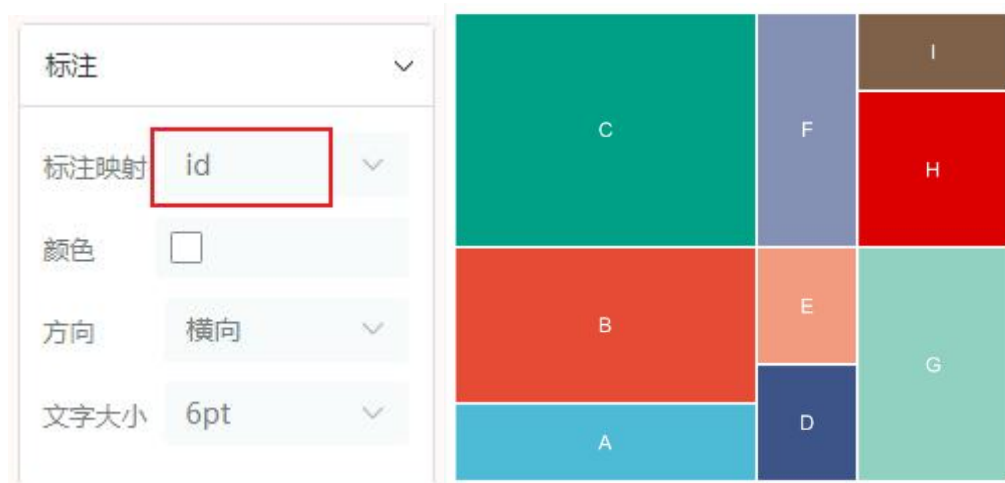
标注映射
id

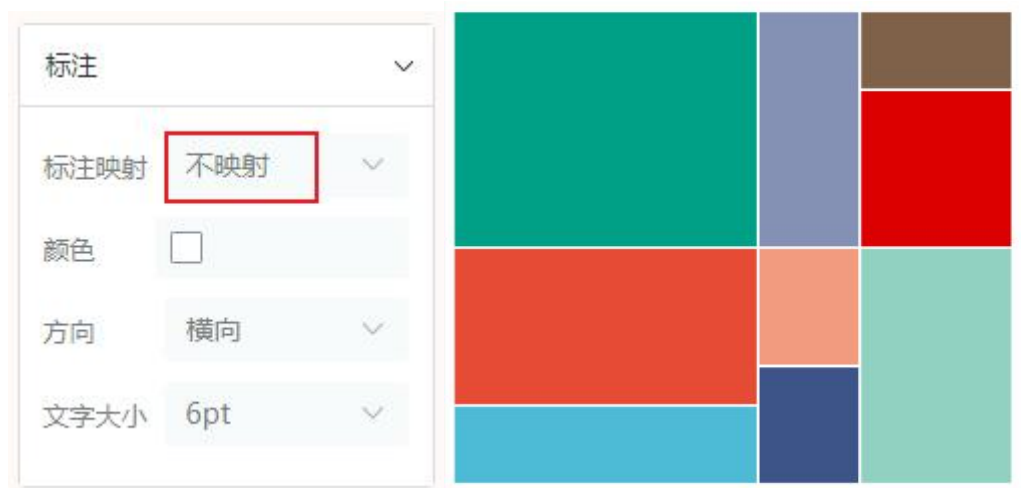
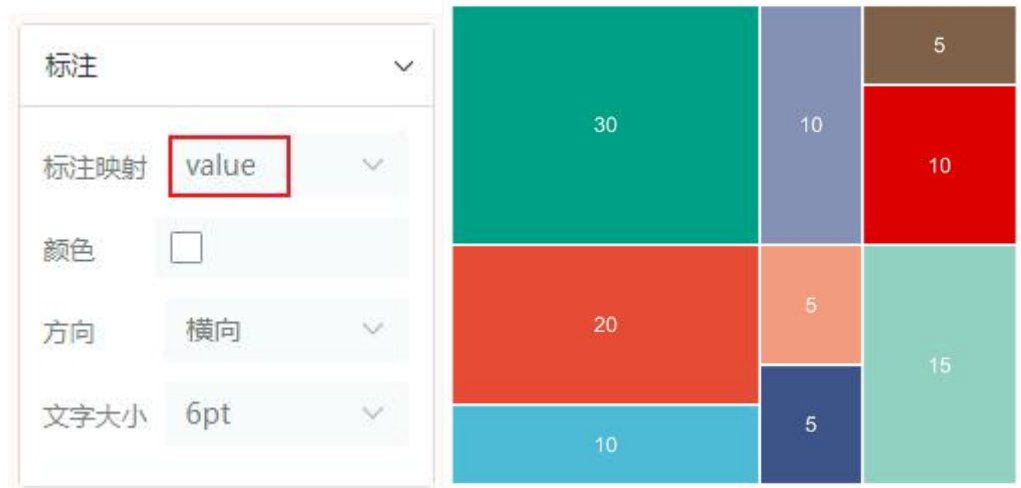
颜色
☐

方向
横向

文字大小
6pt

- 标注映射：可选择“id”、“value”或者“不映射”，默认为上传数据的第一列，即“id”（分类变量）。可选项的内容是上传数据的第一列和第二列的列名，不同数据的映射内容是有差别的。当选择“不映射”时，即为不显示标注。
- 颜色：标注的颜色选项，默认为白色。
- 方向：标注文字的方向，默认为横向。
- 文字大小：标注文字的大小，默认为 6pt。





分面

分面

分面映射

不映射

方向

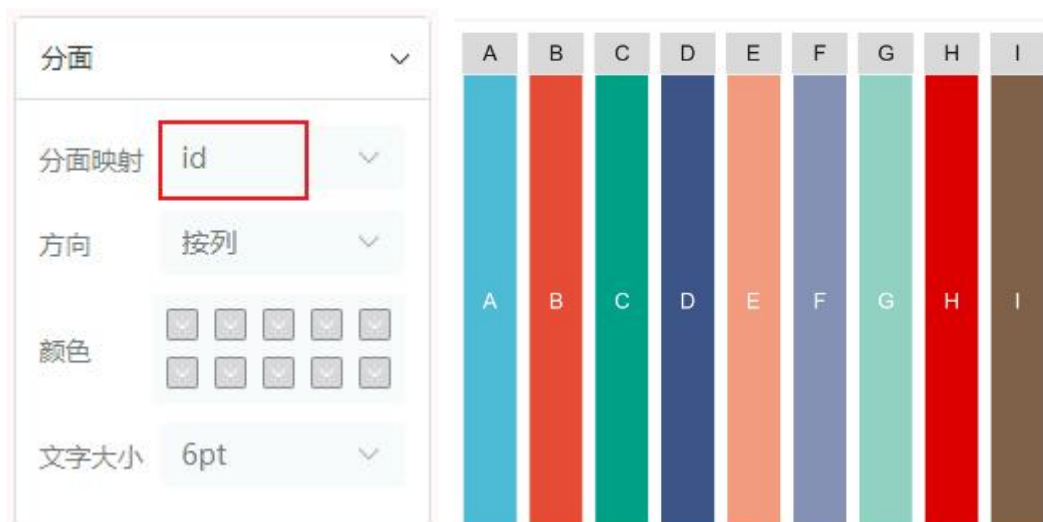
按列

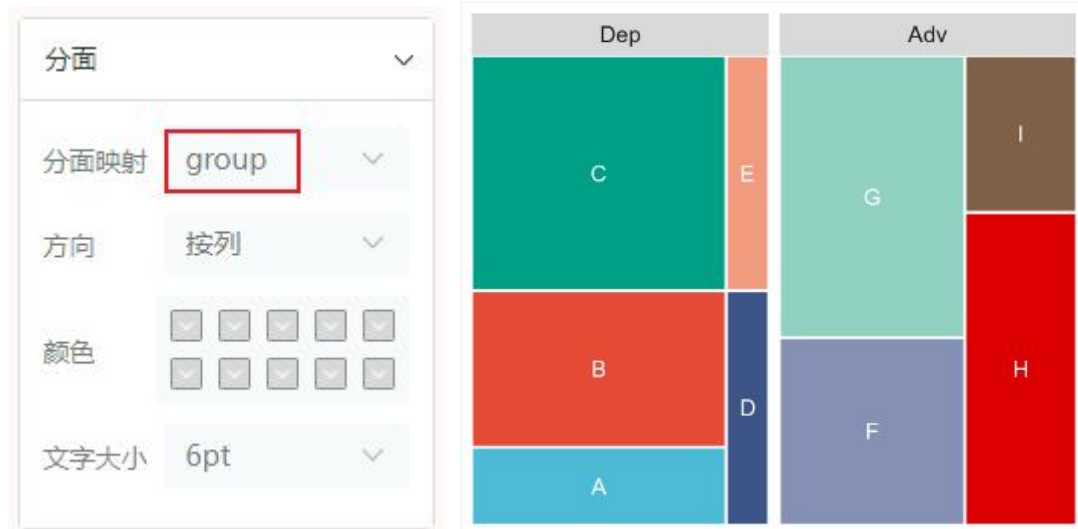
颜色

文字大小

6pt

- 分面映射：可选择“id”、“value”或者“不映射”，默认为“不映射”，即不展示分面结果。可选项的内容是上传数据的列名，不同数据的映射内容是有差别的。
- 方向：分面的方向，默认按列展示。
- 颜色：可以选择并修改分面的颜色。
- 文字大小：分面文字的大小，默认为 6pt。





坐标轴

- 是否显示 x 轴：选择即展示 x 轴。
- 是否显示 y 轴：选择即展示 y 轴。
- x 轴标注旋转：可以选择设置 x 轴标注的倾斜角度。

标题文本



- 大标题：大标题文本。
- x 轴标题：x 轴标题文本。
- 补充：在要换行的中间插入\n。如果需要上标，可以用两个英文输入法下的大括号括住，比如{{2}}；如果需要下标，可以用两个英文输入法下的中括号括住，比如[[2]]



图注 (Legend)



图注

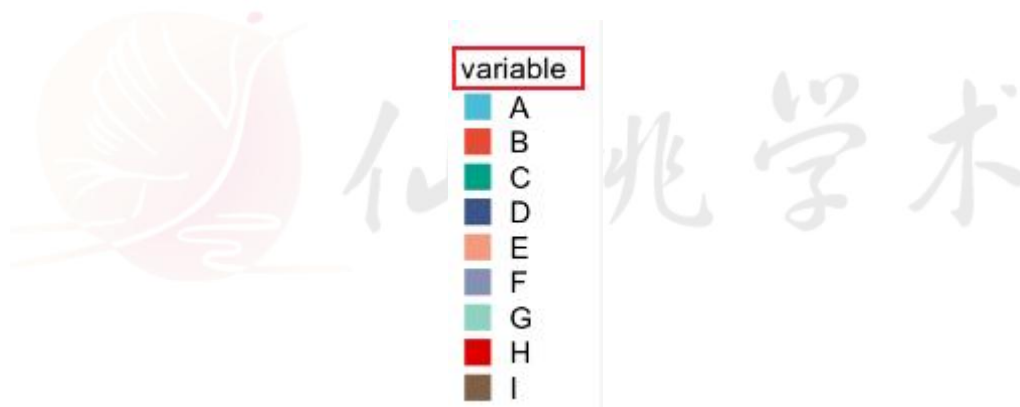
是否展示 ☐

图注标题 图注标题内容

文字大小 6pt

图注位置 默认

- 是否展示：是否展示图注
- 图注标题：可以添加图注标题，如：



- 文字大小：图注标题文字的大小，默认为 6pt。
- 图注位置：可选择默认、右、上、右上、左上。

风格



- 边框：可以选择是否进行添加图形边框的操作
- 文字大小：控制整体文字大小，默认为 7pt



图片

图片	▼
宽度 (cm)	6
高度 (cm)	5
字体	Arial ▼

- 宽度：图片横向长度，单位为 cm
- 高度：图片纵向长度，单位为 cm
- 字体：可以选择图片中文字的字体



结果说明

主要结果



主要结果格式为图片格式，提供 PDF、TIFF、PPTX 格式下载

方法学

软件：R (4.2.1)版本

R 包：ggplot2 包（用于可视化）

处理过程：

(1) 将清洗后的数据用 treemapify 包处理，再使用 ggplot2 包进行堆砌方块图可视化



如何引用

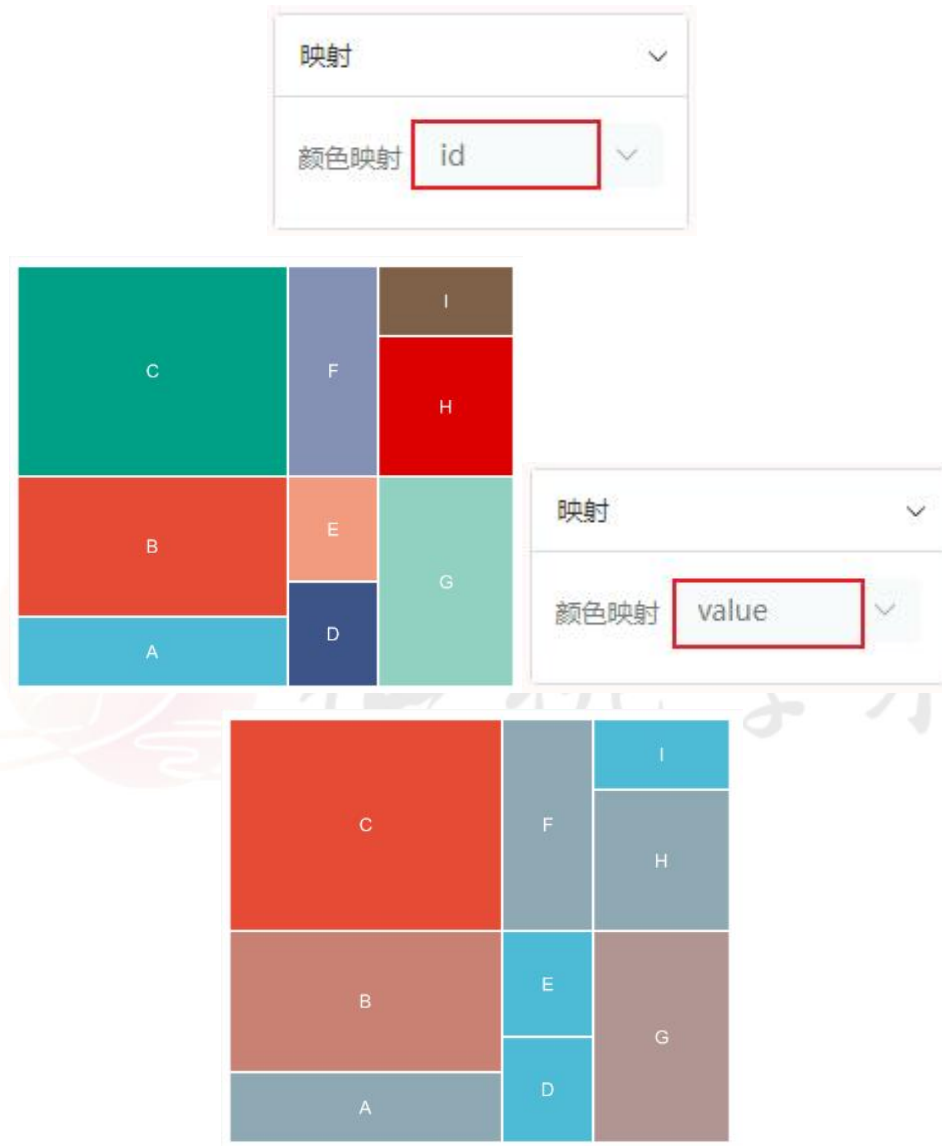
生信工具分析和可视化用的是 R 语言，可以直接写自己用 R 来进行分析和可视化即可，可以无需引用仙桃，如果想要引用仙桃，可以在致谢部分 (Acknowledge) 致谢仙桃学术 (www.xiantao love)。

方法学部分可以参考对应说明文本中的内容以及一些文献中的描述。



常见问题

1. 调整颜色映射的选项后，方块的颜色为什么会改变？



答:此处会识别用于颜色映射的数据的类型,如果输入的是第一列 id(即字符型),则它是离散型的变量,内部会分配单独的颜色给到每个变量;如果输入的是第二列 value(数值型),则它是连续型的变量,内部会分配渐变的颜色给到每个变量,而且此时修改颜色只有第一个和第二个颜色起作用。



2. 如何修改方块排布的顺序?

答：可以找到“方块 - 排序”，下拉选择“从小到大、从大到小或者不排序”来调整；若是均不满足，可以对上传的数据进行调整（调整时注意方块的起始点是从左下往上排的），调整好后选择“不排序”，则方块的顺序就是上传数据的排序情况。