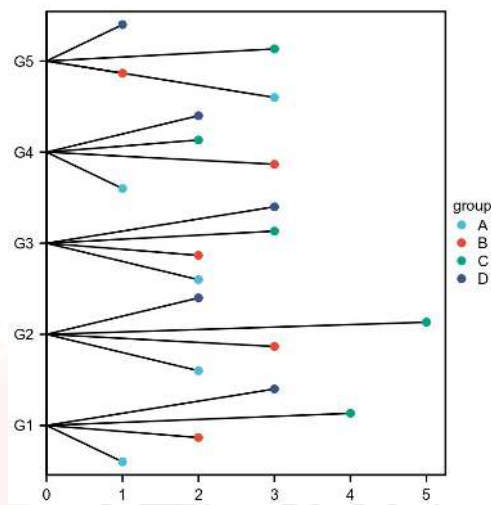


基础绘图 - 分裂棒棒糖图



网址: <https://www.xiantao love>



更新时间: 2023.03.15

目录

基本概念	3
应用场景	3
结果解读	5
数据格式	6
参数说明	7
映射	7
范围	8
点	9
线	11
标注	12
分面	13
标题文本	14
图注	14
坐标轴	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



- 棒棒糖图：通过点图和线图组合成棒棒糖的形式
- 分裂棒棒糖图：绘制同一个 x 对应一个或多个 y 值的棒棒糖图

应用场景

通过点、线组成棒棒糖的形式来展示数据各变量之间的数据分布情况



分析过程

上传数据  数据处理(清洗)  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 列需要提供分类类型，对应分裂棒棒糖图 x 轴（模块为转置之后的结果）
 - 数据第 2 列需要提供数值类型，对应分裂棒棒糖图 y 轴（模块为转置之后的结果）

- 数据第 3 列及以后可以是数值类型也可以是分类类型

	A	B	C	D	E
1	x	y	group	label	facet
2	G1	1	A	44.98496073	group1
3	G2	2	A	292.4500058	group2
4	G3	2	A	105.2045485	group1
5	G4	1	A	30.75092094	group2
6	G5	3	A	293.9174624	group1
7	G1	2	B	152.2163999	group1
8	G2	3	B	85.04235079	group2
9	G3	2	B	184.2218995	group1
10	G4	3	B	34.42079229	group2
11	G5	1	B	81.4023559	group1
12	G1	4	C	145.3496346	group1
13	G2	5	C	167.0746336	group2
14	G3	3	C	240.6165573	group1
15	G4	2	C	208.5968826	group2
16	G5	3	C	289.2315414	group1

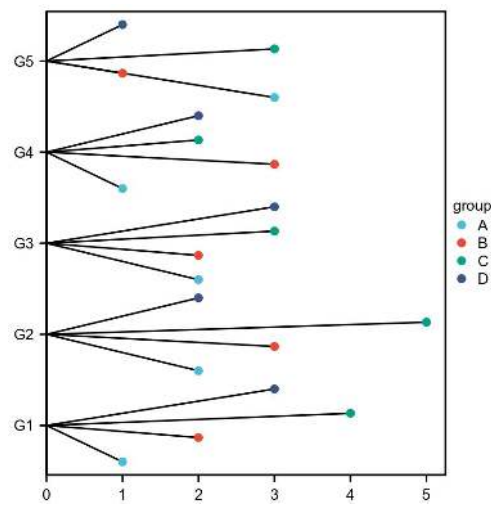
- 数据处理：对第 1 列分类类型数据，第 2 列数值类型的数据及其他列数据进行相应处理

- 数值类型数据只能是纯数值类型数据，不能包含非数值与不规则的值

-

- 可视化：将清洗后的数据进行 ggplot2 包可视化

结果解读



- 横坐标表示分组/变量（上传数据第 1 列）**分类类型的数据**（**图为默认转置后**）
- 纵坐标表示数值（上传数据第 2 列）**数值类型的数据**（**图为默认转置后**）
- 每个点对应的内容（颜色、大小、形状）表示分组/变量具体的值，连接点的横线长度表示其值大小
- 横坐标等于 0 位置的竖线（虚线）表示分割线（图形中与坐标轴重合），一般用以划分正负两个不同方向的值（不影响）

数据格式

	A	B	C	D	E
1	x	y	group	label	facet
2	G1	1	A	44.98496073	group1
3	G2	2	A	292.4500058	group2
4	G3	2	A	105.2045485	group1
5	G4	1	A	30.75092094	group2
6	G5	3	A	293.9174624	group1
7	G1	2	B	152.2163999	group1
8	G2	3	B	85.04235079	group2
9	G3	2	B	184.2218995	group1
10	G4	3	B	34.42079229	group2
11	G5	1	B	81.4023559	group1
12	G1	4	C	145.3496346	group1
13	G2	5	C	167.0746336	group2
14	G3	3	C	240.6165573	group1
15	G4	2	C	208.5968826	group2
16	G5	3	C	289.2315414	group1

数据要求：

- 数据至少 2 列以上，每列至少 1 个观测（即至少 1 行数据），最多支持 10 列和 800 行数据
 - 第 1 列为分类类型数据，对应 x
 - ◆ 该列数据中不能含有缺失的值(空的内容)、不能含有非字符或者无法识别的字符串
 - ◆ 该列所包含的所有分类中不能超过 100 个
 - 第 2 列为数值类型数据，对应 y
 - ◆ 该列数据只能是数值类型的数据，不能含有非数值或者非字符...
 - 数据第 3 列及以后可以是数值类型也可以是分类类型
 - ◆ 这些数据只能是纯数值类型数据或者是纯非数值类型数据(不能混合使用)
 - ◆ 不能含有无法识别的字符
- 每一列列名不能重复.....

参数说明

(说明：标注了颜色的为常用参数。)

映射

映射

颜色映射 group

大小映射 不映射

形状映射 不映射

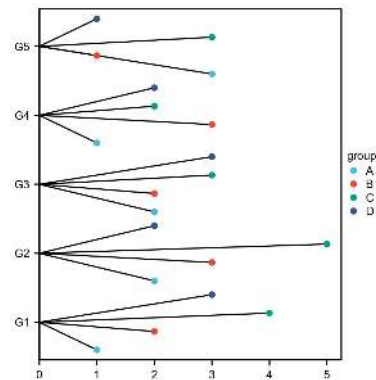
- 颜色映射：可以选择图中点颜色映射的信息，可以是数值类型数据、也可以是分类类型数据，**根据上传数据动态生成**，还可以选择不进行映射，如下：
(第1组为数值类型数据映射，第2组为分类类型数据映射)

映射

颜色映射 group

大小映射 不映射

形状映射 不映射

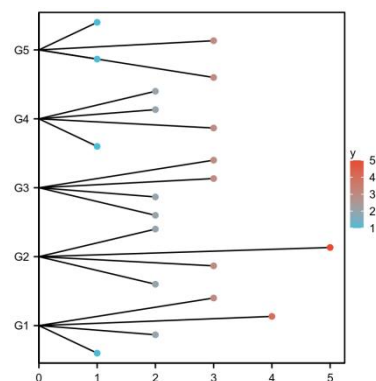


映射

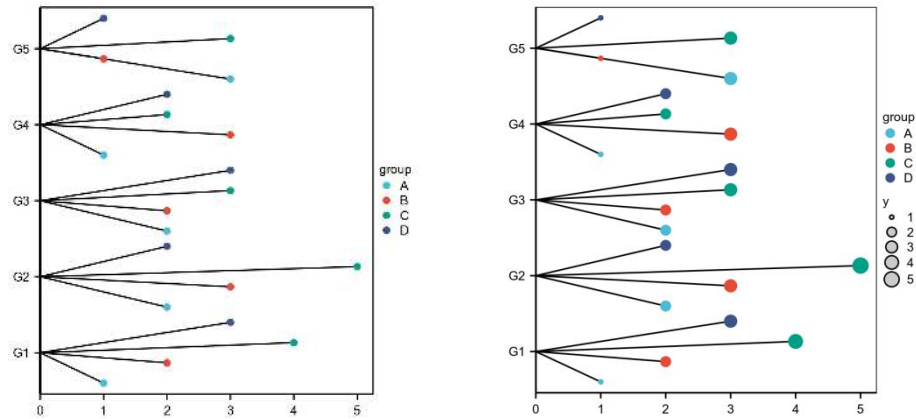
颜色映射 y

大小映射 不映射

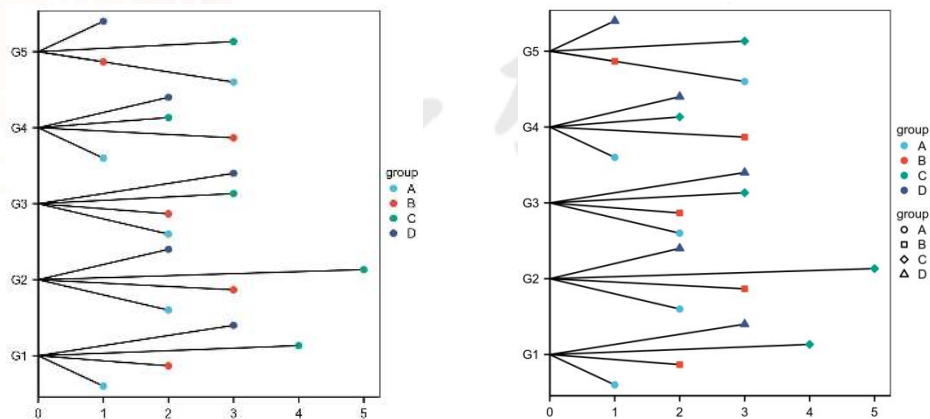
形状映射 不映射



- 大小映射：可以选择图中点大小映射的信息，数值类型数据，根据上传数据动态生成，还可以选择不进行映射（如下：左侧为不映射情况，右侧为有映射）



- 形状映射：可以选择图中点形状映射的信息，分类类型数据，根据上传数据动态生成，还可以选择不进行映射（如下：左侧为不映射情况，右侧为有映射）



间距设置

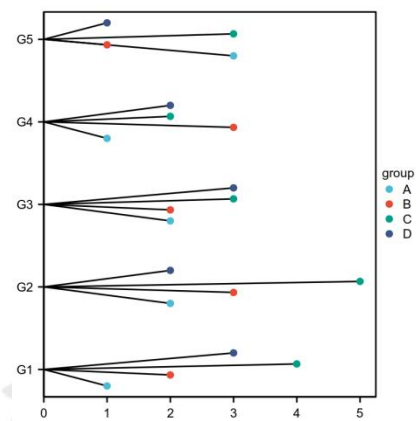
间距设置

组内总宽度 0.8

- 组间范围：可以修改数据第 1 列各分组中样本与样本间的距离总和；组内总宽度：0-1，如下：

间距设置

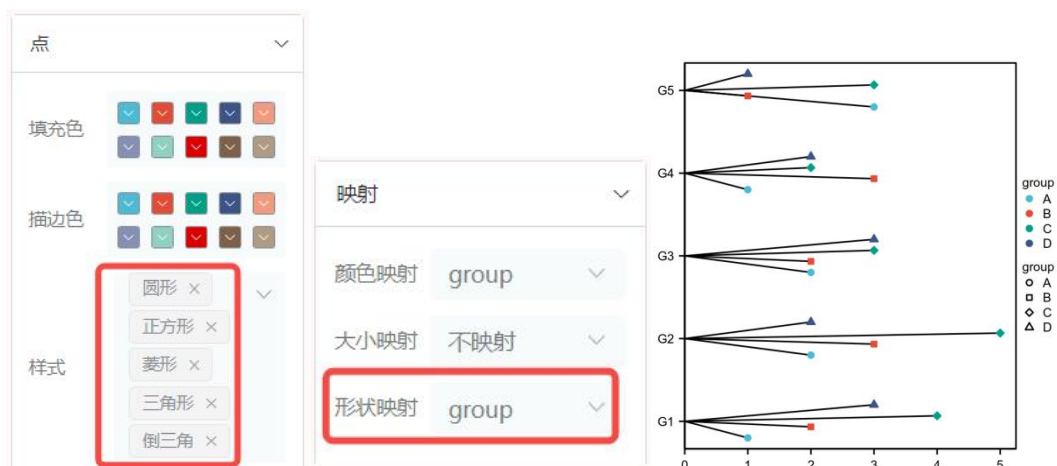
组内总宽度 0.4



点

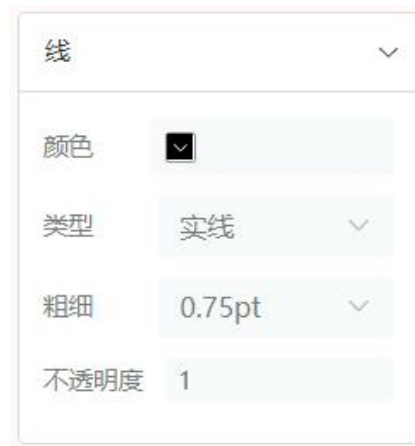


- 填充颜色：可以修改图中各点的填充颜色
- 描边颜色：可以修改图中各点的描边颜色
- 样式（多选）：当进行形状映射的时候，可以修改图中各点的样式（形状），可以选择圆形、正方形、菱形、三角形、倒三角形（可以重复选择），如下：



- 大小比例：可以修改图中个点的大小比例，默认为 1
- 不透明度：可以修改图中各点不透明度，默认为 1，表示完全不透明

线



- 颜色：可以修改图中线的颜色
- 类型：可以选择连接点的线的类型，默认为实线，还可以选择虚线
- 粗细：可以选择修改线条的粗细，默认为 0.75pt
- 不透明度：可以修改线条的不透明度，默认为 1，表示完全不透明

标注

标注

标注映射 不映射

标注大小 6pt

方向 纵向

- 标注映射：根据上传数据特点，可以选择是否对图形进行标注映射，默认为不映射，还可以选择映射，如下：标注大小：10pt



- 标注大小：可以选择标注的大小，默认为 6pt
- 方向：可以选择标注的方向，默认为纵向，还可以选择横向

分面

分面

分面映射

不映射

分面颜色

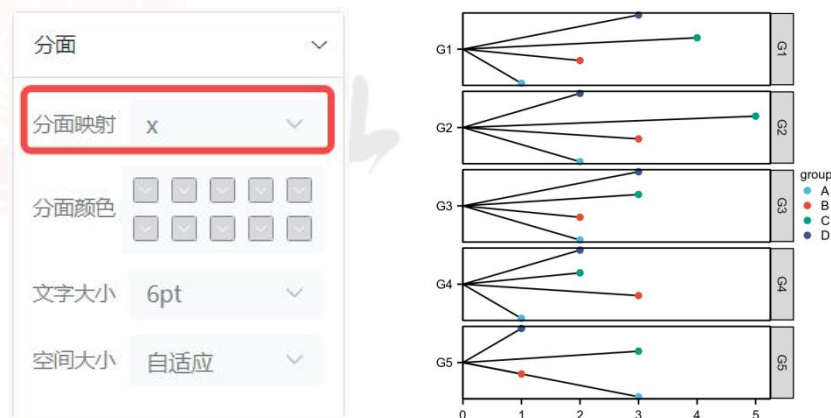
文字大小

6pt

空间大小

自适应

- 分面映射：根据上传数据特点，可以选择是否对图形进行分面映射，默认为不映射，还可以选择映射，如下：



- 分面颜色：可以修改分面图形的分面颜色
- 文字大小：可以选择并修改分面文字的大小，默认为 6pt
- 空间大小：可以选择分面的空间大小，默认为自适应（表示跟随图形变化），还可以选择固定（表示不随图形变化）

标题文本

标题

大标题

大标题内容

x轴标题

x轴标题内容

y轴标题

y轴标题内容

- 大标题：大标题文本
- x 轴标题：x 轴标题文本
- y 轴标题：y 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

图注

是否展示

☒

图注位置

默认

- 是否展示：可以选择是否展示各指标映射的内容，即图注，默认展示
- 图注位置：可以选择图注的位置，默认表示默认展示在右侧，还可以选择上

坐标轴



坐标轴

x轴标注旋转 0

y轴范围+刻度 逗号隔开

- x 轴标注旋转：可选择并修改 x 轴对应刻度文本的旋转角度
- y 轴范围+刻度：可以控制 y 轴范围和刻度，可只提供 2 个值来控制范围。形如 0.1, 0.1, 0.2, 0.3 (最小值和最大值能不能可视化数据范围 20%，如果调整过大可能会无作用)



风格



- 边框：可以选择是否展示图片边框，默认展示
- 网格：可以选择是否展示网格，默认不展示
- 可以选择是否进行 x、y 轴颠倒，默认进行颠倒
- 文字大小：控制整体文字大小，默认为 7pt

图片



图片	▼
宽度 (cm)	7
高度 (cm)	7
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

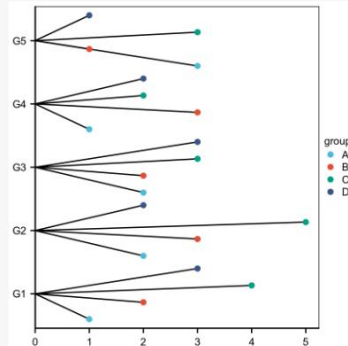


结果说明

主要结果

分裂棒棒糖图

分裂棒棒糖图: 通过点、线的形式来展示数据各变量之间的关系情况



分裂棒棒糖图.pdf

分裂棒棒糖图.tiff



方法学

统计分析和可视化均在 R 4.2.1 版本中进行

涉及的 R 包：ggplot2 包（用于可视化）

处理过程：

(1) 使用 ggplot2 包对上传数据进行可视化



如何引用

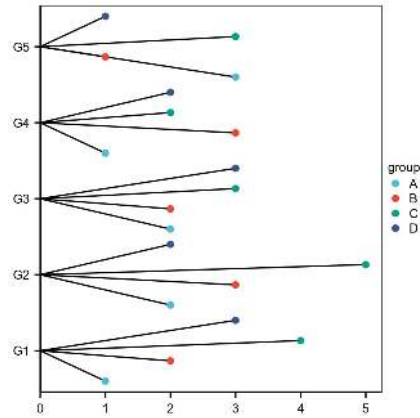
生信工具分析和可视化用的是 R 语言，可以直接写自己用 R 来进行分析和可视化即可，可以无需引用仙桃，如果想要引用仙桃，可以在致谢部分 (Acknowledge) 致谢仙桃学术 (www.xiantao love)。

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. 为什么数据说明与主要结果说明跟图的坐标对应不上?



➤ 横坐标表示变量（第 1 列数据（固定）分类类型的数据）

➤ 纵坐标表示变量（数据第 2 列数值类型的数据）

答：主要结果显示的是默认将图形进行转置之后得到的，所以会跟数据说明与主要结果说明反过来了，如下（左侧为默认转置后，右侧为不转置）

