

中国移动数据可视化分析

一、数据分析与设计

给出的数据由 5 个小区、10 种数据类型、56 个用户的数据使用量组成，实质上由以下 4 个变量组成：

| 变量名 | 类型 | 取值范围 |
|-------|------|-------------|
| 小区 | 类型变量 | 1~5 |
| 数据类型 | 类型变量 | 微信、小说等 10 种 |
| 用户 | 类型变量 | 56 个手机号 |
| 数据使用量 | 定距变量 | 正整数 |

对这些数据可以进行的数据可视化包括：单个变量的绝对数值分析（例如不同小区的数据使用总量展示）、单个变量的比例分析（例如不同种类的数据所占比例）、多个变量的绝对数值分析（例如不同小区内各种数据类型的使用总量展示）、多个变量的比例分析等等。其中，多个变量的分析还要受到变量顺序的影响，例如不同小区内各种数据类型的使用比例，与不同数据类型内各小区的贡献比例，这两者的展示效果是完全不同的。

考虑到可视化展示的多样性，必须增强交互效果，最终决定使用网页的形式，把上面的信息都集中到一个界面之内。主要的展现形式是绘制多重饼图，跟踪鼠标的位置提供不同级别的数据信息，其中包括使用的数据总量以及占总体的比例。

二、可视化实现

1、使用 Python 进行数据统计工作

在原始的 Excel 数据后面再加一列，填入小区编号（1 到 5），这样就可以把五个工作表的数据合并在一起，另存为 CSV 格式，具体见 Python 文件夹下的 DataVisualizationV2.csv 文件。

使用 Python 文件夹下的 main.py 对数据进行分析，统计出不同小区 Community 内不同数据类型 Type 的使用总量，存为 CT.csv；不同小区 Community 内不同数据类型 Type 下各用户 User 的使用总量，存为 CTU.csv；不同数据类型 Type 下的各小区 Community 的使用总量，存为 TC.csv；不同数据类型 Type 下的各小区 Community 的各用户 User 使用总量，存为 TCU.csv。

之所以没有统计各用户 User 在不同社区 Community 中的使用情况，是因为原始数据中每个用户只隶属于一个小区，所以没有必要进行这项分析，上面的 4 种分析顺序已经包含了所有的数据分析情况。

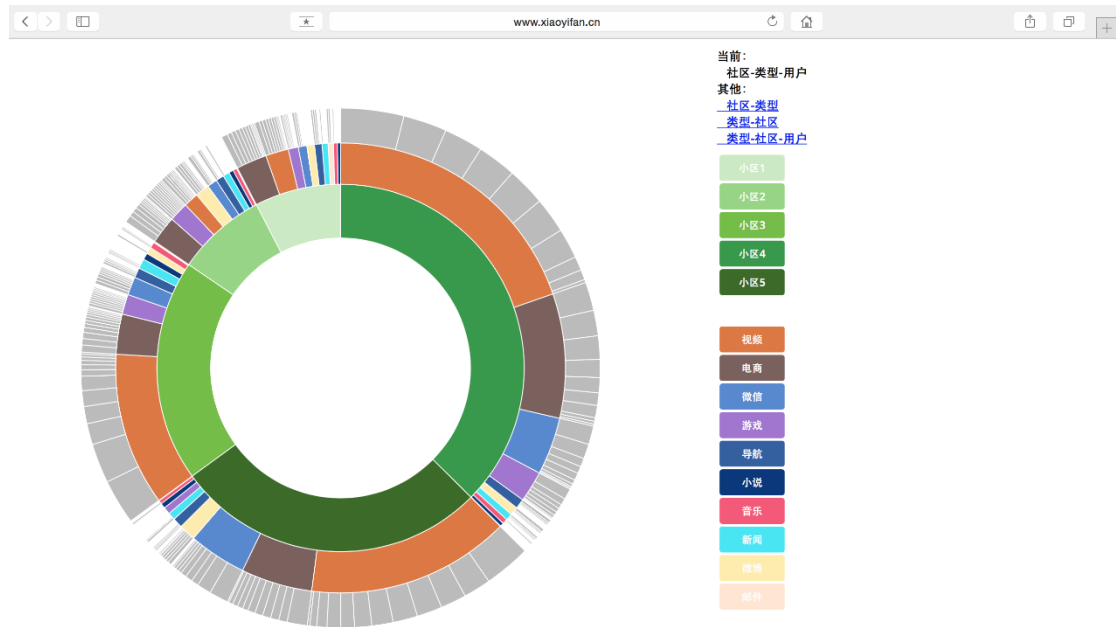
2、搭建 HTML 进行可视化展示

使用 D3 平台编写可视化动画，并进行配色。5 个小区分别用深浅不同的 5 种绿色表示，不同用户均使用灰色，10 种数据类型使用除绿色、灰色之外的 10 种颜色表示。为了保证网页的速度，只有大于 0.005 度的饼图叶会被画出来，而且比例的数据只精确到三个有效数字。

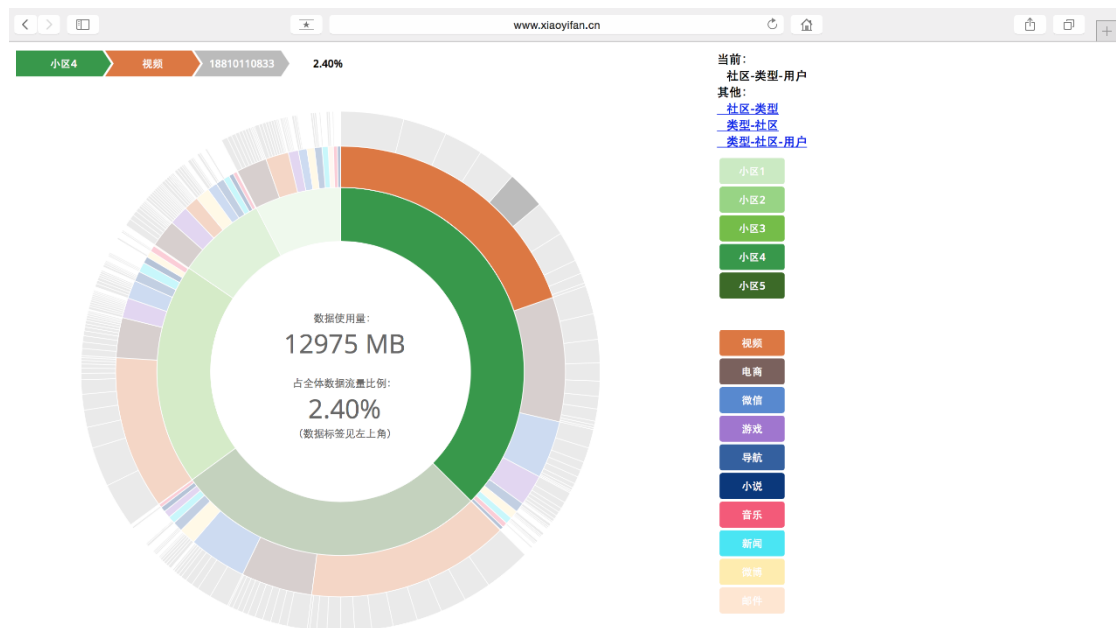
三、成果展示

虽然把网页的源代码已经提交，但是直接打开 index.html 是看不到效果的，因为浏览器不允许直接读取 csv 文件，需要搭建服务器之后才可以打开，本人使用 Apache 服务器，可以访问 <http://166.111.131.98/10086/Spark/DataVis/> 查看。

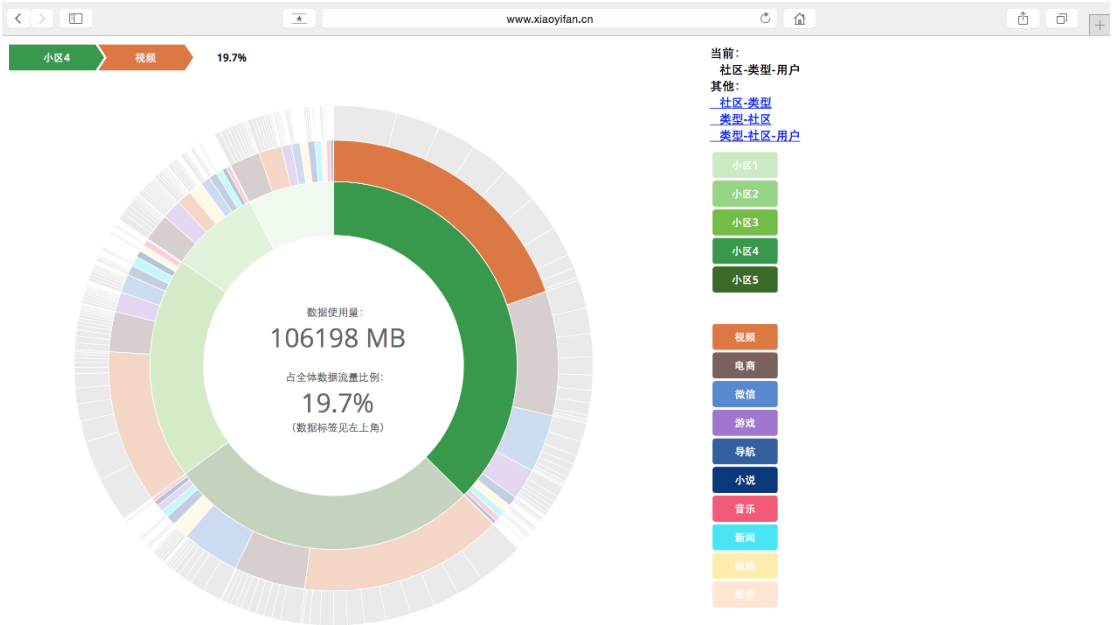
打开网页后的视图如下：



把鼠标移到饼图上，跟随鼠标可以得到数据分析结果，下图得到的信息是：小区 4 中视频类流量里，用户 18810110833 使用了 12975MB，占本次数据分析所有流量的 2.4%。移动鼠标可以得到其他小区、类型、用户的情况。数据标签查看页面左上角即可。

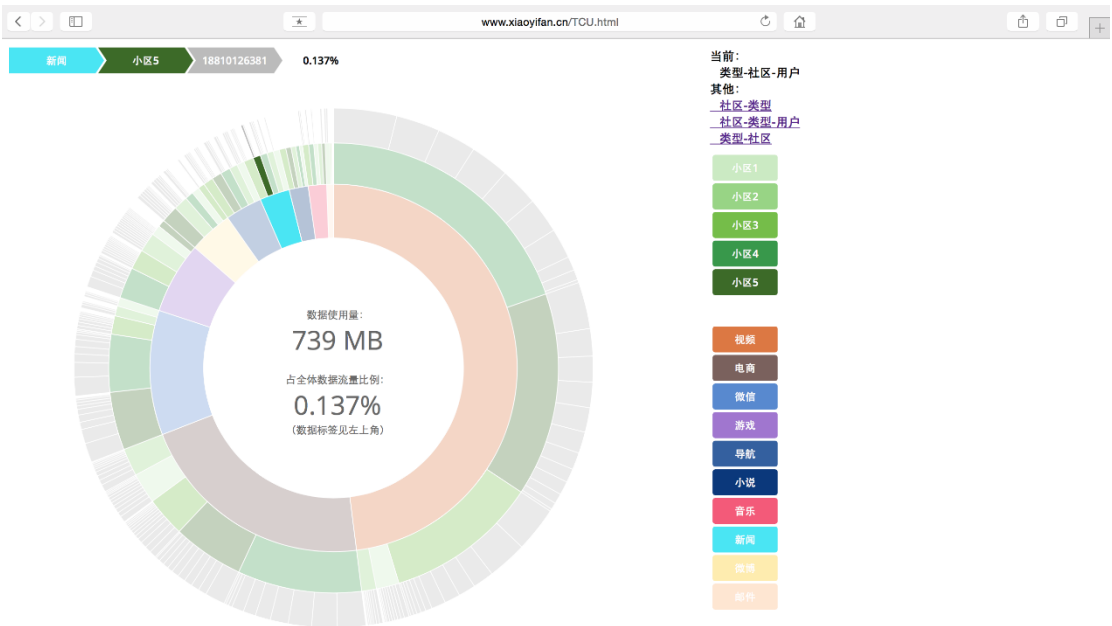


上图中也可以得到前两个变量的情况，把鼠标移到饼图的第二环上即可，结果如下图，信息是小区 4 中使用了 106198MB 视频流量，占总流量的 19.7%。



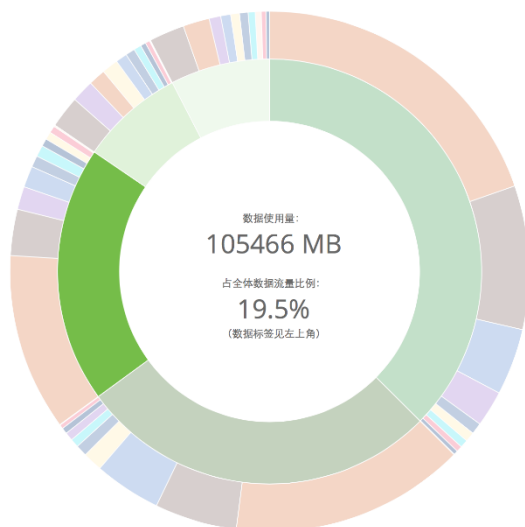
页面右上角可以点击其他的变量分析顺序，默认是社区→类型→用户的顺序。

点击类型-社区-用户的链接，得到下图的分析结果，图中的信息是：新闻类数据流量中，小区 5 的 18810126381 用户使用了 739MB，占总体的 0.137%。



也可以选择不包含用户数据的情况，分别点击社区-类型和类型-社区的链接后，效果分别如下：

小区3 19.5%

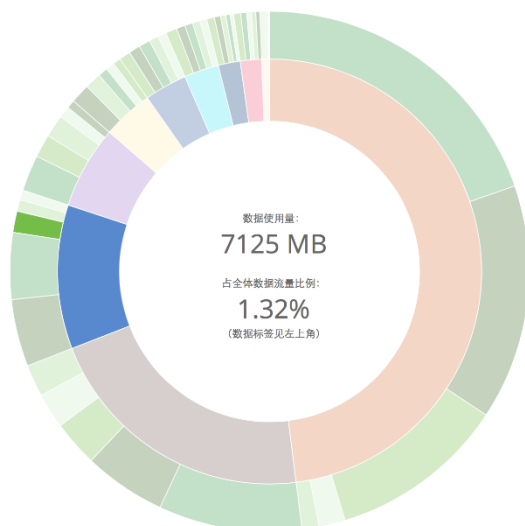


当前:
社区-类型
其他:
[社区-类型-用户](#)
[类型-社区](#)
[类型-社区-用户](#)

- 小区1
- 小区2
- 小区3
- 小区4
- 小区5

- 视频
- 电商
- 微信
- 游戏
- 导航
- 小说
- 音乐
- 新闻
- 微博
- 邮件

微信 小区3 1.32%



当前:
类型-社区
其他:
[社区-类型](#)
[社区-类型-用户](#)
[类型-社区-用户](#)

- 小区1
- 小区2
- 小区3
- 小区4
- 小区5

- 视频
- 电商
- 微信
- 游戏
- 导航
- 小说
- 音乐
- 新闻
- 微博
- 邮件