

Tableau数据可视化

俗语说,一图胜万言。图表在展示数据和揭示信息方面的重要性不言而喻。数据可视化是将数据背后的信息以图表的形式直观地呈现出来,让数据实现自我解释,以此传递各种价值。进一步而言,数据可视化是受数据驱动,利用计算机工具进行图表制作的一门艺术。好的数据可视化要求制作的图表既要简明清晰地反应数据和信息,又要符合人们的审美观念,具有视觉吸引力。为了更深入了解数据可视化,本章将介绍明星工具 Tableau 的基本操作和应用案例。

5.1 Tableau 概述与入门



5.1.1 概述

Tableau 是 Tableau Software 公司研发的软件产品。Tableau Software 公司是一家商业智能软件提供商,由斯坦福大学的三位校友 Patrick M. Hanrahan, Christian Chabot 和 Chris Stole 于 2003 年创立。Hanrahan 是皮克斯(Pixar)动画工作室创始成员之一,在渲染和计算机图形研究方面获得过三项奥斯卡技术奖,他和 Edwin Catmull 由于在计算机图形学领域的卓越贡献,一起获得 2019 年图灵奖(计算机界的全球最高奖项)。

Tableau 是一款敏捷的自助式的数据可视化工具,它能快速灵活地连接和整合数据,提供简单的方式实现从不同的角度观察、计算和展示不同的指标,能马上分享并获得反馈。Tableau 作为轻量级数据可视化工具的优秀代表,在 Gartner(高德纳)咨询公司发布的 2020 年《分析和商业智能平台魔力象限报告》中,连续八年获评领先者象限。Tableau 出色的表现主要得益于以下几个方面的特性。

(1) 高效易用。Tableau 通过内存数据引擎,可以直接查询外部数据库,同时动态地从数据仓库抽取实时数据,极大地提高了数据访问和查询效率。而且,Tableau 提供了友好的可视化界面,用户仅需要通过单击或者拖曳鼠标就可以迅速创建出智能、精美、直观和具有强交互性的报表和仪表盘。

(2) 能够连接多种数据源,轻松实现数据融合。Tableau 支持多种类型的数据源,包括带分隔符的文本文件、Excel 文件、SQL 数据库、Oracle 数据库和多维数据库等。而且,Tableau 支持用户轻松地在多个数据源之间切换,以及整合多个不同的数据源,轻松实现数据融合。

(3) 集成高效的接口,支持多种编程语言。Tableau 提供了多种应用编程 API 接口,支持 C、C++、Java、JavaScript、R 以及 Python 等多种语言。

Tableau 的产品体系非常丰富,主要包括 Tableau Desktop、Tableau Server、Tableau

Online、Tableau Mobile、Tableau Public 以及 Tableau Reader。本章主要围绕 Tableau Desktop 讲授。

Tableau Desktop 是设计和创建美观的视图与仪表板、实现快捷数据分析功能的桌面分析工具,支持 Windows 和 Mac OS 操作系统。它包括 Personal(个人版)和 Professional(专业版)两个版本。Personal 版本仅支持与本地文件和本地数据库连接,分析成果可以发布为图片、PDF 和 Tableau Reader 等格式;而 Professional 版本除了具备 Personal 版本的全部功能之外,还支持更加丰富的数据源,几乎支持与所有类型的数据和数据库系统的连接,还支持将分析结果发布到企业或者个人的 Tableau 服务器、Tableau Online 服务器和 Tableau Public 服务器上,实现移动办公。因此,Professional 版本比 Personal 版本更加通用,但价格更贵。

5.1.2 下载与安装

登录 Tableau 官方网站 <http://www.tableau.com/zh-cn/products/trial> (建议使用 Google Chrome 或者 Firefox 浏览器),如图 5.1 所示,填写“商务电子邮件”后,单击“下载免费试用版”,进入下载过程,即可以下载 Tableau Desktop 最新的免费试用版本。最新的 Tableau Desktop 只提供 64 位版本,而不提供 32 位版本。



图 5.1 Tableau Desktop 下载页面

如果需要下载 Tableau Desktop 的历史版本或者 32 位版本,可以到 <https://www.tableau.com/support/releases> 下载,该链接也可以下载 Tableau Server 等其他产品。Tableau Desktop 安装文件下载完成后,在 Windows 操作系统的安装过程与其他软件基本相同。软件安装结束后,直接进入“激活 Tableau”,用户可以通过产品密钥激活 Tableau。若无产品密钥激活 Tableau,注册完成后只能无限制使用 14 天。

针对一般人员,Tableau 公司要求必须付费购买产品密钥;针对全球高校的老师和学生,Tableau 公司推出学术版,通过申请可以获得 Tableau Desktop 等产品免费一年使用期

限的产品密钥。学生登录网址 <https://www.tableau.com/zh-cn/academic/students> 申请；教师登录网址 <https://www.tableau.com/zh-cn/academic/teaching> 申请。不管是教师还是学生，在申请过程中需要按照指引，详细填写个人信息以及提供相应证明材料。注意要准确填写邮箱地址，以便能接收产品密钥。为更好地证明自己的教师(学生)身份，申请者最好使用教育网邮箱，以及上传自己的教师证(学生证)扫描文件(图片)等材料。当获得产品密钥时，可以选择 Tableau Desktop“帮助”菜单下的“产品激活”命令，在弹出的对话框中输入产品密钥激活产品。

5.1.3 数据类型

通常，数据表中的每一列称为一个字段，代表着一个属性；每一行称为一个实例，代表着一条数据记录。数据表中的每个字段都具有一种特定的数据类型。数据类型反映了该字段存储数据的种类，如字符串、数字和日期时间等。Tableau Desktop 主要有以下五种数据类型：

(1) 字符串(String)类型。字符串是由零个或者多个字符组成的序列，例如，“story book”“abc-12345”都是字符串。字符串通过添加单引号或者双引号进行标识。

(2) 数字(Number)类型。数字可以是整数或浮点数，浮点数即是实数。

(3) 布尔(Boolean)类型。布尔值即是逻辑值，包括 TRUE(真)和 FALSE(假)。例如，表达式 $8 > 6$ 的布尔值为 TRUE，表达式 $8 > 9$ 的布尔值为 FALSE。

(4) 日期/日期时间(Date & Datetime)类型。日期或者日期时间，如“January 12, 2020”或“January 12, 2020 11:45:30 AM”。Tableau 几乎能识别所有格式的日期时间。如果需要将字符串强制识别为日期时间，则在该字符串之前添加一个#符号。

(5) 地图(Map)类型。地图值可以是国家、省和市等，也可以是精确的经度值与纬度值。

在 Tableau Desktop，字段的数据类型在“数据”选项卡中由图标标识，每种数据类型对应的图标如表 5.1 所示。

表 5.1 主要数据类型图标

图 标	类 型	图 标	类 型
Abc	字符串		日期
#	数字		日期时间
T F	布尔		地图

5.1.4 Tableau Desktop 软件界面

Tableau Desktop 的开始界面如图 5.2 所示，左侧的“连接”可以连接 Microsoft Excel、文本文件、JSON 文件等，也可以连接到 Microsoft SQL Server、MySQL、Oracle 等数据库，还可以连接到已保存的数据源。Tableau Desktop 自带有 Superstore、世界发展指标和超市三个数据集。

与数据源建立连接后，Tableau Desktop 进入数据源界面，如图 5.3 所示。该界面给出



图 5.2 Tableau 开始界面



图 5.3 Tableau 数据源界面

了数据源的详细信息,以及其中包含的字段(数据表中的一列变量称为字段)和 1000 行数据。单击左下方的工作表名称可以进入工作簿,即进入 Tableau Desktop 的工作区。在正式介绍工作区环境之前,首先需要了解以下几个基本概念。

- 工作表(worksheet): 又称为视图(visualization),是可视化分析的最基本单元。
- 仪表板(dashboard): 是多个工作表和一些对象(图像、文本、网页和空白等)的组合,可以按照一定的方式对其进行组合和布局,以便揭示数据关系和内涵。
- 故事(story): 是按顺序排列的工作表或者仪表板的集合,故事中各个单独的工作表

或仪表板称为“故事点”。可以使用创建的故事,向用户叙述某些事实,或者以故事的方式揭示各种事实之间的上下文或事件发展的关系。

- 工作簿(workbook): 包含一个或多个工作表,以及一个或多个仪表板和故事,是用户在 Tableau 中工作成果的容器。用户可以把工作表成果组织、保存或发布为工作簿,以便共享和存储。

Tableau Desktop 的工作区包括工作表工作区、仪表板工作区和故事工作区。工作表工作区如图 5.4 所示,该界面的左侧是“数据”和“分析”两个选项卡。“数据”选项卡最顶部显示的是数据源,数据源下方是“维度”和“度量”两个分组,分别用来显示导入的维度字段和度量字段。维度和度量是一种数据角色的划分。该界面的中部从上往下依次为“页面”“筛选器”和“标记”,“标记”中包含了“颜色”“大小”“文本”等按钮。“页面”的右侧是“列”功能区、“行”功能区和视图区。将“数据”选项卡中的字段拖曳到“行”或者“列”功能区时就会在视图区显示相应的轴或者标题。

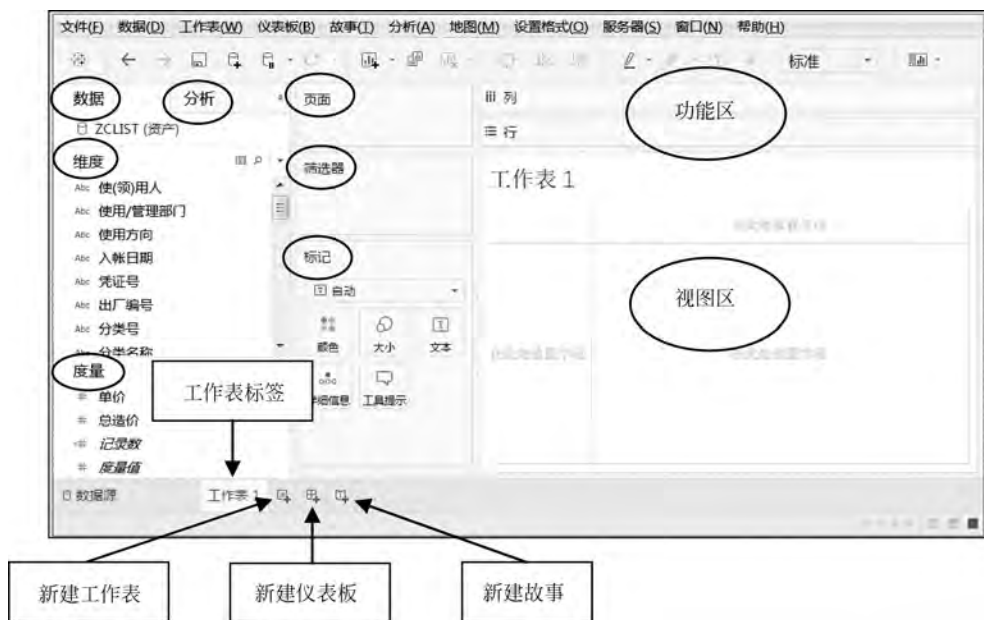


图 5.4 工作表工作区界面

“维度”分组中显示的数据角色被称为维度,往往是具体分类或时间方面定性的离散字段。如果将维度字段拖曳到“行”或“列”的功能区,Tableau 将在视图区创建行或列标题,比如将“使用方向”拖曳到“行”功能区就会出现五种资产的“使用方向”,如图 5.5 所示。

“度量”分组显示的数据角色被称为度量,往往是数值字段,将其拖曳到功能区时,Tableau 默认会进行聚合运算,同时在视图区产生相应的轴,轴上是连续刻度。聚合运算是将多个值聚集为一个数字,如通过求和,求平均数、计数,求最大值或最小值等。

Tableau 连接数据源时会对各个字段进行自动评估,将其分配至“维度”或者“度量”分组。一般而言,这种分配都是正确的,但有时也会出错,比如对于数据源中由一串数字构成的电话号码,Tableau 可能会将这样的字段分配到“度量”中。此时,可以把电话号码从“度量”分组拖曳至“维度”分组,或右击选中“转换为维度”以合理调整数据的角色。



图 5.5 拖曳维度字段到“行”功能区

字段除了可以划分为“维度”和“度量”外,也可以划分为“离散”和“连续”。通常,将字段从“维度”分组拖到“列”或“行”功能区时,该字段的值默认是离散的,Tableau 将创建列或行标题;将字段从“度量”分组拖到“列”或“行”功能区时,该字段的值默认是连续的,Tableau 将创建轴。而且,字段的图标颜色用以区分离散和连续,蓝色代表离散字段,绿色代表连续字段。离散和连续类型可以相互转换,右击字段,在弹出菜单中有离散和连续的选项,单击选择就可实现转换。

5.1.5 文件类型

Tableau 支持以下 6 种专用的文件类型。

(1) 工作簿(twb)文件,它的后缀名为. twb。工作簿可以包含工作表、仪表板和故事等可视化内容,但不包含源数据。

(2) 书签(tbm)文件,它的后缀名为. tbm。书签包含单个工作表,是快速分享所做工作的简便方式。

(3) 打包工作簿(twbx)文件,它的后缀名为. twbx。打包工作簿是一个 zip 压缩文件,包含了所有工作表、连接信息以及所有提供支持的本地资源(本地数据源、背景图像、自定义地理编码等),最适合与不能访问该数据源的其他人分享。

(4) 数据提取(tde)文件,它的后缀名为. tde。该文件是部分或者整个数据源的一个本地副本,可用于共享数据、脱机工作和提供数据库性能。

(5) 数据源(tds)文件,它的后缀名为. tds。该文件是连接经常使用的数据源的快捷方式,不包括实际数据,只包含连接到数据源所必需的信息和在“数据”选项卡中所做的修改。

(6) 打包数据源(tdsx)文件,它的后缀名为. tdsx。该文件是一个 zip 压缩文件,包含数据源文件(tds)和本地文件数据源,可使用此格式创建一个文件,以便与不能访问该数据的其他人共享。

Tableau 默认保存文件至“我的 Tableau 存储库”目录下的关联文件夹中。安装 Tableau 时,在“我的文档”文件夹中自动创建了“我的 Tableau 存储库”。当然,Tableau 的文件也可以指定其他位置保存。



5.2 初级可视化分析

本节将通过实例详细介绍如何使用 Tableau 生成一些简单的图形,如条形图、直方图、饼图、折线图、压力图、树地图和气泡图等。

5.2.1 条形图

条形图又称条状图、柱状图、柱形图,是最常使用的图表类型之一。它通过垂直或水平的条形展示维度字段的分布情况,以每个维度字段的条形长度代表其数据量的大小。水平方向的条形图即为一般意义上的条形图,垂直方向的条形图通常称为柱形图。创建条形图时将维度放在“行”功能区上,并将度量放在“列”功能区上,反之则创建了柱形图。

下面将介绍如何在“资产.xlsx”数据源上创建一个条形图,用于查看每个资产使用人名下的资产情况。打开 Tableau Desktop,连接“资产.xlsx”数据集,在一个空白工作表上按以下步骤操作。

(1) 将“维度”分组中“使(领)用人”拖曳到“列”功能区,“度量”分组中“总造价”拖曳到“行”功能区(显示为“总和(总造价)”),在视图区则生成资产情况垂直条形图(柱形图),如图 5.6 所示。

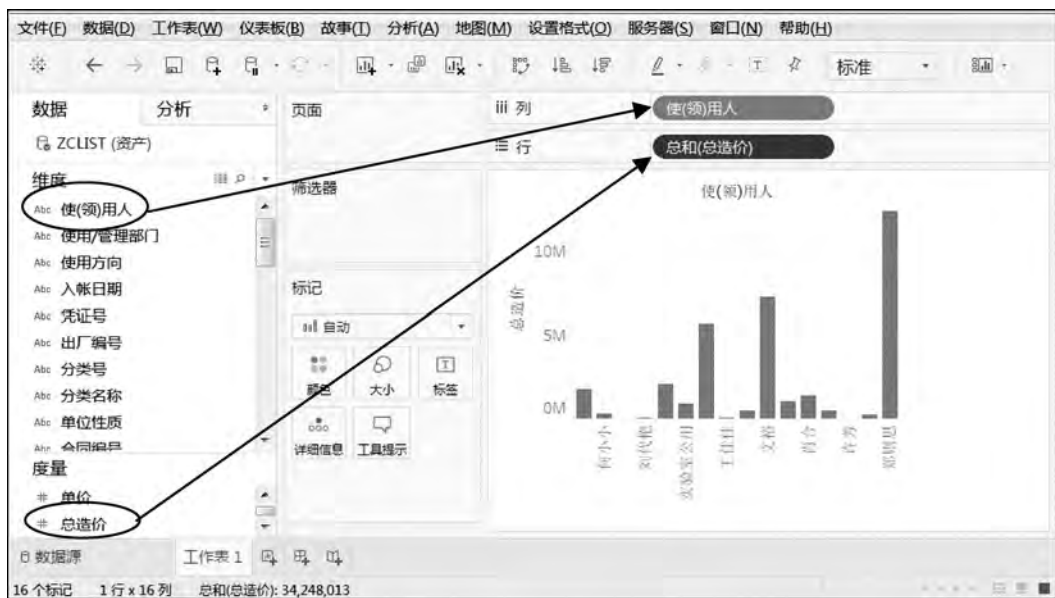


图 5.6 资产情况垂直条形图(柱形图)

(2) 单击工具栏的“交换”按钮,将垂直条形图转置为水平条形图;单击“标记标签”按钮,将显示数据标签;单击“降序”按钮排序,将按总造价的降序排列,如图 5.7 所示。现在可以很直观地看到条形图水平方向显示的是使用人的名称,而垂直方向则是以条



图 5.7 对柱形图进行交换、降序排列

形长度来显示其名下资产总造价的大小。

(3) 若需要进行平均值比较,在“分析”选项卡的“汇总”分组,拖曳“平均线”到视图,放至弹出对话框的“表”位置,此时,在视图中自动生成了一条平均值线,即总造价的平均值为 2 140 501,如图 5.8 所示。平均值线显示了不同使用人的资产造价与平均值的对比情况。右击视图上的平均值线,选择“设置格式”,可以编辑其展示形式。



图 5.8 添加平均值分析

(4) 将“维度”分组中的“生产厂家”拖至“标记”中的“颜色”，生成堆积条形图，可以查看资产使用人按照资产生产厂家统计的造价情况，如图 5.9 所示。



图 5.9 堆积条形图

(5) 当生产厂家字段过多时，生成的堆积条形图不够直观，可以对图中“生产厂家”按照造价的值进行排序。单击“生产厂家”图例的下拉菜单按钮，选择“排序”命令，在弹出的对话框中设置排序，“排序依据”为“字段”，“排序顺序”为“升序”，“字段名称”为“总造价”，“聚合”为“总和”，如图 5.10 所示。



图 5.10 为堆积条形图设置排序

设置完成后，堆积图中的颜色顺序将按照生产厂家总造价的升序排列，如图 5.11 所示。未登记生产厂商供应的资产总造价最大，在条形图的最左方，北泰德公司供应的资产总造价最小；在条形图的最右方，针对某个资产使用人，各个生产厂商向其提供的资产总造价也必须按照总体情况进行排序，而不是单独针对其情况进行排序。



图 5.11 排序后的堆积图

5.2.2 直方图

直方图是一种统计报告图,它是对数据分布情况的图形表示,它的两个坐标分别是统计样本和与该样本对应的某个属性的度量。

直方图与条形图虽然图形效果类似,但是区别在于:条形图的水平轴为单个维度,是通过条形长度来表示度量的多少。而直方图的水平轴是针对某个字段的分组,水平轴宽度表示各组的组距,垂直轴代表每组包含样本数量的多少。条形图往往用于展示离散的且取值较少的维度字段,而直方图则是选取某个字段进行分组统计。分组的原因可能是因为选取的字段是连续的,或者字段虽然离散但是数量过多(可以视为近似于连续),又或者基于某种业务的需要。通常,使用直方图分析的样本数据量最好在 50 个以上。

下面将介绍如何在“资产.xlsx”数据源上创建一个直方图,用于查看总造价小于 10 000 元设备数量的分布情况。为此,首先需要筛选出总造价小于 10 000 元的设备,并创建一个数据桶。数据桶是将需要分析的度量再次细分成同等步长组距并转换为维度的一种方法。打开 Tableau Desktop,连接“资产.xlsx”数据集,在一个空白工作表上按以下步骤操作。

(1) 筛选总造价小于 10 000 的记录,具体操作如下:

① 拖动“度量”分组中的“总造价”到“筛选器”,在弹出的对话框中选择“最大值”,如图 5.12 所示,然后单击“下一步”按钮。

② 在弹出的对话框中,选择“至多”,并在文本框中输入 10 000 作为最大值,如图 5.13 所示,单击“确定”按钮,这样就完成了符合要求的筛选。

(2) 针对总造价小于 10 000 的记录创建数据桶,具体操作如下:

① 右击“度量”分组中的“总造价”,在弹出的菜单中选择“创建”→“数据桶”命令,如图 5.14 所示。

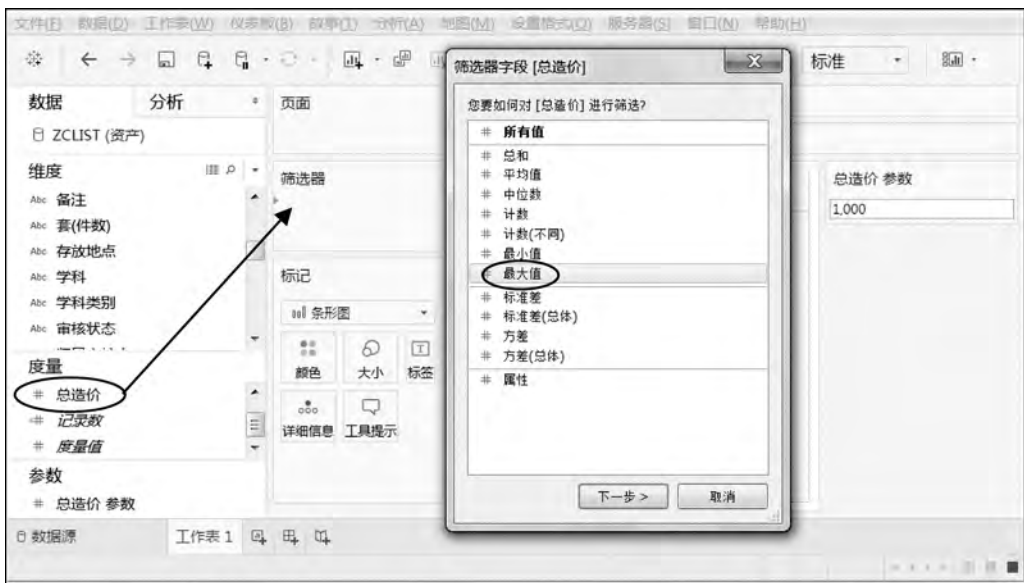


图 5.12 拖动“总造价”到筛选器



图 5.13 设置最大值

② 在弹出的对话框中，“数据桶大小”选择“创建新参数”，如图 5.15 所示。

③ 在弹出的对话框中，“允许的值”选择“全部”，“当前值”填写 1000，如图 5.16 所示，单击“确定”按钮，完成了数据桶参数的设置。

这样就创建了一个符合要求的数据桶，在“维度”分组中新建了“总造价(数据桶)”字段，如图 5.17 所示。

(3) 创建直方图。拖动“维度”分组中“总造价(数据桶)”字段到“列”功能区，拖动度量的“总和(记录数)”到“行”功能区，如图 5.18 所示。从该直方图可知，总造价小于 1000 元的记录有 3007 条，总造价在 1000~1999 元的记录有 1584 条，总造价在 2000~2999 元的记录有 664 条，其他以此类推。



图 5.14 创建数据桶



图 5.15 创建新参数



图 5.16 参数设定



(4) 编辑水平轴刻度的别名。先修改水平轴上第 1 个刻度“0”的别名,右击“0”,在弹出菜单中选择“编辑别名”命令,在弹出的对话框中输入“0-999”,如图 5.19 所示,单击“确定”按钮完成编辑。

完成修改水平轴上第 1 个刻度“0”的别名后,直方图如图 5.20 所示。其他值的修改以此类推。编辑完成后,可以清晰地看到水平轴上的每一个刻度都代表一个区间。

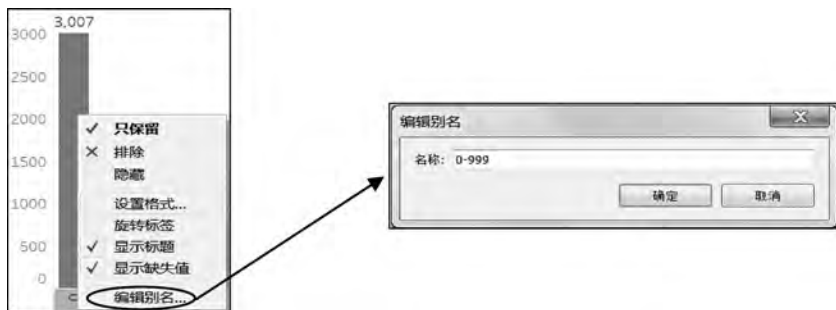


图 5.19 编辑别名

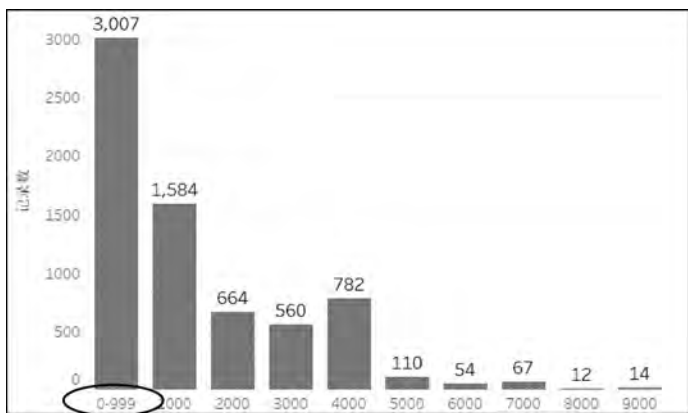


图 5.20 完成编辑第一个值别名

由图 5.20 可知,总造价小于 10 000 元的记录中,绝大部分记录的总造价小于 4000 元,峰值集中于 0~999 元。

5.2.3 饼图

饼图是将数据分类情况表示为不同大小和颜色的圆切片,以此展示各分类所占的比例。使用饼图时,需注意以下几点。

- 圆切片不宜过多,最好不多于 4 块,而且每个切片需占据一定分量比例,这样会显得更为直观。
- 确保各切片的总占比是 100%。
- 避免在切片中使用过多的标签。

下面将介绍如何在“资产.xlsx”数据源上创建一个饼图,用于查看各部门资产总造价的占比情况。打开 Tableau Desktop,连接“资产.xlsx”数据集,在一个空白工作表上按以下步骤操作。

(1) 查看部门。将“维度”分组中的“使用/管理部门”拖曳至“行”功能区展示部门情况,如图 5.21 所示,可以看出部门数量较多,有必要按照部门归属进行分组。

(2) 分组。具体操作如下。

① 右击“维度”分组中的“使用/管理部门”,在弹出菜单中选择“创建”→“组”命令,弹出“创建组”对话框,如图 5.22 所示。

列	
行	使用/管理部门
工作表 1	
使用/管理部门	
电化教育室	Abc
教学实验中心	Abc
实验大楼工程	Abc
实验技术室	Abc
实验技术室(办公)	Abc
网络信息室	Abc
网络学工中心	Abc
学工技术室	Abc
学工技术室(办公)	Abc
综合管理办公室	Abc

图 5.21 使用/管理部门列表



图 5.22 创建组

② 按住 Ctrl 键,选择“电化教育室”“网络信息室”“网络学工中心”,单击“分组”按钮,将它们归为一组,然后输入“网络部门”作为分组名称,如图 5.23 所示。



图 5.23 创建“网络部门”分组

③ 将其余部门按同样操作一一分组,“教学实验中心”“实验技术室”“实验技术室(办公)”归为“实验部门”,“学工技术室”和“学工技术室(办公)”归为“学工部门”,“实验大楼工程”和“综合管理办公室”归为“其他”,如图 5.24 所示。



图 5.24 创建四个分组

④ 单击“确定”按钮,在维度中添加了一个字段“使用/管理部门(组)”,如图 5.25 所示,至此完成了部门分组的创建。



图 5.25 完成创建分组

⑤ 将“维度”分组中的“使用/管理部门(组)”拖曳至“标记”中的“颜色”,并设置标记类型为“饼图”,“标记”中出现了“角度”按钮,如图 5.26 所示。

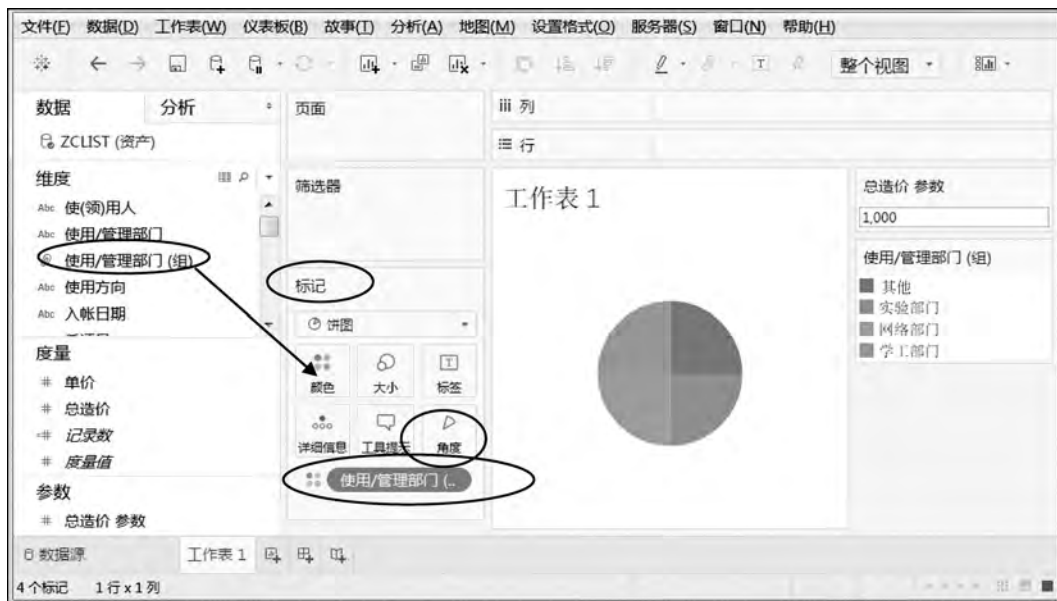


图 5.26 生成饼图样式

⑥ 将“度量”分组中的“总造价”拖至“标记”中的“角度”后,饼图将根据该度量的数值大小改变饼图扇形角度的大小,从而生成占比图。同时,为达到更好的视觉效果,将工具栏中的视图模式由“标准”切换到“整个视图”,如图 5.27 所示。



图 5.27 将“总造价”拖至“角度”

[注意] 创建饼图时“行”与“列”功能区均为空白。

⑦ 为饼图添加标签信息。将“维度”分组中的“使用/管理部门(组)”和“度量”分组中的“总造价”拖动至“标记”中的“标签”,如图 5.28 所示。

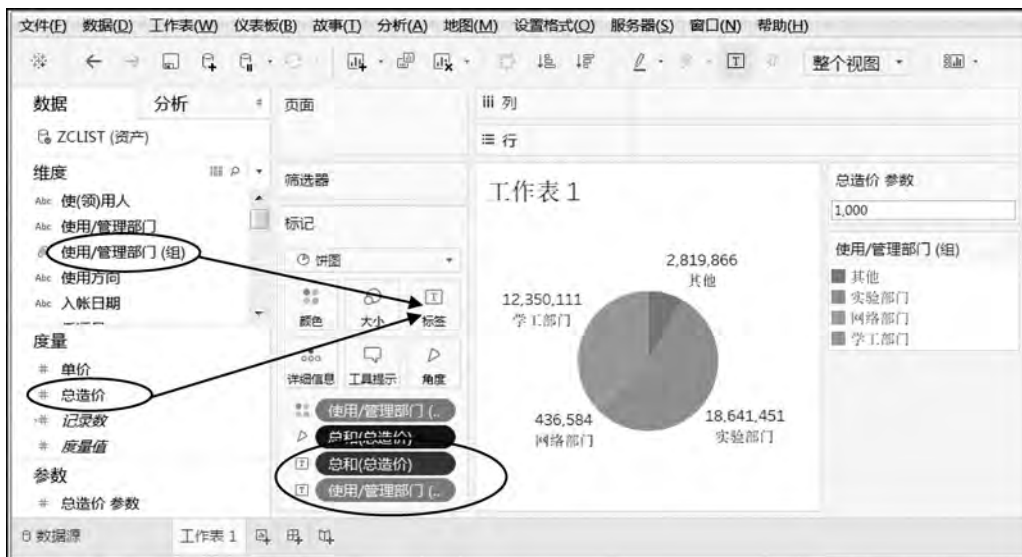


图 5.28 添加标签

⑧ 将“总造价”标签由数值改为百分比。右击“标记”中的“总和(总造价)”，在弹出的菜单中，选择“快速表计算”→“合计百分比”命令，如图 5.29 所示。



图 5.29 设置占比

⑨ 完成以上所有操作后，饼图的效果如图 5.30 所示。此外，为进一步优化展示效果，可在饼图中将各个分组按照总造价的多少排序。

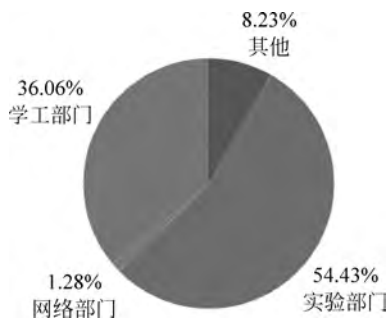


图 5.30 饼图效果

5.2.4 折线图



折线图是用线段将各个数据点连接起来而构成的统计图，它通过折线的上升或下降来表示统计数量的增减变化趋势，适用于描述时间序列数据。与条形图相比，折线图不仅可以表示数量的多少，而且可以直观地反映同一事物随时间序列发展变化的趋势。

下面将介绍如何在“资产.xlsx”数据源上创建一个折线图，用于查看各个年度资产购置情况。打开 Tableau Desktop，连接“资产.xlsx”数据集，在一个空白工作表上按以下步骤操作。

(1) 右击“维度”分组中的“购置日期”，在弹出菜单中选择“更改数据类型”→“日期”命令，将“购置日期”的数据类型由字符串更改为日期，如所图 5.31 示。



图 5.31 更改数据类型

(2) 将维度“购置日期”拖曳至“列”功能区，度量“总造价”拖曳至“行”功能区，生成折线图。将工具栏中的视图模式由“标准”改为“整个视图”，单击工具栏中的“标签”按钮 ，在折线图中显示各年度资产购置的具体数值，如图 5.32 所示。

(3) 部分资产由于种种原因未登记采购日期，在折线图水平轴上表示为 Null，右击该值，在弹出的菜单中选择“排除”将其清除，如图 5.33 所示。



图 5.32 生成折线图

(4) 若要观察各季度或月度的资产购置情况,右击功能区“列”中的“年(购置日期)”,在弹出的菜单中分别选择“季度”或者“月”命令,如图 5.34 所示,这样就可以生成相对应的折线图。

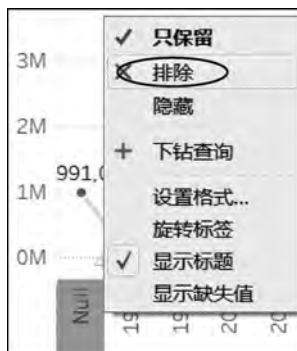


图 5.33 删除 Null 值



图 5.34 设置“季度”或者“月”资产购置情况



图 5.36 添加标签后的压力图



图 5.37 突显表的形式展示数据

上按以下步骤操作。

在“标记”中选择标记类型为“方形”，将“度量”分组中的“总造价”拖曳至“标记”中的“大小”，“记录数”拖曳至“标记”中的“颜色”，“维度”分组中的“使(领)用人”拖曳至“标记”中的“标签”，将“度量”分组中的“总造价”和“记录数”拖曳至“标记”中的“标签”，生成的树地图如图 5.38 所示。

在该树地图中，可以清晰地辨识到矩形的面积大小代表使用人资产总造价多少，而矩形颜色的深浅则表示资产条目数。



图 5.38 树地图的形式展示数据

5.2.7 气泡图

气泡图将数据显示为圆的群集,每个气泡表示维度字段的一个取值,各个气泡的大小及颜色代表了一个或两个度量的值。气泡图的特点是具有视觉吸引力,能够以非常直观的方式展示数据。

下面将介绍如何在“资产.xlsx”数据源上创建一个气泡图,用于查看使用人名下的资产总造价及资产条目数。打开 Tableau Desktop,连接“资产.xlsx”数据集,在一个空白工作表上按以下步骤操作。

(1) “标记”中选择标记类型为“圆”,将“度量”分组中的“总造价”拖曳至“标记”中的“大小”,“记录数”拖曳至“标记”中的“颜色”,“维度”分组中的“使(领)用人”拖曳至“标记”中的“标签”,将“度量”分组中的“总造价”和“记录数”拖曳至“标记”中的“标签”,生成如图 5.39 所示的气泡图。

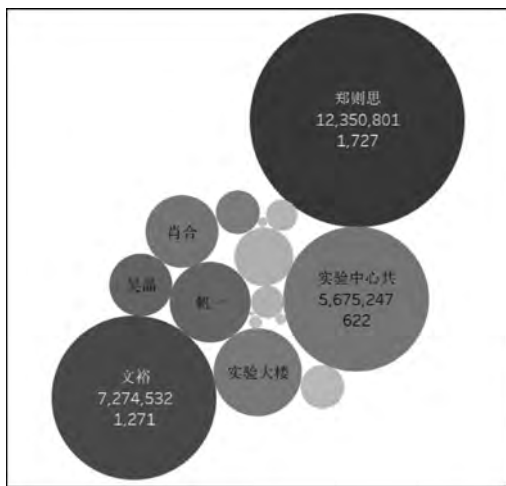


图 5.39 气泡图

(2) 按住 Ctrl 键,单击选择最大的 3 个气泡,然后右击,在弹出菜单中选择“添加注释”→“标记”命令,在弹出的对话框中单击“确定”按钮,如图 5.40 所示,这样就为主要的三个气泡添加了注释。

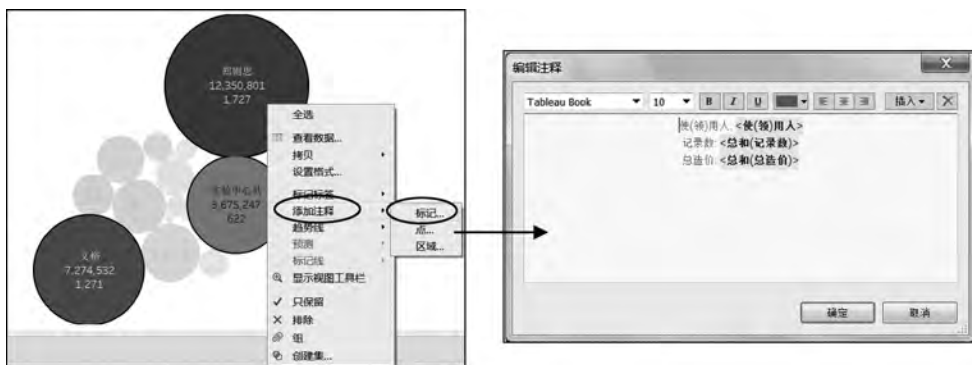


图 5.40 添加注释

完成所有操作后,气泡图效果如图 5.41 所示,在该气泡图中,可以清晰地辨识到气泡的大小代表使用人资产总造价多少,而气泡颜色的深浅则表示资产条目数的多少。

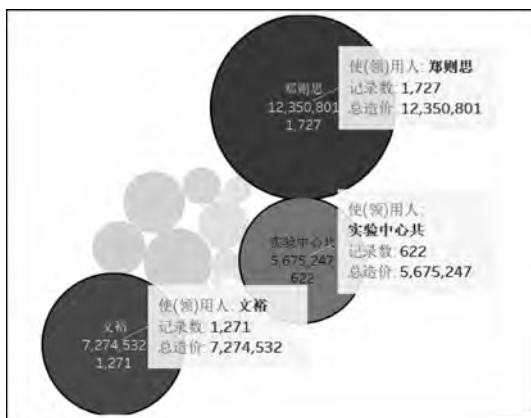


图 5.41 气泡图

5.3 地图分析

5.3.1 认识地图



Tableau 的地图功能十分强大,能够自动识别国家、省/直辖市/自治区、城市等不同等级地理信息的名称、拼音或缩写,且在地图上展示。同时,还可以编辑经纬度信息,实现对地理位置的定制化功能。Tableau 将每一级地理位置信息定义为“地理角色”。“地理角色”包括“国家地区”“省/直辖市/自治区”“城市”“地区”“州”“县”以及“区号”和“邮政编码”。但是,对于中国区域,只有“国家地区”“省”“市/自治区”“城市”有效,如图 5.42 所示。

地理分析的核心是绘制点。地图图像提供背景,坐标绘制在背景上。纬度和经度坐标跟任何其他坐标的点一样。地图上的任何点都可用纬度和经度坐标表示。在 Tableau 中,

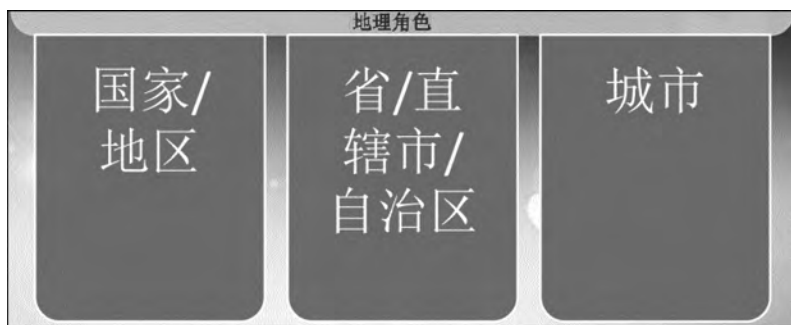


图 5.42 地理角色分类

坐标表示为十进制格式。正纬度表示北半球,正经度表示本初子午线以东。这样,地球上的每个点都有唯一的经纬度坐标。

一般情况,Tableau 连接数据源时,会给地理信息字段自动分配相应的地理角色。此时,该字段在维度窗口显示的图标为 ,表示 Tableau 已自动对该字段进行了地理编码,实现了字段值与纬度值、经度值的关联,并且在度量窗口自动添加了两个字段:“纬度(生成)”和“经度(生成)”。在创建地图时,可以拖曳度量窗口的这两个字段进行展示。但是,Tableau 有时候会把地理信息字段识别为字符串类型,在这种情况下,需要手动为其分配地理角色,具体操作为:在维度窗口中右击该字段,在弹出菜单中选择“地理角色”→“省/市自治区”命令,从而为其分配对应的地理角色,之后该字段的图标将由  变为 ,如图 5.43 所示。



图 5.43 修改地理角色

5.3.2 创建地图

Tableau 创建的地图主要有符号地图、填充地图、多维地图和混合地图 4 种类型,具体介绍如下。

- (1) 符号地图以地图为背景,在对应的地理位置上以多种形状展示信息。
- (2) 填充地图将地理信息作为面积进行填充。
- (3) 多维地图通过对不同维度的信息用多个地图展示,实现信息的分维度比对。多维地图展示要在已经创建好的符号地图或填充地图的基础上进行创建。
- (4) 混合地图是把符号地图和填充地图叠加而形成的一种地图形式。

下面将介绍在“ABC 公司销售数据.xlsx”数据源上创建和设置这四种类型的地图。为方便读者,本书附带的学习资源也包括了该数据集。

1. 符号地图

打开 Tableau Desktop 连接数据源“ABC 公司销售数据.xlsx”,该数据源包含“价位表”“渠道表”“类别表”“销售数据表”4 个数据表。双击打开左方“工作表”窗口的“销售数据表”,如图 5.44 所示。



图 5.44 打开销售数据表

单击左下方的“工作表 1”进入工作表工作区,发现 Tableau 没有将“维度”分组中的“省份”自动识别为“地理角色”,右击“省份”,在弹出菜单中选择“地理角色”→“省市/自治区”命令,如图 5.43 所示。Tableau 将省份转换为地图类型,同时在度量区自动添加“纬度(生成)”和“经度(生成)”两个字段。下面正式创建展示各省毛销售额的符号地图,具体步骤如下。

(1) 初步创建。首先双击“维度”分组中的“省份”,再拖曳“度量”分组中的“毛销售额”至“标记”下的“大小”,在视图区创建一幅符号地图,标识有销售记录的省份,如图 5.45 所示。

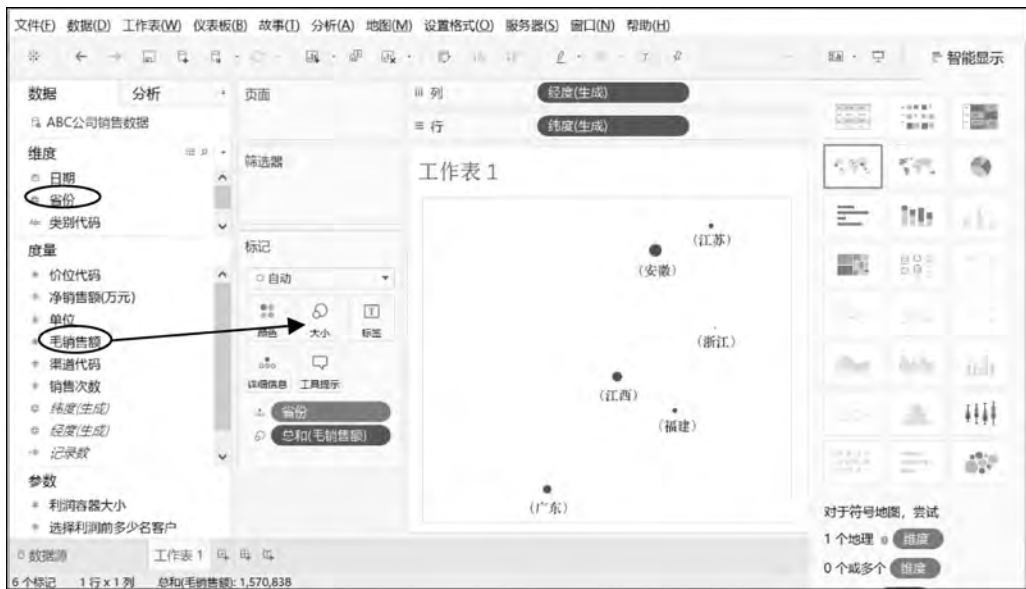


图 5.45 创建符号地图

(2) 添加标签。分别将“维度”分组中的“省份”和“度量”分组中的“毛销售额”拖曳到“标记”中的“标签”，此时各省份的名称及销售量会显示在地图的相应位置，如图 5.46 所示。

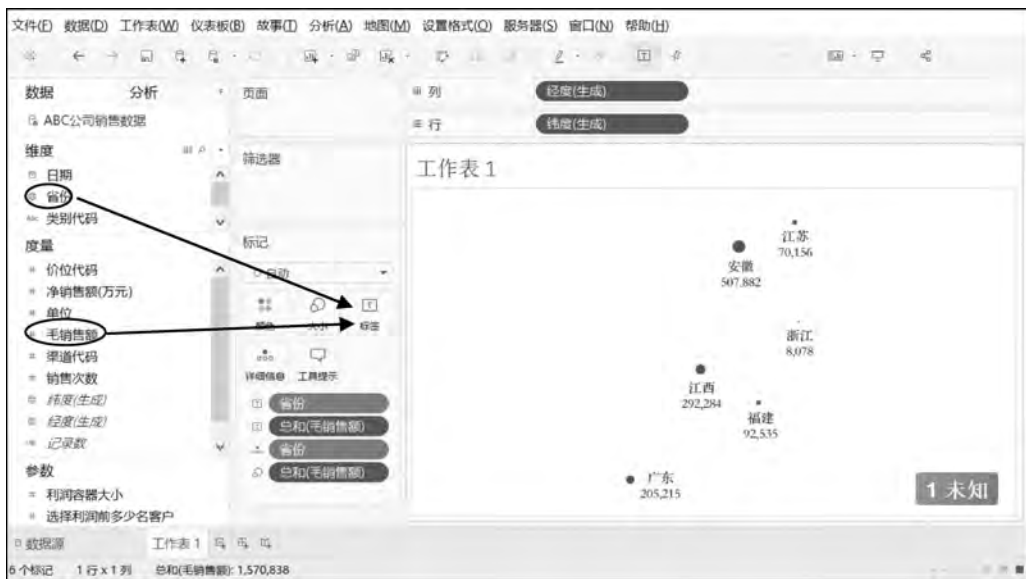


图 5.46 添加标签

(3) 修饰地图层。现在可以对地图上的效果进行设置，如图 5.47 所示。

① 选择菜单栏的“地图”→“地图层”命令，在左侧出现“地图层”窗口，在该窗口的“背景”分组，将地图背景“样式”由默认的“浅色”转换为“普通”。

② 在“地图层”分组，选择“海岸线”和“街道和高速公路”，设置完成后，关闭“地图层”窗格，回到“数据”选项卡。



图 5.47 修饰地图层

(4) 更改“标记”类型。单击“标记”下方的下拉列表框,选择下拉列表中的“形状”选项,则在“标记”中出现了“形状”按钮,单击该按钮,选择三角形 ,则地图上的标签由圆形转换为三角形,如图 5.48 所示。

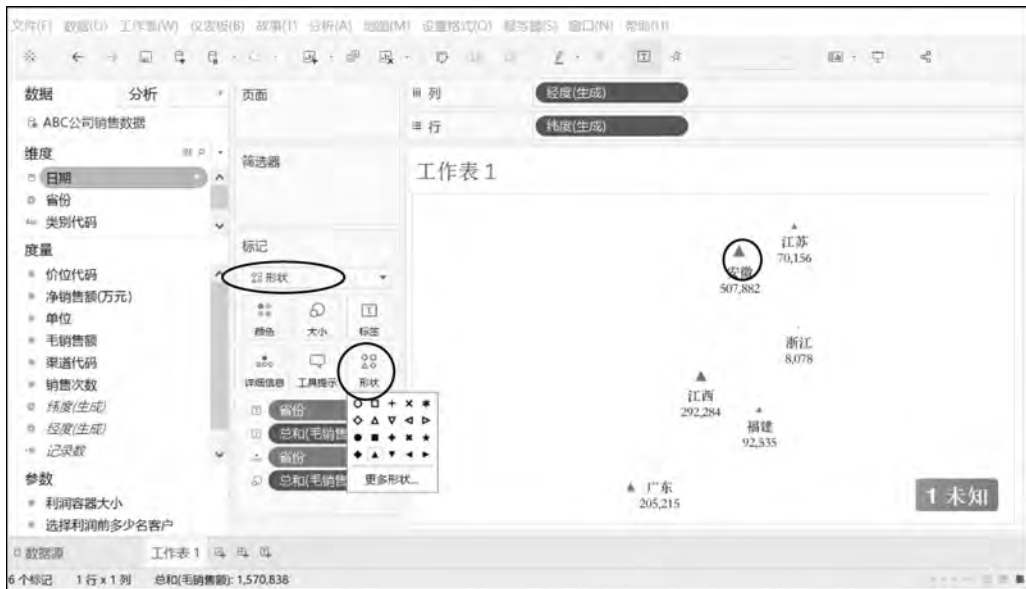


图 5.48 设置标签形状

2. 填充地图

基于创建好的符号地图,可以轻松地实现地图填充。单击“标记”下方的下拉框,选择下拉列表中的“地图”选项,在地图上有销售额的省份被同种颜色填充,如图 5.49 所示。



图 5.49 创建填充地图

如果要按照填充颜色的深浅区分毛销售额的多少,则需要将“度量”分组中的“毛销售额”拖曳到“标记”下的“颜色”,如图 5.50 所示。



图 5.50 区分颜色深浅的填充地图

如果要以不同颜色来填充各个省份,则需要将维度“省份”拖曳到“标记”下的“颜色”,如图 5.51 所示。

观察视图区右下角,发现存在“1 未知”字样,如图 5.51 所示,这意味着存在一个省份未被 Tableau 自动识别。选择菜单栏“地图”→“编辑位置”命令,在弹出的对话框中,发现“齐



图 5.51 颜色不同的填充地图

“鲁”无法识别,未能正确匹配位置,单击“无法识别”,再单击在下拉列表框的倒三角,选择“山东省”与之配对,如图 5.52 所示。

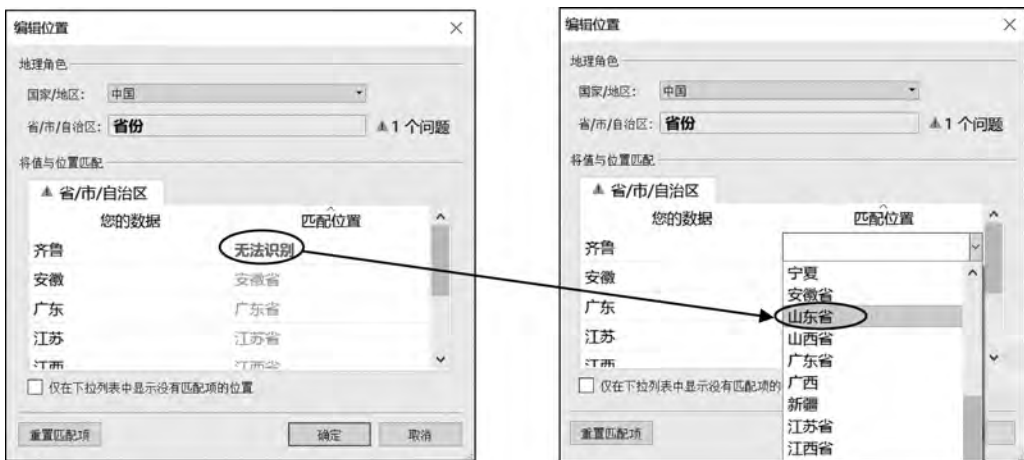


图 5.52 匹配“齐鲁”与“山东省”

手动匹配未自动识别的“齐鲁”后,东部 7 个省份的填充地图成功地展示在视图区,如图 5.53 所示。此外,保存包含该填充地图的工作簿为“符号-填充地图.twb”文件。

3. 多维地图

多维地图需要在符号地图或者填充地图基础上创建。接下来,将以前面创建好的填充地图为基础继续创建多维地图,对比各省份逐年的毛销售额。打开“符号-填充地图.twb”文件,拖曳“维度”分组中的“日期”至“列”功能区,视图区自动创建三幅地图,展示 2009 年、2010 年和 2011 年三年的毛销售额,如图 5.54 所示。



图 5.53 正确匹配所有省份的填充地图



图 5.54 展示 2009—2011 年的毛销售额

根据视图中的三幅地图,可以看出江苏和广东两个省份是在 2010 年打开市场,福建与浙江两个省份是在 2011 年打开市场,但广东的销售增长非常快,2011 年广东省的毛销售额达到七省之首。

如果需要查看每个季度的销售情况,右击“列”功能区中的“年(日期)”,在弹出的菜单中,选择“季度 第 2 季度”,则展示每个季度的毛销售额,如图 5.55 所示。

4. 混合地图

混合地图同样需要在符号地图或者填充地图基础上创建。接下来,将以之前创建好的



图 5.55 展示每个季度的毛销售额

填充地图为基础继续创建混合地图,展示各省份逐年的毛销售额和销售次数。具体操作如下。

(1) 打开“符号-填充地图.twb”文件,将“度量”分组中的“纬度(生成)”再次拖曳到“行”功能区,此时在视图区出现了两幅地图(下方地图为副本),并在“标记”中添加了“纬度(生成)(2)”折叠面板,如图 5.56 所示。



图 5.56 创建地图副本

(2) 调整地图副本设置。单击“标记”中“纬度(生成)(2)”,如图 5.56 所示,展开针对地图副本的“标记”面板,执行以下操作:首先将类型由“地图”修改为“圆”;接着右击“总和(毛销售额)”标签,在弹出的菜单中选择“移除”删去该标签,用同样的方法删去“省份”标签;然后将度量“销售次数”分别拖曳到“标记”下的“颜色”“大小”和“标签”;最后单击“颜色”,在弹出窗口中选择“编辑颜色”,选择更为醒目的“紫色”,如图 5.57 所示。



图 5.57 编辑地图副本颜色

完成此步操作后,视图区如图 5.58 所示。



图 5.58 修改地图副本设置

(3) 合并两个地图。右击“行”功能区中右方的“纬度(生成)”,在弹出的菜单中选择“双轴”,视图中的两个地图合二为一,如图 5.59 所示,从而实现了在一幅地图上同时展示各省份逐年的毛销售额和销售次数。



图 5.59 混合地图

5.4 高级数据操作

本节主要介绍 Tableau 的高级数据操作,如创建和使用分层结构、组、集、参数、计算字段、参考线和参考区间等。本节使用的是 Tableau 自带的“超市”数据源,该数据源文件的默认存储路径为“库\文档\我的文档\我的 Tableau 存储库\数据源\10.5\zh_CN-China\示例-超市.xls”,通过该路径可以找到并打开它,也可以启动 Tableau 时,在图 5.60 界面直接打开它。

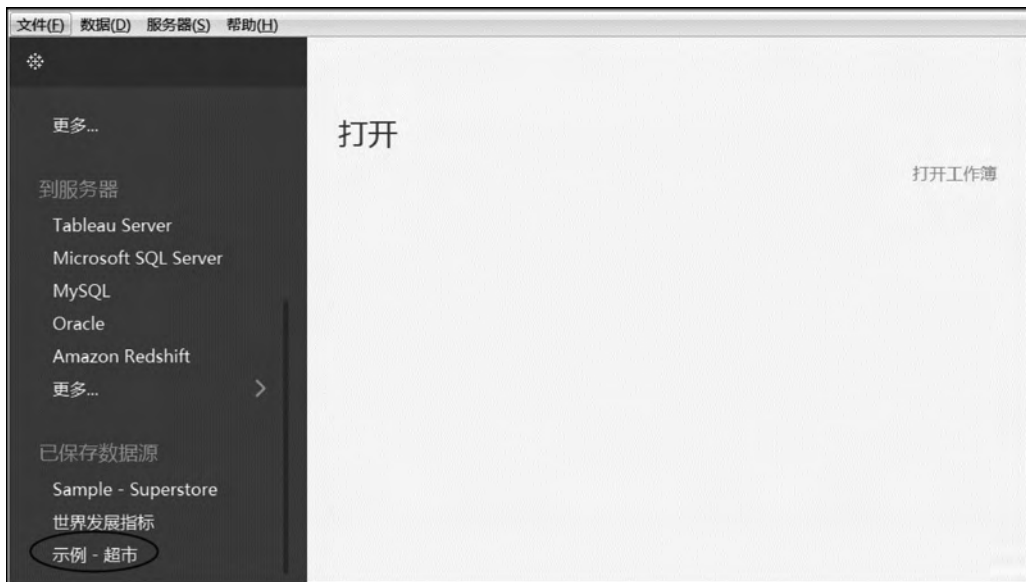


图 5.60 打开“超市”数据源

5.4.1 分层结构

分层结构(hierarchy)是一种维度之间自上而下的组织形式。Tableau 默认了一些字段存在分层结构,比如日期类型的字段有着“年-季度-月-日”的分层结构。除此之外,Tableau 允许用户针对维度字段自定义分层结构,分层结构将显示在维度窗口,字段图标为 。

打开“超市”数据源,发现已经有“产品”和“地点”两个分层结构。“地点”分层包括“国家-省/自治区-城市”的结构,如图 5.61 所示。“国家”包含了若干个“省/自治区”,“省/自治区”又包含了若干个“城市”。“国家”是大类,“省/自治区”是“国家”下的小类,“城市”是“省/自治区”下的小类,这样的分层结构体现了数据粒度的逐步细化和深入。



图 5.61 “地点”分层结构

分层结构通过重新组合维度字段之间上下层关系,进而实现向上钻取(drill up/roll up)和向下钻取(drill down/roll down)功能。例如,在查看不同地点的利润时,单击“列”功能区的  按钮,可以下钻查看各个省/自治区的利润;单击“列”功能区上的  按钮,上钻查看各个国家的利润,如图 5.62 所示。

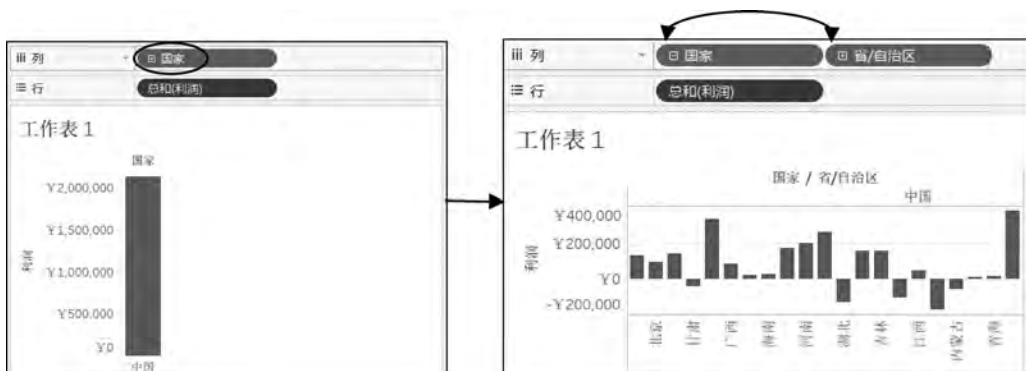


图 5.62 “地点”分层结构向下钻取

下面观察“产品”分层结构,该分层包括“类别-子类别-制造商-产品名称”的结构,如图 5.63 所示。

为更好地观察“产品”分层中的各个字段的含义,分别拖动“类别”“子类别”“制造商”和“产品名称”到“行”功能区,视图区呈现的内容分别如图 5.64 所示。

“类别”包含了若干个“子类别”,“子类别”包含了若干个“制造商”,“制造商”包含了若干个“产品名称”,如此分层结构体现了分类的逐步细化和深入,是一种金字塔的结构。

在实际中,当不需要“产品”分层结构时,只需要右击维度“产品”,在弹出的菜单中选择“移除分层结构”即可删去。

分层结构可以实现向上钻取和向下钻取的功能,那么如何创建一个分层结构呢?下面以重建“产品”分层结构来说



图 5.63 “类别”分层结构

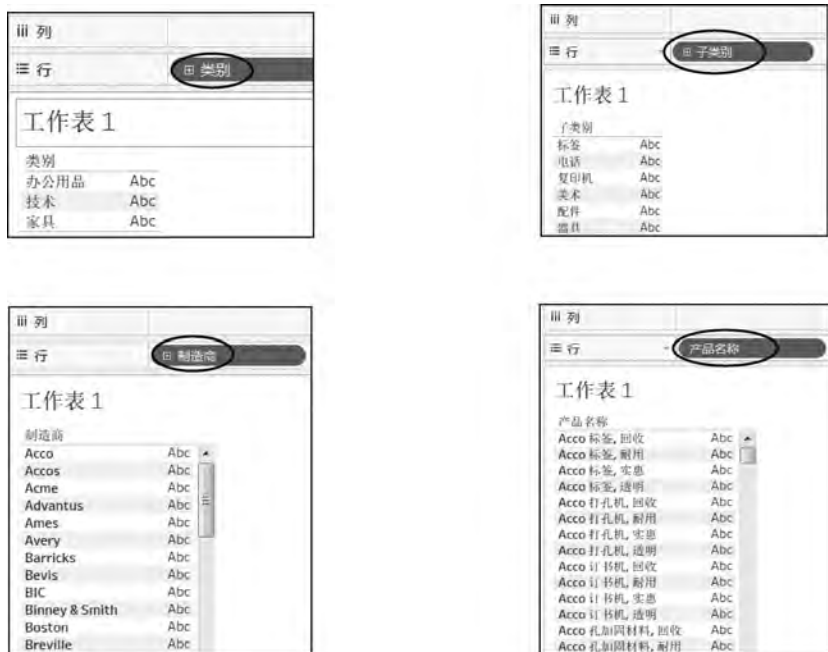


图 5.64 “产品”分层中的各个字段的取值

明,步骤如下。

(1) 右击维度“类别”,在弹出菜单中选择“分层结构”→“创建分层结构”命令,在弹出的对话框中输入分层结构的名称“产品”,再单击“确定”按钮,在维度中新添加了“产品”分层,如图 5.65 所示。

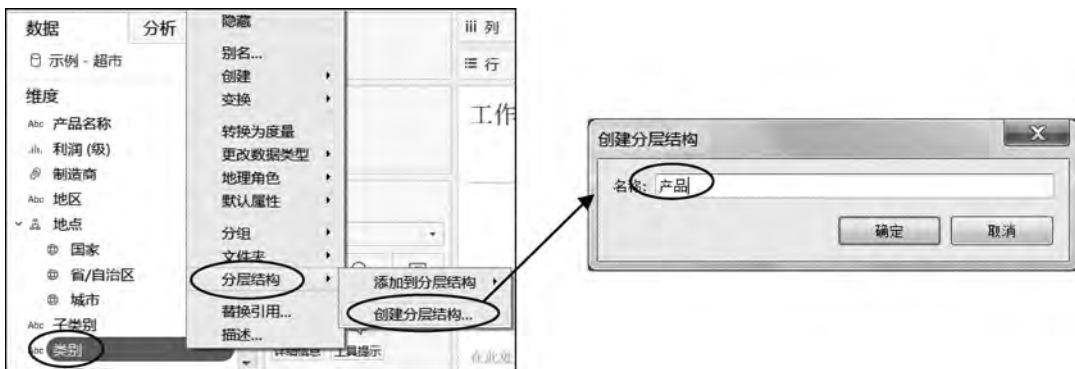


图 5.65 创建“产品”分层结构

(2) 右击“维度”分组中的“子类别”,在弹出的菜单中选择“分层结构”→“添加到分层结构”→“产品”命令,那么“子类别”就添加到了“产品”分层,如图 5.66 所示。以此方法,将“制造商”和“名称”添加到“产品”分层中。

根据创建的分层结构,单击功能区或“标记”中的  或者  按钮可以轻松完成向下钻取或者向上钻取的工作。根据“产品”分层结构可以实现由第一层“类别”一直下钻到第 4 层“产品名称”,如图 5.67 所示。



166



图 5.68 创建“北方”分组

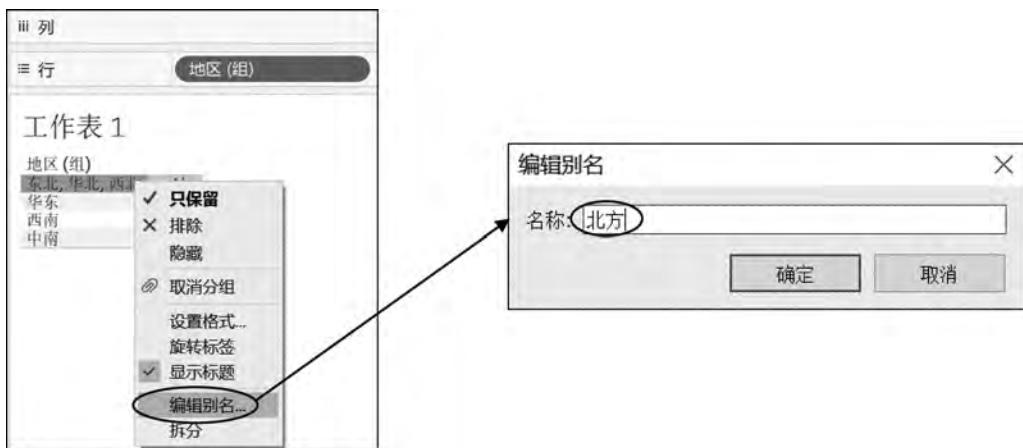


图 5.69 编辑别名

了“地区(分组)”字段,右击该字段,在弹出的菜单中选择“重命名”,输入“北方 or 南方”,完成后如图 5.70 所示。

(4) 单击工具栏中的“清除工作表”按钮 。在空白的工作表上,首先依次双击“维度”分组中的“销售额”“利润”“北方 or 南方”;然后在“标记”中将类型由“自动”改选为“圆”;再按住 Ctrl 键,分别拖曳“标记”下方的“北方 or 南方”到“标签”和“颜色”;进而单击“标记”中的“大小”,向右侧拖动弹出的滑块,使得视图区中标记数据点的圆变大一些,如图 5.71 所示。

依据创建的可视化图,可以清楚观察到南方在销售额和利润两方面都在一定程度上优于北方。

下面将介绍在“超市”数据源上对“维度”分组中的“地区”利用拖曳实现分组,用于比较旺区和淡区的销售额和利润,旺区是销售额和利润较好的区域,淡区则是销售额和利润较差的区域。在一张空白的工作表上,进行如下步骤操作。

(1) 第(1)步的操作类似于图 5.71 展示北方和南方的销售额和利润的操作,只是操作的维度由分组“北方 or 南方”改变为字段“地区”。具体操作如下:首先依次双击“维度”分

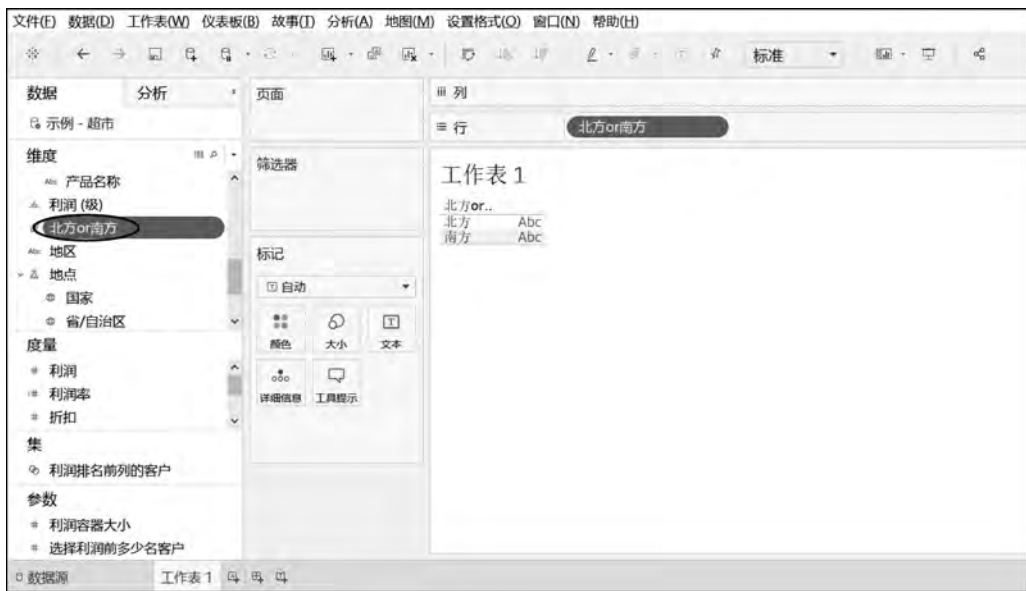


图 5.70 完成创建“北方 or 南方”分组

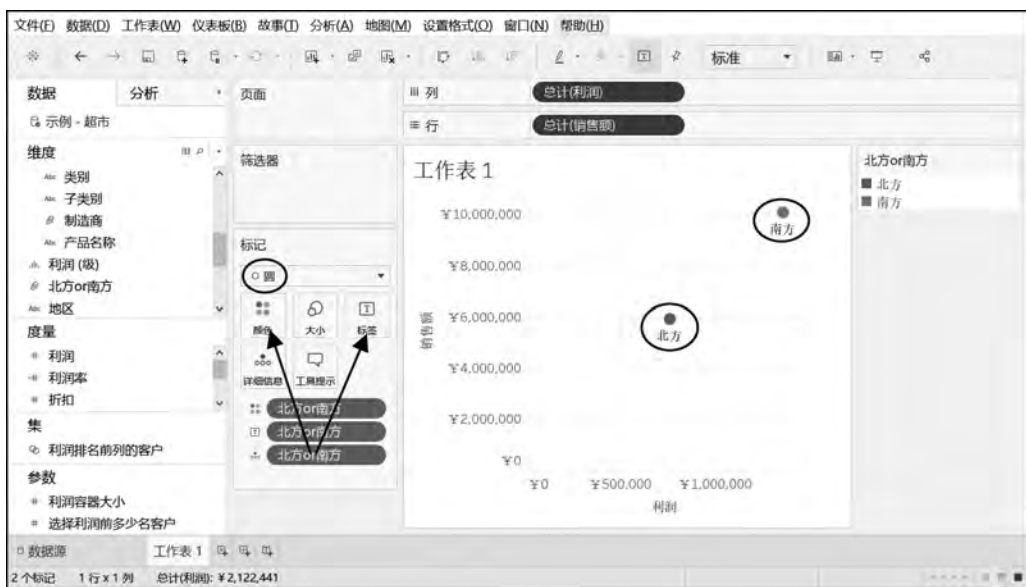


图 5.71 展示北方和南方的销售额和利润

组中的“销售额”“利润”“地区”；接着将标记类型由“自动”改选为“圆”；然后按住 Ctrl 键，分别拖曳“标记”下方的“地区”到“标签”和“颜色”；最后单击“标记”中的“大小”，向右方拖动弹出的滚动条，使得视图区中标记数据点的圆变大一些。完成这一系列操作后，视图区如图 5.72 所示。

(2) 经过观察，“华东”“中南”“东北”“华北”4 个区域具有较好的销售额和利润，这是“旺区”，按住左键拖曳出一个矩形框选定以上 4 个区域，右击，在弹出的菜单中选择“组”命

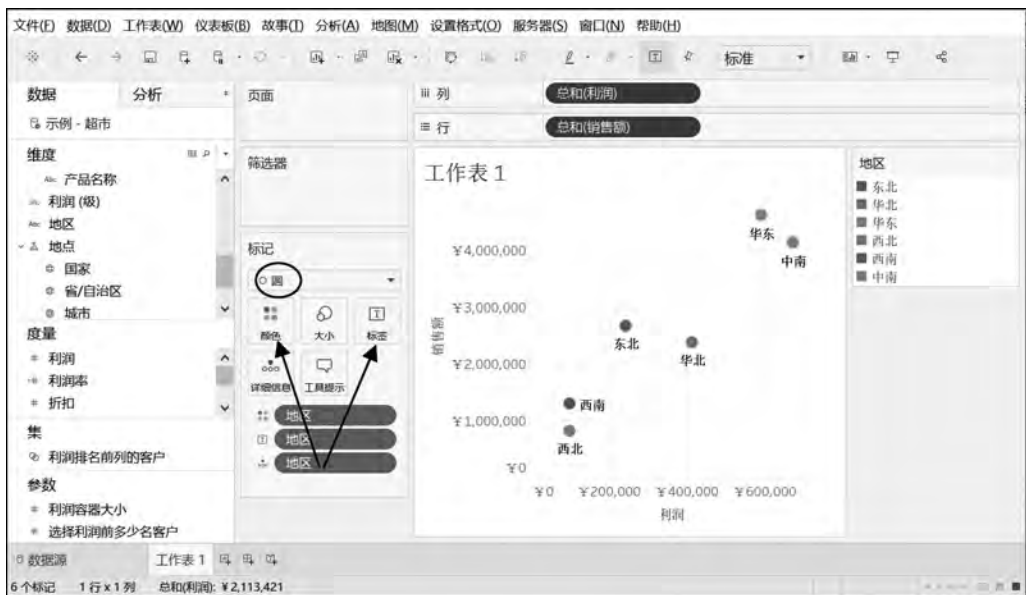


图 5.72 展示各个地区的销售额和利润

令,在维度中添加了“地区(组)1”,这样就创建了“旺区”组别,其他两个区域属于“淡区”组别,两个分组在视图区中标记数据的圆的颜色也有深浅的分别,如图 5.73 所示。

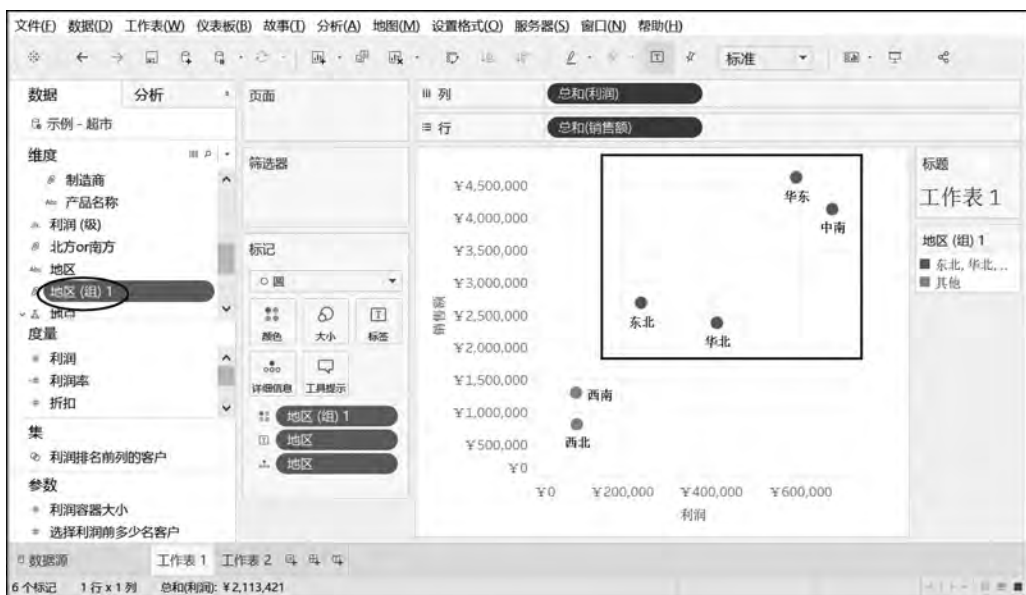


图 5.73 创建“旺区”和“淡区”分组

(3) 右击“维度”分组中的“地区(组)1”,在弹出的菜单中选择“编辑组”命令,在弹出的对话框中,将“字段名称”修改为“旺区 or 淡区”,右击“东北,华北,华东 和 1 以上”,在弹出的菜单中选择“重命名”将其重命名为“旺区”。同样,将“其他”重命名为“淡区”。单击“确定”按钮,这样就完成了利用拖曳创建分组,如图 5.74 所示。



图 5.74 编辑组

5.4.3 集

集(set)是根据某些字段条件定义的数据子集。集统一显示在“数据”选项卡底部,使用  作为图标。集可以基于计算添加建立,也可以参与计算字段的编辑。

创建集类似于进行数据筛选,用于选择满足给定字段条件的记录作为数据子集,以实现对不同记录的选取。创建的集主要有以下两方面作用。

(1) 对比和分析集内外记录。Tableau 提供了集的一对特性——内/外(in/out),通过选择“在集内/外显示”可以直接对集内和集外记录通过聚合进行对比分析。

(2) 分析集内记录。当只分析集内记录时,可选择“在集内显示成员”,集的作用就是筛选,只展示属于集内的记录。

下面将介绍在“超市”数据源上分析和比较高利润客户和低利润客户的销售额。根据业务情况,高利润客户定义为利润大于等于 250 元的客户,其他则是低利润客户。具体步骤如下。

(1) 右击“维度”分组中的“订单日期”,在弹出的菜单中选择“创建”→“集”命令,如图 5.75 所示。



图 5.75 创建集

(2) 在弹出的“创建集”对话框中,首先在“名称”中输入“高利润客户”,然后选择“条件”选项卡,在该选项卡中选择“按字段”,再选择“利润”“平均值”“>=”,以及在文本框中输入“250”,这样就设定了挑选高利润客户的条件“平均利润大于等于 250 元”,如图 5.76 所示。

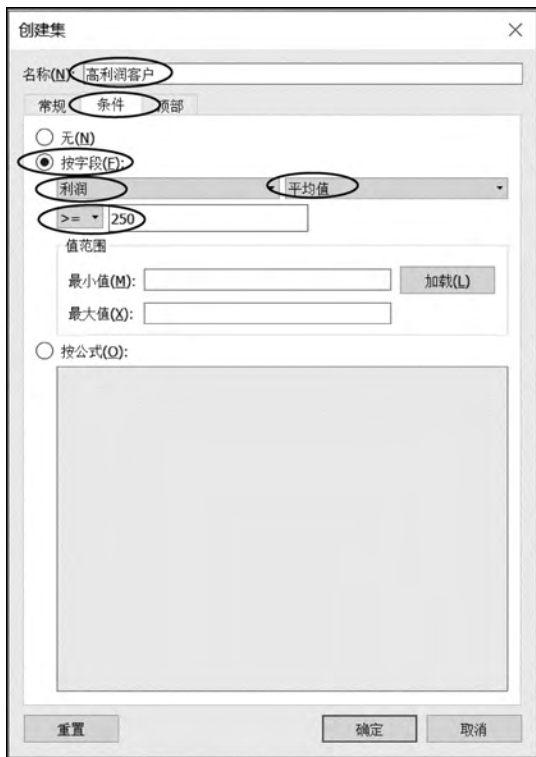


图 5.76 “创建集”对话框

(3) 双击“维度”分组中的“订单日期”,以及拖动“销售额”到“行”功能区;然后右击“列”功能区的“年(订单日期)”,在弹出的菜单中选择“月 2015 年 5 月”,如图 5.77 所示。



图 5.77 选择“月”

(4) 最后,在“数据”选项卡的“集”分组中双击“高利润客户”,产生的折线图如图 5.78 所示。

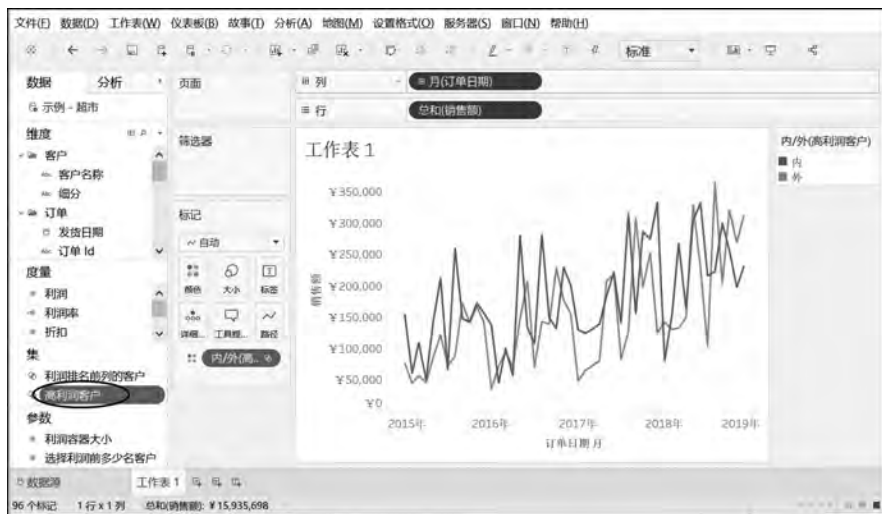


图 5.78 内/外(高利润客户)销售额折线图

分析图 5.78 视图区的“内”表示高利润客户，“外”表示低利润客户。总体而言，高利润客户和低利润客户的销售额整体都是逐步递增的，但高利润客户的销售额逐步低于低利润客户的销售额，说明了其销售额逐步上升，但利润在逐步下降。

5.4.4 参数



参数(parameter)是由用户自定义的,可在集、筛选器、计算集、计算字段中替换常量值的动态值。用户通过控制和调整参数,能够快速和轻松地与工作表视图进行交互,从而实现图表动态分析。参数在工作簿中是全局变量,可以同时运用于多个工作表。

下面介绍在“超市”数据源上展示排名前 N 名的销售总额, N 是 1~10 的整数。步骤如下。

(1) 在“数据”选项卡的“参数”分组,右击任意一个字段,在弹出的菜单中选择“创建”→“参数”命令,如图 5.79 所示。



图 5.79 创建参数

在弹出的“创建参数”对话框中,首先在“名称”中输入“top”,然后“数据类型”选择“整数”,“当前值”输入 10,“允许的值”选择“范围”,“最小值”输入 1,“最大值”输入 10,“步长”输入 1,这样就设定了“创建参数”对话框,如图 5.80 所示。



图 5.80 设置“创建参数”对话框

在图 5.80 所示的对话框中,主要是在“创建参数”时需要提供的一些设置,尤其需要如下说明。

- 注释: 输入对参数意义的描述,以帮助理解所设参数的含义,此处非强制项,可不填写。
- 属性: “数据类型”用于设置参数值接受的数据类型;“当前值”用于设置参数的默认值;“显示格式”用于设置在参数字段中数值的显示格式。“允许的值”用于设置参数接受值的方式,包括 3 种类型:“全部”表示参数可以调整为任意值;“列表”表示参数只能设置为列表内的值,有 3 种设置方法,分别是“手动输入”“从字段中添加”或“从剪切板粘贴”;“范围”表示参数可在指定范围内进行调整,可设置最小值、最大值和每次调整的步长。

(2) 创建“销售额 TOP10”集。右击“维度”分组中的“订单日期”,在弹出的菜单中选择“创建”→“集”命令,在弹出的“创建集”对话框中,首先在“名称”中输入“销售额 TOP10”,然后单击选择“顶部”选项卡,在该选项卡中选择“按字段”,再选择“顶部”“top”,选择依据的字段是“销售额”,聚合方式“总和”,这样就设定了挑选销售总额排名前 top 位客户的条件,单击“确定”按钮即成功创建了集,如图 5.81 所示。

(3) 在一张空白工作表中,首先双击“维度”分组中的“订单日期”,以及拖动“销售额”到“行”功能区;然后右击“列”功能区的“年(订单日期)”,在弹出的菜单中选择“月 2015 年 5 月”;下一步,双击“集”分组中的“销售额 TOP10”,则在视图区生成一个折线图,如图 5.82 所示。

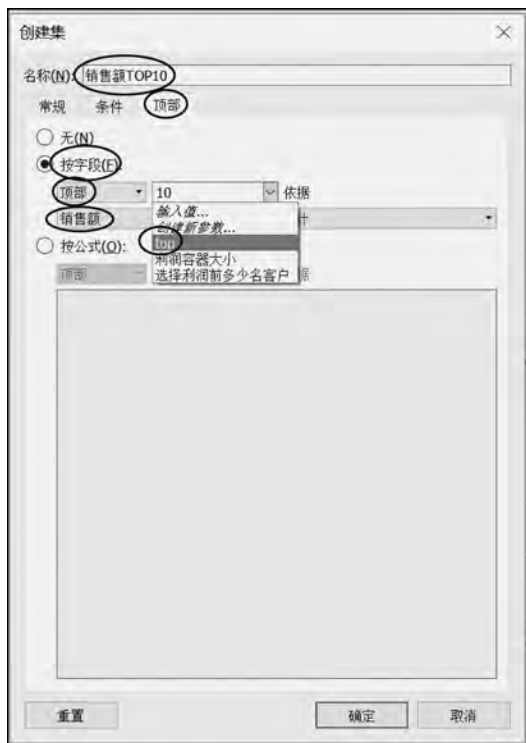


图 5.81 “销售额 TOP10”集对话框

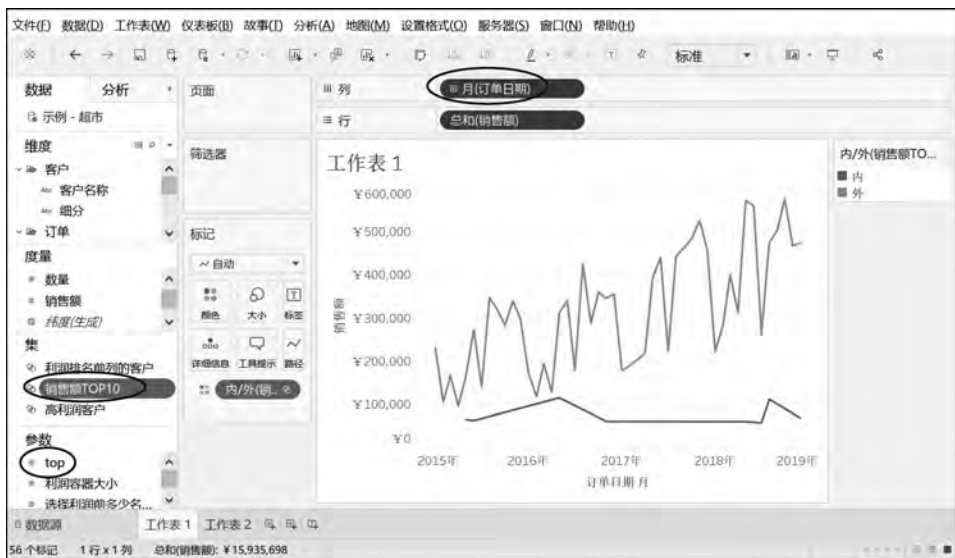


图 5.82 内/外(销售额 TOP10)销售额折线图

(4) 在“数据”选项卡的“参数”分组,右击参数“top”,在弹出的菜单中选择“显示参数控件”命令,此时参数控件将显示在视图区域的右上角。通过在参数控件中调整 top 的值,可动态观察销售额折线的变化。图 5.83 是 top=5 时所呈现的折线图。

图 5.84 是 top=9 时所呈现的折线图。

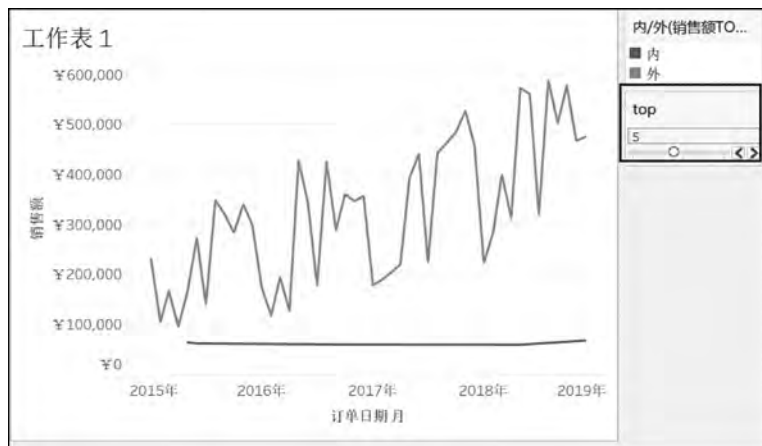


图 5.83 top=5 时的折线图

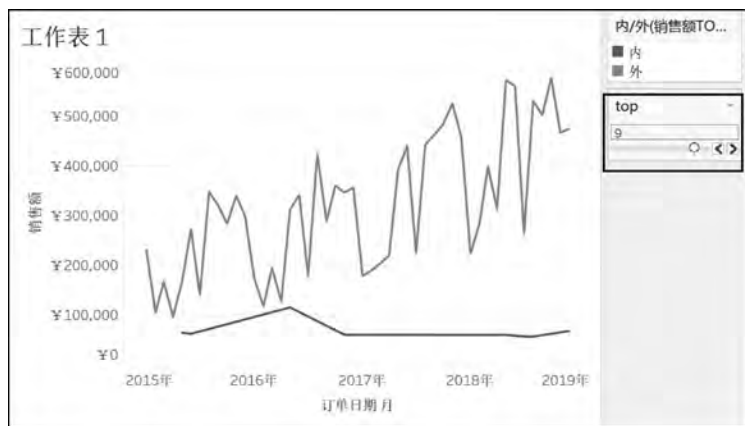


图 5.84 top=9 时的折线图

5.4.5 计算字段

计算字段(calculated field)是根据维度、度量、参数等,使用函数和运算符构建公式来定义的新字段,它的计算结果(字段值)也具有某种数据类型。与其他字段一样,计算字段也能拖曳到各功能区去创建视图,还能用于创建新的字段。

下面将介绍在“超市”数据源上创建一个计算字段“让利金额”,用于展示销售额前 N 名的让利总额, N 是 1~10 的整数,按照业务逻辑,让利金额定义为销售额 $\times$ 折扣,具体步骤如下。

(1) 右击“度量”分组中的“折扣”,在弹出的菜单选择“创建” $\rightarrow$ “计算字段”命令,打开对话框,如图 5.85 所示。

在图 5.85 所示界面的左方是输入窗口,右方是函数窗口。在输入窗口的上方文本框中输入计算字段名称。在输入窗口的下方文本框中,可输入计算公式,包括运算符、计算字段和函数。运算符支持算数运算符[加(+)、减(-)、乘($\times$)、除(/)等]、关系运算符和逻辑运算符等。字符、数字、日期/时间、集、参数等字段均可作为计算字段。函数是 Tableau 自带



图 5.85 创建计算字段

的,实现某种特定功能。包括数字、字符串、日期、类型转换、逻辑、聚合和表计算 7 大类型,大部分 Tableau 函数的功能与 Excel 函数功能类似。

所有的函数按照类型呈现函数窗口,在函数窗口中双击函数即可以出现在输入窗口,也可以直接在输入窗口中输入。Tableau 具有自动填写功能,即在输入时会自动提示和填充可使用计算的字段名称和函数名称。在图 5.85 对话框中,首先在上方的文本框中输入计算字段名称“让利金额”。然后在下方文本框的“[折扣]”之后输入乘号“*”,再拖动度量“销售额”到乘号之后。最后,单击“应用”或者“确定”按钮,度量窗口添加了一个“让利金额”字段。

(2) 首先双击维度“订单日期”,以及拖动“让利金额”到“行”功能区;然后右击“列”功能区的“年(订单日期)”,在弹出菜单中选择“月 2015 年 5 月”;双击“集”分组中“销售额 TOP10”,则在视图区生成一个折线图。

(3) 在“数据”选项卡的“参数”分组,右击参数“top”,在弹出菜单中选择“显示参数控件”命令,此时参数控件将显示在视图区域的右上角。通过在参数控件中调整 top 的值,可动态观察让利总额折线的变化,top=10 时的视图如图 5.86 所示。

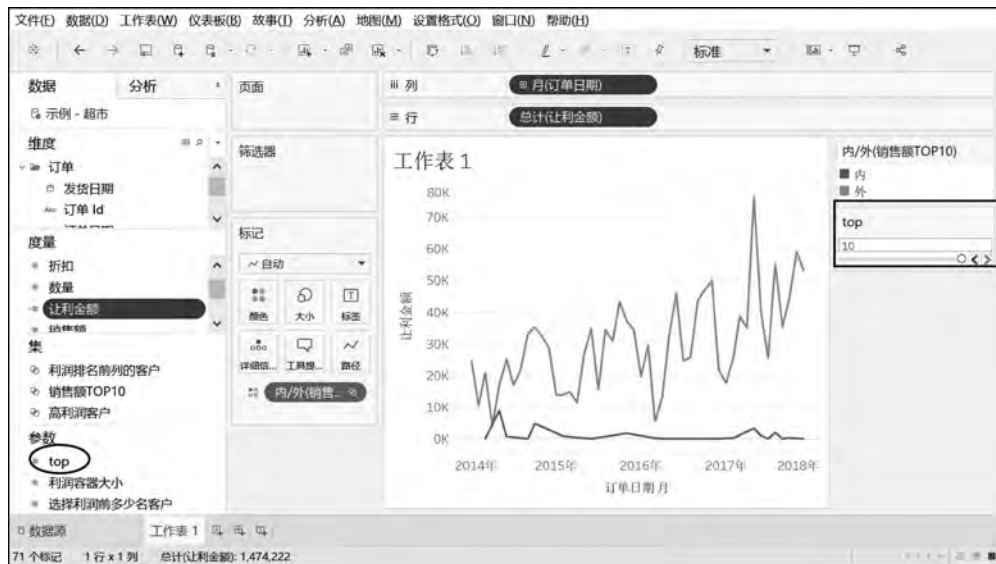


图 5.86 显示参数控件

由于销售额 Top N 的销售总额比较小,因此让利总额也小。如果右击“行”功能区“总计(让利金额)”,在弹出菜单中选择“度量(总和)”→“平均值”命令,“行”功能区就变成“平均值(让利金额)”,如图 5.87 所示。

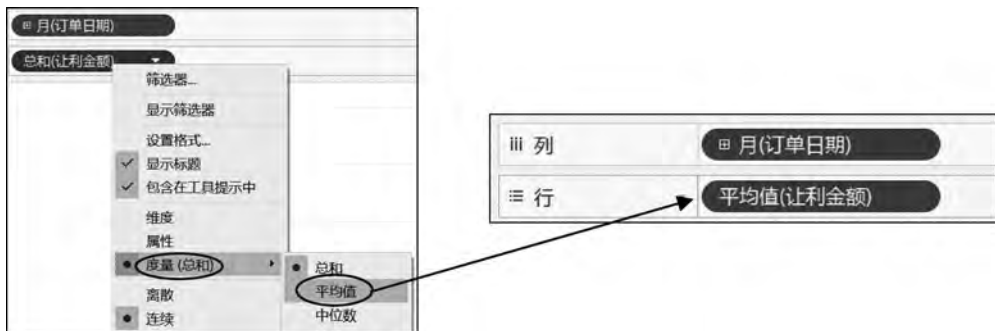


图 5.87 设置“平均值(让利金额)”

通过图 5.88,可以观察到销售总额排名前 10 位客户的平均让利金额波动非常大,而其他客户的平均让利金额波动很小,非常平稳,而且平均让利金额都比较低。

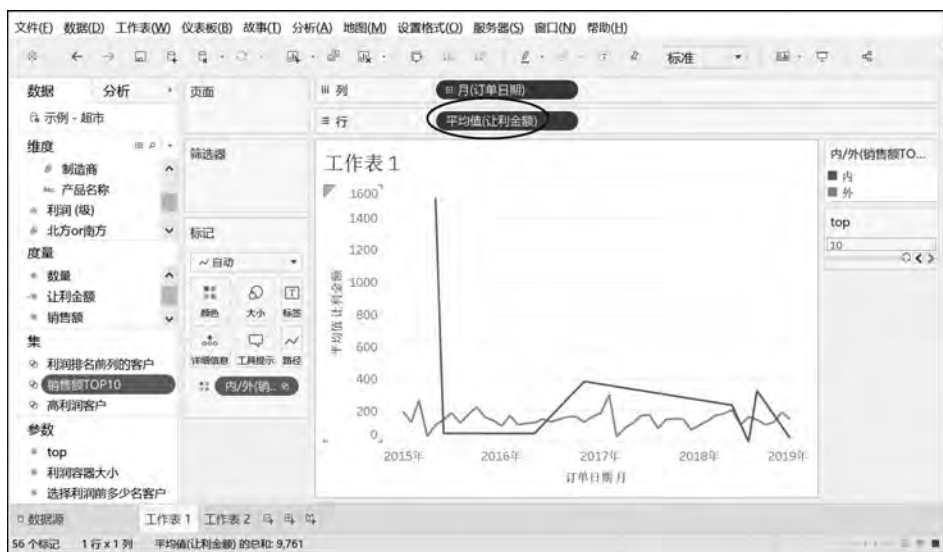


图 5.88 平均让利金额折线图

5.5 分析图表整合

本节将简单介绍 Tableau 仪表板的功能。仪表板能按照一定的方式组合和布局多个工作表以及支持对象(图像、文本、网页和空白等),并可完成添加表间筛选、网页链接、突出显示等交互式操作,以便实现关键数据的对比和分析结论的展示。仪表板的运用非常广泛,上至高管驾驶舱,下至日常工作汇报,仪表板都是一种常用的展示形式。

单击 Tableau 工作界面左下方的“新建仪表板”按钮 ,即可创建一个空白的仪表板,



如图 5.89 所示,仪表板工作区左侧上方是“仪表板”和“布局”选项卡;“仪表板”选项卡包括“大小”“工作表”“对象”3 个分组;“布局”选项卡包括“选定项”和“顶分层结构”两个分组。右侧大片区域是视图区。各个部分具体介绍和基本操作如下。

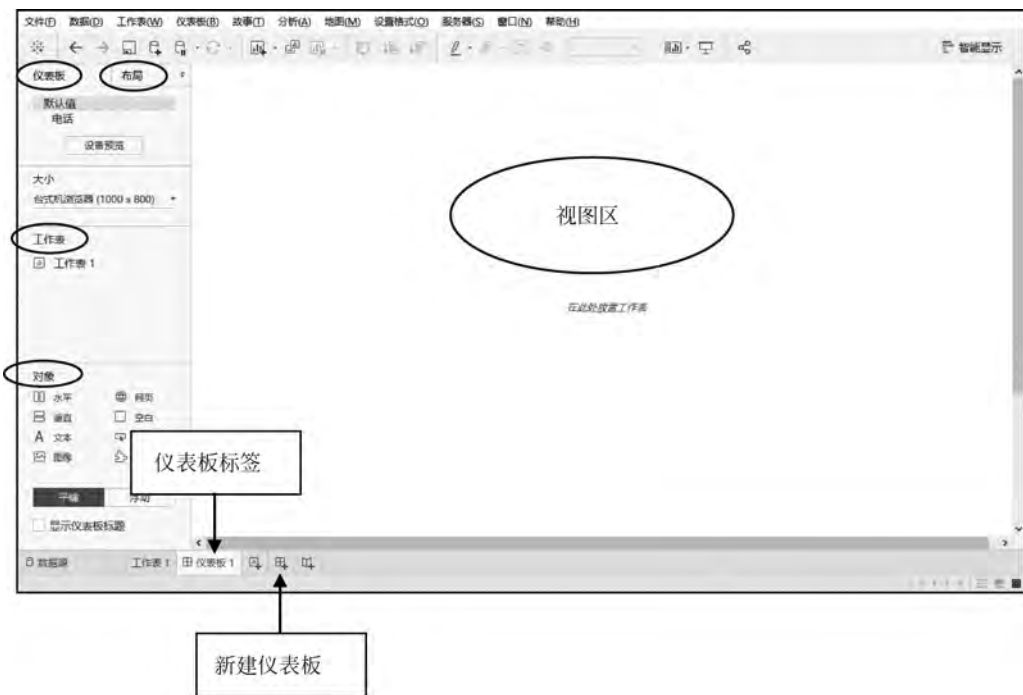


图 5.89 仪表板工作区

(1) “大小”分组。该分组可用于调整仪表板整体的大小,以及视图中各工作表或对象的大小和位置。仪表板默认大小为“台式浏览器(1000×800)”,即宽度 1000px、高度 800px。如图 5.90 所示,确定仪表板大小有 3 种方式:第 1 种为“固定大小”,指仪表板宽度和高度始终保持固定,这需要预先获知展示设备屏幕的大小;第 2 种为“自动”,指仪表板自动填充整个窗口,实现自动排版;第 3 种为“范围”,指设定仪表板中所有对象缩放展示的最大值和最小值。

(2) “工作表”分组。该分组列出在当前工作簿中的所有工作表,新建工作表后,仪表板分组会自动更新。

(3) “对象”分组。该分组包括容器和对象的创建和设置。容器是仪表板布局的框架,分为“水平”和“垂直”两种,可用于组织仪表板中的工作表对象,新增容器会在仪表板中创建一个区域。“对象”是除工作表外可用于展示的元素,包括“图像”“文本”“网页”“空白”等。



图 5.90 确定仪表板大小的 3 种方式

此外,该分组还用于调整工作表或对象的布局方式,包括“平铺”和“浮动”两种。“平铺”是指所选工作表或者对象平行分布而不相互覆盖,Tableau 会根据仪表板大小自动分配工作表和对象的大小;“浮动”是指所选工作表或者对象会相互覆盖地展示,用户可手动调整它们的大小和位置。

(4) “选定项”分组。该分组的“位置”指的是工作表或者对象的左上角在仪表板坐标中的坐标位置,该坐标轴以视图区坐上角为原点,x 代表水平坐标的值,y 代表垂直坐标的值;布局窗口的“大小”指的是工作表或者对象的宽度和高度,“宽”代表的是长度,“高”代表的是高度。图 5.91 所示的图表位置在 $x=0,y=0$; 宽为 330px,高为 190px。除了输入 x 和 y 坐标值改变图表在仪表板中的位置外,还可以单击选择图表,把鼠标放至图表上边框图标  处,待鼠标变成十字架时,就可以拖动它到仪表板的其他位置。此外,单击选择图表,把鼠标放至图表右下角,当鼠标变成双箭头斜杠时,拖动鼠标也能改变图表的大小。单击图表右上角的  按钮可以将其清除。

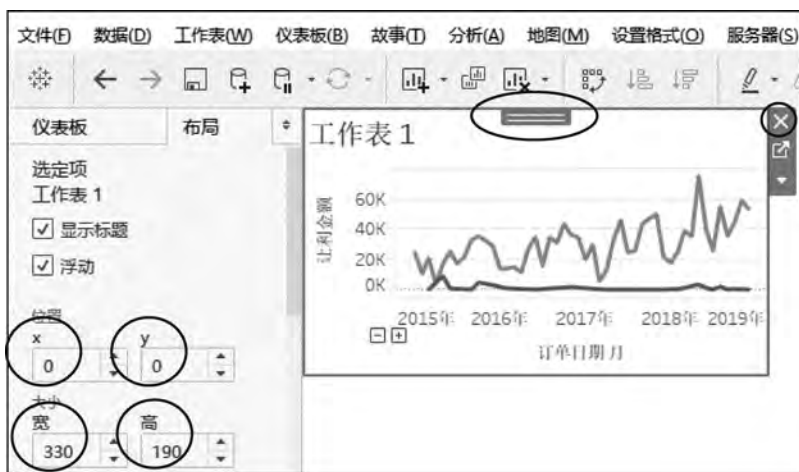


图 5.91 位置和大小

(5) “顶分层结构”分组。该分组展示了视图中各个工作表和对象的层级树形结构,如图 5.92 所示。

(6) 视图区。该区域是用于创建和修饰仪表板的工作区域,可以添加工作表以及图像、文本、网页和空白等对象。之前介绍“计算字段”时创建了一个“销售额排名前 N 名的让利总额”折线图,该折线图所在工作表名称为“工作表 1”,下面以此为基础简单介绍仪表板的创建过程,具体步骤如下。

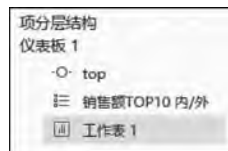


图 5.92 层级树形结构

① 单击左下方的“新建仪表板”按钮 , 创建仪表板。仪表板大小固定,自定义为 1709×902,即宽度为 1709px,高度为 902px。

② 在“对象”分组设置以“浮动”的方式添加对象,以方便在仪表板上排版。

③ 将“工作表”的“工作表 1”拖动到仪表板视图区,如图 5.93 所示,切换到“布局”选项卡,设定“位置”中的 x 值为 400,y 值为 300,“大小”中的“宽”值为 800,“高”值为 300。

④ 将“图例”和“显示参数控件”拖曳至图表左上角合适的位置。

完成以上操作,单击工具栏中的演示模式按钮 , 屏幕如图 5.94 所示。

本节简单介绍了仪表板的基本操作和功能,后续结合实际案例将会进行更详细地介绍。

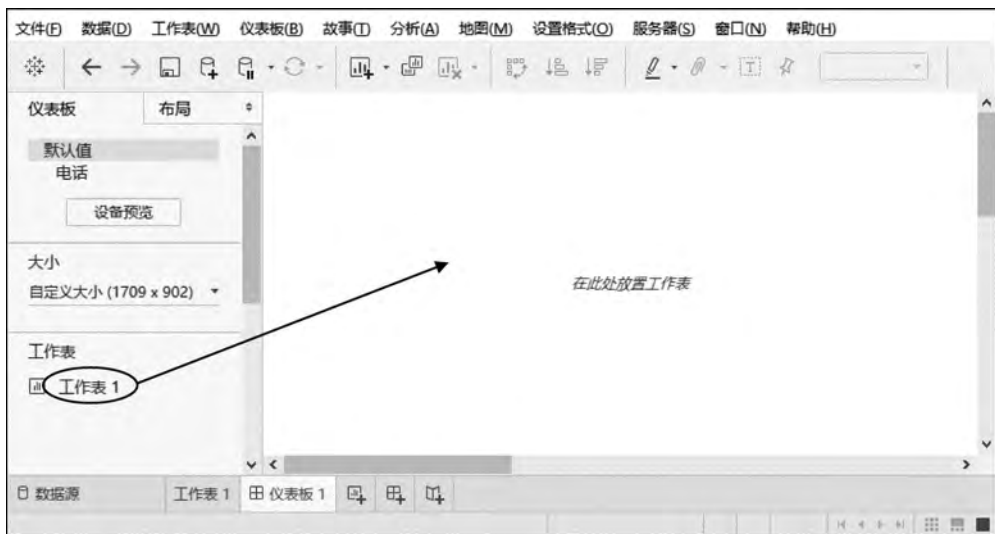


图 5.93 拖动工作表至视图区

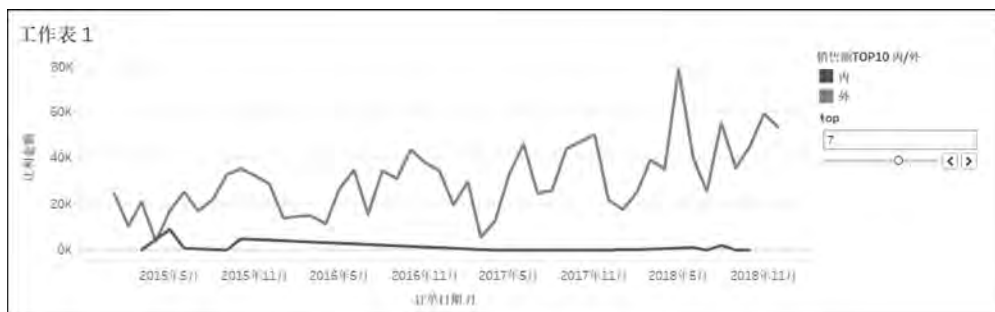


图 5.94 演示仪表板

5.6 案例一：无锡市宜居时间分析



本节将介绍一个综合案例,该案例是分析无锡市的宜居情况,数据来自无锡市气象局,选取时间是 2017 年。为方便读者,本书附带的学习资源也包括了该数据集,即“2017 年无锡市天气数据. hyper”文件。在进行数据分析时,往往需要注意两点。

- 综合考虑分析的目的、获取的数据、业务逻辑,确定分析的思路 and 角度。
- 根据分析思路 and 角度,选取并使用数据集中的相关数据,注意并非一定用到所有获取到的数据。

打开 Tableau Desktop,连接“2017 年无锡市天气数据. hyper”数据集。该数据集包含日期、年份、平均 AQI(空气质量指数)、天气状况、节假日、季节、风向、风向变化、风力、平均气温、最低气温、最高气温等字段。

经过观察获取的所有数据,同时依据宜居的一般标准,确定分析思路为从空气质量和气温(人体感知)两个角度确定无锡市宜居的时间段。

5.6.1 创建计算字段

本小节将分别创建“空气质量等级”“气温感知”“温差”3个计算字段。

1. 创建“空气质量等级”计算字段

数据集中,衡量空气质量的是“平均 AQI”字段,但该字段的取值是连续的,不方便统计。为此,根据我国空气质量指数 AQI 分级标准,先在“平均 AQI”字段的基础上创建离散的计算字段“空气质量等级”。我国空气质量指数 AQI 具体分级标准如表 5.2 空气质量等级标准所示。

表 5.2 空气质量等级标准

平均 AQI(空气质量指数)	空气质量等级	平均 AQI(空气质量指数)	空气质量等级
0~50	I 级,优	201~300	IV 级,重度污染
51~100	II 级,良	大于 300	V 级,严重污染
101~200	III 级,轻度污染		

在创建“空气质量等级”字段时,需要使用 Tableau 的 If 函数。该函数提供了一种多分支的结构,建立 3 个分支的语法格式为: IF test1 THEN value1 ELSEIF test2 THEN value2 ELSEIF test3 THEN value3 END。在这行简单的代码中,test1、test2 和 test3 为逻辑表达式,运算结果为 TRUE 或者 FALSE。首先测试逻辑表达式 test1,如果 test1 的运算结果为 TRUE,则整个 IF 函数的结果为 value1;否则进一步测试逻辑表达式 test2,如果 test2 的运算结果为 TRUE,则整个 IF 函数的结果为 value2;否则进一步测试逻辑表达式 test3,如果 test3 的运算结果为 TRUE,则整个 IF 函数的结果为 value3。如果需要建立更多的分支,格式以此类推。

Tableau 的函数与 Excel 的函数在定义和使用上非常相似,因此,基本上可以按照 Excel 的函数去理解和使用 Tableau 的函数。使用 Tableau 函数时有两点需要特别说明:一是不区分英文字母的大小写;二是必须使用英文输入状态下的各种符号。

在明确创建“空气质量等级”计算字段思路后,下面进行具体操作,步骤如下。

(1) 右击左下方“工作表 1”,在弹出的菜单中选择“重命名工作表”命令,将“工作表 1”重命名为“空气质量”,如图 5.95 所示。

(2) 右击“度量”分组中的“平均 AQI”字段,在弹出的菜单中选择“创建”→“计算字段”命令,如图 5.96 所示。

(3) 在弹出对话框的文本框中输入计算字段名称“空气质量等级”,在输入窗口的下方文本框中输入以下语句:

```
IF AVG([平均 AQI])<= 50 THEN 'I 级'
ELSEIF 50 < AVG([平均 AQI]) and AVG([平均 AQI])<= 100 THEN 'II 级'
ELSEIF 100 < AVG([平均 AQI]) and AVG([平均 AQI])<= 200 THEN 'III 级'
ELSEIF 200 < AVG([平均 AQI]) and AVG([平均 AQI])<= 300 THEN 'IV 级'
ELSE 'V 级'
END
```



图 5.95 重命名工作表

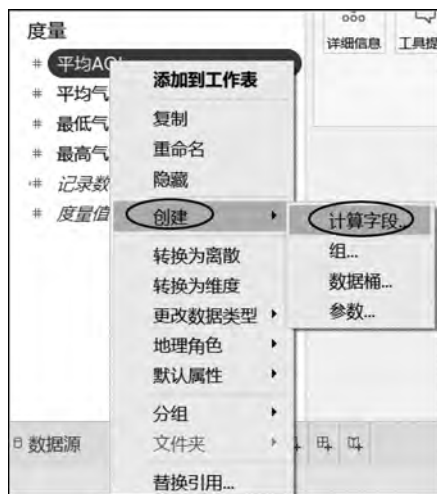


图 5.96 创建计算字段

完成语句输入且检查无误后,单击“确定”按钮,如图 5.97 所示,则在度量中产生了一个新的字段“空气质量等级”。

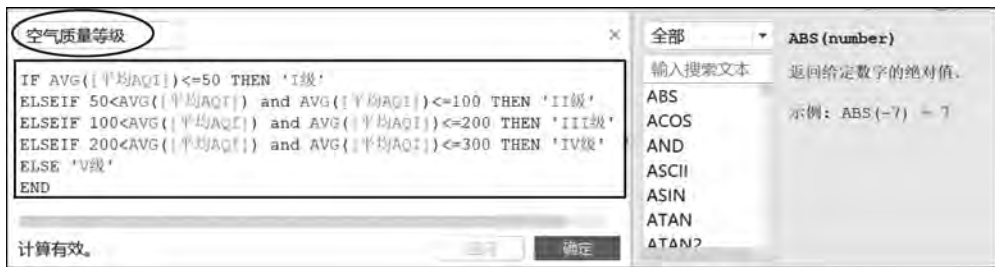


图 5.97 输入空气质量等级公式

2. 创建“气温感知”计算字段

数据集有“平均气温”“最高气温”“最低气温”3 个与气温有关的字段,而气温与人体感知往往有对应关系(如表 5.3 所示)。为了将气温与宜居建立联系,先在“平均气温”字段的基础上创建计算字段“气温感知”。

表 5.3 气温与人体感知对应关系

气温/℃	人体感知	气温/℃	人体感知
< 0	冻	>= 18 且 < 25	适宜
>= 0 且 < 10	冷	>= 25 且 < 30	热
>= 10 且 < 18	凉	>= 30	炎热

在明确创建“气温感知”计算字段思路后,下面进行具体操作,步骤如下。

(1) 右击“度量”分组中的“平均气温”字段,在弹出的菜单中选择“创建”→“计算字段”命令,如图 5.98 所示。



图 5.98 创建计算字段

(2) 在弹出对话框的文本框中输入计算字段名称“气温感知”，在输入窗口的下方文本框中输入以下语句：

```
IF AVG([平均气温])<0 THEN '冻'
ELSEIF 0<= AVG([平均气温]) and AVG([平均气温])<10 THEN '冷'
ELSEIF 10<= AVG([平均气温]) and AVG([平均气温])<18 THEN '凉'
ELSEIF 18<= AVG([平均气温]) and AVG([平均气温])<25 THEN '适宜'
ELSEIF 25<= AVG([平均气温]) and AVG([平均气温])<30 THEN '热'
ELSE '炎热'
END
```

完成语句输入且检查无误后，单击“确定”按钮，如图 5.99 所示，则在“度量”分组中产生了一个新的字段“气温感知”。



图 5.99 输入气温感知公式

3. 创建“温差”计算字段

温差也是一个影响宜居的重要因素。数据集有“最高气温”和“最低气温”两个字段。温差=最高气温-最低气温。依照此公式，下面将创建“温差”计算字段，步骤如下：

(1) 右击“度量”分组中的“最低气温”字段，在弹出的菜单中选择“创建”→“计算字段”命令。

(2) 在弹出对话框的文本框中输入计算字段名称“温差”，在输入窗口的下方文本框中输入以下公式：

$AVG([最高气温]) - AVG([最低气温])$

完成公式输入且检查无误后，单击“确定”按钮，如图 5.100 所示，则在“度量”分组中产生了一个新的字段“温差”。



图 5.100 输入温差公式

5.6.2 空气质量

本节将创建折线图，展示无锡市空气质量在各个月份的总体情况，步骤如下：



图 5.101 重命名工作表

(1) 连接“2017 年无锡市天气数据. hyper”数据集，进入工作表工作区，右击左下方“工作表 1”，在弹出的菜单中选择“重命名工作表”命令，将“工作表 1”重命名为“空气质量”，如图 5.101 所示。

(2) 生成折线图。拖曳“维度”分组中的“日期”到“列”功能区，右击“年(日期)”，在弹出的菜单中选择“月 2015 年 5 月”，修改日期的粒度为月；拖曳“度量”分组中的“平均 AQI”到“行”功能区，右击“平均值(平均 AQI)”，在弹出菜单中选择“度量(总和)”→“平均值”命令，如图 5.102 所示。

完成此步骤操作后，产生的折线图如图 5.103 所示。

(3) 修改垂直轴结束点。双击视图区中的垂直轴，在弹出的对话框中，“范围”选择“固定”，“固定结束”的值设定为 120，如图 5.104 所示。关闭对话框，使得折线图在坐标轴上稍微下移，从而方便后续的操作，也更为美观。

[注意] “固定结束”的值设定为 120 是经过调试得来。

(4) 修饰折线的颜色和粗细。具体操作如下。

- ① 按住 Ctrl 键，拖曳“行”功能区的“平均值(平均 AQI)”到“标记”中的“颜色”。
- ② 单击“标记”中的“颜色”，在弹出的窗口中选择“编辑颜色”，然后在弹出对话框的“色板”中选择“红色”，如图 5.105 所示。
- ③ 单击“标记”中的“大小”，向右拖动滑块到合适位置，把折线粗细调整合适，如图 5.106 所示。



图 5.102 修改为平均值

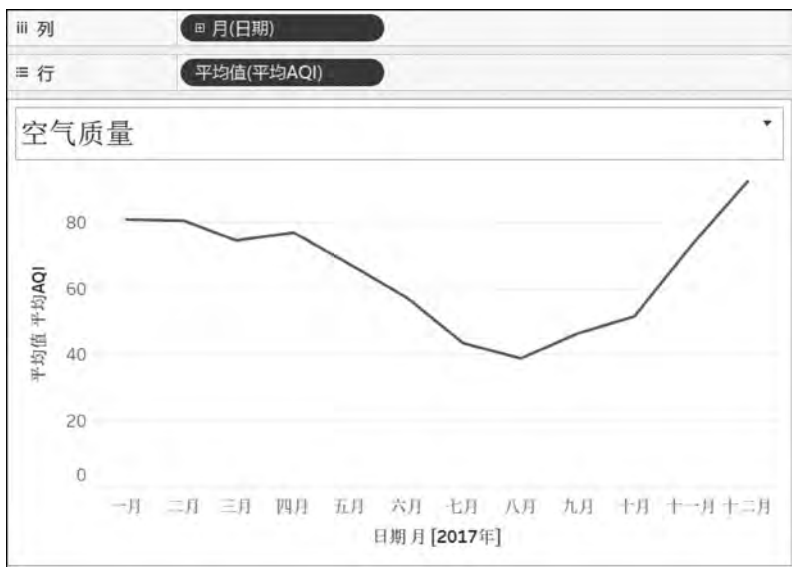


图 5.103 平均值(平均 AQI)折线图

(5) 添加趋势线。单击“分析”选项卡,拖动“模型”分组下的“趋势线”到视图区,放置在悬浮窗口中的“多项式”处,如图 5.107 所示。

完成此步骤操作后,在视图区中增加了一条虚线形式的趋势线,如图 5.108 所示。

(6) 显示关键点标签。具体操作如下。

① 按住 Ctrl 键,拖动“行”功能区的“平均值(平均 AQI)”到“标记”中的“标签”。

② 单击“标签”按钮,在弹出窗口中选择“标签标记”中的“线末端”,那么只在折线的两端出现标签,如图 5.109 所示。



图 5.104 修改垂直轴结束点

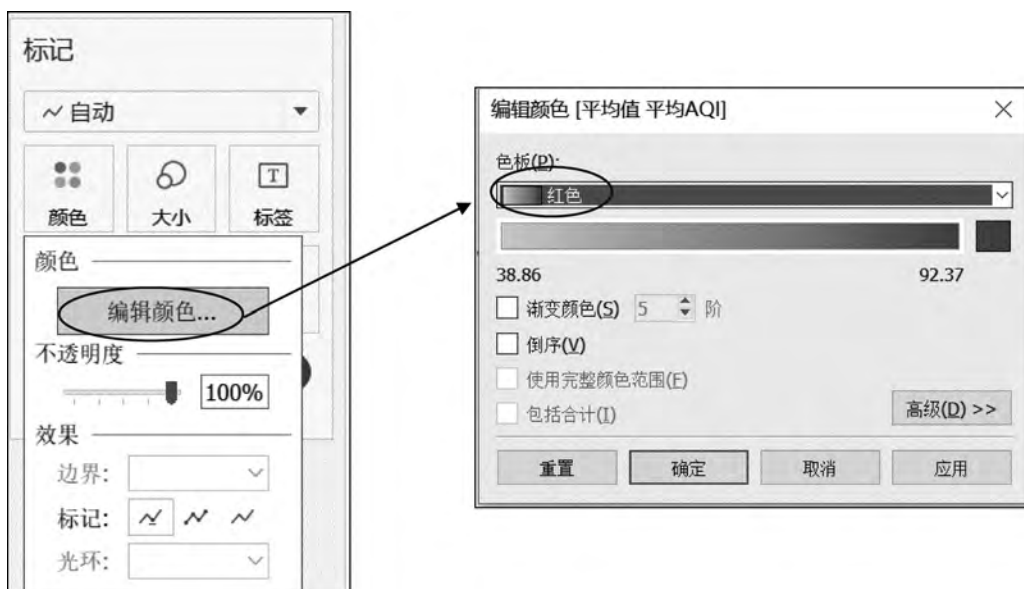


图 5.105 编辑颜色



图 5.106 改变折线粗细



图 5.107 添加趋势线

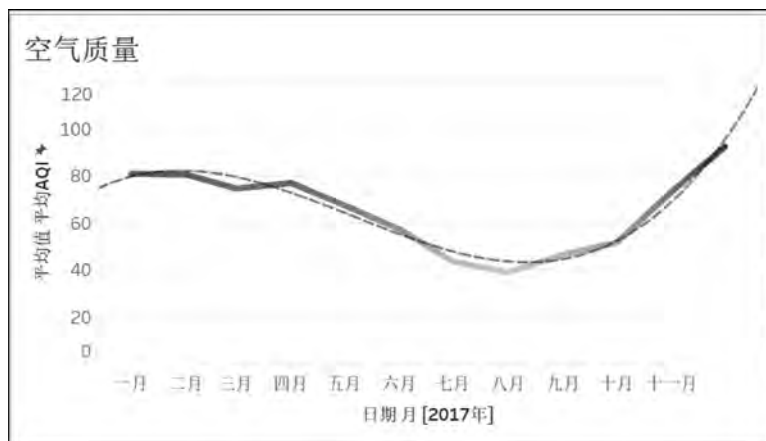


图 5.108 添加趋势线的折线图

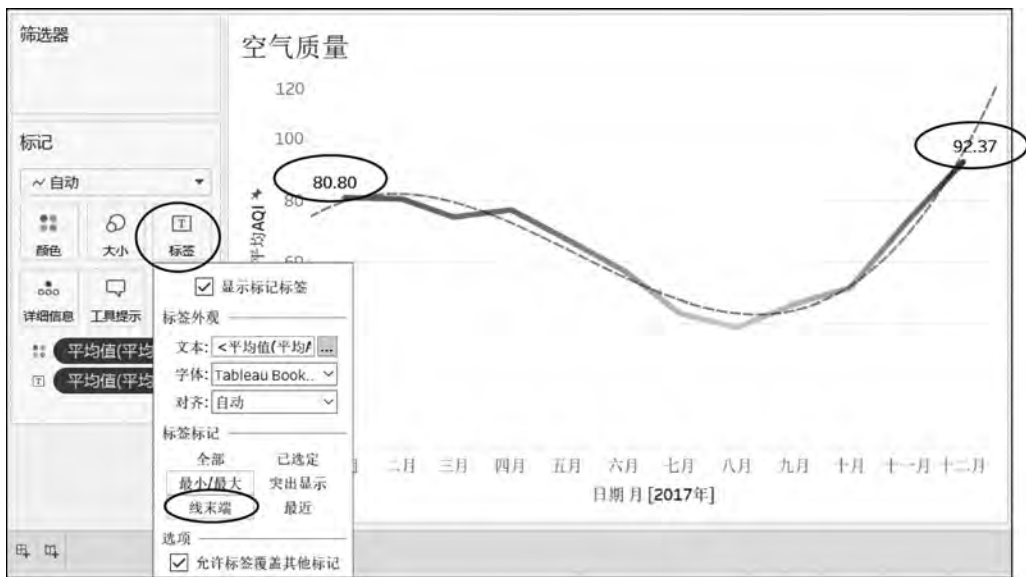


图 5.109 添加标签

(7) 设置折线图线格式。右击折线图的任意空白位置,在弹出的菜单中选择“设置格式”命令,在弹出的对话框中选择最右方的按钮设置线格式,首先单击“行”,将“网格线”和“轴标尺”都设置为无,这样就去掉了网格线和纵坐标轴;单击“列”,“轴标尺”设置为实线,粗细为倒数第 2 个,颜色为黑色,这样将横坐标轴设置为醒目的黑色粗实线,如图 5.110 所示。

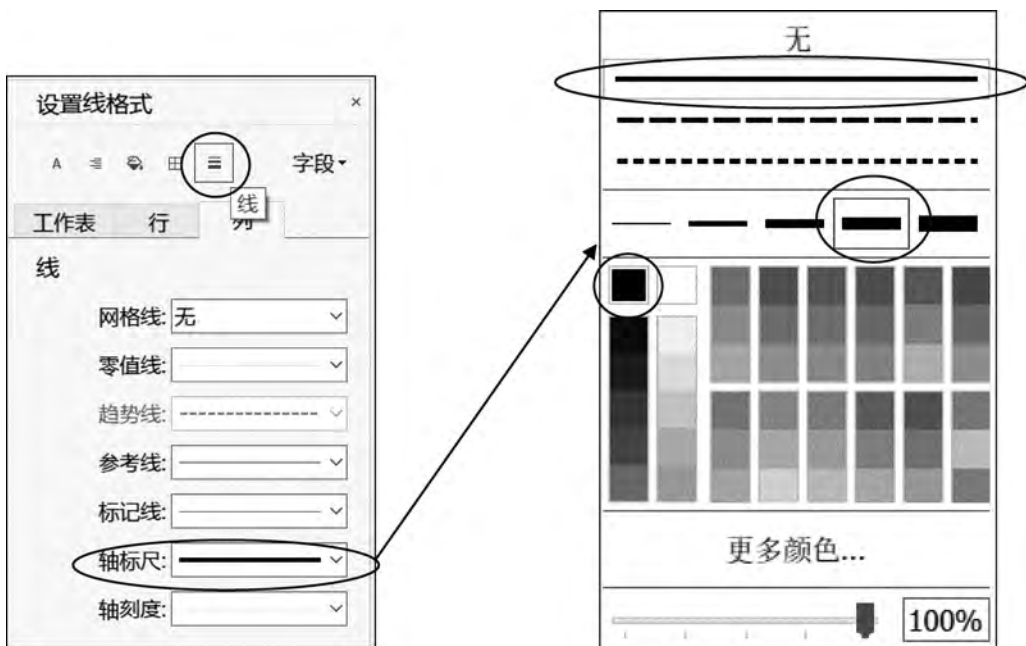


图 5.110 设置线格式

完成此步骤操作后,折线图如图 5.111 所示。

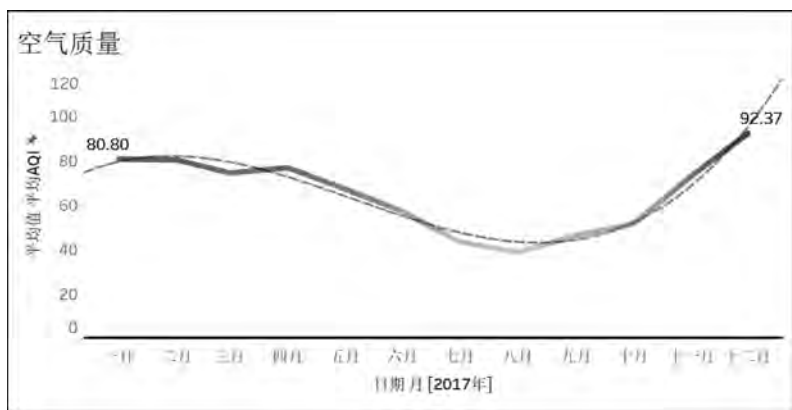


图 5.111 设置线格式后的折线图

(8) 添加参考线,设置参考线格式。具体操作如下。

① 添加优秀空气质量参考线。右击垂直轴刻度,在弹出的菜单中选择“添加参考线”命令,在弹出的“编辑参考线、参考区间或框”对话框中选择“区间”,“区间开始”选择“常量”,“值”为 0,“标签”选择“自定义”,内容为“空气质量,优”;“区间结束”选择“常量”,“值”为 50 (根据我国空气质量指数 AQI 分级标准,AQI 小于等于 50 时,空气质量等级为优秀),“标签”选择“无”;“格式”中的“填充”选择淡蓝色,如图 5.112 所示。

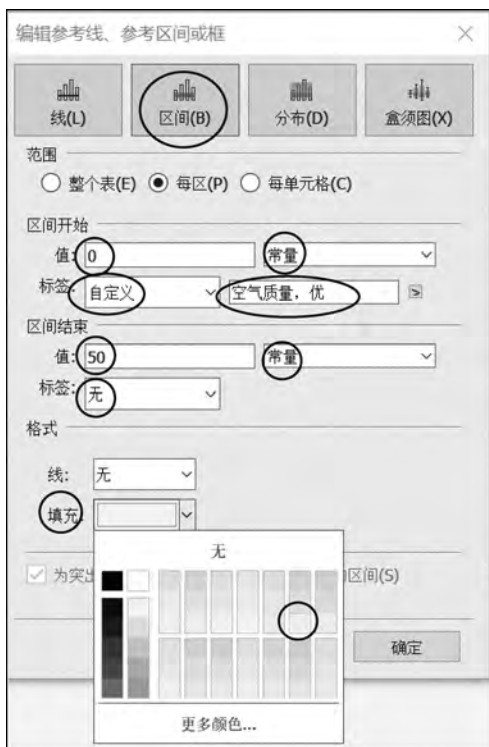


图 5.112 添加优秀空气质量参考线

② 添加良好空气质量参考线。右击垂直轴刻度,在弹出的菜单中选择“添加参考线”命令,在弹出的“编辑参考线、参考区间或框”对话框中选择“区间”,“区间开始”选择“常量”,“值”为 50,“标签”选择“自定义”,内容为“空气质量,良好”;“区间结束”选择“常量”,“值”为 100(根据我国空气质量指数 AQI 分级标准,AQI 大于 50 小于等于 100 时,空气质量为良),“标签”选择“无”;“格式”中的“填充”选择浅蓝色,如图 5.113 所示。

③ 添加轻度污染空气质量参考线。右击垂直轴刻度,在弹出菜单中选择“添加参考线”命令,在弹出的“编辑参考线、参考区间或框”对话框中,选择“区间”,“区间开始”选择“常量”,“值”为 100,“标签”选择“自定义”,内容为“空气质量,轻度污染”;“区间结束”选择“常量”,“值”为 150(根据我国空气质量指数 AQI 分级标准,AQI 大于 100 小于等于 150 时,空气质量为轻度污染),“标签”选择“无”;“格式”中的“填充”选择蓝色,如图 5.114 所示。

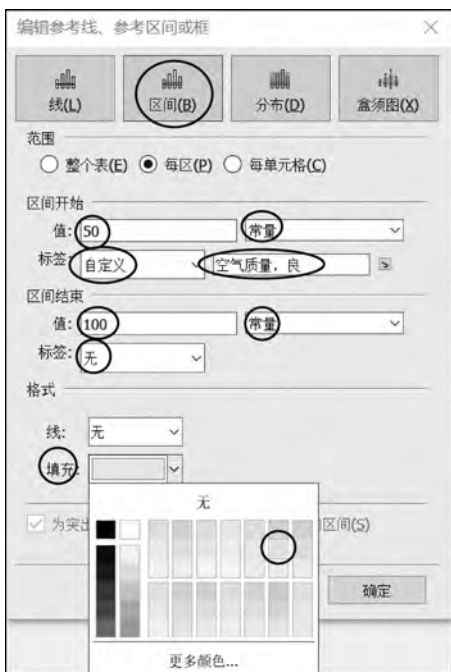


图 5.113 添加良好空气质量参考线

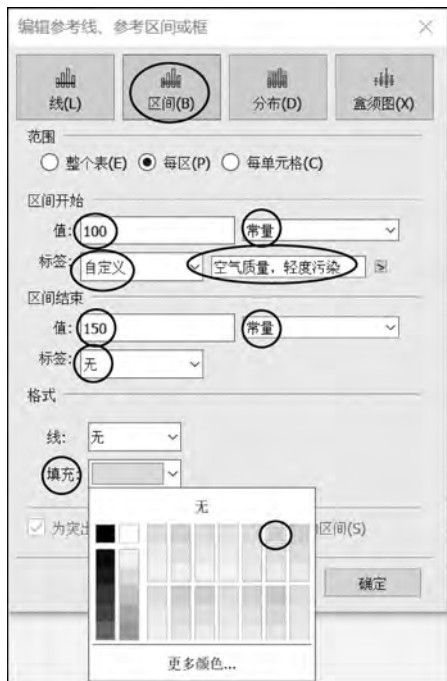


图 5.114 添加轻度污染空气质量参考线

(9) 隐藏垂直轴刻度。由于柱形图上已经有标签,可以隐藏垂直轴上的刻度使得柱形图更为简洁。右击垂直轴,在弹出菜单中选择“显示标题”,从而去掉其前面的 $\checkmark$,隐藏垂直轴刻度。

(10) 隐藏标题。右击视图区中的标题“工作表 1”,在弹出菜单中选择“隐藏标题”。

完成此步骤操作后,视图区的折线图如图 5.115 所示。

通过观察图 5.115,可以很直观地发现,6—10 月无锡市空气质量为优,其他月份为良。



5.6.3 气温

创建“气温感知”计算字段后,接下来建立柱形图展示无锡市一年四季各个季节的平均温度,具体步骤如下:

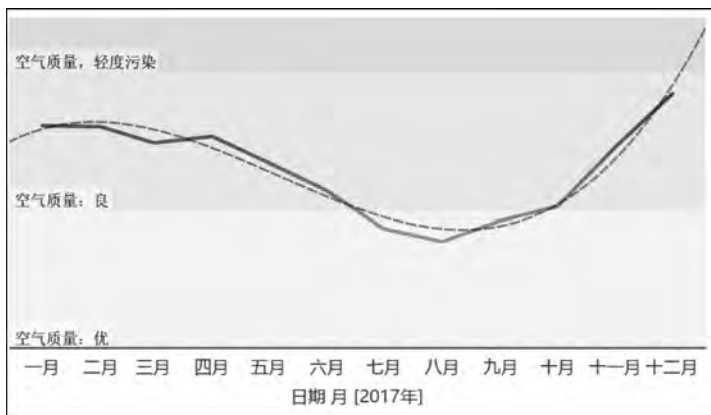


图 5.115 添加参考线后的折线图

(1) 新建一个工作表,命名为“气温”。

(2) 生成柱形图。拖动“维度”分组中的“季节”到“列”功能区,度量“平均气温”到“行”功能区,右击“平均值(平均气温)”,在弹出的菜单中选择“度量(总和)”→“平均值”命令,如图 5.116 所示。



图 5.116 修改为平均值

完成此步骤操作后,视图区生成的柱形图如图 5.117 所示。

(3) 将柱形按“春夏秋冬”排序。右击“列”功能区的“季节”,在弹出的菜单中选择“排序”命令,弹出“排序”对话框,将“排序依据”选为“手动”,然后拖动框中的四季调整成“春夏秋冬”次序,如图 5.118 所示。

(4) 修饰颜色。具体操作如下。

① 按住 Ctrl 键,拖曳“列”功能区的“平均值(平均气温)”到“标记”中的“颜色”。

② 单击“标记”中的“颜色”,在弹出的窗口中选择“编辑颜色”,进而在弹出的“编辑颜色”对话框中,将“色板”选为“温度发散”,如图 5.119 所示。

(5) 添加标签。具体操作如下。

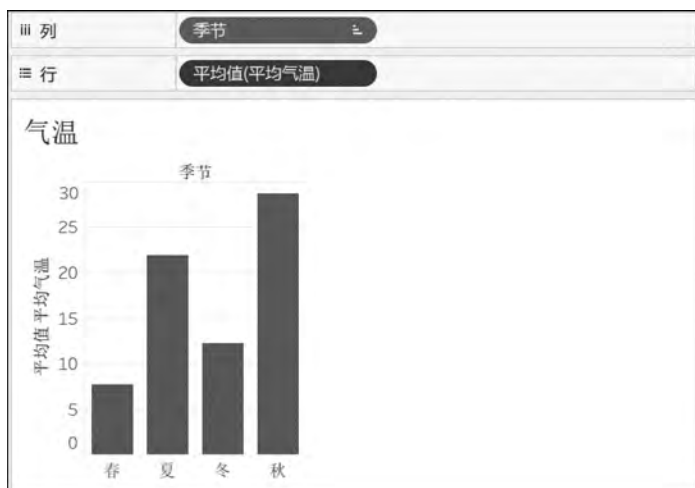


图 5.117 柱形图

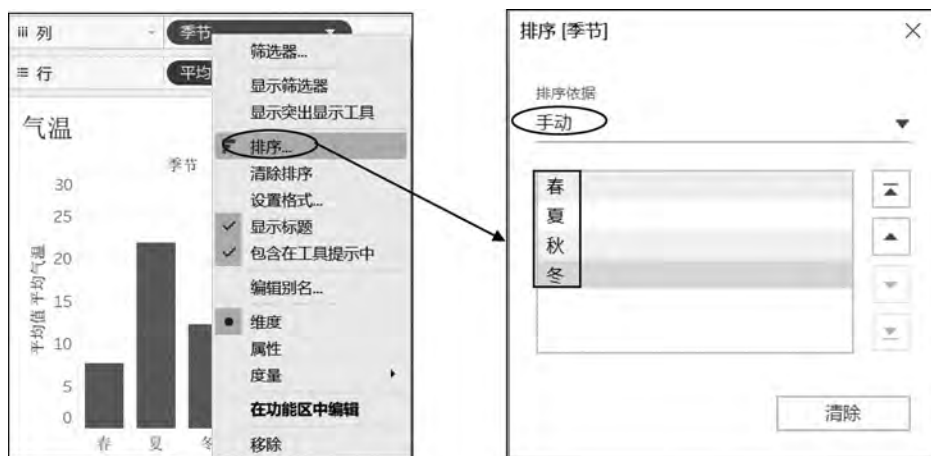


图 5.118 调整次序

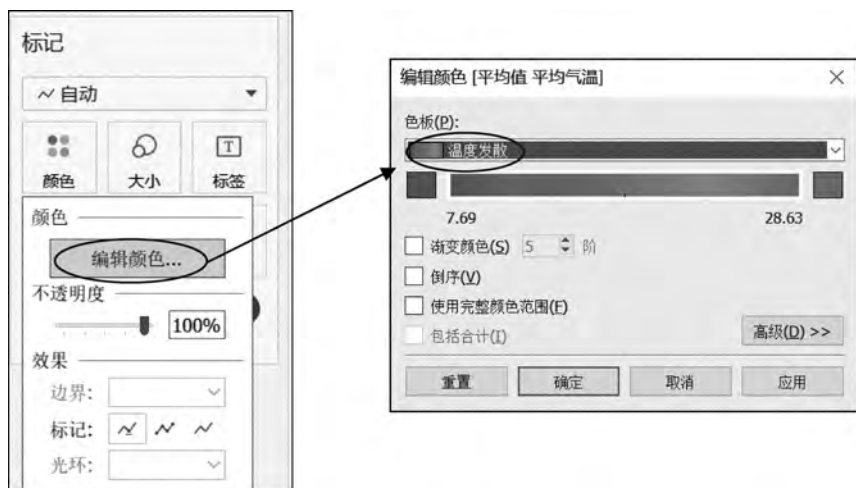


图 5.119 修饰颜色

① 按住 Ctrl 键,拖曳“列”功能区的“平均值(平均气温)”到“标记”中的“标签”,4 个季节的平均气温呈现在对应柱子的上方。

② 右击“列”功能区的“平均值(平均气温)”,在弹出的菜单中选择“设置格式”命令,在“区”选项卡下的“默认值”分组中,首先“数字”选择“数字(自定义)”,然后“小数位数(E)”设置为 0(温度显示为整数),“后缀”设置为“℃”(该符号可以从百度获取),如图 5.120 所示。



图 5.120 设置标签格式

(6) 修饰标签。为使记录数呈现在柱形下方,单击“标记”中的“标签”,在弹出的窗口中进行如下操作。

① 单击“对齐”右方的倒三角,在弹出窗口中,“水平”选择“居中”,“垂直”选择“底部”。

② 单击“文本”右方的倒三角,在弹出的“编辑标签”对话框中,按 Ctrl+A 快捷键全选框内文本,再选择“微软雅黑”字体和 8 号字号,如图 5.121 所示。

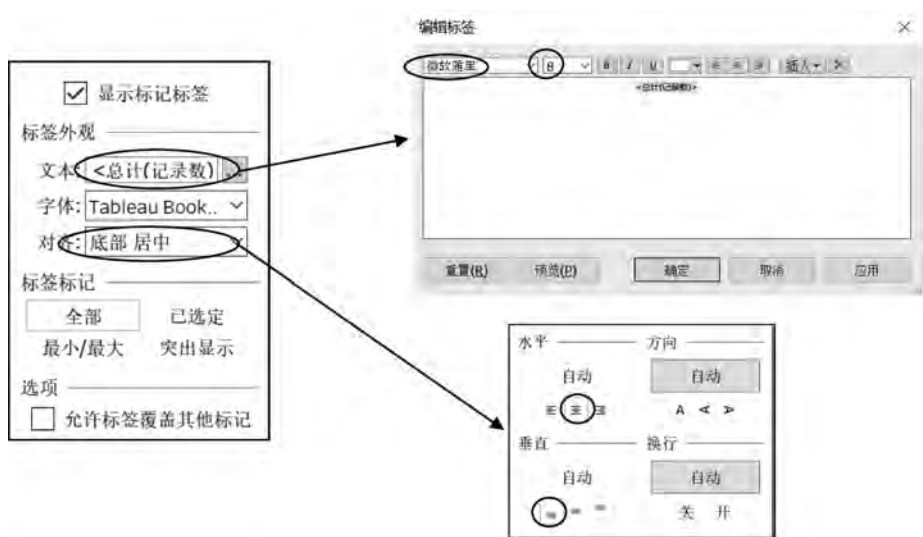


图 5.121 设置标签格式

(7) 设置线格式。右击柱形图背景的任意空白位置,在弹出的菜单中,选择“设置格式”命令,在左方出现的窗口中,单击按钮设置线格式,接下来的操作中,首先单击设置“行”“网格线”和“轴标尺”都为无,这样就去掉了网格线和纵坐标轴;再单击设置“列”“轴标尺”为实线,粗细为倒数第2个,颜色为黑色,如图 5.122 所示,这样将横坐标轴设置为醒目的黑色粗实线。

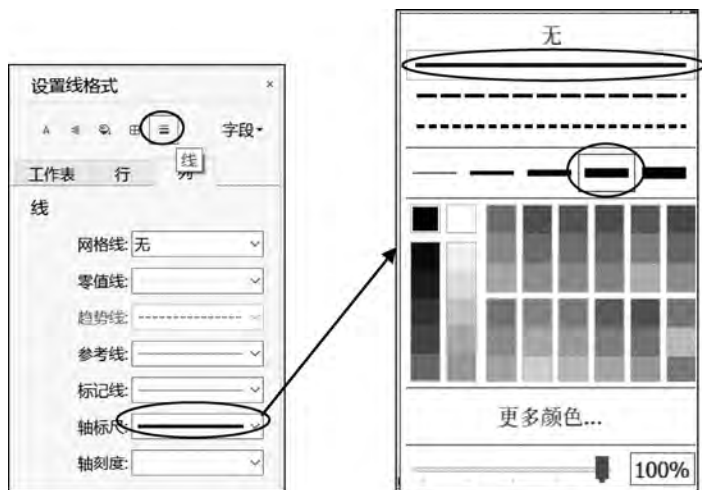


图 5.122 设置线格式

(8) 隐藏垂直轴刻度。右击垂直轴刻度,在弹出的菜单中单击“显示标题”,从而去掉其前面的符号√,隐藏垂直轴刻度。

(9) 设置列标题。右击“列”功能区的“季节”,在弹出的菜单中选择“设置格式”命令,在“标题”选项卡下的“默认值”分组中设置字体为“微软雅黑”,12pt 字号,“黑色”。

(10) 隐藏标题。右击视图区中的标题“气温”,在弹出的菜单中选择“隐藏标题”。

完成以上所有操作后,视图区的柱形图如图 5.123 所示。

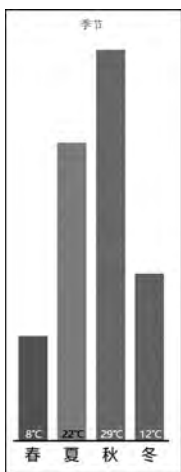


图 5.123 柱形图

5.6.4 宜居时间



针对宜居时间,主要从天气状况、气温、空气质量和天数四个角度考虑。下面制作图表展示无锡宜居的时间段,具体步骤如下。

(1) 新建一个工作表,重命名为“宜居时间”。

(2) 生成柱形图。拖动“维度”分组中的“天气状况”到“列”功能区,度量分组中的“平均气温”到“行”功能区,右击“平均值(平均气温)”,在弹出菜单中选择“度量(总和)”→“平均值”命令。

(3) 按住 Ctrl 键,把“行”功能区的“平均气温”拖曳至它的右侧,形成上下两个一样的柱形图,同时,在“标记”处,出现了“全部”“平均值(平均气温)”“平均值(平均气温)(2)”3 个折叠面板,分别对应设置视图区的全部柱形图、上方柱形图和下方柱形图,如图 5.124 所示。

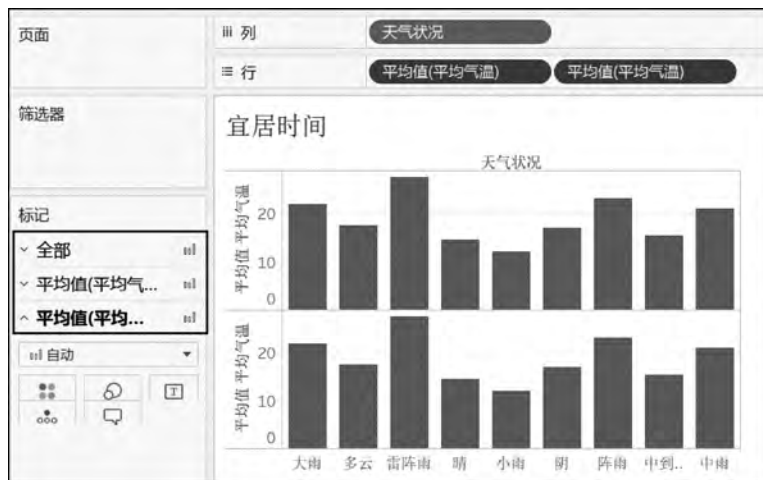


图 5.124 生成上下两个柱形图

(4) 设置上方图表。单击“标记”的“平均值(平均气温)”展开面板,进行如下操作。

① 将“类型”由“自动”更改为“条形图”,使得柱形变细。单击“大小”,向右拖动滑块到合适位置,如图 5.125 所示。

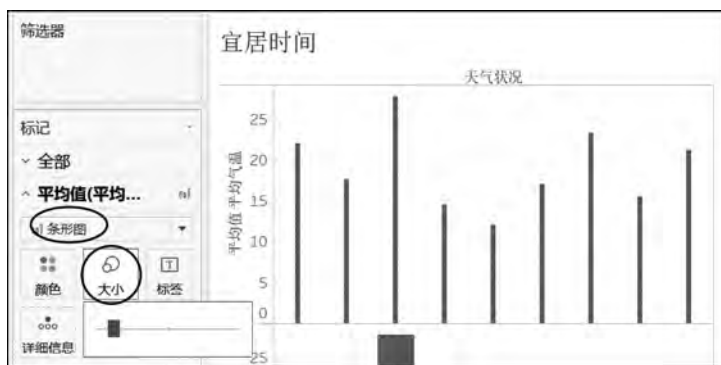


图 5.125 设置上方图表

② 拖动“度量”分组中的“气温感知”至“标记”中的“颜色”。

③ 双击上方图表的垂直轴,在弹出的“编辑轴”对话框中,将“范围”选为“固定”,“固定结束”的值设定为 45,如图 5.126 所示,再关闭对话框,使得图表在坐标轴上稍微下移,方便后续的操作,也更为美观。

(5) 设置下方图表。单击“标记”中的“平均值(平均气温)(2)”展开面板,进行如下操作。

① 将“类型”由“自动”更改为“圆”。单击“大小”,向右拖动滑块到合适位置,如图 5.127 所示。

② 拖动度量“气温感知”至“标记”中的“颜色”,在“标记”最下方显示出“聚合(气温感知)”。

③ 统一上下方图表垂直轴高度。双击上方图表的垂直轴,在弹出的对话框中,“范围”

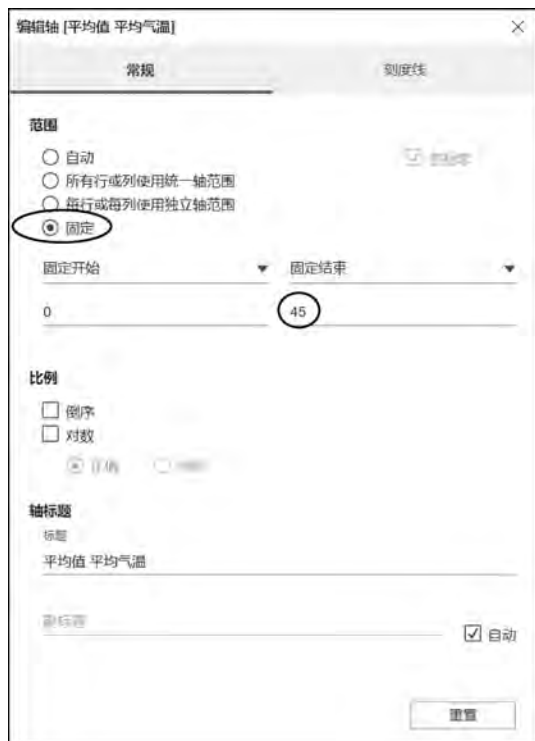


图 5.126 修改垂直轴结束点

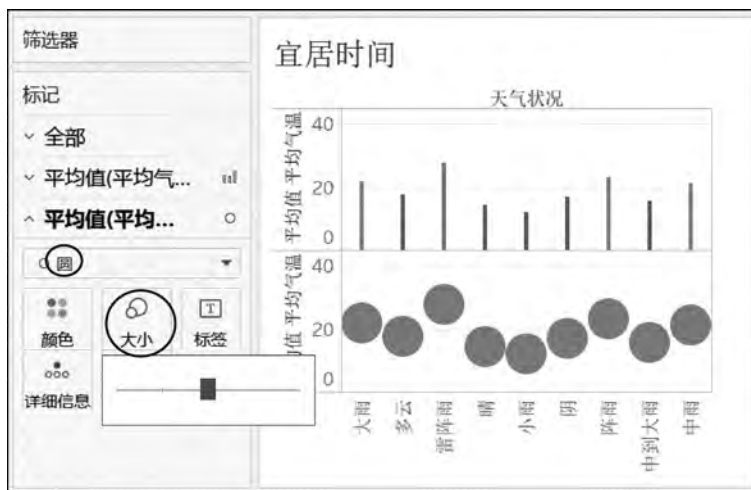


图 5.127 设置下方图表的类型和大小

选择“固定”，与上方图表一样，“固定结束”的值设定为 45。

完成此步骤操作后，视图区如图 5.128 所示。

(6) 合并轴。右击“行”功能区右方的“平均值(平均气温)”，在弹出的菜单中选择“双轴”，实现上下方两个图表合并，然后，在工具栏中将视图由“标准”改为“整个视图”。完成此步骤操作后，视图区如图 5.129 所示。

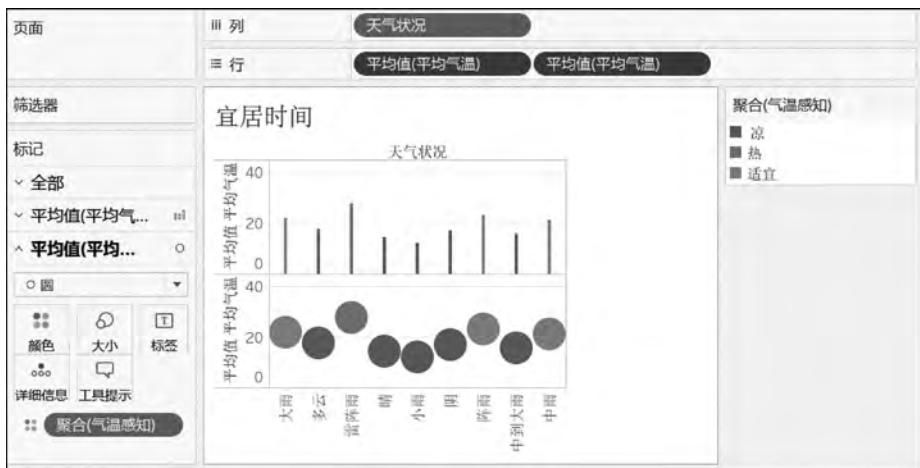


图 5.128 设置下方图表

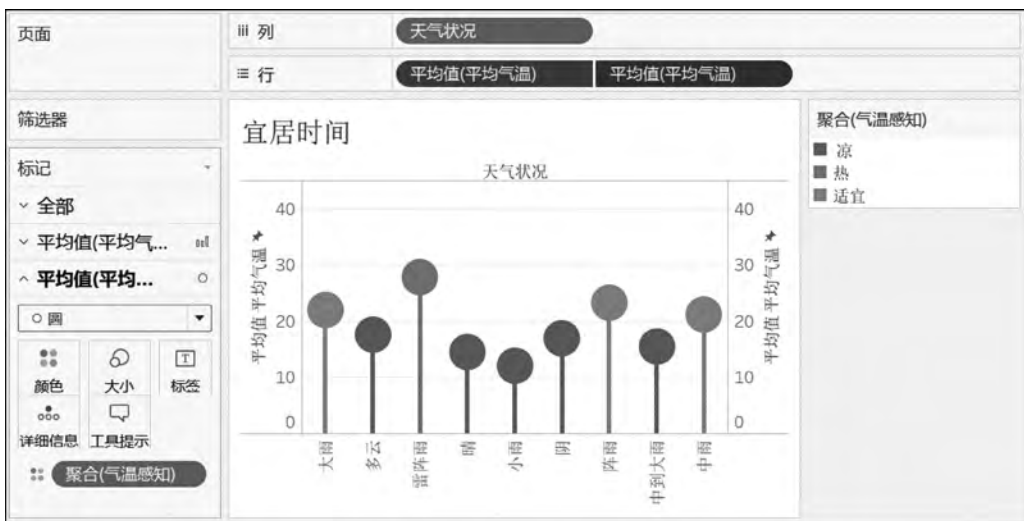


图 5.129 合并上下方图表

(7) 添加“平均气温”标签。具体操作如下。

- ① 单击“标记”的“平均值(平均气温)(2)”展开面板。
- ② 按住 Ctrl 键,拖动“行”功能区的任一“平均值(平均气温)”到“标记”中的“标签”。
- ③ 右击“标记”下方的“T 平均值(平均气温)”,在弹出菜单中选择“设置格式”命令。
- ④ 在左方窗口“区”选项卡下“默认值”分组中,首先将“数字”选为“数字(自定义)”,然后“小数位数”设置为 0(温度度数显示为整数),“后缀”设置为“℃”,如图 5.130 所示。

⑤ 单击“标记”中的“标签”,在弹出窗口中单击“对齐”右方的 ▾,在弹出窗口中,“水平”选择“居中”[居],“垂直”选择“居中”,如图 5.131 所示,从而使得平均气温数值呈现在圆的中间。然后单击“文本”右方的 [文],在弹出的对话框中,按住 Ctrl+A 快捷键全选框内文本,再选择“微软雅黑”字体、8 号字号、“加粗”“白色”,如图 5.131 所示。

(8) 添加其他标签。具体操作如下。



图 5.130 设置“平均气温”标签格式

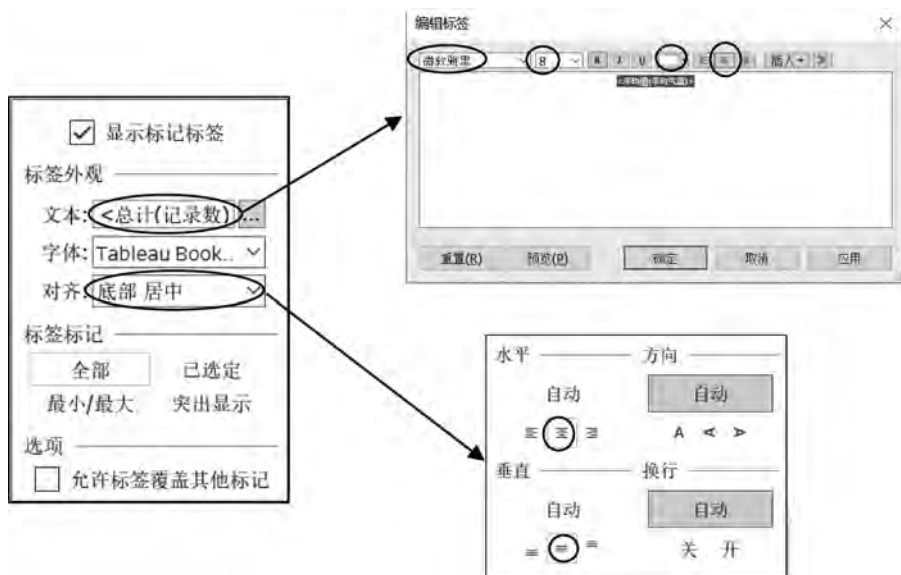


图 5.131 设置标签格式

- ① 单击“标记”的“平均值(平均气温)”展开面板。
- ② 拖动“度量”分组中的“记录数”至“标记”中的“标签”。
- ③ 右击“标记”下方的“**T** 总和(记录数)”，在弹出菜单中选择“设置格式”命令。
- ④ 在左方出现的窗口中，在“区”选项卡下的“默认值”分组，首先将“数字”选为“数字(自定义)”，然后“小数位数(E)”设置为 0(温度显示为整数)，“前缀”设置为“天数”，如图 5.132 所示。
- ⑤ 分别拖动度量“空气质量等级”和“温差”至“标记”中的“标签”。
- ⑥ 右击“标记”下方的“**T** 聚合(温差)”，在弹出菜单中选择“设置格式”命令。

⑦ 在左方出现的窗口中,在“区”选项卡下的“默认值”分组,首先将“数字”选为“数字(自定义)”,然后“小数位数”设置为0(温度度数显示为整数),“前缀”设置为“温差”,“后缀”设置为“°C”,如图 5.133 所示。

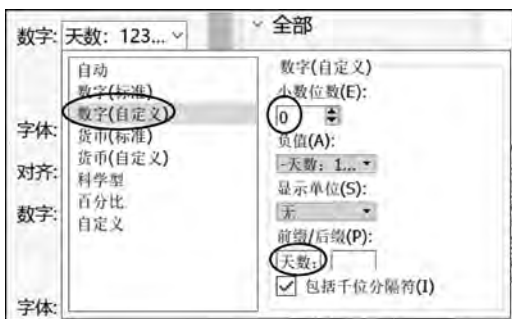


图 5.132 设置“天数”标签格式

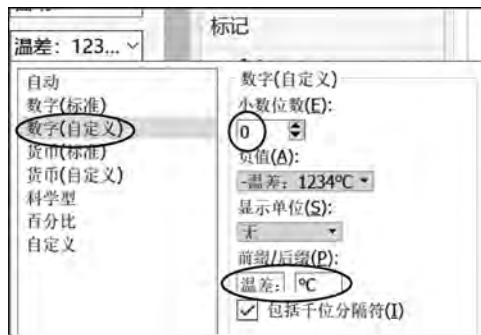


图 5.133 设置“温差”标签格式

⑧ 单击“标记”中的“标签”,在弹出窗口中,单击“文本”右方的 ,在弹出的“编辑标签”对话框中,执行如下操作:首先在文本框中的“<聚合(空气质量等级)>”之前输入“空气质量:”;接着调整文本框每行文字的次序,第1行为“<总和(记录数)>”,第2行为“<聚合(温差)>”,第3行为“空气质量: <聚合(空气质量等级)>”,第四、五和六行添加为空行(通过回车实现);然后按 Ctrl+A 快捷键全选框内文本,再选择“微软雅黑”字体、8 号字号、“淡黑色”“左对齐”,如图 5.134 所示。



图 5.134 设置标签格式

完成此步骤操作后,视图区如图 5.135 所示。

(9) 设置列标题。右击“列”功能区的“所属区域”,在弹出的菜单中选择“设置格式”命令,在“标题”选项卡下的“默认值”分组,设置字体为“微软雅黑”,11 号字,“黑色”。

(10) 隐藏网格线。右击图表的任意空白位置,在弹出的菜单中选择“设置格式”命令,在弹出的对话框选择最右方的按钮  设置线格式,首先单击“行”,然后将“网格线”设置为无。

(11) 隐藏左右两边垂直轴刻度。由于柱形图上已经有标签,可以隐藏垂直轴上的刻度使得柱形图更为简洁。分别右击左右两边垂直轴刻度,在弹出的菜单中选择“显示标题”从

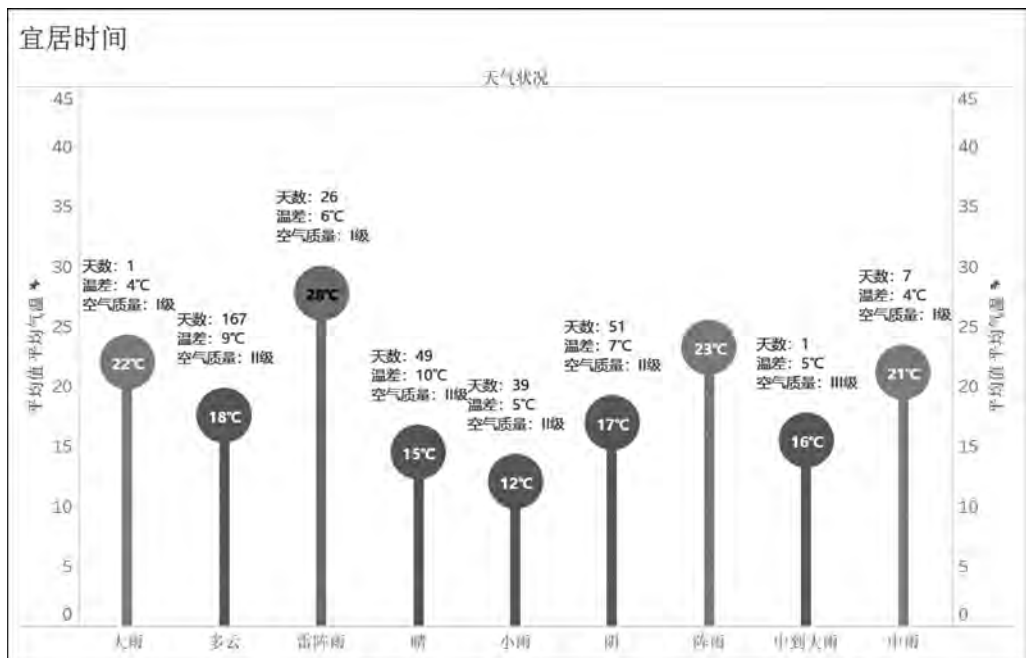


图 5.135 图表

而去掉其前面的√,隐藏两边的垂直轴刻度。

(12) 筛选 6—10 月数据。根据前面生成的空气质量折线图可知,6—10 月期间无锡市的空气质量为优,其他月份为良,因此重点观察这五个月的宜居情况。筛选的具体操作为:拖动维度“日期”至“筛选器”,在弹出的对话框中选择“年/月”,单击“下一步”按钮,在弹出的对话框中勾选“2017 年 6 月”至“2017 年 10 月”5 个月份,如图 5.136 所示。



图 5.136 筛选月份

(13) 编辑标题。右击视图区中的标题“宜居时间”，在弹出的菜单中选择“编辑标题”命令，弹出“编辑标题”对话框，在文本框内输入“6—10 月宜居情况”，再按 Ctrl+A 快捷键全选刚刚输入的标题，设置字体为“微软雅黑”，字号为 15，“加粗”“淡黑”“居中”，如图 5.137 所示。

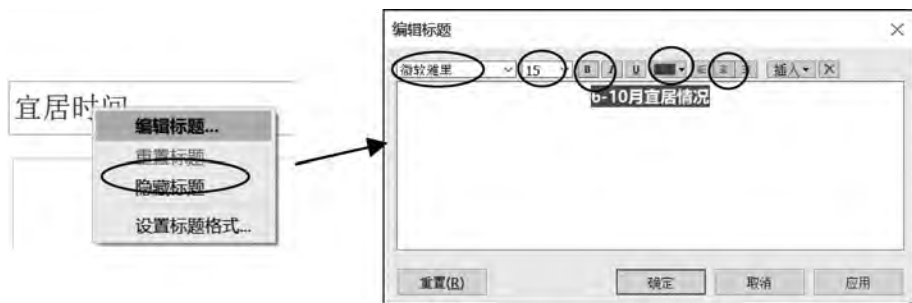


图 5.137 编辑标题

完成以上所有步骤后，视图区中创建的图表如图 5.138 所示。

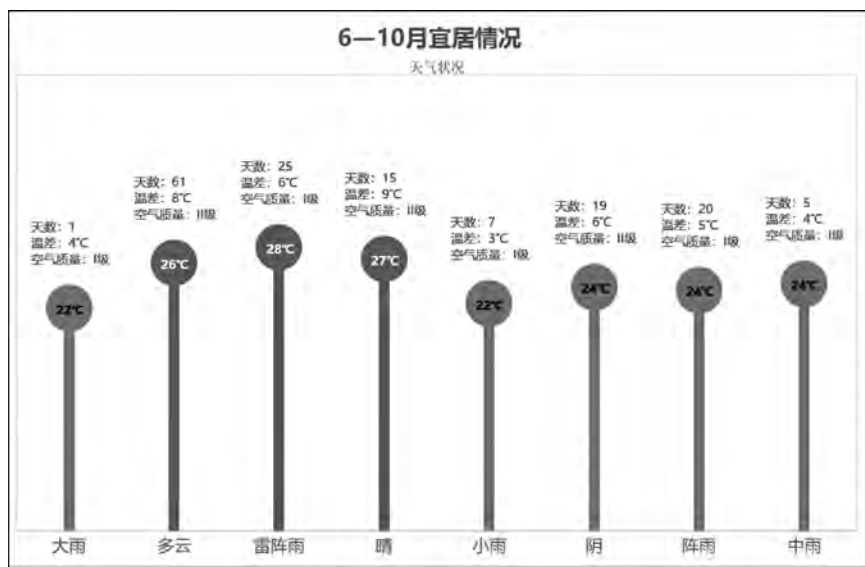


图 5.138 6—10 月宜居情况展示图表

5.6.5 制作仪表板

本节将基于前面创建的“空气质量”“气温”“宜居时间”3 张工作表制作仪表板。具体步骤如下。

(1) 单击左下方按钮 创建仪表板。仪表板大小选择“固定大小”，自定义为 1638×1017，即是宽度为 1638px，高度为 1017px，如图 5.139 所示。

(2) 排版对象(图表)。具体操作如下。

① 在“对象”分组设置以“浮动”的方式添加对象，以方便在仪表板上排版，如图 5.140 所示。

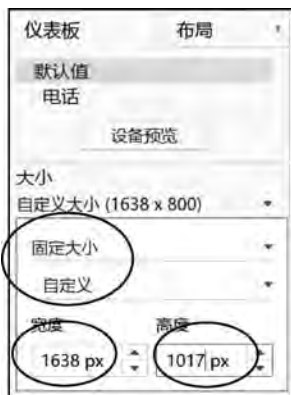


图 5.139 仪表板大小

② 将“工作表”分组的“宜居时间”工作表拖动到仪表板视图区；切换到“布局”选项卡,设定“位置”中的 x 值为 43,y 值为 145,“大小”中的“宽”值为 1105,“高”值为 802,如图 5.141 所示。把光标放置在图例上方的 ,当光标成十字架形状时拖动图例到该图表的左上角,关闭其他图例。

③ 将“工作表”分组的“气温”工作表拖动到仪表板视图区右上方空白区域；切换到“布局”选项卡,清除“显示标题”复选框,设定“位置”中的 x 值为 1213,y 值为 202,“大小”中的“宽”值为 394,“高”值为 346。

④ 将“工作表”分组的“空气质量”工作表拖动到仪表板视图区右下方空白区域；切换到“布局”选项卡,取消“显示标题”,设定“位置”中的 x 值为 1204,y 值为 622,“大小”中的“宽”值为 416,“高”值为 367。



图 5.140 设定浮动方式



图 5.141 设置宜居时间图表布局

完成这一步骤的操作后,单击工具栏上的  按钮,或者按 F7 键,进入演示模式查看总体布局情况,如图 5.142 所示。

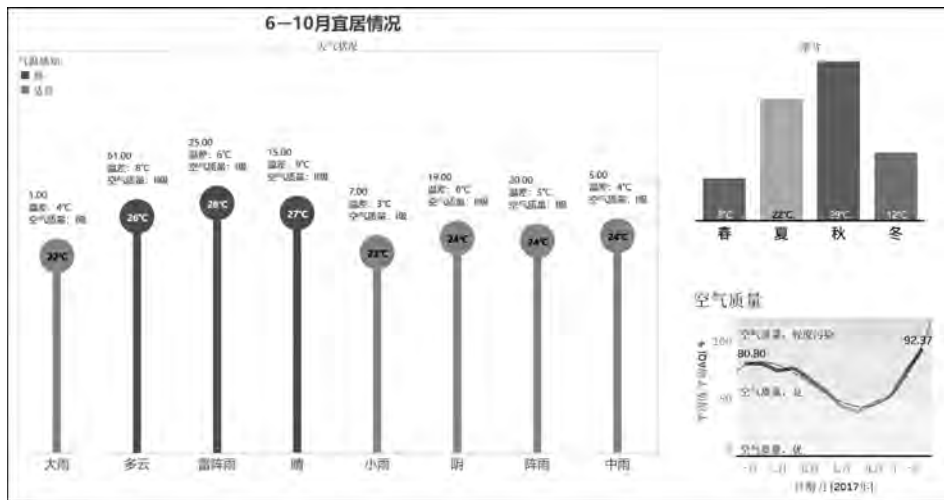


图 5.142 整体布局

(3) 插入背景。具体操作如下。

① 选择“布局”选项卡,先在“对象”分组中选择“平铺”方式添加对象,以方便在仪表板上添加背景图像,再双击“图像”按钮,在弹出的对话框中选择“88.jpg”作为背景图像,如图 5.143 所示。

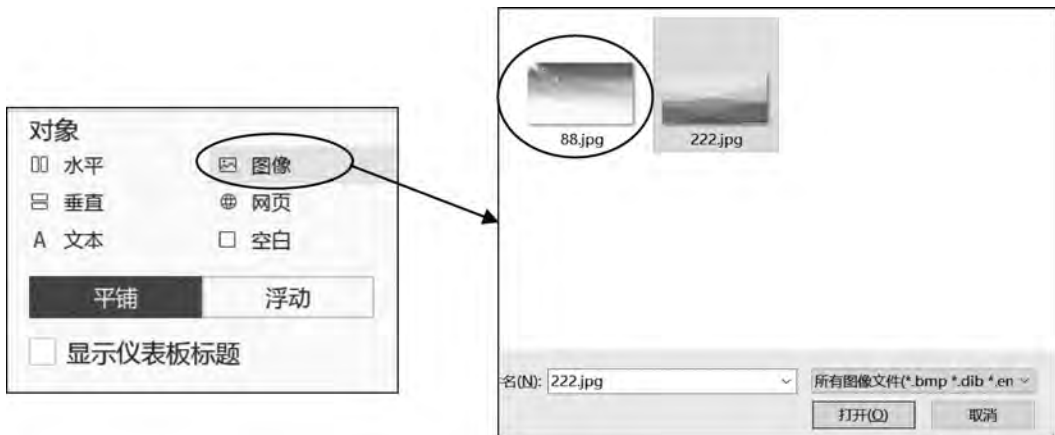


图 5.143 选择背景图像

② 在“顶分层结构”分组选择“仪表板 1”→“平铺”→“88.jpg”,再右击“88.jpg”,在弹出的菜单中选择“适合图像”,再次右击“88.jpg”,在弹出的菜单中选择“使图像居中”。

③ 选择视图区的“宜居时间”图表,再选择工具栏下“设置格式”→“阴影”,在窗口左方出现“设置阴影格式”窗格,在“工作表”选项卡设定“默认值”分组下的“工作表”为“无”;再逐一选择视图区的“空气质量”和“气温”图表,在“工作表”选项卡,分别设定“默认值”分组下的“工作表”为“无”,从而实现所有图表的背景透明。

[注意] 此操作仅在 Tableau 2018.3 以及之后版本可用。

(4) 添加文本。选择“仪表板”选项卡,拖动“对象”分组中的“文本”至视图上方中央,在弹出的对话框中输入“无锡市”,再按 Ctrl+A 快捷键全选刚刚输入的文字,设置字体为“微软雅黑”,字号为 48,“加粗”“蓝色”“居中”,如图 5.144 所示。设置完成后,可以拖动微调文本框至更合适的位置。



图 5.144 添加文本

该步骤操作完成后,仪表板在演示模式下如图 5.145 所示。

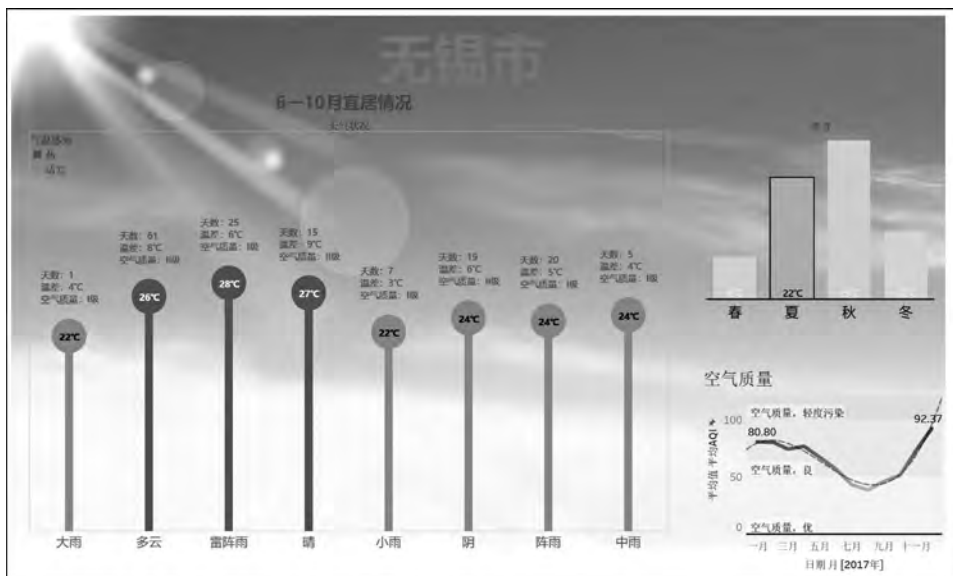


图 5.145 演示仪表板

(5) 实现图表之间联动。单击选择仪表板中的柱形图,再单击右上方的“用作筛选器”漏斗状按钮 ,成功设置后,空漏斗将变成实漏斗,实现了以柱形图作为筛选器联动其他两个图表。借此可以进行如下一些分析。

① 当在柱形图中选择代表“秋季”的柱子时,左方的宜居图表呈现了无锡市在秋季(7月、8月和9月3个月份)的宜居情况,折线图呈现了无锡市在秋季(7月、8月和9月3个月份)的空气质量,如图 5.146 所示。通过观察该图,可以看到无锡市在秋季时多云天气多达 37 天,平均气温达到 29°C,较为炎热,整个季节的平均空气质量都为优。

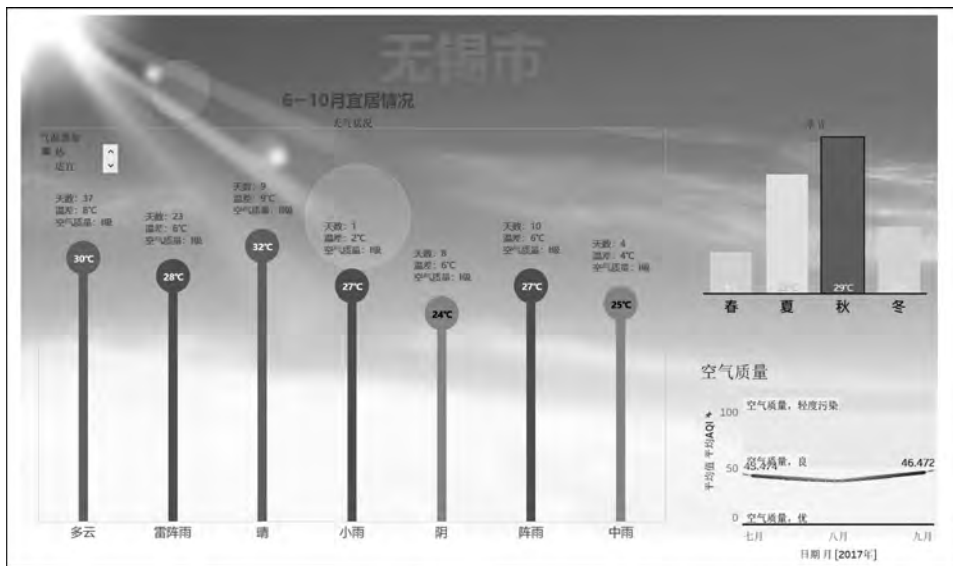


图 5.146 秋季

② 当在柱形图中选择代表“夏季”的柱子时,左方的宜居图表呈现了无锡市在6月份(由于只考虑6—10月五个月份,那么只把6月份一个月归属于夏季)的宜居情况,折线图呈现了无锡市在6月份的空气质量,如图5.147所示。经过观察该图,可以看到无锡市在6月份时多云天气达到9天,平均气温是22度,比较舒适,但整个六月的平均空气质量为良。

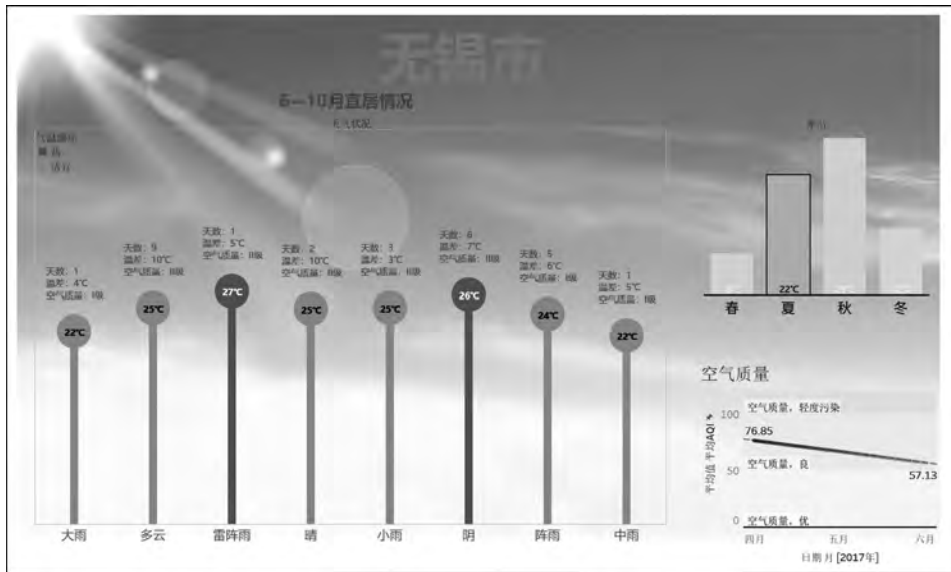


图 5.147 夏季(6月份)

③ 当在柱形图中选择代表“冬季”的柱子时,左方的宜居图表呈现了无锡市在10月份(由于只考虑6—10月五个月份,那么只把10月份一个月归属于冬季)的宜居情况,折线图呈现了无锡市在10月份的空气质量,如图5.148所示。经过观察该图,可以看到无锡市在10月份时多云天气也是达到9天,平均气温只有12度,比较凉,整个10月份的平均空气质量为良。

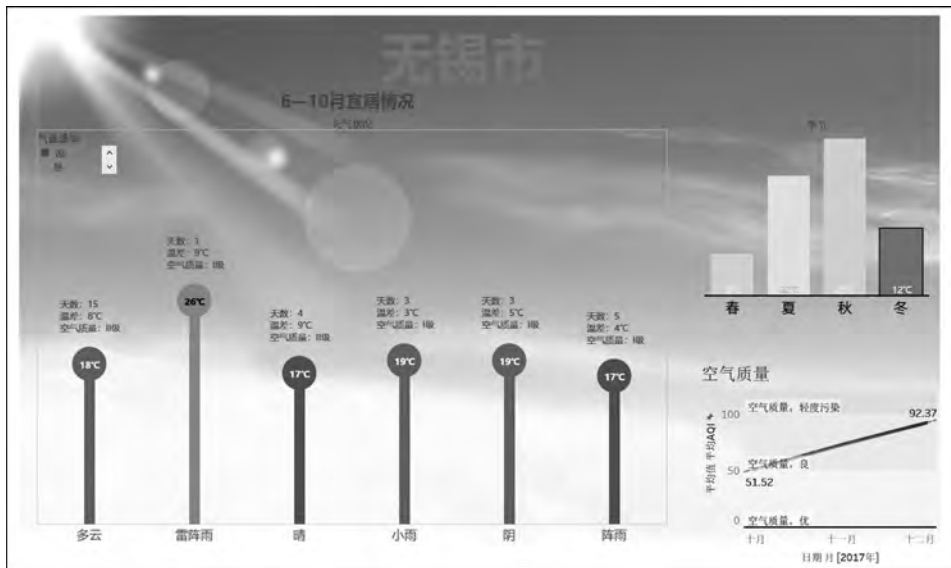


图 5.148 冬季(10月份)

通过以上季节的分析,可以看出无锡市整体而言以多云天气为主。7—9 月份无锡市的空气质量较好,但天气有些炎热。

5.7 案例二: 佛山市纳税企业增长情况分析



本节将介绍第 2 个综合案例,该案例是分析佛山市纳税企业增长情况,纳税企业数据来自佛山市数据开放平台,数据来源详细地址为 http://www.foshan-data.cn/data/catalog/catalogDetail.htm?cata_id=43270,数据选取的时间段是 2015 年和 2016 年两年。为方便读者,本书附带的学习资源也包括了该数据集,即“佛山市纳税企业增长数据.hyper”文件。

打开 Tableau Desktop,连接“佛山市纳税企业增长数据.hyper”文件,打开数据集。该数据集包含组织机构代码、企业注册号、纳税人识别号、纳税人名称、纳税人税务经营状态、法定代表人、注册地址、注册类型、税务登记日期、税务登记机关、行业、纳税人登记类别以及实际经营地址等字段。除了税务登记日期字段的数据类型为日期型外,其他所有字段的数据类型都为字符型。数据集中一条记录对应一个企业,记录数即为企业数。数据集中一共有 128121 条记录,即意味着 2015—2017 年佛山市新增 128121 个纳税企业。

数据集中绝大多数字段都为字符串型,那该如何去分析纳税企业的增长情况呢?为让分析进行得更为充分,拟采用以下几个角度进行分析。

- 考察全量的增长分布和预测未来的纳税企业增长情况。
- 展示各个行业和佛山各个区域的纳税企业预测增长情况。

在上述角度中需要针对企业所属城区进行分析,但在数据中不存在“所属区域”字段,因此,需首先创建“所属区域”字段。

5.7.1 创建字段

本小节将介绍如何创建“所属区域”字段,思路如下。

- 佛山市下辖禅城、南海、顺德、三水和高明 5 个区,可以照此划分每个企业所属的区域。
- 经过观察数据,可以发现每个企业的“实际经营地址”字段,或者“注册地址”字段,或者“税务登记机关”字段,又或者“纳税人名称”字段记录了该企业所属区域的信息,因此可以从这 4 个字段中抽取信息,从而建立“所属区域”字段。

为实现有用信息的抽取,除了使用 If 函数外,还需要使用 Tableau 的 Contains 函数。该函数的语法格式为 Contains(string, substring)。该函数包括两个字符串类型的参数 string 和 substring。如果字符串 string 包含了子字符串 substring,则 Contains 函数的结果为 TRUE,否则为 FALSE。在 Tableau 中会给数据添加双引号以表示它是字符串,例如,“book”、“中国”等。

明确创建“所属区域”字段的思路后,下面进行具体操作,步骤如下。

(1) 单击选定“维度”分组中的“实际经营地址”,右击,在弹出的菜单中选择“创建”→“计算字段”命令,如图 5.149 所示。

(2) 在弹出的对话框的上方文本框中输入计算字段名称“所属区域”,在输入窗口的下方文本框中输入以下语句:



图 5.149 创建计算字段

IF

CONTAINS([实际经营地址], "禅城") or CONTAINS([注册地址], "禅城") or CONTAINS([税务登记机关], "禅城") or CONTAINS([纳税人名称], "禅城") THEN "禅城"

ELSEIF

CONTAINS([实际经营地址], "南海") or CONTAINS([注册地址], "南海") or CONTAINS([税务登记机关], "南海") or CONTAINS([纳税人名称], "南海") THEN "南海"

ELSEIF

CONTAINS([实际经营地址], "顺德") or CONTAINS([注册地址], "顺德") or CONTAINS([税务登记机关], "顺德") or CONTAINS([纳税人名称], "顺德") THEN "顺德"

ELSEIF

CONTAINS([实际经营地址], "三水") or CONTAINS([注册地址], "三水") or CONTAINS([税务登记机关], "三水") or CONTAINS([纳税人名称], "三水") THEN "三水"

ELSEIF

CONTAINS([实际经营地址], "高明") or CONTAINS([注册地址], "高明") or CONTAINS([税务登记机关], "高明") or CONTAINS([纳税人名称], "高明") THEN "高明"

END

完成公式输入且检查无误后,单击“确定”按钮,如图 5.150 所示,则在“维度”分组中产生了一个新的字段“所属区域”。



图 5.150 输入所属区域公式

(3) 分别双击“维度”分组中的“所属区域”和“度量”分组中的“记录数”,如图 5.151 所示,观察视图区,可以发现有一条记录(一个企业)未能识别出它所属的区域。右击视图区的“Null”,在弹出的菜单中选择“只保留”,从而筛选出这条记录(这个企业)。



图 5.151 未能识别所属区域的记录

为查验该记录未能识别所属区域的原因,依次双击“实际经营地址”字段、“注册地址”字段、“税务登记机关”字段和“纳税人名称”字段,视图区如图 5.152 所示。经过观察,发现该记录的上述 4 个字段值的确不包含“区”的信息,确实无法准确提取该企业的所属区域。由于只有一条这样的记录,为不影响后续的分析,可以直接排除它。



图 5.152 查验未识别原因

5.7.2 预测



本节将介绍如何预测佛山市纳税企业的增长情况。具体步骤如下。

(1) 右击左下方“工作表 1”,在弹出的菜单中选择“重命名工作表”命令,重命名工作表为“预测”,如图 5.153 所示。

(2) 查看“维度”分组中的“税务登记日期”前面的图标是否为 。若是,即“税务登记日期”字段的数据类型为日期型,否则,右击“维度”分组中的“税务登记日期”,在弹出的菜单中,选择“更改数据类型”→“日期”命令。

(3) 生成折线图。拖曳“维度”分组中的“税务登记日期”到“列”功能区,右击“年(税务登记日期)”,在弹出的菜单中选择“月 2015 年 5 月”,修改日期的粒度为月;拖曳度量“记

录数”到“行”功能区,产生的折线图如图 5.154 所示。

(4) 修改垂直轴结束点。双击视图区中的垂直轴,在弹出的“编辑轴”对话框中,“范围”选择“固定”,“固定结束”的值设定为 7200,如图 5.155 所示,再关闭对话框,使得折线图在坐标轴上稍微下移,从而方便后续的操作,也更为美观。

(5) 显示关键点标签。具体操作如下。

① 按住 Ctrl 键,拖“行”功能区的“总计(记录数)”到标记中的“标签”。

② 单击“标签”按钮,在弹出窗口中,选择“标签标记”中的“线末端”,在折线的两端出现了记录数的值;然后,分别右击折线中的最低点和最高点,并且在各自的弹出菜单中选定“添加标记”→“始终显示”命令。

③ 按住 Ctrl 键,拖曳“行”功能区的“年(税务登记日期)”到“标记”中的“标签”。



图 5.153 重命名工作表

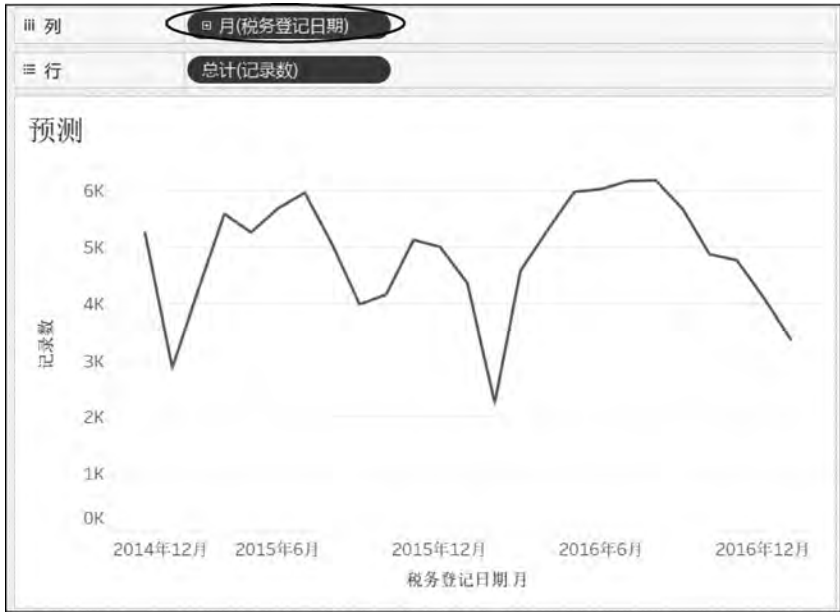


图 5.154 折线图

经过这一系列操作后,折线图上的起点、终点、最高点和最低点上显示了记录数的数值和具体日期,如图 5.156 所示。

(6) 实现预测。从“分析”标签项中的“模型”分组拖动“预测”到视图区,在视图区出现了预测折线(具有蓝色背影),但它与原有折线不连续。为此,右击“标记”下方“属性(预测指示器)”,在弹出的菜单中选择“属性”命令,则两条不连续的折线自动连接,如图 5.157 所示。

(7) 设置预测时间长度。具体操作如下。

① 右击预测折线,在弹出的菜单中,选择“预测”→“预测选项”命令,在弹出的“预测选项”对话框中,选择“预测长度”为“精确”,数值为 2,单位为“季度”,这样就可以将默认预测

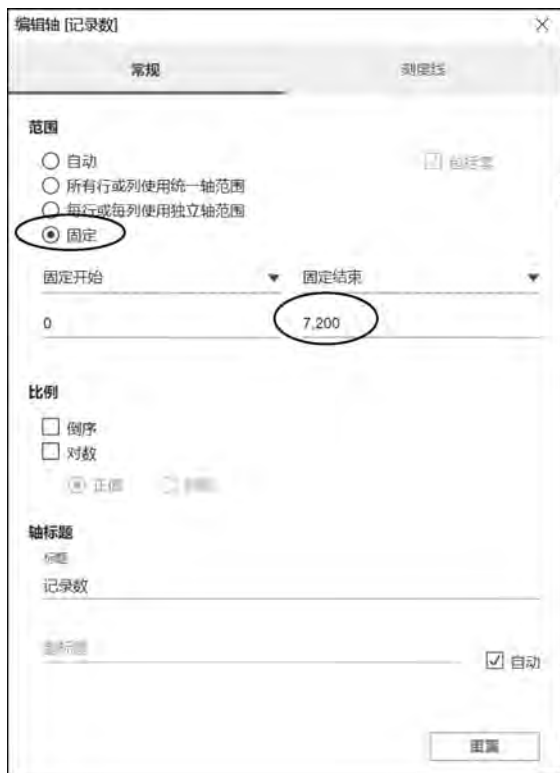


图 5.155 修改垂直轴结束点

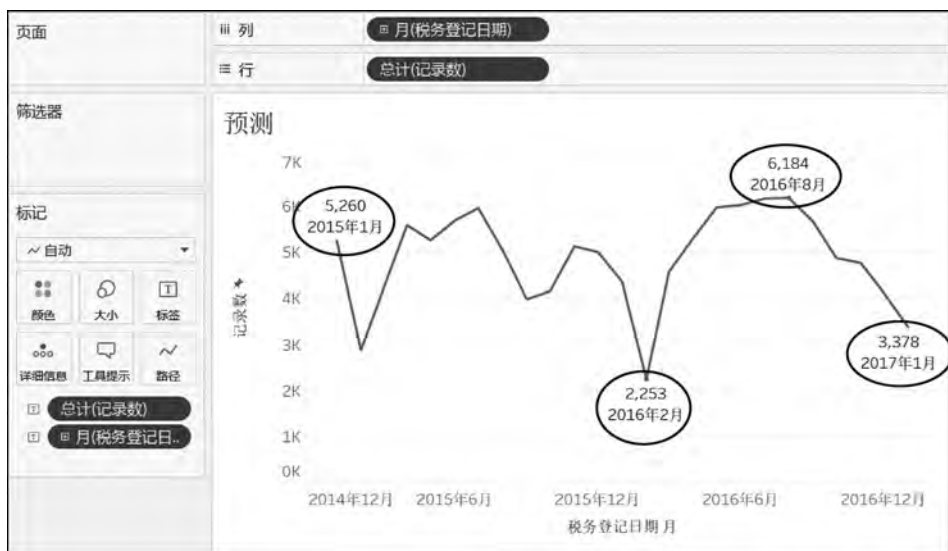


图 5.156 显示关键点标记标签的折线图

时间由 12 个月改为 6 个月(两个季度),如图 5.158 所示,再单击“确定”按钮,可以看到预测时间只有 6 个月的折线图。

② 采用同样的方法,设置显示关键点的标记标签,产生的折线图如图 5.159 所示。

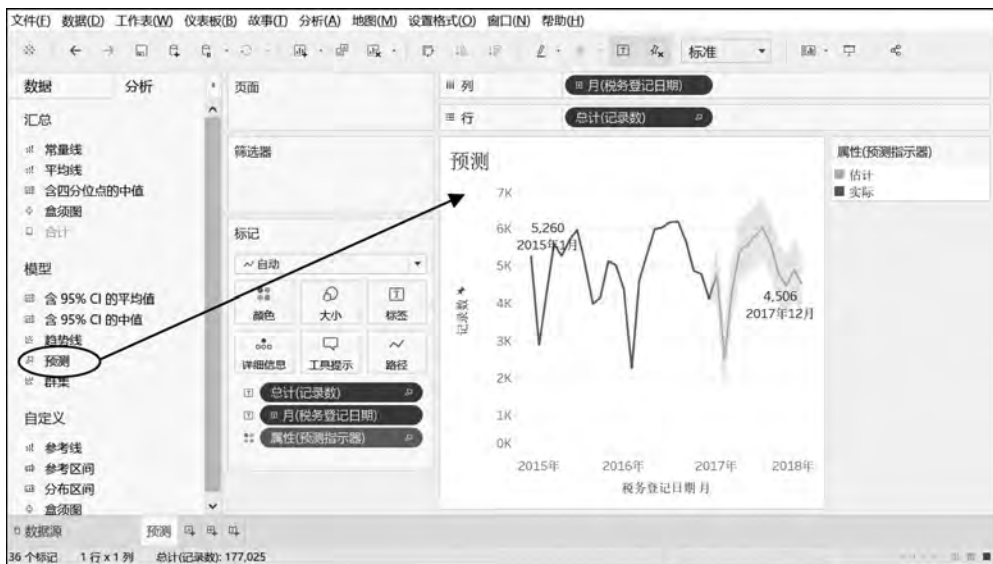


图 5.157 实现预测的折线图

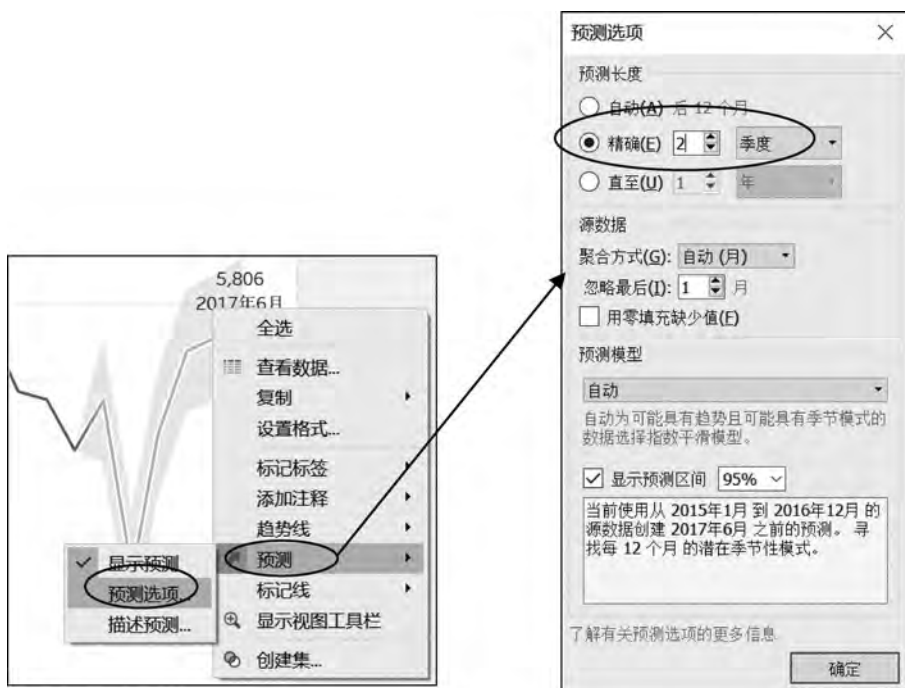


图 5.158 修改预测时间为 2 个季度

(8) 修饰折线的颜色和粗细。具体操作如下。

① 单击右方“属性(预测指示器)”旁边的倒三角形 , 选择弹出菜单中的“编辑颜色”命令, 在弹出的对话框中, 先选择左方的“估计”, 再选择右方的“橙色”, 从而设定了“估计”的折线为橙色。然后为左方的“实际”选择右方“红色”, 从而设定了“实际”的折线为红色, 如图 5.160 所示。

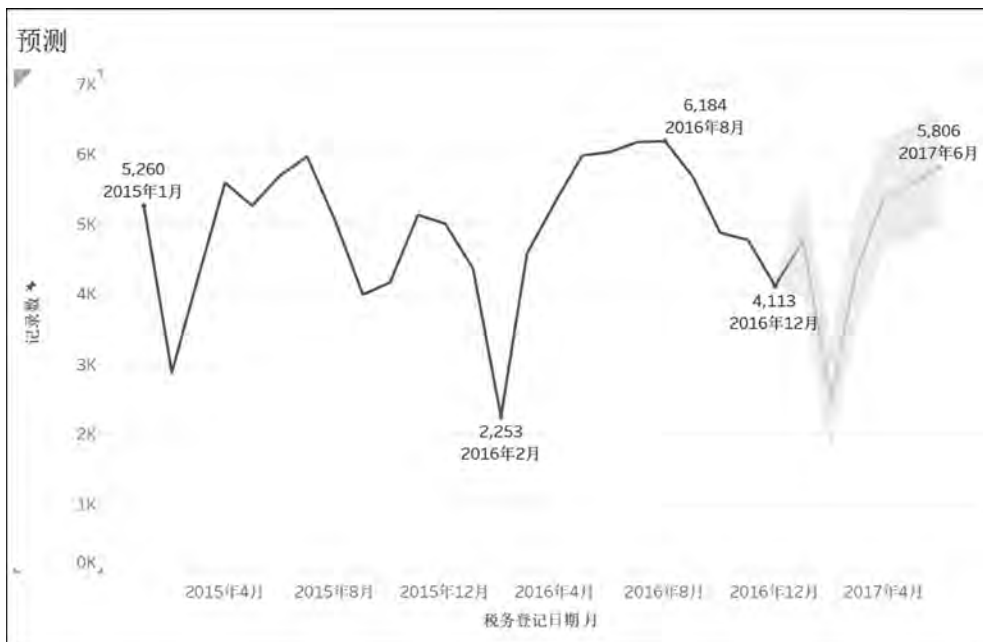


图 5.159 预测时间为两个季度的折线图

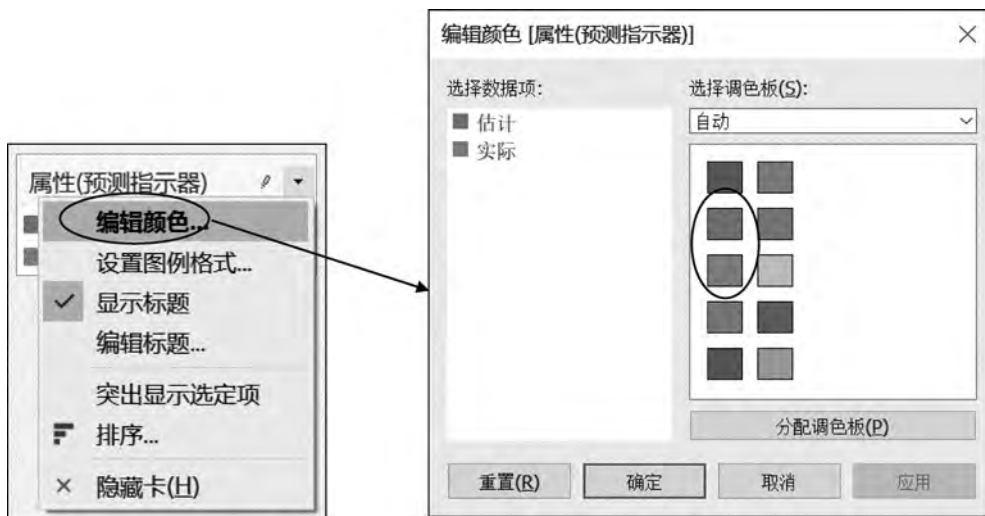


图 5.160 改变折线颜色

② 单击“标记”中的“大小”，向右拖动滑块到合适位置，使得折线变为合适粗细，如图 5.161 所示。

(9) 设置折线图线格式。右击折线图的任意空白位置，在弹出菜单中选择“设置格式”命令，在弹出的“设置格式”对话框中选择最右方的按钮设置线格式，首先单击“行”，将“网格线”和“轴标尺”都设置为无，这样就去掉了网格线和纵坐标轴；再单击“列”，将“轴标尺”设置为实线，粗细为倒数第 2 个，颜色为黑色，这样将横坐标轴设置为醒目的黑色粗实线，如图 5.162 所示。



图 5.161 改变折线粗细

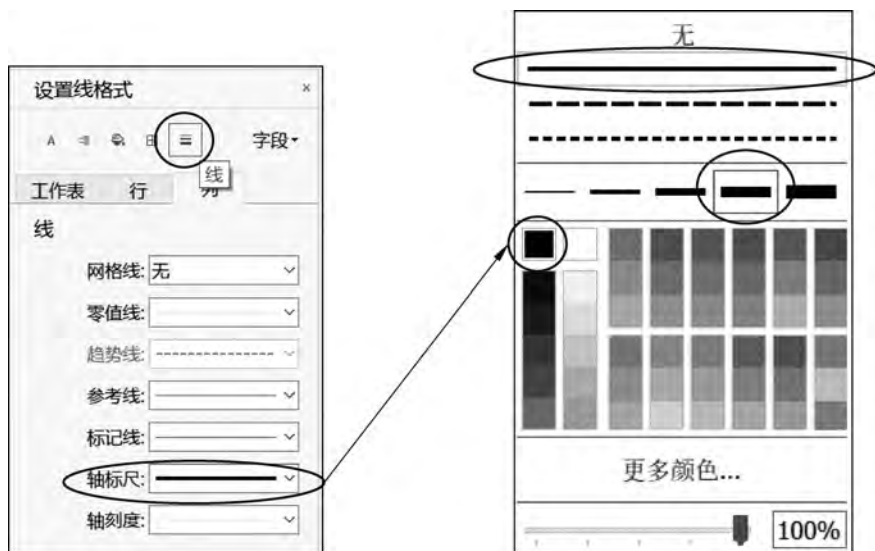


图 5.162 设置折线图线格式

完成该步骤设置后,视图区的折线图如图 5.163 所示。

(10) 添加参考线,并设置参考线格式。具体操作如下。

① 右击纵坐标轴,在弹出菜单中选择“添加参考线”命令,在弹出的“编辑参考线、参考区间或框”对话框中,“线”的“标签”选择“自定义”,在右方文本框中输入“均值,<值>”(注意是英文状态下的符号),如图 5.164 所示。

单击“确定”按钮完成设置,则在视图区中出现了一条均值参考线,如图 5.165 所示。

② 设置参考线格式:首先,右击视图中的参考线,在弹出菜单中选择“设置格式”命令,在弹出的“设置参考线格式”对话框中,“字体”选择“微软雅黑,8pt”;“对齐”中水平方向选择“居中”,垂直方向选择“顶部”;“数字”选择“数字(自定义)”→“小数位数”设置为 0(即为整数),如图 5.166 所示。

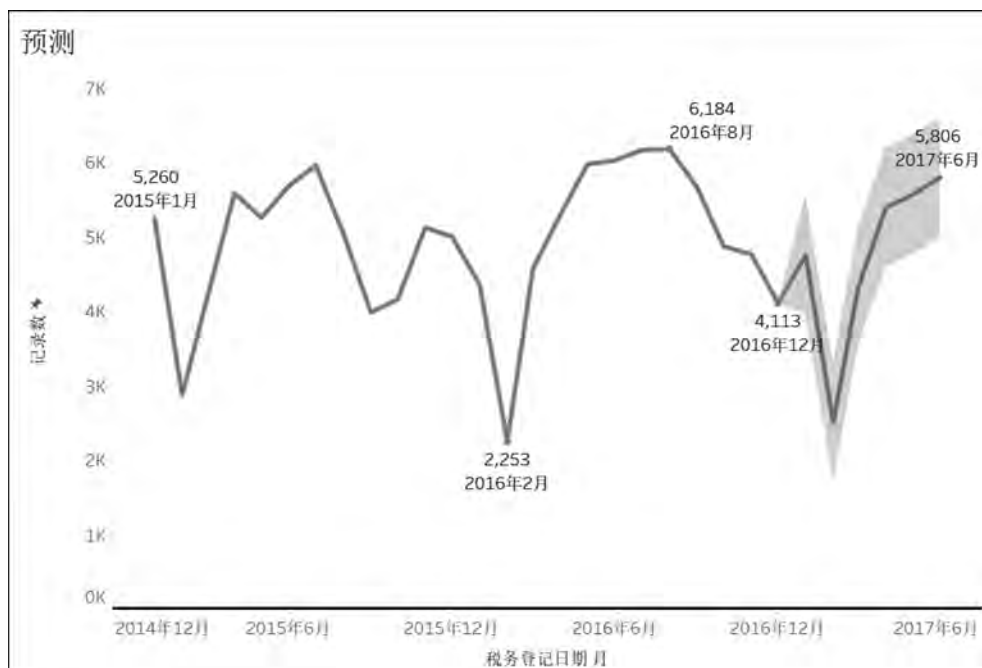


图 5.163 设置线格式后的折线图

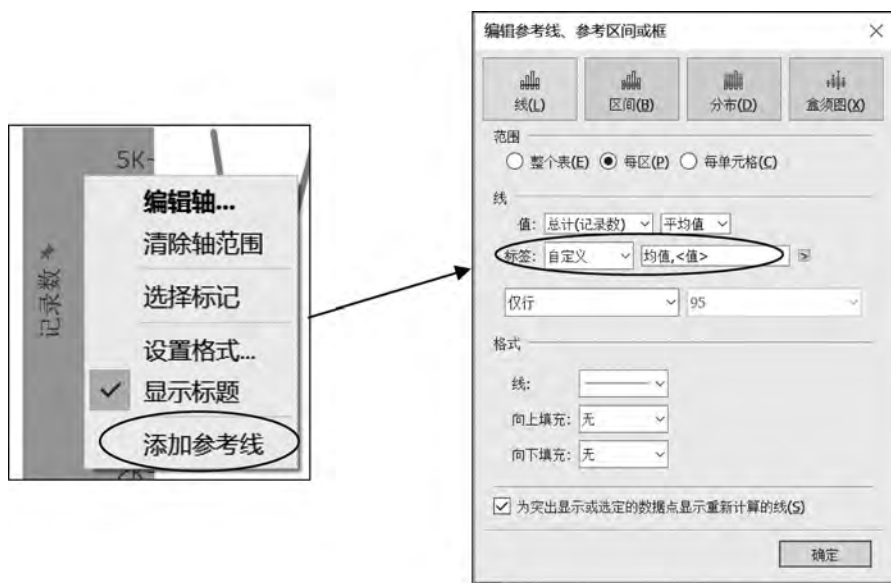


图 5.164 添加参考线



图 5.165 均值参考线

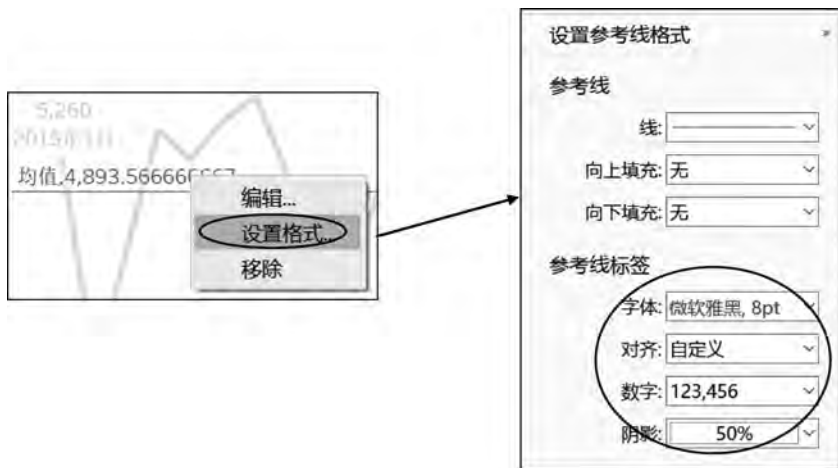


图 5.166 设置参考线格式

(11) 编辑标题。右击视图区中的标题“剖析-行业”，在弹出菜单中选择“编辑标题”命令，在弹出的“编辑标题”对话框中，文本框内输入“佛山市纳税企业增长情况分析(含：预测)”，按 Ctrl+A 快捷键全选刚刚输入的标题，设置字体为“微软雅黑”字号为 16，“加粗”“黑色”“居中”，如图 5.167 所示。

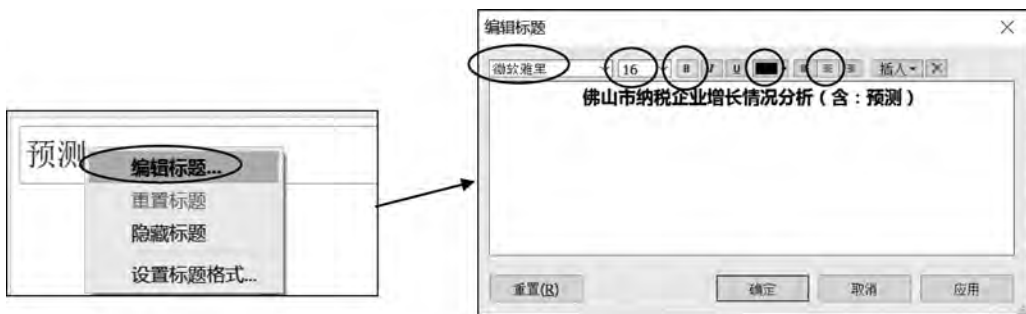


图 5.167 编辑标题

(12) 隐藏垂直轴刻度。由于折线图上的关键点已经有标签，可以隐藏垂直轴上的刻度使得整个视图更为简洁。右击垂直轴上的任一刻度，在弹出的菜单中取消选择“显示标题”，垂直轴刻度被隐藏。

那么，完成以上所有操作步骤后，佛山市纳税企业增长情况分析(含：预测)折线图创建完毕，如图 5.168 所示。

5.7.3 剖析-行业

本小节展示佛山市各行业纳税企业占比的情况。具体步骤如下。

(1) 新建一个工作表，重命名为“剖析-行业”。

(2) 生成气泡图。具体操作如下。

① 分别双击“维度”分组中的“行业”和度量“记录数”；再选择“智能显示”右下角的“填充气泡图”，如图 5.169 所示。

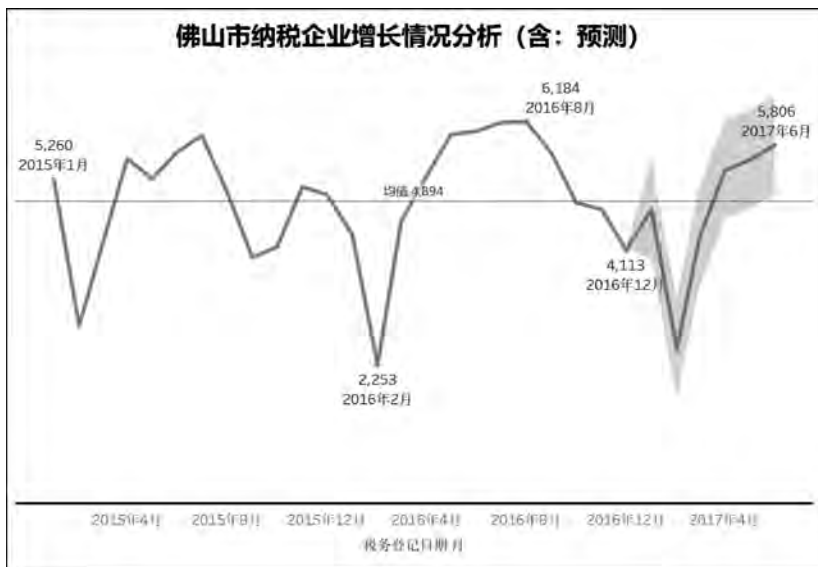


图 5.168 添加参考线后的折线图

② 按住 Ctrl 键,拖动“标记”下方的“总计(记录数)”到“标记”中的“颜色”,如图 5.170 所示。



图 5.169 选择气泡图



图 5.170 气泡颜色

③ 右击视图区空白区域,在弹出菜单中单击“标题”去掉它前面的√,从而隐藏视图区中的标题“剖析-行业”。

④ 将工具栏中的视图方法由“标准”改为“整个视图”,从而让气泡图铺满视图区。

(3) 产生记录数占比标记,具体操作如下。

① 拖动度量“记录数”到“标记”中的“标签”,那么视图区的一些气泡中出现了该行业具有的记录数(企业数)。

② 右击“标记”下方的“T 总计(记录数)”,在弹出的菜单中,选择“快速表计算”→“总额百分比”命令,如图 5.171 所示,视图区气泡中的记录数数值将变为记录数百分比。

③ 再次右击“标记”下方的“T 总计(记录数)”,在弹出的菜单中,选择“设置格式”,在左方“默认值”分组中,将“数字”选择“百分比”,且“小数位数”为 2,如图 5.172 所示。

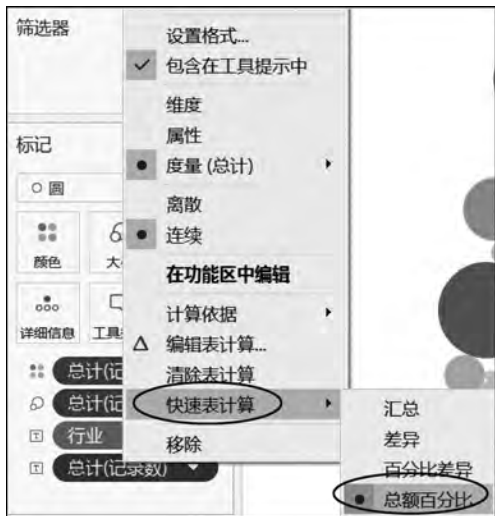


图 5.171 修改为总额百分比

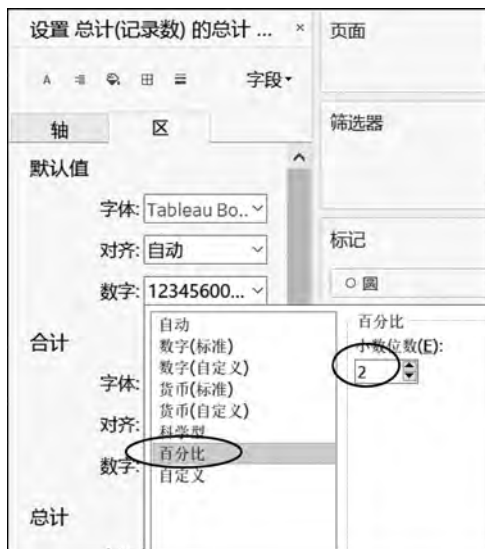


图 5.172 修改百分比小数位数

④ 单击“标记”中的“标签”按钮,在弹出的窗口中,单击“文本”右方的 [] 按钮,在弹出的“编辑标签”对话框中,使用 Ctrl+A 快捷键全选文本后,选择“微软雅黑”字体和 8 号字号,如图 5.173 所示,最后单击“确定”按钮。

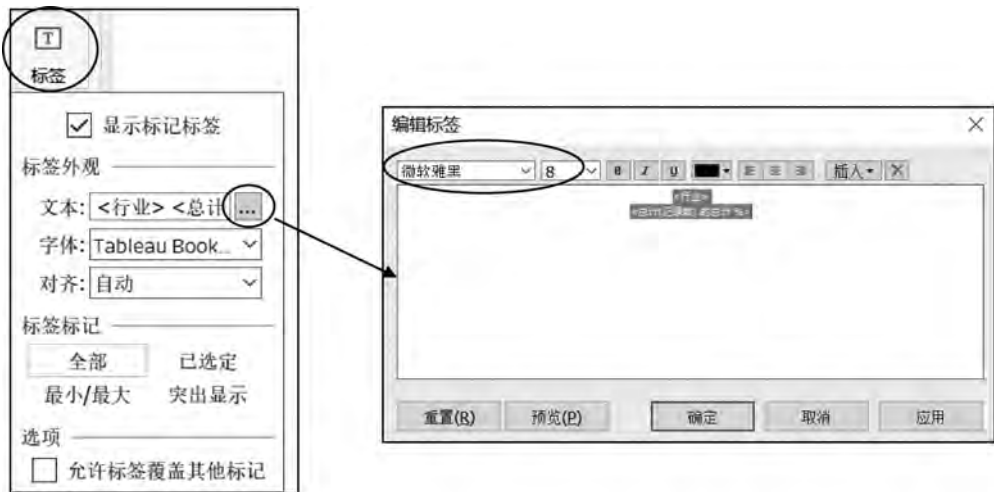


图 5.173 设置标签格式

该步骤操作完成后,视图区如图 5.174 所示。

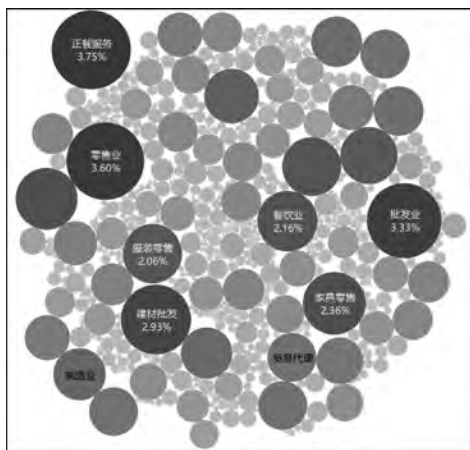


图 5.174 气泡图

[注意] 视图区中只有比较大的气泡才会显示标签,小气泡的标签则被隐藏。

5.7.4 剖析-区域

本节展示佛山市纳税企业在各个区域的数量情况。具体步骤如下。

(1) 新建一个工作表,重命名为“剖析-区域”。

(2) 生成柱状图,具体操作如下。

① 分别双击“维度”分组中的“所属区域”和“度量”分组中的“记录数”。

② 视图区中的表显示有一条记录“所属区域”值为 Null,这意味未准确识别出企业所属区域(创建“所属区域”字段时提过此问题)。右击 Null,在弹出菜单中选择“排除”命令,在分析时忽略这一条记录。

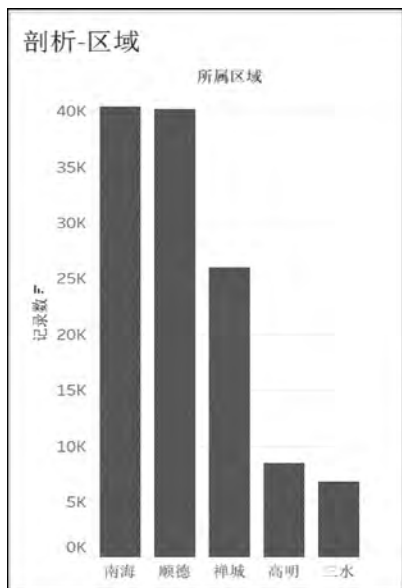


图 5.175 垂直柱形图

③ 选择“智能显示”第 3 行第 1 列的“水平条”,则在视图区产生了一个水平柱形图。

④ 单击工具栏中的“交换行和列”按钮 ,从而将视图区中的水平柱形图转换为垂直柱形图。

⑤ 单击工具栏中“降序”按钮 ,从而将视图区中柱形按记录数从大到小排序,即左方柱形较高。

该步骤操作完成后,视图区如图 5.175 所示。

(3) 修饰柱形图。具体操作如下。

① 编辑颜色。按住 Ctrl 键,拖动“行”功能区中的“总计(记录数)”到“标记”中的“颜色”,视图区中柱子的颜色按照记录数的大小着深浅色,即颜色深浅呈现梯度变化。为使得颜色更美观,右击“标记”中的“颜色”按钮,在弹出的窗口中单击“编辑颜色”,弹出“编辑颜色”对话框,将“色板”由“自动”改变为“温度发散”,如图 5.176 所示。

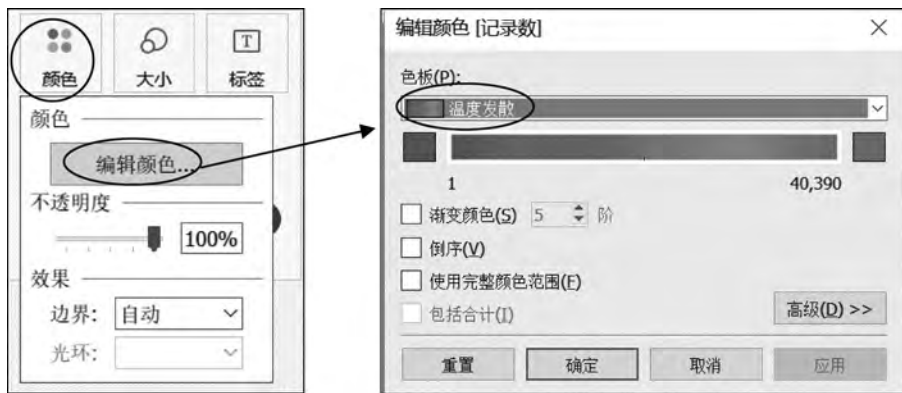


图 5.176 编辑颜色

② 展示标签。按住 Ctrl 键，拖动“行”功能区中的“总计(记录数)”到“标记”中的“标签”，各个区域的记录数呈现在对应柱形的上方。为使记录数呈现在柱形下方，单击“标记”中的“标签”，在弹出的窗口中，首先单击“对齐”右方的 ▾，在弹出的窗口中，“水平”选择“居中”☐，“垂直”选择“底部”☐，然后单击“文本”右方的 ⋮，在弹出的对话框中，按 Ctrl+A 快捷键全选框内文本，再选择“微软雅黑”字体和 8 号字号，“颜色”选择“白色”，如图 5.177 所示。

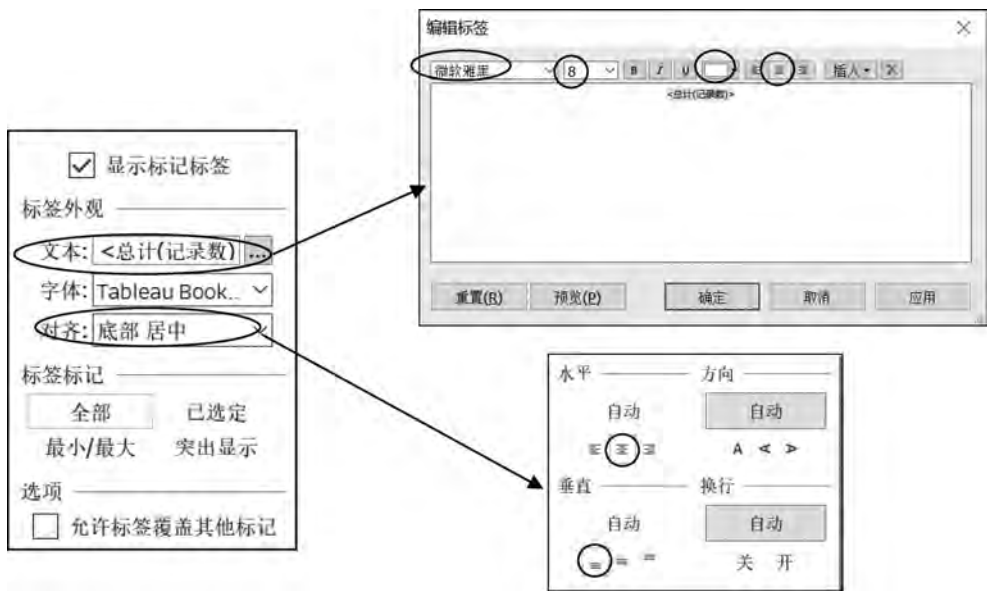


图 5.177 设置标签格式

③ 设置线格式，此处操作与前面设置折线图的线格式的操作类似。右击柱形图背景的任意空白位置，在弹出的菜单中，选择“设置格式”命令，在左方出现的窗口中，单击按钮 ☐ 设置线格式。接下来的操作中，首先单击设置“行”，将“网格线”和“轴标尺”都设置为无，这样就去掉了网格线和纵坐标轴；再单击设置“列”，将“轴标尺”设置为实线，粗细为倒数第 2 个，颜色为黑色，如图 5.178 所示，这样将横坐标轴设置为醒目的黑色粗实线。

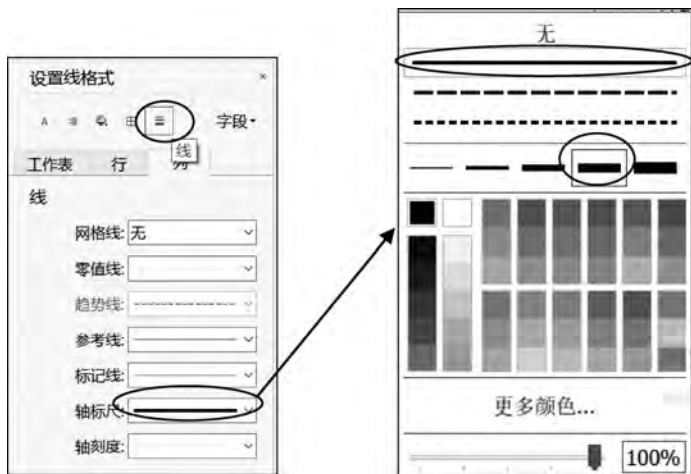


图 5.178 设置线格式

④ 隐藏垂直轴刻度。右击垂直轴中的某个刻度,在弹出的菜单中取消选择“显示标题”,所有垂直轴刻度被隐藏。

⑤ 隐藏标题。右击视图区中的标题“剖析-区域”,在弹出的菜单中选择“隐藏标题”命令。

⑥ 设置列标题。右击“列”功能区的“所属区域”,在弹出的菜单中选择“设置格式”命令,在“标题”选项卡下的“默认值”分组,设置字体为“微软雅黑”,12pt 字号,“黑色”。

⑦ 将工具栏中的视图方式由“标准”改为“整个视图”。

该步骤操作完成后,视图区如图 5.179 所示。

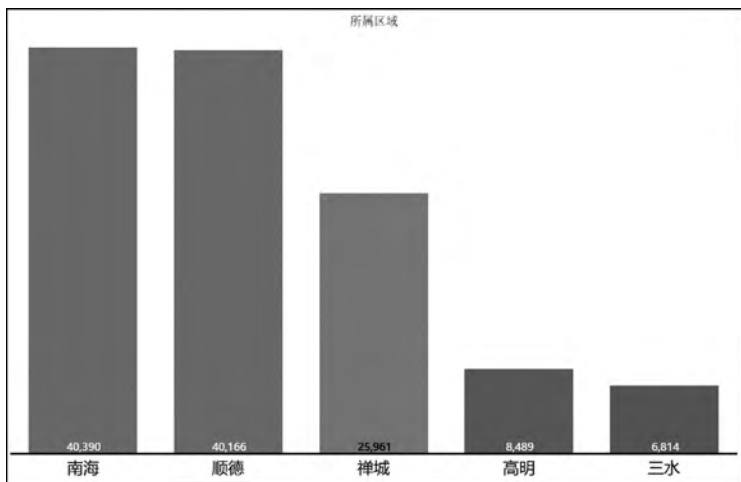


图 5.179 柱形图



5.7.5 制作仪表板

本节将基于前面创建的“预测”“剖析-行业”和“剖析-区域”3 张工作表制作仪表板,具体步骤如下。

(1) 单击左下方按钮  创建仪表板。仪表板大小固定,自定义为 1800×1000 ,即宽度为 1800px ,高度为 1000px ,如图 5.180 所示。

(2) 排版对象(图表),具体操作如下:

① 在“对象”分组设置以“浮动”的方式添加对象,以方便在仪表板上排版,如图 5.181 所示。

② 将“工作表”分组的“剖析-行业”工作表拖动到仪表板视图区;切换到“布局”选项卡,取消“显示标题”,设定“位置”中的 x 值为 91 , y 值为 210 ,”大小”中的“宽”值为 670 ,”高”值为 621 ;单击选定气泡图图例,再单击右上角的“ $\times$ ”移除图例,如图 5.182 所示。

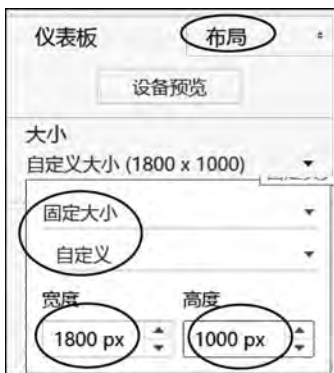


图 5.180 仪表板大小



图 5.181 设定浮动方式



图 5.182 设置气泡图布局

③ 将“工作表”分组的“预测”工作表拖动到仪表板视图区空白区域;切换到“布局”选项卡,设定“位置”中的 x 值为 848 , y 值为 144 ,”大小”中的“宽”值为 832 ,”高”值为 423 ;单击选定折线图图例,再右击,在弹出菜单中取消选择“显示标题”,然后把光标放置在图例上方,当光标变成十字架形状时拖动图例到折线图右上角边上,最后利用鼠标拖动图例边角,将图例调整为合适大小。

④ 将“工作表”分组的“剖析-区域”工作表拖动到仪表板视图区空白区域;切换到“布局”选项卡,取消“显示标题”,设定“位置”中的 x 值为 848 , y 值为 607 ,”大小”中的“宽”为 831 ,”高”为 265 ;单击选定柱形图图例,再单击右上角的“ $\times$ ”移除图例。

完成这一步骤的操作后,单击工具栏上的按钮  或者按下 $F7$ 键,进入演示模式查看总体布局情况,如图 5.183 所示。

(3) 插入背景。具体操作如下。

① 先在“对象”分组中选择“平铺”方式添加对象,以方便在仪表板上添加背景图像。再双击“图像”按钮,在弹出对话框中选择“222.jpg”作为背景图像,如图 5.184 所示。

② 在“顶分层结构”分组选择“仪表板 1”→“平铺”→“222.jpg”,再右击“222.jpg”,在弹出菜单中选择“适合图像”,再次右击“222.jpg”,在弹出菜单中选择“使图像居中”命令。

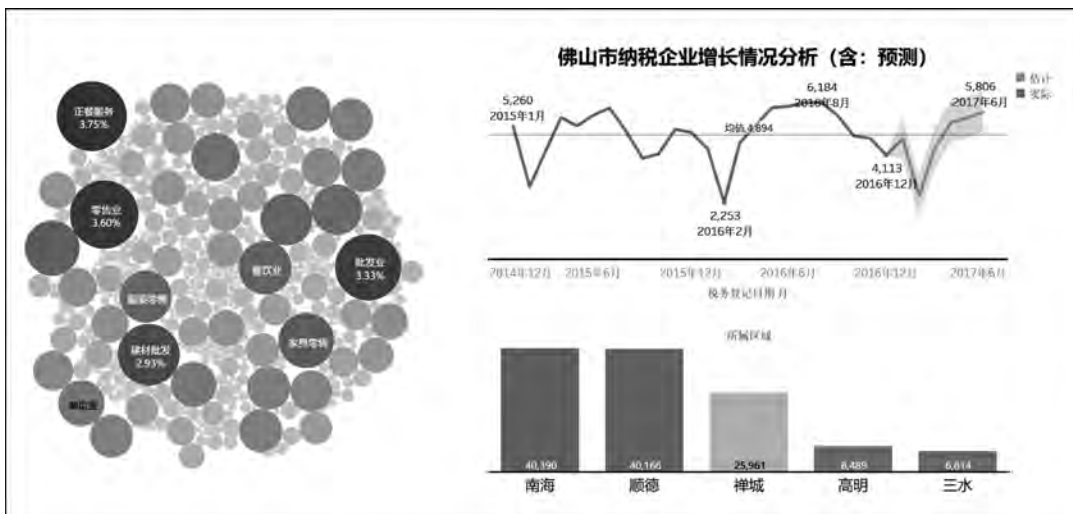


图 5.183 整体布局

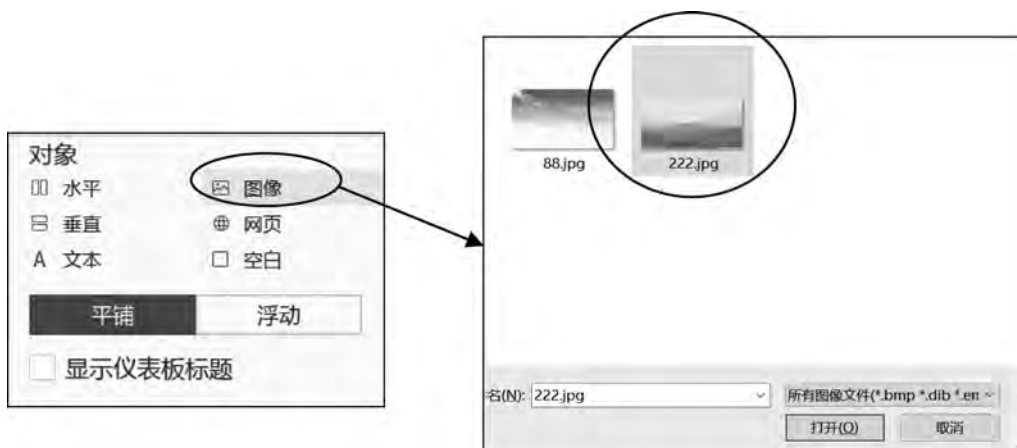


图 5.184 选择背景图像

③ 选择视图区的“预测”图表,再选择工具栏下“设置格式”→“阴影”命令,在窗口左方出现“设置阴影格式”窗格,在“工作表”选项卡,设定“默认值”分组下的“工作表”为“无”;再逐一选择视图区的“剖析-行业”和“剖析-区域”图表,在“工作表”选项卡,分别设定“默认值”分组下的“工作表”为“无”,从而实现所有图表的背景透明。

[注意] 此操作仅在 Tableau 2018.3 以及之后版本可用。

该步骤操作完成后,仪表板在演示模式下如图 5.185 所示。

(4) 实现图表之间联动。单击选择仪表板中的柱形图,再单击右上方的“用作筛选器”漏斗状按钮 ,成功设置后,空漏斗将变成实漏斗,实现了以柱形图作为筛选器联动其他两个图表。当在柱形图中选择代表“南海”的柱子时,气泡图呈现了南海区各行业企业数量的占比情况,折线图呈现了南海区纳税企业增长情况以及预测,如图 5.186 所示。在柱形图中选择代表不同区域的柱子,气泡图和折线图也会发生相应的变化。

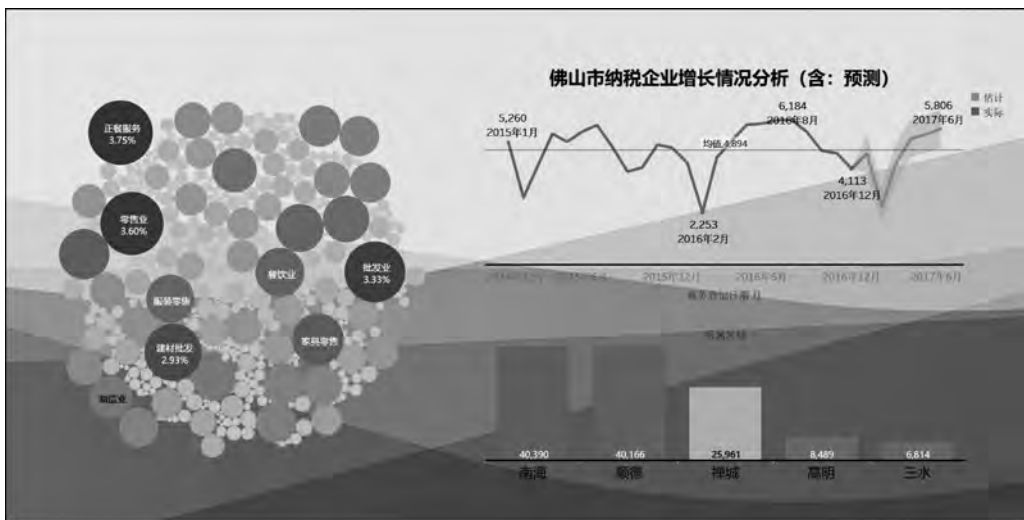


图 5.185 演示仪表板

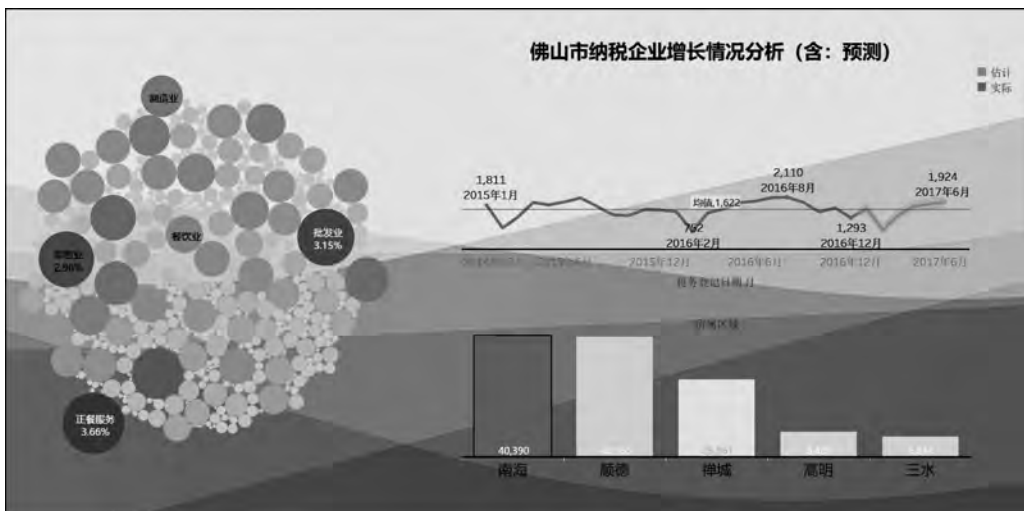


图 5.186 南海区纳税企业分析

5.7.6 分析

基于 5.7.5 节完成的仪表板,本节对佛山以及它下面各个区域的纳税企业增长情况进行一个简单的演示或者报告。

1. 佛山市

2015 年和 2016 年佛山市纳税企业新增情况如图 5.187 所示。依据折线图可以看出,佛山市的纳税企业平均每月增长 4894 家,每年年初和年末的新增企业数较少,年中较多,预计 2017 年 6 月新增纳税企业数是 5806 家。依据气泡图可以看出,佛山市新增的纳税企业中,正餐服务的占比最高,这是由于佛山是粤菜的发源地,人们很注重日常饮食;其次是零售业;第三是批发业;第四是建材批发;其他行业的占比相对较小。

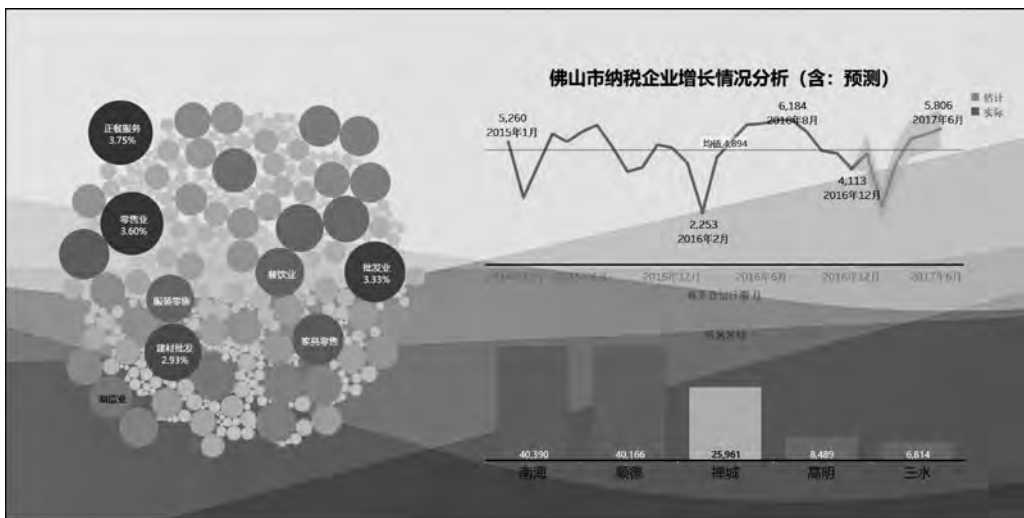


图 5.187 佛山市纳税企业增长情况

2. 南海

单击柱形图中代表“南海”的柱形,仪表板视图区呈现了在 2015 年和 2016 年南海地区纳税企业新增情况,如图 5.188 所示。依据柱形图可以看出,南海两年一共新增纳税企业 40390 家。依据折线图可以看出,南海纳税企业平均每月增长 1622 家;预计 2017 年 6 月新增纳税企业数是 1924 家。依据气泡图可以看出,南海新增的纳税企业中,正餐服务的占比也是最高;其次是批发业;第三是零售业。

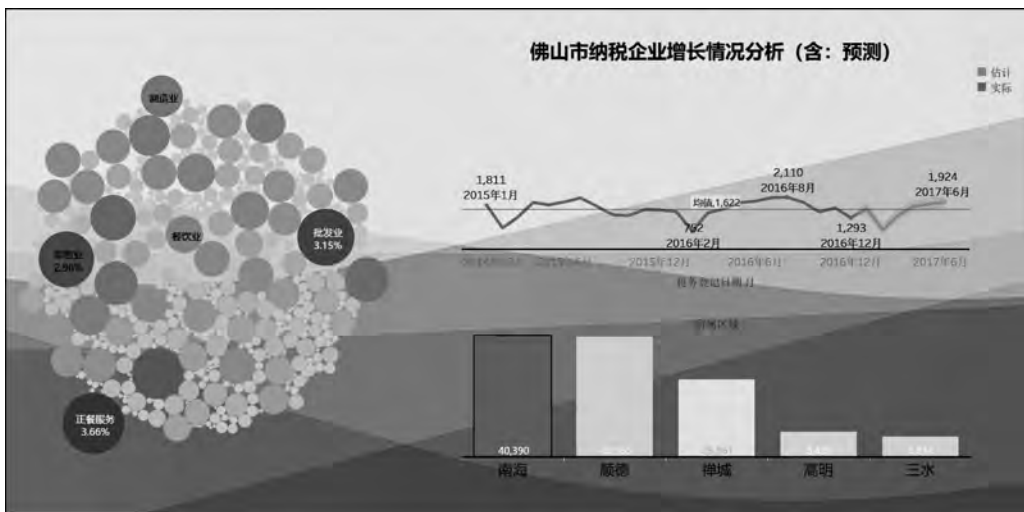


图 5.188 南海纳税企业增长情况

3. 顺德

单击柱形图中代表“顺德”的柱形,仪表板视图区呈现了在 2015 年和 2016 年顺德地区纳税企业新增情况,如图 5.189 所示。依据柱形图可以看出,顺德两年一共新增纳税企业

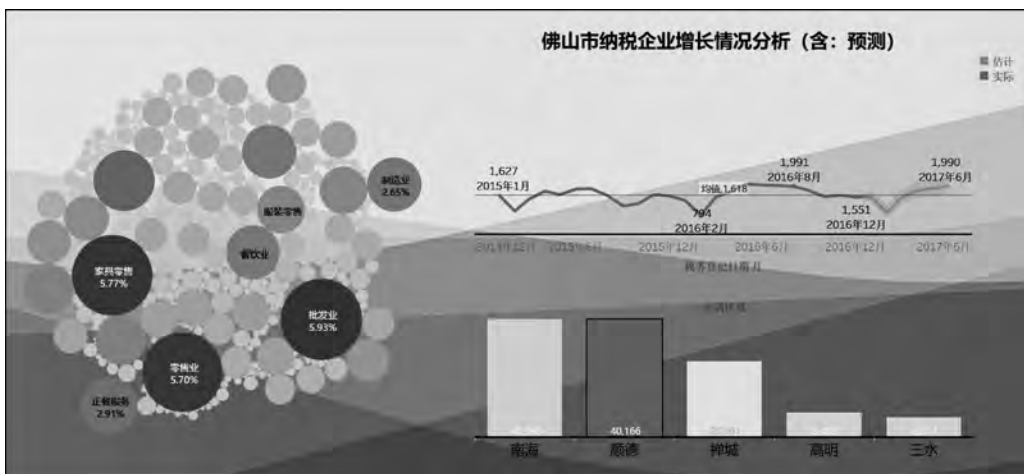


图 5.189 顺德纳税企业增长情况

40166 家。依据折线图可以看出,顺德纳税企业平均每月增长 1618 家;预计 2017 年 6 月新增纳税企业数是 1990 家。依据气泡图可以看出,顺德新增的纳税企业中,批发业占比最高;其次是家具零售;第三是零售业;第四是正餐服务。

4. 禅城

单击柱形图中代表“禅城”的柱形,仪表板视图区呈现了在 2015 年和 2016 年禅城地区纳税企业新增情况,如图 5.190 所示。依据柱形图可以看出,禅城两年一共新增纳税企业 25961 家。依据折线图可以看出,禅城纳税企业平均每月增长 1033 家,预计 2017 年 6 月新增纳税企业数是 1185 家。依据气泡图可以看出,禅城新增的纳税企业中,建材批发占比最高;其次是贸易代理;第三是正餐服务。

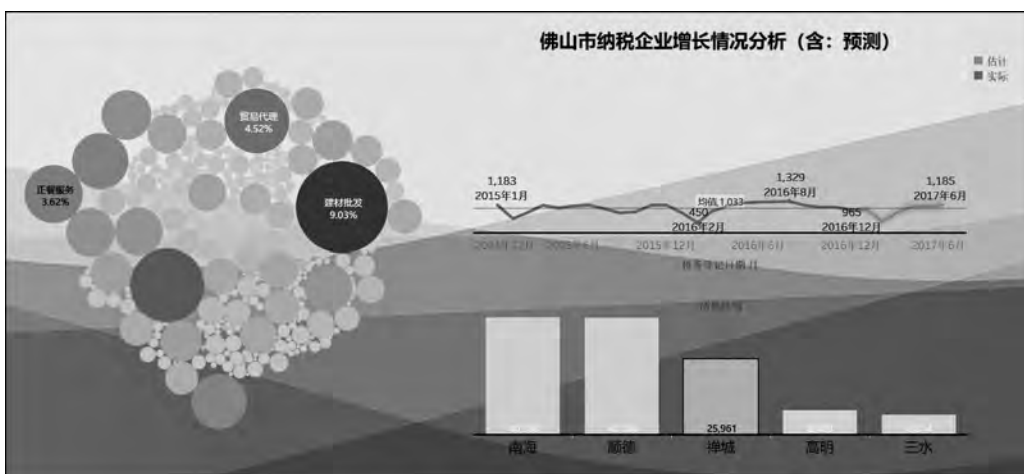


图 5.190 禅城纳税企业增长情况

5. 高明

单击柱形图中代表“高明”的柱形,仪表板视图区呈现了在 2015 年和 2016 年高明地区纳税企业新增情况,如图 5.191 所示。依据柱形图可以看出,高明两年一共新增纳税企业

8489 家。依据折线图可以看出,高明纳税企业平均每月增长 316 家,跟前面 3 个地区相比差距明显,预计 2017 年 6 月新增纳税企业数是 179 家。依据气泡图可以看出,高明新增的纳税企业中,餐饮业占比最高;其次是其他综合零售;第三是零售业,第四是正餐服务,前四位的总占比接近 30%,说明高明的产业集中度比较高,主要依赖这几个行业。

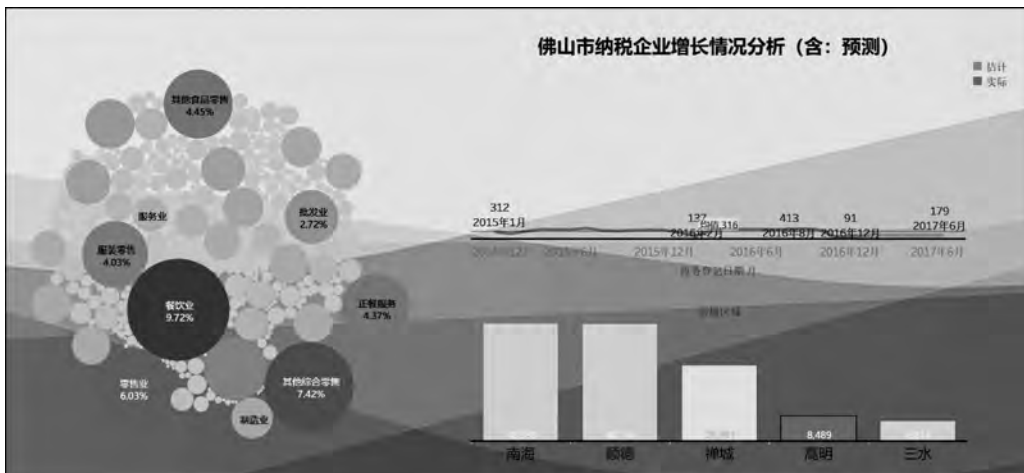


图 5.191 高明纳税企业增长情况

6. 三水

单击柱形图中代表“三水”的柱形,仪表板视图区呈现了在 2015 年和 2016 年三水地区纳税企业新增情况,如图 5.192 所示。依据柱形图可以看出,三水两年一共新增纳税企业 6814 家。依据折线图可以看出,高明纳税企业平均每月增长 278 家,预计 2017 年 6 月新增纳税企业数是 324 家,跟同期预测高明新增 179 家企业相比较,三水的增长势头更强。依据气泡图可以看出,三水新增的纳税企业中,正餐服务占比最高;其次是零售业;第三是服装零售,三水的企业主要集中于居民日常服务领域,缺乏工业和支撑工业的产业。

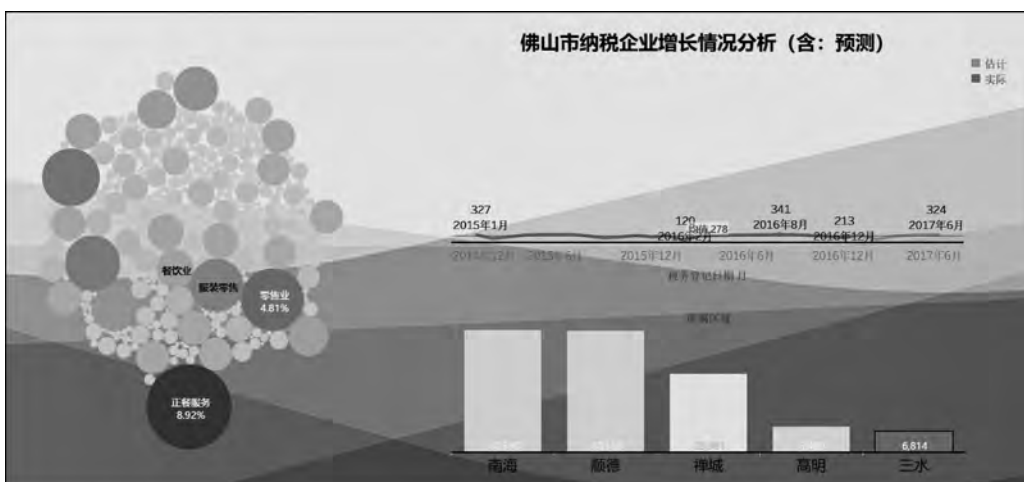


图 5.192 三水纳税企业增长情况