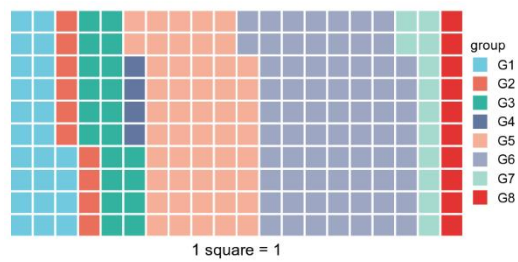


基础绘图 - 华夫饼图



网址: <https://www.xiantao love>



更新时间: 2023.05.26

目录

基本概念	3
应用场景	3
分析过程	3
主要结果	5
数据格式	6
参数说明	7
映射	7
格子系数	8
华夫饼（颜色）	8
标题	9
图注(Legend)	10
风格	10
图片	11
结果说明	12
主要结果	12
方法学	13
方法学	13
如何引用	14
常见问题	15

基本概念

- 华夫饼：用于展示每组数据占总体的比例情况。

应用场景

华夫饼图可以快速看出数据中不同类别的分布，同时也可以知道不同分组之间的分布差异。

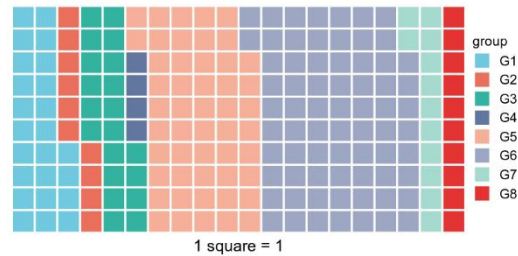
分析过程

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 列为分类类型，对应华夫饼图中的格子**分组**
 - 数据第 2 列必须为数值类型，对应华夫饼图的格子**数量**和**高度**
- 计算过程：
 - 数据第 2 列数值 除以 格子系数，获得**各分类**的格子数，并计算绘图 **总格子数**
 - 总格子数 除以 行数，获得对应的 **列数**
 - 格子按分组从下到上，依次往右侧排列，最终获得 **华夫饼图**



主要结果



- 图中一个格子代表单位 1，该系数可根据具体数据调整。
- 不同颜色代表分类，颜色映射对应数据的第一列。
- 格子默认按照分类顺序，从下到上，从左到右排列。



数据格式

	A	B	C
1	group	value	other
2	G1	24	A
3	G2	10	A
4	G3	20	B
5	G4	4	B
6	G5	50	C
7	G6	70	C
8	G7	12	D
9	G8	10	D

数据要求：

- 数据至少 2 列，2 行。第 1 列需要是 id（分类类型），第 2 列为对应的计数（数值类型），第 3 列及以后为想要提供映射的列，**根据所选列的数据类型对颜色进行映射。**
 - 颜色映射：数据类型可以为**数值型** 或 **分类型**。本示例取 group 列信息作为格子颜色。
- 第 1 列作为分类数据，不能含有重复、空值。
- 第 2 列数值作为计数，不能含有空值、负数、全是 0 的情况。
- **最多支持 10 列，500 行。**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。

参数说明

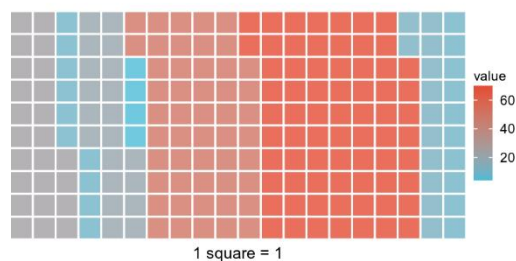
(说明：标注了颜色的为常用参数。)

映射



- 颜色映射：主要影响格子的颜色范围，**注意映射内容的数值类型，数值型数据为渐变色，分类型数据为单个颜色**。可选择 分类型、数值型 数据列信息或不映射，根据具体情况映射格子的颜色。

- 以下为 数值类型 映射颜色的展示，可以看到同一分类（第1列）的颜色相同：



格子系数



格子参数

格子系数 1

行数 10

- 格子系数：设置格子所代表的数值，即单位系数。建议是可整除一定的倍数，如 10、100 等。
- 行数：影响格子排列的行数, 1-25 之间的整数。与格子具体个数有关。

华夫饼（颜色）



华夫饼

填充色

描边色

描边粗细 0.75pt

不透明度 0.8

- 填充色：格子的填充色颜色选项，取决于 [颜色映射](#) 参数所选择的内容，展示数值型内容时，修改第一和第二色卡作为数值从小到大的渐变色；展示分类型内容（如 类别）时，有多少个功能类别会提取多少个颜色，最多支持修改 10 个颜色。受配色方案全局性修改。

- 描边色：格子的描边色颜色选项，取决于 **颜色映射** 参数所选择的内容，展示数值型内容时，修改第一和第二色卡作为数值从小到大的渐变色；展示分类型内容（如 类别）时，有多少个功能类别会提取多少个颜色，最多支持修改 10 个颜色。默认 白色。**不受配色方案全局性修改**。
- 描边粗细：柱子描边的粗细，默认为 0.75pt。
- 不透明度：柱子的透明度。0 为完全透明，1 为完全不透明。

标题



- 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(Legend)



图注

是否展示 ☒

图注标题 图注标题内容

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
- 图注位置：可选择 默认、右、上、下。

风格



风格

边框 ☐

文字大小 6pt

- 外框：是否添加外框
- 文字大小：针对图中文字的大小控制，默认为 6pt

图片

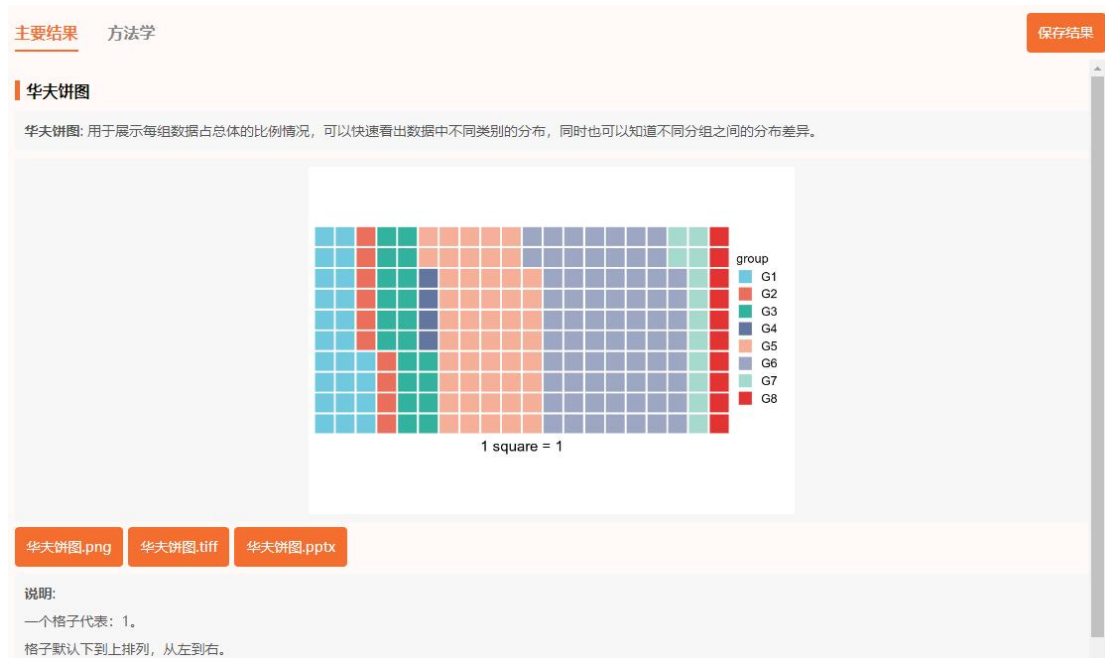
图片	▼
宽度 (cm)	7
高度 (cm)	5
字体	Arial ▼

- 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 和 PPTX 格式下载。

方法学

方法学

所有分析和可视化均在 R 4.2.1 中进行

涉及的 R 包: ggplot2 包 (用于可视化)

处理过程: 将清洗后的数据进行相应统计, 用 ggplot2 包绘制华夫饼图。



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](http://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。

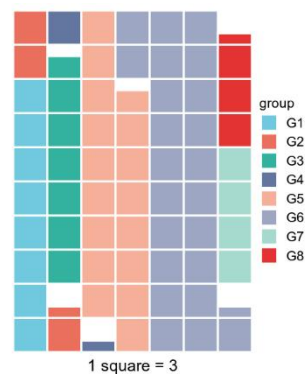


常见问题

1. 格子系数是什么?

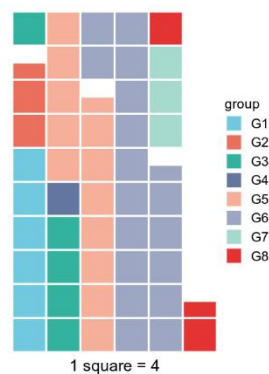
答:

格子系数，是指定一个格子所代表的数值，即单位。如示例数据中，设置系数为 3 时，一个格子则代表 3，而数据第 2 列整体除以 3 才能获得真正的绘图格子数量。如下，1 square = 3:



2. 为什么有的格子不全?

答:



这是由于格子系数无法整除数据第 2 列导致的，如下:

1) 当格子系数 = 4 时，各分类格子数的计算即为:

第 2 列分类的原始数值 / 格子系数

- 2) 当数值无法被系数整除时，将出现小数位的情况导致无法绘制完整格子，此时，将余数数值赋值给对应分类的最后一个格子作为**相对高度**绘图。
- 3) 如分组“G2”，最后一个格子的相对高度为 0.5，代表数值 $0.5 * 4 = 2$ ；结合该分类的剩余格子数可以推算， $2 * 4 + 0.5 * 4 = 10$ ，即原始数值。

计算格子数

	A	B	C	D	编辑栏	F
1	group	value	other	格子数		
2	G1	24	A	6		
3	G2	10	A	2.5		
4	G3	20	B	5		
5	G4	4	B	1		
6	G5	50	C	12.5		
7	G6	70	C	17.5		
8	G7	12	D	3		
9	G8	10	D	2.5		

无法整除时，该分类的最后一个格子即为余数

注意：格子系数的设置一定与原始数值（第 2 列）相关，出现格子不全时，可以由计算过程推断结果。