

第 1 章

Tableau 快速入门

作为本书的第 1 章，本章主要介绍在 Tableau Desktop 中进行数据分析所需了解的一些重要概念和术语，以及 Tableau Desktop 的界面组成和工作区的使用方法。无论是否使用过 Tableau Desktop，都建议仔细阅读本章内容，从而更好地了解和使用 Tableau Desktop。

1.1 Tableau 产品体系和功能简介

Tableau 有一套完整的产品体系，其中的各个产品分别适用于数据处理和分析过程中的各个环节。Tableau 全套产品包括 Tableau Prep、Tableau Desktop、Tableau Server、Tableau Cloud（即原来的 Tableau Online）、Tableau Public、Tableau Mobile、Tableau Reader 等。下面简要介绍其中的几个主要产品。

1. Tableau Prep

Tableau Prep 包括 Tableau Prep Builder 和 Tableau Prep Conductor 两个子产品，在分析数据之前，可以先使用 Tableau Prep 对数据进行合并、调整和整理，以便在 Tableau Desktop 中更轻松地分析数据。

2. Tableau Desktop

Tableau Desktop 是 Tableau 的核心产品，它是一个桌面端应用程序。使用 Tableau Desktop 可以连接来自文件、数据库等不同来源的数据，并构建数据模型，然后以文本、图表等多种形式呈现数据分析结果。使用 Tableau Desktop 中的仪表板可以同时查看和分析多个结果，使用 Tableau Desktop 中的故事可以将分析结果以生动的形式完整呈现给业务决策者和相关人员。本书主要介绍使用 Tableau Desktop 进行数据可视化分析的方法。

3. Tableau Server

Tableau Server 是服务器端应用程序，需要在本地服务器中安装该产品，使用该产品可以发布和管理仪表板和数据源。将在 Tableau Desktop 中制作好的仪表板发布到 Tableau Server 中，网络中的其他用户就可以使用浏览器查看仪表板的内容，便于多人交流和互动。

4. Tableau Cloud

Tableau Cloud 原名为 Tableau Online，该产品是 Tableau Server 的云托管版本，具有与 Tableau Server 类似的功能，但是其访问范围是世界级的，而非 Tableau Server 限于本地网络之内。这意味着将仪表板发布到 Tableau Cloud 后，在任何可以连接 Internet 的地方，都可以使用浏览器访问 Tableau Cloud 中的仪表板。

1.2 Tableau 中的重要概念和术语

在开始真正接触 Tableau 之前，首先需要了解 Tableau 中的几个重要概念和术语，它们将为继续学习并与其他用户进行交流提供帮助。这些概念在 Tableau 中的实现方法和具体操作，将在本书后续章节中进行介绍。

1.2.1 数据源

简单来说，数据源是指数据的来源。数据可以来自于一个文本文件或 Excel 文件，也可以来自于一个数据库，甚至是云端数据，它们是要在 Tableau 中进行分析的源数据。

将“数据源”一词放到 Tableau 环境中，其含义将变得有所不同。Tableau 中的数据源是指用户的源数据与 Tableau 之间的链接，其内容包括以实时连接或数据提取方式连接到的源数据、连接信息（例如数据的本地位置或网络位置、数据库服务器的登录信息等）、包含数据的表或工作表的名称，以及在 Tableau 中对数据进行的自定义设置（例如重命名字段、创建计算字段等）。

在第 2 章实际操作数据源时，会对上述概念有更深刻的理解。

1.2.2 字段、维度和度量

字段是指表中的列，一个表有几列数据，在 Tableau 中就会有几个字段，列和字段本质上是可互换的术语，在 Tableau 中通常使用“字段”描述列数据。由同一行中各个字段组成的数据称为记录，每一行数据都是一条记录。

用户使用的字段是由连接到 Tableau 中的数据源自动提供的。然而，Tableau 默认会创建以下几个字段：度量名称、度量值、记录数或表名称（计数）。如果数据源包含地理字段，则还会创建“纬度（生成）”和“经度（生成）”两个字段。

列中的数据称为字段的值或成员，不同列中的数据可以是文本、数值、日期等不同的数据类型。数值类型的数据可以是正值或负值，这取决于字段本身的含义。例如，“销售利润”字段的值可以为正也可以为负，“销售额”字段的值只能为正值。

由于可以对数值类型的数据进行计算，例如求和或计数，而通常不会对文本类型的数据进行计算，因此，可以将字段分为度量和维度两类。维度主要用于描述事物而非计算，名称、类别、颜色、日期等字段都是维度；度量主要用于对数值进行计算，销量、销售额、浏览量、人数等字段都是度量。

除了维度和度量之外，字段还分为离散和连续两种。离散是指各自分离且不同，范围是有限的；连续是指一个不间断的整体，范围是无限的。维度通常是离散的，度量通常是连续的，但也并非必须如此，这意味着维度可以是连续的，度量可以是离散的。在 Tableau Desktop 中，离散的维度和度量显示为蓝色，连续的维度和度量显示为绿色。

在 Tableau Desktop 中，离散字段在图表中以标题的形式出现，连续字段在图表中以轴的形式出现。标题意味着信息是分段不连续的，轴意味着刻度从 0 开始的一系列连续的值。

在图 1-1 (a) 中，由于“数量”是一个离散的维度字段，并且位于“列”功能区中，因此将该字段创建为水平标题，并使用柱形图分段对比不同数量的利润。在图 1-1 (b) 中，由于“数量”是一个连续的维度字段，并且位于“列”功能区中，因此将该字段创建为水平轴，并使用折线图表示一个连续的趋势。

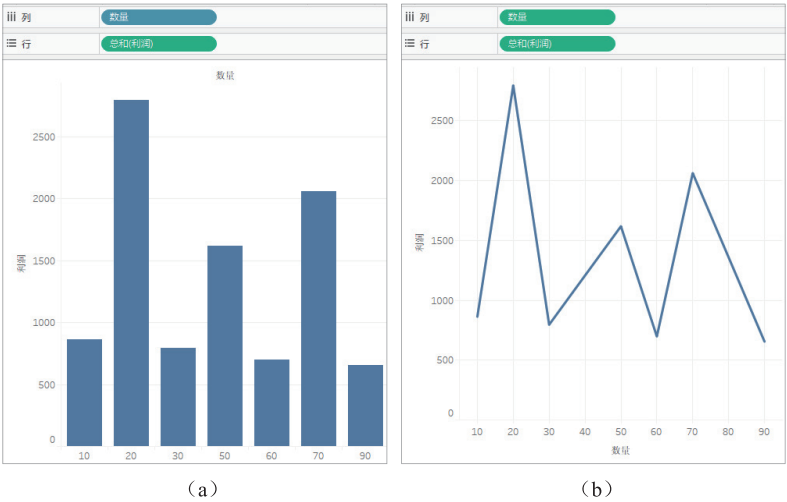


图 1-1 离散字段和连续字段

提示：在 Tableau Desktop 中，将图表和其他可视化对象统称为视图。

1.2.3 聚合和粒度

聚合是指将多个值经过计算组合为单一值，例如计算多个数值之和或求它们的平均值。Tableau 自动对度量字段中的值进行聚合，默认的聚合方式是求和。可以根据需要，将聚合方式更改为平均值、计数、最大值、最小值等。也可以对维度字段进行聚合，但是聚合方式只有计数、最大值、最小值几种。

粒度是指数据的详细程度，它由维度定义。一个表中的维度字段越多，表中数据所表达的信息越详细。比较图 1-2 所示的两张表，图 1-2 (a) 显示每个月的数量，图 1-2 (b) 显示每一天的数量，图 1-2 (b) 中的数据显然更详细。

日期	数量	日期	数量
1月	60	1月1日	10
2月	90	1月2日	20
3月	30	1月3日	30
4月	50	2月1日	10
5月	80	2月2日	30
6月	20	2月3日	50

(a) (b)
图 1-2 不同粒度的数据

1.2.4 数据类型

数据源中的每一个字段都具有一种数据类型。数据类型表示字段中存储的数据种类，并为 Tableau 提供有关如何格式化、解释数据以及可以对该数据执行哪些操作的信息。例如，可以对数值字段执行数学运算，为地理字段创建地图。表 1-1 列出了 Tableau 中用于标识数据类型的图标及其对应的数据类型。

表 1-1 Tableau 中的数据类型

图 标	数 据 类 型
Abc	字符串（文本）
#	数值
白	日期
白	日期和时间
T/F	布尔值
地球	地理值

1.2.5 数据模型

“数据模型”是数据分析领域最常出现的术语之一，它是指通过关系使数据相互关联在一起的一组表。数据模型中表的数量可以有两个或多个，它们之间两两相关，最后形成一张关系网。实际上，单张表也是一个数据模型，只要其中的数据结构符合规范并有利于分析，就可以认为它是一个数据模型。

1.3 Tableau Desktop 界面组成

Tableau Desktop 的界面整体分为“开始”页面、“数据源”页面、“工作区”页面 3 个部分，每个页面用于执行特定的任务。

1.3.1 “开始”页面

启动 Tableau Desktop 后，首先显示的是“开始”页面，如图 1-3 所示。“开始”页面是 Tableau Desktop 的中心位置，它由“连接”“打开”和“探索”3 个窗格组成，可以在“开始”页面中执行以下任务：

- “连接”窗格：连接新的数据源或打开已保存的数据源。数据源可以是存储在 Excel 文件、文本文件、Access 文件、Tableau 数据提取文件中的数据，也可以是存储在 SQL Server、Oracle 等数据库中的数据。根据连接频率，“连接”窗格中显示的名称会自动调整。
- “打开”窗格：曾经使用过的工作簿会以缩略图的形式显示在该窗格中，单击缩略图即可在 Tableau Desktop 中打开相应的工作簿。
- “探索”窗格：浏览 Tableau 社区中的博客文章、新闻、培训视频和教程等内容。



图 1-3 “开始”页面

提示：已保存的数据源是指存储在“我的 Tableau 存储库”文件夹中的数据源。如果手头没有可用的数据源，则可以使用 Tableau 自身提供的几个示例数据源来熟悉 Tableau Desktop 的功能。

1.3.2 “数据源” 页面

在 Tableau Desktop 中连接数据源后，将自动显示“数据源”页面。在“工作区”页面中单击“数据源”选项卡，也会显示“数据源”页面。“数据源”页面分为左侧窗格、画布、数据网格 3 个区域，如图 1-4 所示。

- 左侧窗格：已连接的数据源名称及其中包含的表将显示在左侧窗格中。图 1-4 中连接到的数据源名称为“销售数据”，其中包含“商品”和“订单”两张表。
- 画布：画布位于“数据源”页面右侧的上半部分，为了对数据源中的数据进行分析，需要将左侧窗格中的一个或多个表添加到画布中，以此来构建数据模型，为数据分析做好准备。图 1-4 中已将“订单”表添加到画布中。
- 数据网格：将一个表添加到画布后，该表中的数据会显示在画布下方的数据网格中。如果画布中有多个表，则会显示当前选中的表中的数据。如需查看表中的元数据（即组成表的各个字段的名称和数据类型），可以单击数据上方工具栏中的“管理元数据”按钮。单击“浏览数据源”按钮，将重新显示表中的数据。

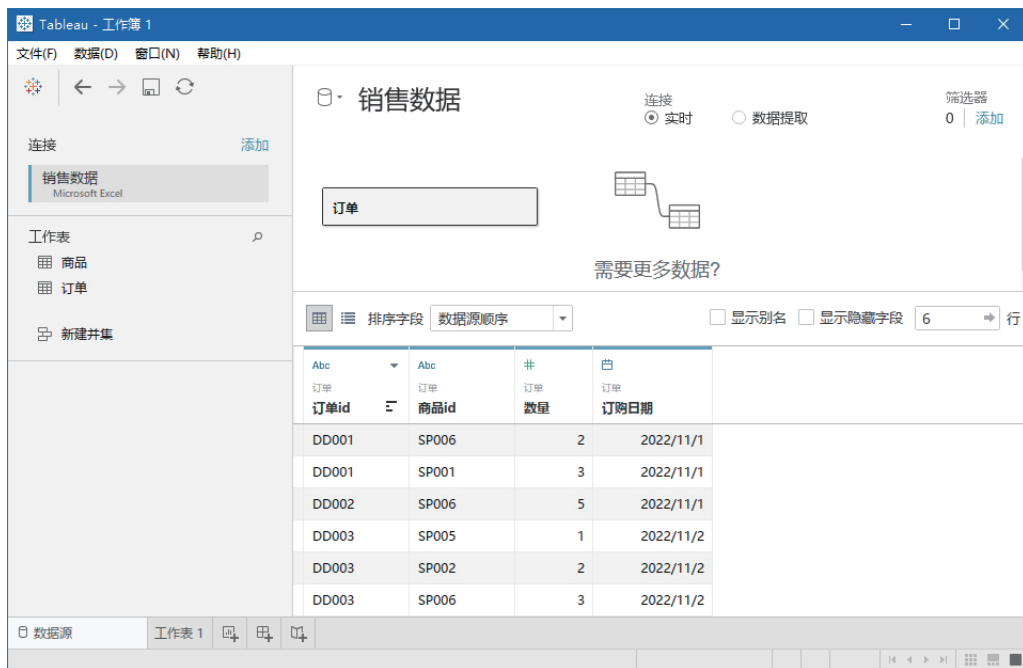


图 1-4 “数据源” 页面

1.3.3 “工作区” 页面

连接好数据源，接下来就可以在“工作区”页面中分析数据了。在 Tableau Desktop 窗口底部单击“工作表 1”选项卡，将切换到“工作区”页面，如图 1-5 所示。“工作区”页面由“数据/分析”窗格、视图、“行”和“列”功能区、“页面”功能区、“筛选器”功能区、“标记”卡、状态栏等部分组成。

“工作区”页面的下方是用于切换工作表、仪表板和故事的标签，单击这些标签可以在工作表、仪表板或故事之间切换。在 Tableau Desktop 中创建的文件称为“工作簿”，每个工作簿至少要包含一个工作表，当然，也可以包含多个工作表，还可以包含零个或多个仪表板和故事。

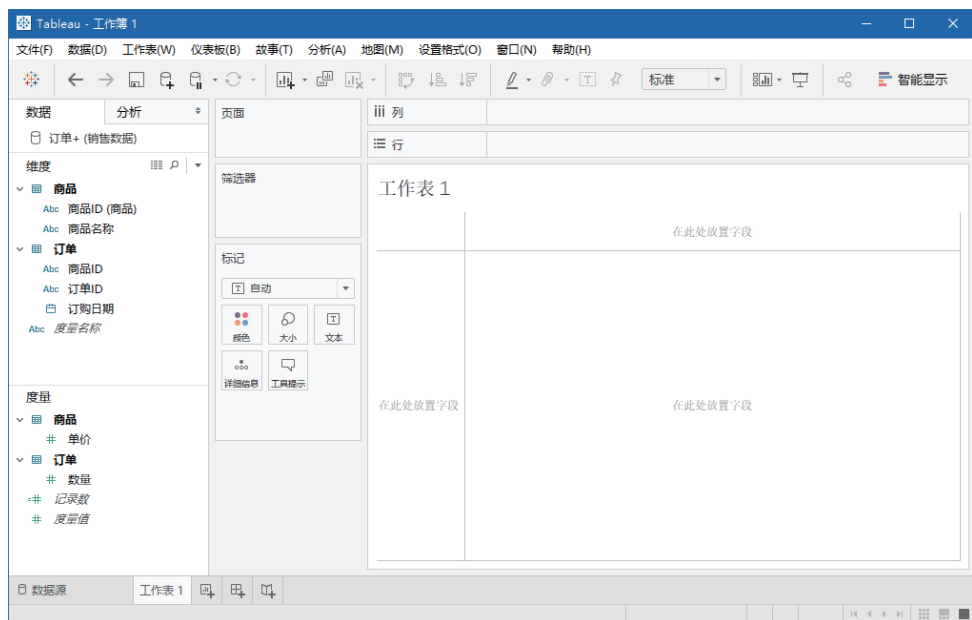


图 1-5 “工作区”页面

“工作区”页面的各部分的功能如下：

- “数据 / 分析”窗格：如果在“数据源”页面的画布中添加了一个或多个表，则会在“数据”窗格中显示这些表中的字段，并自动将所有字段划分为维度和度量。“分析”窗格中的选项用于为视图添加高级分析功能，例如参考线、平均线、盒须图等。
- 视图：视图是用户在 Tableau Desktop 中创建的一种可视化项，可以由文本和数字组成的表格，也可以是由标题、轴、形状和颜色组成的图表，还可以是表示地理位置的地图。
- 功能区和卡：功能区和卡包括“行”和“列”功能区、“页面”功能区、“筛选器”功能区、“标记”卡几个部分，将字段放置到这些位置上即可创建视图。某些功能区和卡只有在执行特定操作时才会显示，例如“度量值”功能区。
- 状态栏：状态栏位于 Tableau Desktop 窗口的底部，其中显示当前视图的相关信息。

1.4 熟悉 Tableau Desktop 工作区

使用 Tableau Desktop 分析数据的绝大多数工作都是在“工作区”页面中进行的，所以非常有必要熟悉该页面的布局和功能。

1.4.1 “数据”窗格

“数据”窗格位于“工作区”页面的左侧，其中显示在“数据源”页面的画布中添加的表中的所有字段，并自动按照表名以及维度和度量对字段进行分类。每个字段名称左侧的图标识字段的数据类型。

在一些 Tableau Desktop 版本的“数据”窗格中显示“维度”和“度量”两个标签，所有表中的字段以维度和度量作为分类依据，以表为单位分别显示在“维度”和“度量”标签下，“维度”和“度量”两类字段之间以灰色横线分隔，如图 1-6 (a) 所示。

在另一些 Tableau Desktop 版本的“数据”窗格中不显示“维度”和“度量”两个标签，每个表中的所有维度字段和度量字段都以本表为单位进行分类，同一个表中的两类字段之间以灰色横线分隔，维度位于横线上方，度量位于横线下方，如图 1-6 (b) 所示。

已连接的数据源的名称显示在“数据”窗格的顶部，如图 1-7 所示。如果连接了多个数据源，则需要要在它们之间选择一个当前要使用的数据源，选中的数据源中的字段会显示在“数据”窗格中。

如需更改所有字段在“数据”窗格中的分组和排列方式，可以单击该窗格中的下拉按钮，在弹出的菜单中选择如图 1-8 所示的选项。



图 1-6 Tableau Desktop 不同版本的“数据”窗格



图 1-7 在“数据”窗格的顶部显示已连接的数据源名称

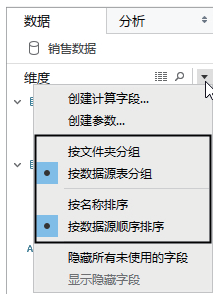


图 1-8 选择字段的分组和排列方式

如果不想显示“数据”窗格，则可以将其隐藏起来，有以下两种方法：

- 单击“数据”窗格右上角的 或 按钮，该按钮在 Tableau Desktop 的不同版本中具有不同的外观。
- 单击 Tableau Desktop 窗口菜单栏中的“窗口”，在弹出的菜单中选择“显示边条”命令，取消该命令的选中状态，如图 1-9 所示。

提示：“显示边条”命令在有的 Tableau Desktop 版本中的名称为“显示边栏”。

如需重新显示“数据”窗格，只需重复执行以上两种操作之一即可。

在“数据”窗格执行的都是与字段有关的操作，包括但不限于以下任务：

- 将“数据”窗格中的字段添加到视图或各个功能区中，构建可视化项。
- 更改字段的名称、数据类型，为字段设置别名。
- 设置字段的默认属性。
- 为字段创建分层结构。
- 将字段在维度和度量之间转换。

上述内容将在本书后续章节中进行详细介绍。



图 1-9 取消“显示边条”命令的选中状态

1.4.2 视图

视图是 Tableau Desktop 中的可视化元素的集合，在“工作区”页面中占据最大的区域，如图 1-10 所示。在一个视图中可以包含标题、标签、轴、单元格、标记、图例、说明、工具提示等多种元素，用户可以灵活调整和设置这些元素的显示状态和外观格式。

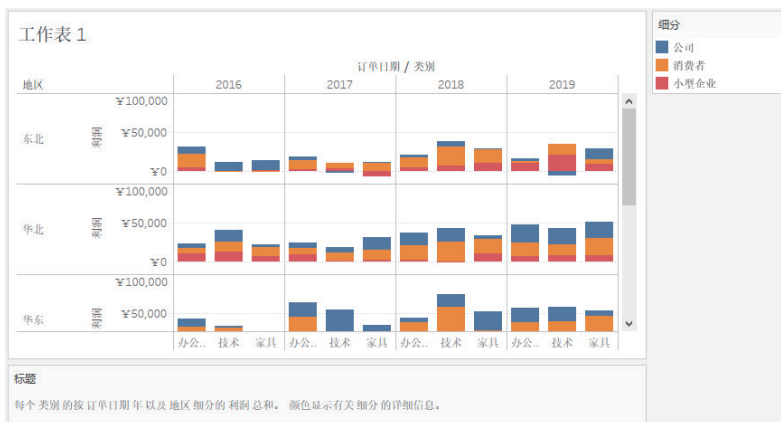


图 1-10 视图

图 1-10 中的视图包含以下元素：

- 工作表标题：“工作表 1”是工作表标题，默认为工作表选项卡标签上的名称，用户修改工作表标签名称时，工作表标题会同步更新。
- 字段标签：“订单日期”“类别”和“地区”都是字段标签，它们是添加到“行”和“列”功能区中的离散字段的名称。
- 标记：图中的所有矩形条都是标记，标记以文本或图形的方式表示视图中的数据。除了矩形条，标记还可以是其他形状、线条、文本、地图等对象。
- 标题：“订单日期”字段中的年份，“类别”字段中的“办公”“技术”和“家具”，“地区”字段中的“东北”“华北”和“华东”，这些都是维度字段中的成员，它们在视图中显示为水平方向或垂直方向上的标题，为视图提供信息的详细级别。
- 轴：每行第一个矩形条左侧的 3 个数字所在的位置就是轴，其上的数字是轴的刻度，此处显示的是垂直轴。将连续字段添加到“行”或“列”功能区时，将创建垂直轴或水平轴，每个轴的刻度从 0 开始。
- 图例：视图右上角有一个标题为“细分”的卡，其中的 3 项内容就是图例。在视图中使用形状或颜色时，图例用于说明每个形状或颜色代表的内容。
- 说明：视图底部的文字，用于说明视图的整体情况。

工具提示也是视图中的一种元素，它是将鼠标指针悬停在视图中的一个或多个标记上时显示的信息和相关选项，如图 1-11 所示。

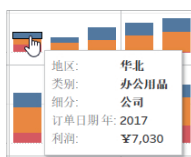


图 1-11 工具提示

Tableau 为用户创建不同类型的视图提供了极大的灵活性，将字段放置在“工作区”页面中的功能区或卡上，并使用颜色、大小、形状、文本等改变标记的显示效果，可以构建出不同结构的视图，从而以不同的角度洞察数据，获得所需的商业见解。

1.4.3 “行”和“列”功能区

“行”和“列”功能区默认位于视图的上方，可以将任意数量的字段添加到这两个功能区中，如图 1-12 所示。添加的字段数量越多，表中就会包含越多的标题和轴，视图中展示的信息也会越详细。同一个功能区中的多个维度字段的标题会形成层级结构，字段之间的排列顺序决定视图中各级标题的显示级别。

如果将维度字段添加到“行”或“列”功能区，则会使用维度字段中的成员名称在视图中创建行标题或列标题。如图 1-13 所示是将“订单日期”字段添加到“列”功能区时在视图中创建的列标题。标题下方的 Abc 表示文本标记值将出现的位置，以后在视图中添加度量字段后，度量字段中的值会替换 Abc。



图 1-12 “行”和“列”功能区



图 1-13 “订单日期”字段在视图中创建为列标题

如果同时在“行”和“列”功能区中添加维度字段，则会在视图中创建行标题和列标题，如图 1-14 所示。

如果将度量字段添加到“行”或“列”功能区，则会根据度量字段中的值在视图中创建垂直轴或水平轴。如图 1-15 所示是将“数量”字段添加到“行”功能区时在视图中创建的垂直轴。

如果在“行”和“列”功能区中添加维度字段和度量字段，则会根据两个功能区中的内层字段的类型，在视图中创建柱形图、条形图、折线图等图表。如图 1-16 所示是将“数量”字段添加到“行”功能区，将“地区”字段添加到“列”功能区后创建的柱形图。

地区	订单日期			
	2016	2017	2018	2019
东北	Abc	Abc	Abc	Abc
华北	Abc	Abc	Abc	Abc
华东	Abc	Abc	Abc	Abc
西北	Abc	Abc	Abc	Abc
西南	Abc	Abc	Abc	Abc
中南	Abc	Abc	Abc	Abc

图 1-14 在视图中创建行标题和列标题

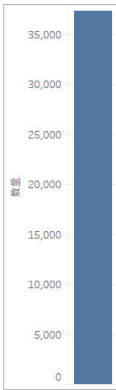


图 1-15 创建的垂直轴

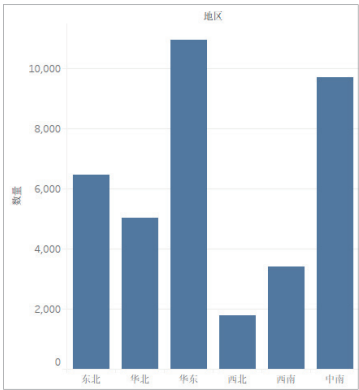


图 1-16 创建的柱形图

1.4.4 “页面”功能区

“页面”功能区默认位于“行”和“列”功能区的左侧，将维度字段添加到“页面”功能区时，将自动在视图的右侧显示页面卡，使用该卡中的控件可以在多个视图之间切换，这些视图是基于该维度字段中的每个成员创建的。

例如，将“地区”字段添加到“页面”功能区，在视图的右侧会显示标题为“地区”的页面卡，如图 1-17 所示。

使用页面卡中的控件可以按照不同的地区查看视图，有以下几种方法：

- 跳转到特定页面：单击页面卡上方的下拉按钮，在打开的列表中选择要查看的地区，如图 1-18 所示。
- 手动翻阅页面：在页面卡中单击地区名称左右两侧的箭头按钮◀▶，或者拖动该名称下方的滑块，将逐一在各个地区之间切换。
- 自动翻阅页面：单击页面卡中“显示历史记录”复选框上方的箭头按钮◀▶，将按照正序或倒序自动在各个地区之间切换。单击两个箭头按钮中间的方块按钮■，将停止自动切换。这 3 个按钮右侧的几个按钮，用于控制自动切换页面的速度。



图 1-17 “页面”功能区及其关联的页面卡

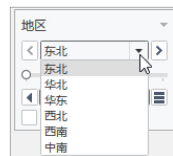


图 1-18 选择要查看的地区

1.4.5 “筛选器”功能区

“筛选器”功能区默认位于“页面”功能区的下方，使用该功能区可以筛选要在视图中包含或排除的数据，以便重点查看感兴趣的内容。对已添加到视图中的字段进行筛选称为“内部筛选”，对未添加到视图中的字段进行筛选称为“外部筛选”。无论是内部筛选还是外部筛选，所有设置了筛选的字段都会显示在“筛选器”功能区中。

假设当前视图如图 1-19 所示，显示了各个地区的数量。如果只想在视图中显示“东北”“华北”和“华东”3 个地区的数量，则可以对“列”功能区中的“地区”字段进行筛选，此时是内部筛选。

如需筛选视图中的“地区”字段，可以右击“列”功能区中的“地区”字段，在弹出的菜单中选择“筛选器”命令，在打开的对话框中取消“西北”“西南”和“中南”3 项的选中状态，如图 1-20 所示。单击“确定”按钮，视图将显示筛选后的数据，在“筛选器”功能区中也会出现一个“地区”字段，如图 1-21 所示。

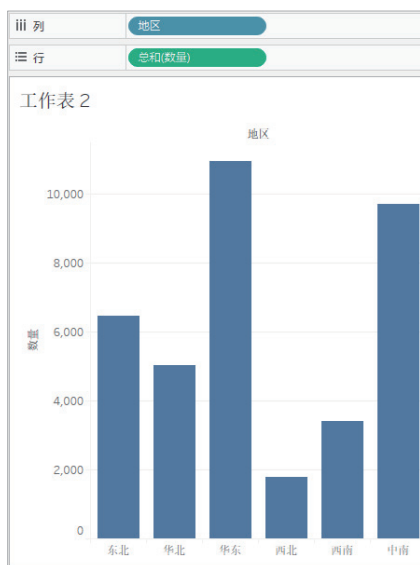


图 1-19 各个地区的数量



图 1-20 取消不想出现在视图中的选项的选中状态

如需执行外部筛选，可以在“数据”窗格中将要筛选的字段拖动到“筛选器”功能区中，然后在打开的对话框中选择要在视图中显示的数据。如图 1-22 所示是对“类别”字段执行的外部筛选，因为没有将该字段添加到“行”或“列”功能区中。

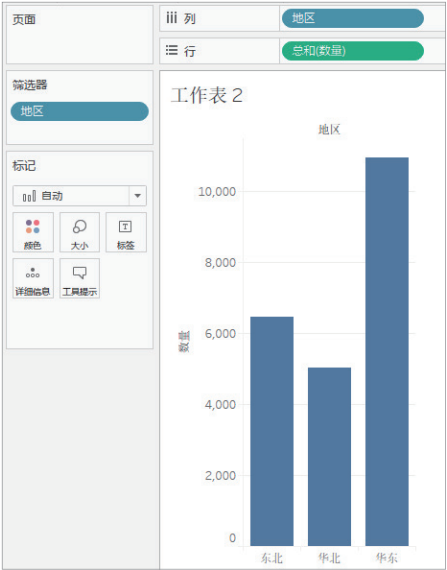


图 1-21 筛选“地区”字段后的视图



图 1-22 对“类别”字段执行外部筛选

1.4.6 “标记”卡

使用“标记”卡可以控制视图中标记的颜色、大小、形状、文本等属性。如需使用“标记”卡，只需将一个或多个字段拖动到“标记”卡中的按钮上，即可在不改变视图整体结构的情况下，通过颜色、大小、形状、文本等不同形式展示数据。

如图 1-23 所示是将“数量”字段拖动到“标记”卡的“文本”按钮上创建的视图，该视图是一个文本表，第一行和第一列是标题，其他行和列是数据，该文本表类似于 Excel 中的数据透视表。

如图 1-24 所示是将“数量”字段的标记类型改为“颜色”后的效果，现在可以通过颜色的深浅度来表示数量的多少。



图 1-23 在视图中使用文本标记

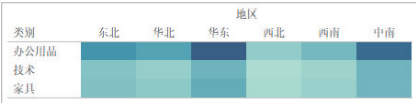


图 1-24 在视图中使用颜色标记

1.4.7 工作表

Tableau Desktop 中的工作表与 Excel 工作表类似，它们都是组成工作簿的基本单元，一个工作簿可以包含多个工作表。它们的不同之处在于，Tableau Desktop 中的工作表不是用于输入数据的，而是利用已导入的数据创建可视化视图，使数据分析更直观。

在 Tableau Desktop 中，一个工作表包含单个视图、功能区、卡、图例、“数据”窗格和“分析”窗格。当工作簿包含多个工作表时，可以单击 Tableau Desktop 窗口下方的工作表标签，在不同的工作表之间切换。如果单击 Tableau Desktop 窗口状态栏右侧的“显示幻灯片”按钮 ，则所有工作表将以缩略图的形式显示，如图 1-25 所示。



图 1-25 以缩略图形式显示工作表

无论工作表是以选项卡形式还是以缩略图形式显示，在工作表的选项卡或缩略图上右击，都会弹出一个菜单，其中包含与工作表操作相关的命令，如图 1-26 所示。在 Tableau Desktop 中对工作表的很多操作与 Excel 工作表类似，例如新建、移动、复制、重命名、删除等。



图 1-26 与工作表相关的操作

在 Tableau Desktop 中删除一个工作表后，可以按 Ctrl+Z 组合键恢复刚删除的工作表，而在 Excel 中只能通过关闭但不保存工作簿，才能恢复已删除的工作表。

1.4.8 重新组织工作区

本章前面介绍工作区中的各个功能区和卡的位置时使用了“默认”一词，这意味着用户可以调整它们的位置，以便按照个人习惯重新组织工作区。

如需移动功能区和卡，可以将鼠标指针移动到功能区或卡的标题区，当鼠标指针变为十字箭头  时，单击并将功能区或卡拖动到目标位置。拖动过程中会显示一条黑线，它指示要放置的目标位置，如图 1-27 所示。

对于不想显示的功能区或卡，可以单击工具栏中的“显示 / 隐藏卡”按钮，然后在弹出的菜单中取消选择不想显示的功能区或卡，如图 1-28 所示。重新显示处于隐藏状态的功能区和卡的操作方法与此类似，只需在该菜单中将它们选中即可。



图 1-27 黑线指示要放置的位置



图 1-28 设置功能区和卡的显示或隐藏状态

如需恢复 Tableau Desktop 中的所有功能区和卡的默认位置，可以单击工具栏中的“显示 / 隐藏卡”按钮，然后在弹出的菜单中选择“重置卡”命令。

第 2 章

连接和管理数据源

开始分析数据前，必须先连接到要使用的数据并为 Tableau Desktop 设置数据源，然后才能在 Tableau Desktop 中使用这些数据。连接数据并设置数据源的工作需要在“数据源”页面中完成。本章将介绍连接和管理数据源，以及编辑数据源中的数据等内容，为数据源构建数据模型的内容将在第 3 章进行介绍。

2.1 适合分析的数据格式

由于数据的来源多种多样，因此，不同来源的数据会以不同的格式存储在表中。然而，并非所有格式的数据都适合在 Tableau 中进行分析。适合分析的数据至少需要符合以下格式要求：

- 以一维表结构存储数据。
- 尽可能细化数据，而非聚合数据。
- 表中的每一列都有列标题。
- 表中没有空行和空单元格。

1. 以一维表结构存储数据

表中的数据不能存储为像数据透视表那样的二维表，而应该像数据库中的表结构一样，即所谓的一维表，它是指每一列只存储同一类数据，表中不能存在包含同类内容的多个列。

如图 2-1 所示是一个二维表的示例，表中的最后 3 列虽然名称不同，但是它们都同属于商品，它们都是商品中的其中一种。为了使表中的数据适合分析，需要将最后 3 列转换为两列，其中一列存储商品的名称，另一列存储商品的销量，合并后的表如图 2-2 所示。

编号	日期	面包	牛奶	果汁
1	6月6日	19	19	26
2	6月7日	50	23	38
3	6月8日	32	16	25

图 2-1 以二维表结构存储的数据

编号	日期	名称	销量
1	6月6日	面包	19
2	6月6日	牛奶	19
3	6月6日	果汁	26
4	6月7日	面包	50
5	6月7日	牛奶	23
6	6月7日	果汁	38
7	6月8日	面包	32
8	6月8日	牛奶	16
9	6月8日	果汁	25

图 2-2 将二维表转换为一维表

2. 尽可能细化数据，而非聚合数据

表中的数据应该尽可能细化，而不是聚合后的数据。如图 2-3 所示的最后一行对每种商品的销量进行求和，求和运算是聚合的一种，聚合还包括求平均值、计数、求最大值、求最小值等。在 Tableau 中分析数据之前，应该删除表中的聚合数据。

另一种情况是表中的数据应尽可能详细。例如，表中包含按天记录的数据要好于按月记录的数据，数据级别越详细，在 Tableau 中就可以进行更多的分析。

3. 表中的每一列都有列标题

应确保表中的每一列都有列标题，以便在将这些数据导入 Tableau Desktop 时能够显示正确的字段名称。

4. 表中没有空行和空单元格

表中的所有数据之间不能有空行，以便使所有数据位于一个连续的范围之内。更重要的是，表中不能有空单元格。如图 2-4（a）所示，即使几种商品具有相同的分类名称，但是在“类别”列中也不能使单元格为空，而必须重复填入相同的类别名称，如图 2-4（b）所示。

编号	日期	面包	牛奶	果汁
1	6月6日	19	19	26
2	6月7日	50	23	38
3	6月8日	32	16	25
总销量		101	58	89

图 2-3 表中不应该包含聚合数据

编号	类别	名称	库存
1	烘焙	吐司面包	36
2		手撕包	27
3		蛋糕	33
4	饮料	矿泉水	32
5		可乐	27
6		雪碧	23

(a)

编号	类别	名称	库存
1	烘焙	吐司面包	36
2	烘焙	手撕包	27
3	烘焙	蛋糕	33
4	饮料	矿泉水	32
5	饮料	可乐	27
6	饮料	雪碧	23

(b)

图 2-4 表中不能有空单元格

2.2 连接数据并创建数据源

在 Tableau Desktop 中分析数据时，首先要连接到外部文件或数据库。成功连接到一个文件或数据库后，会自动创建一个数据源，用户可以在该数据源中继续连接其他的文件或数据库。这意味着在一个数据源中可以包含多个数据连接，每个数据连接都有自己的名称，并指向不同的文件或数据库。除了连接数据时自动创建的数据源之外，用户可以在同一个工作簿中手动创建一个或多个新数据源，并在这些数据源中连接文件或数据库。

2.2.1 连接文件

在 Tableau Desktop 中可以连接多种类型的文件，例如 Excel 文件、文本文件、Access 文件、PDF 文件、JSON 文件等。本小节以连接 Excel 文件为例，介绍在 Tableau Desktop 中连接文件的方法。连接 Excel 文件的操作步骤如下：

Step01 启动 Tableau Desktop，在“开始”页面的“连接”窗格的“到文件”类别下单击 Microsoft Excel，如图 2-5 所示。

Step02 打开“打开”对话框，导航到 Excel 文件所在的文件夹，然后双击要连接的 Excel 文件，如图 2-6 所示。

Step03 稍后，将在“数据源”页面中显示连接到的 Excel 文件的名称和其中包含的工作表。本例连接到的 Excel 文件的名称为“产品销售数据”，其中包含“客户信息”“订单信息”和“订单明细”3 个工作表，如图 2-7 所示。

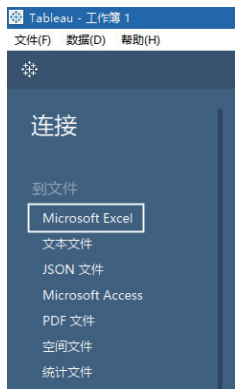


图 2-5 单击 Microsoft Excel

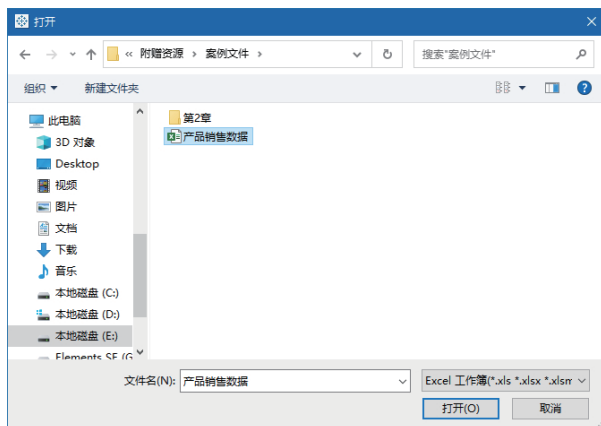


图 2-6 双击要连接的 Excel 文件

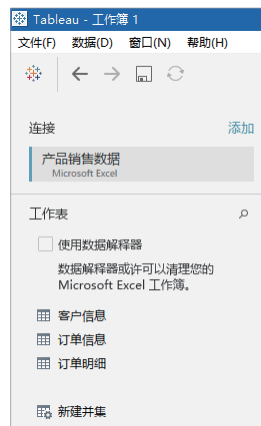


图 2-7 显示已连接的文件名和其中包含的表

2.2.2 连接数据库

在 Tableau Desktop 中可以连接多种类型的数据库，例如 SQL Server、MySQL、Oracle、DB2、MongoDB 等。本小节以连接 SQL Server 数据库为例，介绍在 Tableau Desktop 中连接数据库的方法。连接 SQL Server 数据库的操作步骤如下：



图 2-8 单击 Microsoft SQL Server

Step01 启动 Tableau Desktop，在“开始”页面的“连接”窗格中的“到服务器”类别下单击 Microsoft SQL Server，如图 2-8 所示。

Step02 打开如图 2-9 所示的界面，在“服务器”文本框中输入要连接到的 SQL Server 数据库所在的服务器的名称，在“数据库”文本框中输入数据库中的某个表，如果不知道表名，可以不填写该项。在“输入数据库登录信息”中选择登录服务器的方式，可以选择“使用 Windows 身份验证”方式或“使用特定用户名和密码”两种方式之一。

Step03 设置完成后，单击“登录”按钮，如果通过了身份验证，则会在 Tableau Desktop 的“数据源”页面中显示连接到的服务器名称，以及在 Step02 中指定的数据库和其中包含的表，如图 2-10 所示。

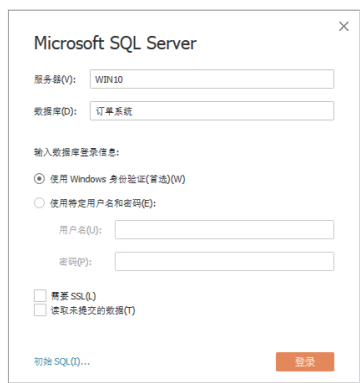


图 2-9 设置数据库服务器的登录信息



图 2-10 连接到 SQL Server 数据库

提示：如果连接到的服务器中有多个数据库，则可以在“数据库”下拉列表中选择一个数据库，下方会显示该数据库中的表。

2.2.3 在一个数据源中建立多个数据连接

在 Tableau Desktop 中成功连接到一个文件或数据库之后，会自动创建一个数据源，数据的连接信息包含在该数据源中。如果使用的数据存储在不同的文件或数据库中，则可以在当前数据源中继续添加新的数据连接，以便连接到所需使用的每一个文件或数据库。在一个数据源中建立多个数据连接的操作步骤如下：

- Step01** 按照 2.2.1 小节或 2.2.2 小节中的方法，先在 Tableau Desktop 中连接到一个文件或数据库。
- Step02** 如需在当前数据源中连接到第二个文件或数据库，可以在“数据源”页面中单击“添加”，如图 2-11 所示。
- Step03** 选择要连接的文件或数据库，然后参照 2.2.1 小节或 2.2.2 小节中的步骤完成数据连接。如图 2-12 所示为连接到 Excel 文件和 SQL Server 数据库，按照建立连接的先后顺序，两个连接的名称依次排列在“数据源”页面中。单击某个连接，下方将显示该连接中包含的表。

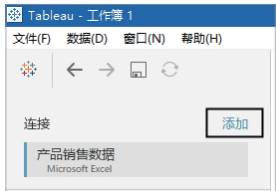


图 2-11 单击“添加”

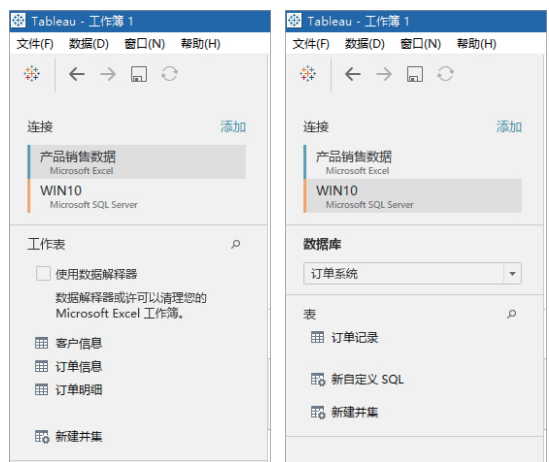


图 2-12 在一个数据源中建立多个数据连接

2.2.4 创建多个数据源

有时可能需要创建多个数据源，然后在每个数据源中连接不同的数据，而不是在一个数据源中建立多个数据连接。启动 Tableau Desktop 并成功连接到一个文件或数据库后，会自动创建一个数据源。此时如需创建一个新的数据源，可以在“数据源”页面中使用以下两种方法：

- 单击菜单栏中的“数据”|“新建数据源”命令，如图 2-13 所示。
- 在画布上方单击数据源图标右侧的下拉按钮，然后在弹出的菜单中选择“新建数据源”命令，如图 2-14 所示。

使用以上两种方法之一，将显示如图 2-15 所示的界面，从中选择要连接的数据类型，并完成数据连接即可。



图 2-13 使用菜单栏中的命令



图 2-14 使用画布中的命令



图 2-15 为新建的数据源选择要连接的数据类型



图 2-16 选择要显示的数据源

在 Tableau Desktop 中创建多个数据源之后，在“数据源”页面中只会显示其中一个数据源中的数据连接和表。可以单击画布上方的下拉按钮，在弹出的菜单中选择要显示的数据源，如图 2-16 所示。

2.2.5 在画布中添加表

完成前面几节操作后，还并未完成数据源的设置工作，此时 Tableau Desktop 还不知道用户要使用哪些数据开展分析工作。只有将已建立数据连接的一个或多个表添加到“数据源”页面的画布中，才真正为 Tableau 准备好要使用的数据。

如需在画布中添加表，只要在“数据源”页面的左侧窗格中将一个或多个表拖动到右侧上方的画布中即可，如图 2-17 所示。



图 2-17 在画布中添加表

2.3 管理数据源

成功连接数据后，可以对数据源执行一些管理操作，例如更改数据源的位置和名称、刷新数据源、复制和导出数据源等，从而使数据源可以更好地工作。

2.3.1 更改数据源的位置和名称

如果在 Tableau Desktop 中建立的数据连接所指向的数据源的位置或名称发生改变，则 Tableau Desktop 中的数据连接会自动失效，此时将中断 Tableau Desktop 与数据源之间的连接，在 Tableau Desktop 中不会显示数据源中的数据。

如需解决此问题，可以在“数据源”页面的左侧窗格中右击数据连接，然后在弹出的菜单中选择“编辑连接”命令，如图 2-18 所示。在打开的对话框中找到并双击更改位置或名称后的数据源，即可使数据源中的表重新显示在“数据源”页面中。

在 Tableau Desktop 中分析数据时，使用的数据源名称显示在“数据源”页面右侧的画布上方，单击即可使其进入编辑状态，如图 2-19 所示，输入新名称后单击页面中的空白处，完成名称的修改。



图 2-18 选择“编辑连接”命令



图 2-19 修改数据源的名称

2.3.2 更改数据连接的名称

连接到一个文件或数据库后，默认使用该文件或数据库的名称作为在 Tableau Desktop 中显示的数据连接的名称。有时可能需要修改数据连接的名称，使其含义更清晰，尤其在连接到多个文件或数据库时可能更需要执行此操作。

更改数据连接的名称有以下两种方法：

- 在“数据源”页面的左侧窗格中右击一个数据连接，然后在弹出的菜单中选择“重命名”命令，如图 2-20 所示。
- 双击一个数据连接。

无论使用哪种方法，数据连接的名称都将进入编辑状态，输入所需的名称，然后单击页面中的其他位置即可。

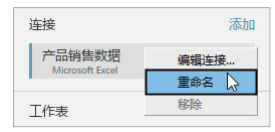


图 2-20 选择“重命名”命令

2.3.3 刷新数据源

如果在 Tableau Desktop 中连接到的数据源本身的数据发生改变，为了将所有更改及时反映到 Tableau Desktop 中，可以在“数据源”页面的工具栏中单击“刷新数据源”按钮，如图 2-21 所示。

提示：如果无法单击“刷新数据源”按钮，则可能是因为还没有在画布中添加表。

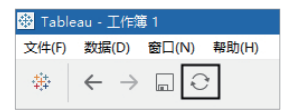


图 2-21 单击“刷新数据源”按钮

2.3.4 复制数据源

如需对数据源进行一些修改，但是又不想影响使用该数据源的现有工作表和视图，那么可以通

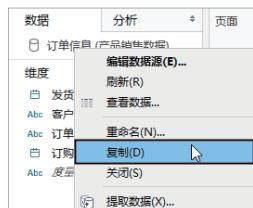


图 2-22 选择“复制”命令

过复制数据源获得其副本，然后修改数据源的副本。复制数据源有以下两种方法：

- 在“数据源”页面中单击菜单栏中的“数据”|“复制数据源”命令。
- 在“工作区”页面中右击“数据”窗格顶部的数据源，然后在弹出的菜单中选择“复制”命令，如图 2-22 所示。

2.3.5 导出数据源

如果在工作簿中创建好的数据源可能会在以后重复使用，则可以将该数据源以文件的形式导出，导出的数据源有以下两种文件格式：

- 数据源（.tds）：包含连接到文件或数据库所需的信息。如果要使用数据源的其他用户具有数据连接所指向的文件或数据库的访问权限，则可以使用该文件格式。
- 打包数据源（.tdsx）：包含数据源（.tds）文件中的所有信息，以及数据连接所指向的数据源中的数据。如果要使用数据源的其他无权访问数据源，则可以使用该文件格式。

默认将数据源导出到以下路径中，假设 Windows 操作系统安装在 C 盘。

C:\Users\<用户名>\Documents\我的 Tableau 存储库\数据源

如需导出数据源，需要单击 Tableau Desktop 窗口底部的“工作表 1”（或其他工作表），切换到“工作区”页面，右击“数据”窗格顶部的数据源名称，然后在弹出的菜单中选择“添加到已保存的数据源”命令，如图 2-23 所示。

打开如图 2-24 所示的对话框，将自动定位到上面给出的数据源的默认存储位置，设置好文件名和文件类型，单击“保存”按钮，即可将数据源导出到默认位置。



图 2-23 选择“添加到已保存的数据源”命令

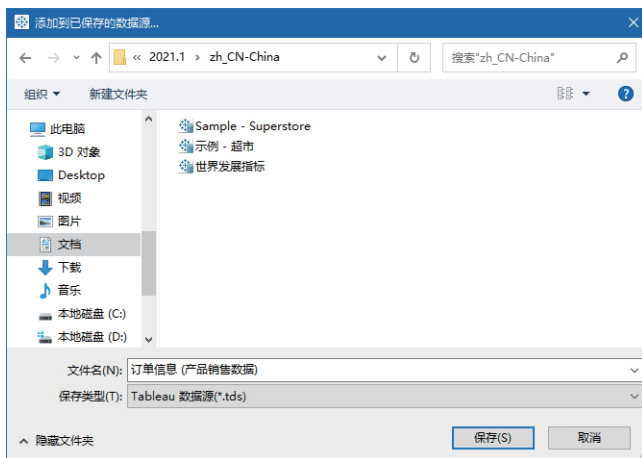


图 2-24 设置数据源的文件名和文件类型

将数据源导出到默认路径后，在 Tableau Desktop 的“开始”页面的“连接”窗格底部会显示该数据源，如图 2-25 所示。选择该数据源，即可在 Tableau Desktop 中自动连接到数据源所指向的文件或数据库。

用户可以更改“我的 Tableau 存储库”文件夹的默认位置，只需单击菜单栏中的“文件”|“存储库位置”命令，在打开的对话框中选择一个文件夹，然后单击“选择文件夹”按钮，如图 2-26 所示。



图 2-25 导出到默认位置的数据源显示在“开始”页面中

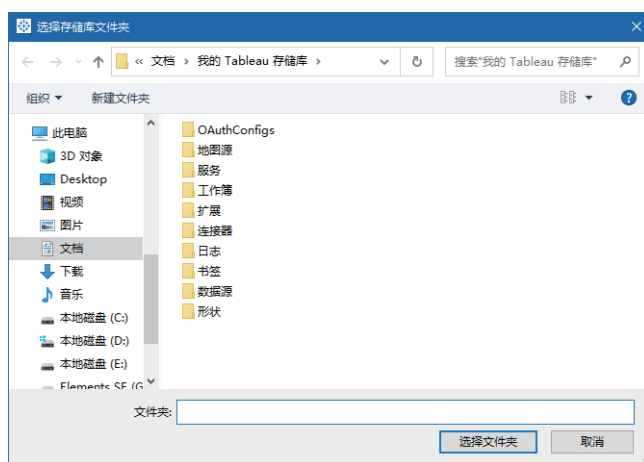


图 2-26 设置存储库的默认位置

2.3.6 关闭数据源

如果不再使用已连接到 Tableau Desktop 中的数据源，则可以将其关闭，即可断开 Tableau Desktop 与数据源之间的连接。关闭数据源时，将自动删除使用该数据源的所有工作表。关闭数据源有以下两种方法：

- 在“数据源”页面中单击菜单栏中的“数据”|“关闭数据源”命令。
- 在“工作区”页面中右击“数据”窗格顶部的数据源，然后在弹出的菜单中选择“关闭”命令。

提示：如果意外关闭了数据源，则可以单击工具栏中的“撤销”按钮或按 Ctrl+Z 组合键，使 Tableau Desktop 重新连接到刚刚关闭的数据源。

2.4 编辑数据源中的数据

成功连接数据后，可以在 Tableau Desktop 中对数据源进行一些调整，例如更改字段名称、隐藏和拆分字段、排序和筛选数据、转置数据等。这些操作不会修改数据连接所指向的数据源本身，而只作用于显示在 Tableau Desktop 中的数据源。本节介绍的内容都是在“数据源”页面中操作的，其中的很多操作也可以在“工作区”页面的“数据”窗格中完成。



图 2-27 选择“重命名”命令

提示：更改字段名称后，以后如需恢复字段的默认名称，可以右击字段，在弹出的菜单中选择“重置名称”命令。

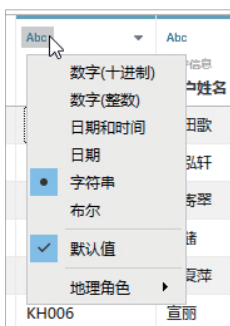


图 2-28 更改字段的数据类型

2.4.2 更改字段的数据类型

成功连接数据后，Tableau 会自动检测和设置表中每个字段的数据类型，有时检测结果可能有误，此时可以手动更改字段的数据类型。如需更改字段的数据类型，可以在“数据源”页面的数据网格中单击表示数据类型的图标，然后在弹出的菜单中选择所需的数据类型，如图 2-28 所示。

提示：如果在菜单中的“默认值”选项的开头显示勾选标记，则说明该字段的数据类型没有被更改过。

2.4.3 隐藏字段

如果不想显示某些字段，则可以将其隐藏，但是字段中的数据仍然存在。如需隐藏字段，可以在数据网格中右击要隐藏的字段，然后在弹出的菜单中选择“隐藏”命令。

如需使隐藏的字段重新显示出来，可以在数据网格中勾选“显示隐藏字段”复选框，如图 2-29 所示。处于隐藏状态的字段的字体颜色较浅，如需使其恢复为正常显示状态，可以右击隐藏状态的字段，在弹出的菜单中选择“取消隐藏”命令，如图 2-30 所示。

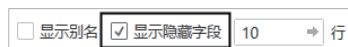


图 2-29 勾选“显示隐藏字段”复选框

2.4.1 更改字段名称

将表添加到“数据源”页面中的画布后，会在下方的数据网格中显示当前选中的表中的数据，字段自动加粗显示。如需更改某个字段的名称，可以使用以下两种方法：

- 双击字段名称。
- 右击字段名称所在的区域或单击字段右上角的下拉按钮，在弹出的菜单中选择“重命名”命令，如图 2-27 所示。

无论使用哪种方法，字段名称都会进入编辑状态，输入所需的名称，然后单击字段名称之外的其他位置。



图 2-30 选择“取消隐藏”命令

2.4.4 拆分字段

如果字段中的每项数据由多个值组成，为了便于分析该字段中的数据，可以使用 Tableau Desktop 中的拆分功能将多个值拆分为单个值，拆分后的这些值分别存储在新建的多个字段中。

如图 2-31 所示，“客户信息”字段中的每项数据由姓名、性别、等级 3 部分组成，各部分之间以“-”分隔。

为了在数据分析时分别处理客户的姓名、性别和等级，需要将“客户信息”字段中的数据按照姓名、性别、等级拆分为 3 个字段，每个字段只存储一类信息，拆分字段的操作步骤如下：

Step01 在数据网格中右击“客户信息”字段名称所在的区域，然后在弹出的菜单中选择“自定义拆分”命令，如图 2-32 所示。

Step02 打开“自定义拆分”对话框，在“使用分隔符”文本框中输入“-”，然后在“拆分”下拉列表中选择“全部”，如图 2-33 所示。

客户信息 (2)	客户信息
客户编号	客户信息
KH001	肖田歌-男-特级
KH002	赵泓轩-男-特级
KH003	米寄琴-女-一级
KH004	姚储-男-一级
KH005	黄夏萍-女-一级
KH006	宣丽-女-一级
KH007	周芹-女-二级
KH008	余严-男-二级

图 2-31 “客户信息”字段中的每项数据包含多个值

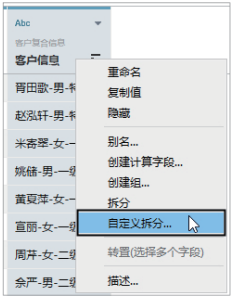


图 2-32 选择“自定义拆分”命令

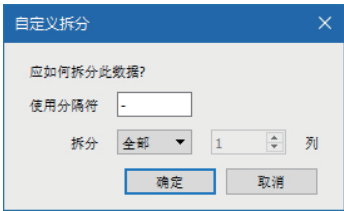


图 2-33 设置拆分选项

Step03 单击“确定”按钮，将“客户信息”字段拆分为 3 个字段，如图 2-34 所示。

客户信息 (2)	客户信息	计算	计算	计算
客户编号	客户信息	客户信息 - 拆分 1	客户信息 - 拆分 2	客户信息 - 拆分 3
KH001	肖田歌-男-特级	肖田歌	男	特级
KH002	赵泓轩-男-特级	赵泓轩	男	特级
KH003	米寄琴-女-一级	米寄琴	女	一级
KH004	姚储-男-一级	姚储	男	一级
KH005	黄夏萍-女-一级	黄夏萍	女	一级
KH006	宣丽-女-一级	宣丽	女	一级
KH007	周芹-女-二级	周芹	女	二级
KH008	余严-男-二级	余严	男	二级

图 2-34 拆分字段

Step04 为拆分后得到的 3 个字段设置合适的名称，然后将原来的“客户信息”字段隐藏即可。

2.4.5 转置数据

如图 2-35 所示为 2.1 节介绍过的二维表，如需在 Tableau Desktop 中对该表中的数据分析，需要先将其转换为一维表。

使用 Tableau Desktop 中的转置功能可以将二维表转换为一维表，操作步骤如下：

编号	日期	面包	牛奶	果汁
1	6月6日	19	19	26
2	6月7日	50	23	38
3	6月8日	32	16	25

图 2-35 二维表

Step01 在数据网格中选择要转置的多个列，本例为表中的最后 3 列。

Step02 右击 Step01 中选择的任意一列的字段名称所在的区域，在弹出的菜单中选择“转置”命令，如图 2-36 所示，将创建两个新列，在其中一列中存储上述 3 列的字段名称，另一列存储这 3 列中的值，如图 2-37 所示。最后为转换后的各列设置合适的名称。

图 2-36 选择“转置”命令

图 2-37 转置为一维表

2.4.6 排序数据

可以对 Tableau Desktop 中的数据源按照行或列排序。如需按行排序，可以在数据网格中单击字段标题右侧的 按钮，有以下 3 种效果：

- 第一次单击时，升序排列数据。
- 第二次单击时，降序排列数据。
- 第三次单击时，恢复数据的原始顺序。

如图 2-38 所示是对“客户编号”字段中的数据升序排列后的效果，客户编号从小到大依次排列。

图 2-38 升序排列数据

图 2-39 选择排序方式

按列排序是指调整表中各列的前后次序。如需按列排序，可以在数据网格中单击“排序字段”右侧的下拉按钮，然后在打开的下拉列表中选择排序方式，如图 2-39 所示。

2.4.7 筛选数据

如果只想对数据源中的部分数据进行分析，则可以在开始分析之前，先从数据源中筛选出所需使用的数据，从而减少数据源中的数据量。筛选数据源的操作步骤如下：

Step01 在“数据源”页面中单击画布右上方的“添加”，如图 2-40 所示。

Step02 打开“编辑数据源筛选器”对话框，单击“添加”按钮，如图 2-41 所示。



图 2-40 单击“添加”

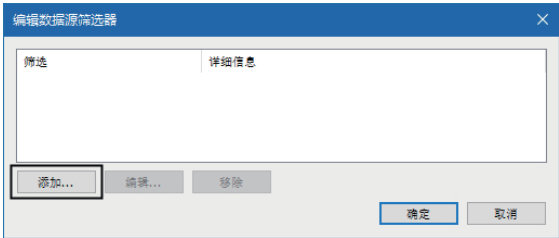


图 2-41 单击“添加”按钮

Step03 打开“添加筛选器”对话框，选择要筛选的字段，然后单击“确定”按钮，如图 2-42 所示。

Step04 打开“筛选器”对话框，为 Step03 选择的字段设置筛选条件，此处为“性别”字段设置的筛选条件是只筛选女性客户，如图 2-43 所示。

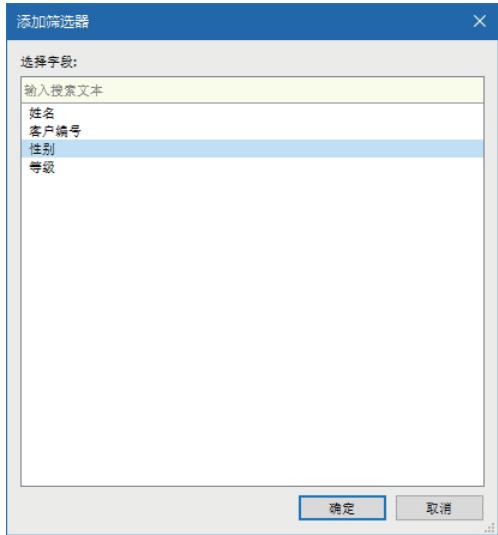


图 2-42 选择要筛选的字段



图 2-43 设置筛选条件

Step05 单击“确定”按钮，返回“编辑数据源筛选器”对话框，为“性别”字段设置的筛选条件显示在该对话框中，如图 2-44 所示。

Step06 单击“确定”按钮，将在数据网格中只显示女性客户的数据，如图 2-45 所示。

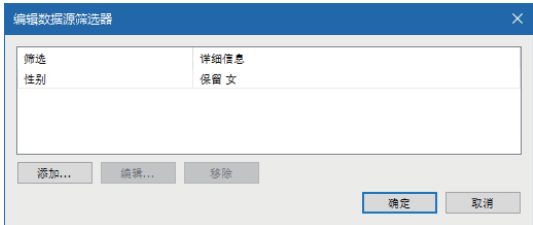


图 2-44 显示已设置的筛选条件

客户编号	姓名	性别	等级
KH003	米嘉翠	女	一级
KH005	黄夏萍	女	一级
KH006	宣丽	女	一级
KH007	周芹	女	二级
KH009	楚诗夏	女	二级

图 2-45 筛选后的数据

上面介绍的是筛选文本类型字段的一般操作方法。筛选数值和日期两种类型的字段时具有更

多、更灵活的方式。如图 2-46 所示为筛选日期类型字段时打开的对话框，其中包含“相对日期”和“日期范围”两种筛选方式。“日期范围”操作方法与前面介绍的筛选文本类型数据的方法类似，也是通过勾选或取消对复选框的勾选来筛选字段中的数据。

如果选择“相对日期”，然后单击“下一步”按钮，则将打开如图 2-47 所示的对话框，在上方选择一种日期设置方式，然后在下方对该方式进行具体设置。

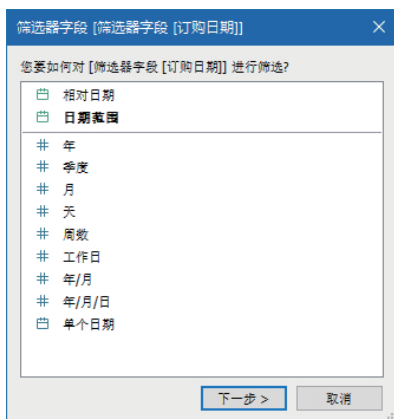


图 2-46 筛选日期类型的字段



图 2-47 设置相对日期

无论筛选哪种类型的数据，如需删除已创建的筛选器，使数据恢复到完全显示的状态，可以单击画布右上方的“编辑”，如图 2-48 所示，在打开的对话框中选择要删除的筛选器，然后单击“移除”按钮。



图 2-48 单击“编辑”

提示：“编辑”左侧的数字表示当前已创建的筛选器的数量。

2.4.8 查看和管理元数据

元数据定义了一个表中包含的字段的名称和数据类型。将一个表添加到画布后，可以单击“管理元数据”按钮，查看该表的元数据，如图 2-49 所示。第 1 列显示字段的名称和数据类型，第 2 列显示该字段所属的表的名称，第 3 列显示字段在源数据中的原始名称。在 Tableau Desktop 中可以修改第一列中的字段名称和数据类型，操作方法请参考 2.4.1 小节和 2.4.2 小节。

字段名称	表	远程字段名称
Abc 订单编号	订单信息	订单编号
订购日期	订单信息	订购日期
发货日期	订单信息	发货日期
Abc 客户编号	订单信息	客户编号

图 2-49 查看表的元数据

第 3 章

构建数据模型

第 1 章简要介绍了数据模型的概念，在 Tableau Desktop 中进行分析的数据可能来自于一个或多个表，如需使这些数据成为一个内在关联的整体，需要将这些数据合并起来，即构建数据模型。在 Tableau Desktop 中通过为两个或多个表创建关联、联接或并集，可以将不同表中的数据合并到一起，本章将介绍合并数据的几种方式的概念和实现方法。

3.1 Tableau 数据模型简介

在 Tableau Desktop 中创建的每个数据源都有一个数据模型，每个数据模型由添加到画布中的一个或多个表组成。可以将数据模型看作是一个由多个表以及它们之间的关系组成的关系图，通过此关系图，Tableau 可以获悉如何查询这些表中的数据。

在 Tableau Desktop 2020.2 之前的版本中，数据模型只有一个物理层，将表添加到画布时，实际上是将表添加到了数据模型中的物理层，物理层中的表称为物理表。在物理层中可以通过联接或并集来合并表中的数据。如果将两个表添加到画布中，则默认使用联接的方式合并两个表中的数据，联接会根据两个表中的相关字段，将两个表中的数据合并到单个表中，在单个表中横向整合两个表中的字段。

从 Tableau Desktop 2020.2 版本开始，在数据模型中除了包含物理层，还新增了一个逻辑层。在 Tableau Desktop 2020.2 及更高版本中将表添加到画布时，默认是将表添加到了数据模型中的逻辑层。逻辑层中的表称为逻辑表，每个逻辑表在物理层中都有一个与其对应的物理表，每个逻辑表可以包含一个或多个物理表。如果将两个表添加到画布中，则默认认为两个表创建关系，以此来合并两个表中的数据，此时并未真正将两个表中的数据合并到单个表中，而是通过两个表中的匹配字段为表中数据建立关联，匹配字段就是关系数据库软件中的主键和外键。

在 Tableau 中构建数据模型需要注意以下几个问题：

- 添加到画布中的第一个表将成为数据模型中的根表，然后可以按任意顺序添加其他表。
- 如需创建星形模型，需要先将事实表添加到画布中，然后添加维度表。
- 如需创建雪花模型，需要将维度表连接到另一个维度表。
- 关联表时设置的关系必须至少由一对匹配的字段组成，这些字段必须具有相同的数据类型。

3.2 关联表中的数据

关联是根据两个表中的匹配字段为两个表创建关系。关联不会真正将两个表中的数据合并到单个表中，而是通过设置的关系告知 Tableau 两个表中的数据是如何相关的，具体来说，就是如何在两个表之间连接数据行。只有在创建视图时，Tableau 才会分析两个表之间的关系，并根据关系中指定的匹配字段在相关表中查找适合的数据。在 Tableau Desktop 中关联表的行为类似于在关系数据库软件（例如 Access）中为表设置关系。

3.2.1 关系类型

关系是指两个表之间存在的某种内在联系。两个表之间的关系有 3 种类型：一对一、一对多（或多对一）、多对多，关系的类型决定两个表中的数据关联方式。

1. 一对一

“一对一”关系是指第一个表中的每条记录在第二个表中只有一个匹配的记录，而第二个表中的每条记录在第一个表中也只有一个匹配的记录。

一对一关系不太常见，因为在大多数情况下，应该将具有一对一关系的两个表中的数据合并到一个表中。在某些特定情况下，可能需要为两个表创建一对一关系。例如，在一个包含客户的个人信息、账号和密码的表中，为了提高账号的安全性，可以将账号和密码单独存储在一个表中，该表与用户个人信息所在的表就是“一对一”关系。

2. 一对多

“一对多”关系是指第一个表中的每条记录在第二个表中有一条或多条匹配的记录，而第二个表中的每条记录在第一个表中只有一条匹配的记录。多对一与一对多是同一种关系类型，只是由于两个表的位置不同而导致相反的关系方向。

客户和订单之间是一对多关系。每个客户可以提交一个或多个订单，但是每个订单只属于一个客户。如图 3-1 所示，“客户信息”表中的“客户编号”字段能够唯一标识每一个客户，在“订单信息”表中也包含该字段，但是客户编号可能会出现多次，这样就可以通过“客户信息”表中的“客户编号”字段，在“订单信息”表中找到每个客户提交了哪些订单，或者通过“订单信息”表中的“客户信息”字段在“客户信息”表中找到对应的客户信息。

客户编号	姓名	性别	等级	订单编号	订购日期	发货日期	客户编号
KH001	肖田歌	男	特级	DD001	2022/7/6	2022/7/8	KH002
KH002	赵泓轩	男	特级	DD002	2022/7/6	2022/7/8	KH007
KH003	米寄翠	女	一级	DD003	2022/7/6	2022/7/7	KH001
KH004	姚储	男	一级	DD004	2022/7/6	2022/7/7	KH007
KH005	黄夏萍	女	一级	DD005	2022/7/6	2022/7/7	KH004
KH006	宣丽	女	一级	DD006	2022/7/6	2022/7/7	KH001
KH007	周芹	女	二级	DD007	2022/7/7	2022/7/8	KH005
KH008	余严	男	二级	DD008	2022/7/7	2022/7/8	KH003
KH009	楚诗夏	女	二级	DD009	2022/7/7	2022/7/10	KH003
KH010	张恭	男	二级	DD010	2022/7/7	2022/7/10	KH005
				DD011	2022/7/7	2022/7/10	KH007
				DD012	2022/7/7	2022/7/9	KH003
				DD013	2022/7/7	2022/7/10	KH010
				DD014	2022/7/7	2022/7/10	KH009
				DD015	2022/7/7	2022/7/8	KH002
				DD016	2022/7/8	2022/7/11	KH008
				DD017	2022/7/8	2022/7/11	KH006
				DD018	2022/7/8	2022/7/9	KH007
				DD019	2022/7/8	2022/7/10	KH005
				DD020	2022/7/8	2022/7/11	KH010

图 3-1 一对多关系

例如，在“客户信息”表中客户编号为 KH001 的客户，在“订单信息”表中该客户编号出现在订单编号为 DD003 和 DD006 的两个订单中，这意味着这两个订单是客户编号为 KH001 的客户提交的。

3. 多对多

“多对多”关系是指第一个表中的每条记录在第二个表中有一条或多条匹配的记录，而第二个表中的每条记录在第一个表中也有一个或多个匹配的记录。

商品和订单之间是多对多关系，同一种商品可以出现在多个订单中，而一个订单也可以包含多种商品。

3.2.2 在逻辑层中关联表

如需为两个或多个表创建关联，需要使用 Tableau Desktop 2020.2 及更高版本。关联在一起的表可以是同一个数据连接或不同数据连接中的表，无论哪种情况，这些表都必须位于同一个数据源中。

在 Tableau Desktop 中建立数据连接后，如需关联两个表中的数据，可以在“数据源”页面中将一个表拖动到画布中，然后将另一个表拖动到画布中，Tableau 会自动检测两个表中的匹配字段。如果发现匹配字段，将为两个表创建关系，并在两个表之间显示一条关系线，如图 3-2 所示。



图 3-2 自动为两个表创建关系

提示：将第二个表添加到画布后，Tableau 为两个表创建关系的同时会自动打开“编辑关系”对话框，单击对话框右上角的✕按钮可以关闭该对话框。

如果 Tableau 检测不到两个表中的匹配字段，则会在两个表之间显示警告图标，将鼠标指针移动到该图标上，会显示如图 3-3 所示的信息，此时需要手动设置用于创建关系的匹配字段。

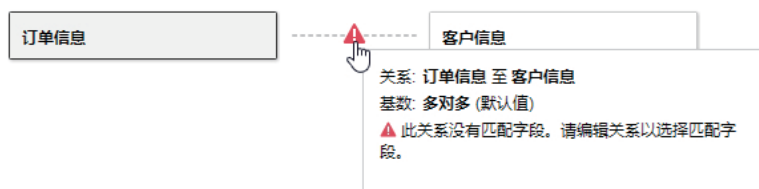


图 3-3 检测不到匹配字段时显示警告图标

提示：即使没有显示警告图标，也可以将鼠标指针移动到关系线上，此时会显示关联的基本信息。

3.2.3 更改或添加匹配字段

无论是需要更改用于确定表关系的匹配字段，还是由于未检测到匹配字段而需要手动对其进行设置，都可以单击两个表之间的关系线或警告图标，打开“编辑关系”对话框，左、右两侧显示当前关系涉及的两个表的名称和已指定的匹配字段，如图 3-4 所示。由于本例中的匹配字段在两个表中的名称相同，为了便于区分，在其中一个字段名称的右侧显示了表名。



图 3-4 “编辑关系”对话框

如需删除当前设置好的匹配字段，可以将鼠标指针移动到字段所在的行，此时会在该行的右侧显示垃圾桶图标，单击该图标将删除当前的匹配字段，如图 3-5 所示。



图 3-5 单击垃圾桶图标删除当前的匹配字段

如需更改当前设置好的匹配字段，可以单击左、右两个字段中的任意一个字段，此时会在字段的下方显示表中的所有字段，分别在左、右两侧选择要在关系中使用的匹配字段，即可替换当前的匹配字段，如图 3-6 所示。



图 3-6 更改关系中的匹配字段

提示：如果表中包含的字段数量较多，则可以通过在搜索框中输入字段的名称快速找到所需的字段。

如需为关系添加多组匹配字段，可以在“编辑关系”对话框中单击“添加更多字段”，然后分别在左、右两侧选择新的一组匹配字段，如图 3-7 所示。



图 3-7 添加新的匹配字段

注意：如果当前正在“编辑关系”对话框中显示表中的所有字段，则需要先单击对话框右下角的“关闭”，然后在“编辑关系”对话框才会显示“添加更多字段”。

3.2.4 优化关系性能

通过设置关系的基数和引用完整性，可以优化 Tableau 在关联的表中查询数据的性能。在为表创建关联时，Tableau 会自动设置关系的基数和引用完整性，所以即使不额外设置这两项，关联的表中的数据也能正常工作。如果了解两个表中的记录之间的唯一性和匹配方式，则可以更改这两项设置，从而更加准确地描述数据，使 Tableau 发挥最佳性能。

如需设置基数和引用完整性，可以在“编辑关系”对话框中单击“性能选项”，然后在展开的对话框中进行设置，如图 3-8 所示。



图 3-8 设置基数和引用完整性

“基数”选项用于设置关系的类型，即一对一、一对多（多对一）和多对多。由于“订单信息”和“客户信息”两个表是多对一的关系，所以需要将图 3-8 中右侧的“基数”选项设置为“一个”，如图 3-9 所示。



图 3-9 更改“基数”选项

“引用完整性”选项用于设置一个表中的每条记录在其他表中是否存在匹配项。如果一个表中的每条记录始终与另一个表中的一条或多条记录匹配，则应该将“引用完整性”选项设置为“所有记录匹配”，否则设置为“某些记录匹配”。

例如，“订单信息”表中的任意一条记录肯定与“客户信息”表中的某条记录匹配，因为每个订单都必须由一个客户提交，所以该订单必然在“客户信息”表中有一个对应的客户。反之则未必如此，“客户信息”表中的任意一个客户未必提交过订单，这样在“订单信息”表中也就没有对应的订单记录了。

提示：如果在更改“基数”和“引用完整性”两个选项之后，想要使它们恢复到 Tableau 最初的默认设置，则可以单击“编辑关系”对话框中的“恢复为默认值”按钮。“基数”选项的默认值为“多对多”，“引用完整性”选项的默认值为“某些记录匹配”。

3.2.5 移动或删除关系中的表

画布中的每个表都至少与另一个表具有关系。如需为一个表创建与另一个表的关系，可以使用

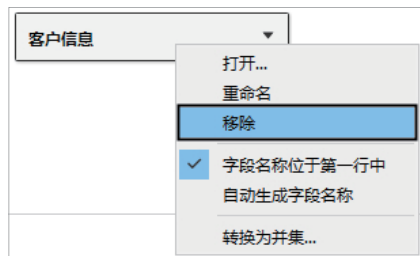


图 3-10 选择“移除”命令

鼠标将该表拖动到另一个表的附近，当在两个表之间显示关系线时，释放鼠标左键，将在两个表之间创建关系，并自动删除该表原来具有的关系。

如需删除关系中的表，可以在画布中右击该表，然后在弹出的菜单中选择“移除”命令，如图 3-10 所示。

注意：删除画布中的表时，将自动删除与该表关联的下级表。如果删除根表，则会自动删除数据模型中的所有其他表。

3.3 联接表中的数据

与在逻辑层中创建关联不同，联接是在物理层中将两个表中的数据合并到单个表中，而非像关联那样可以保持两个表中的数据各自独立。联接不如关联灵活和智能，并且还存在着一些缺陷，所以在大多数情况下，关联是在 Tableau 中合并数据的首选方案。

3.3.1 联接类型

联接的类型决定两个表中数据的合并方式。在 Tableau 中有 4 种联接类型：内联接、左联接、右联接和完全外部联接。

1. 内联接

使用内联接合并两个表时，在合并后的单个表中只包含与两个表都匹配的值，如图 3-11 所示。

名称	产地		名称	数量		名称	产地	数量
A	北京	+	A	1	=	A	北京	1
B	天津		B	2		B	天津	2
C	上海		D	3				

图 3-11 内联接

2. 左联接

使用左联接合并两个表时，在合并后的单个表中将包含左侧表中的所有值，以及右侧表中的匹配值。如果左侧表中的值在右侧表中没有匹配值，则在合并后的单个表中的相应单元格中显示 null，如图 3-12 所示。

名称	产地		名称	数量		名称	产地	数量
A	北京	+	A	1	=	A	北京	1
B	天津		B	2		B	天津	2
C	上海		D	3		C	上海	null

图 3-12 左联接

3. 右联接

使用右联接合并两个表时，在合并后的单个表中将包含右侧表中的所有值，以及左侧表中的匹配值。如果右侧表中的值在左侧表中没有匹配值，则在合并后的表中的相应单元格中显示 null，如图 3-13 所示。

名称	产地		名称	数量		名称	数量	产地
A	北京	+	A	1	=	A	1	北京
B	天津		B	2		B	2	天津
C	上海		D	3		D	3	null

图 3-13 右联接

4. 完全外部联接

使用完全外部联接合并两个表时，在合并后的单个表中将包含两个表中的所有值，两个表中互不匹配的值将在单元格中显示 null，如图 3-14 所示。

名称	产地		名称	数量		名称	产地	数量
A	北京	+	A	1	=	A	北京	1
B	天津		B	2		B	天津	2
C	上海		D	3		C	上海	null
						D	null	3

图 3-14 完全外部联接

提示：如果不确定使用哪种联接类型来合并多个表中的数据，则可以选择为表创建关联，这是因为关联表时 Tableau 会自动设置联接类型，无须用户额外设置。

3.3.2 在物理层中联接表

在 Tableau Desktop 2020.2 之前的版本中，将表添加到画布时，表默认被添加到物理层。物理层中的表默认使用联接方式合并数据，这意味着如果将两个表添加到画布中，Tableau 会自动为它们创建联接，并在两个表之间使用维恩图表示联接类型，如图 3-15 所示。



图 3-15 联接两个表

为两个表创建联接后，在数据网格中将显示两个表合并为单个表之后的数据，如图 3-16 所示。

客户信息 客户编号 (客户信息)	客户信息 姓名	客户信息 性别	客户信息 等级	订单信息 订单编号	订单信息 订购日期	订单信息 发货日期	客户信息 客户编号
KH002	赵弘轩	男	特级	DD001	2022/7/6	2022/7/8	KH002
KH007	周芹	女	二级	DD002	2022/7/6	2022/7/8	KH007
KH001	胥田歌	男	特级	DD003	2022/7/6	2022/7/8	KH001
KH007	周芹	女	二级	DD004	2022/7/6	2022/7/8	KH007
KH004	姚储	男	一级	DD005	2022/7/6	2022/7/9	KH004
KH001	胥田歌	男	特级	DD006	2022/7/6	2022/7/7	KH001
KH005	黄夏萍	女	一级	DD007	2022/7/7	2022/7/8	KH005
KH003	米春翠	女	一级	DD008	2022/7/7	2022/7/8	KH003

图 3-16 联接将两个表中的数据合并到单个表中

如需在 Tableau Desktop 2020.2 及更高版本中创建联接，可以先将一个表添加到画布中，此时该表默认位于逻辑层。然后在画布中双击该表进入其物理层，其中包含该表在物理层中对应的物理表。此时将另一个表添加到当前的物理层中，即可为两个表创建联接，如图 3-17 所示。

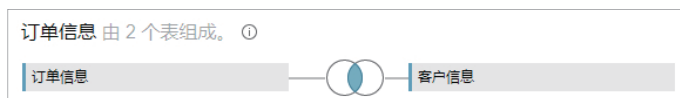


图 3-17 在逻辑表的物理层中

提示：单击物理层右上角的  按钮，将关闭物理层，并返回逻辑层。

3.3.3 更改联接类型

创建联接时，Tableau 会根据两个表中的匹配字段自动设置联接类型，用户可以在创建联接后更改表之间的联接类型。

如果使用的是 Tableau Desktop 2020.2 之前的版本，则可以在“数据源”页面的画布中单击两个表之间的维恩图，然后在打开的“联接”对话框中选择联接类型，如图 3-18 所示。

如果使用的是 Tableau Desktop 2020.2 及更高版本，则可以先在画布中双击包含联接在一起的物理表的逻辑表，进入物理层后单击维恩图，然后选择联接类型，如图 3-19 所示。

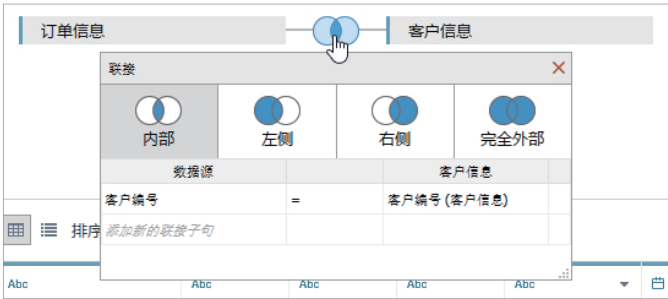


图 3-18 更改联接类型 1

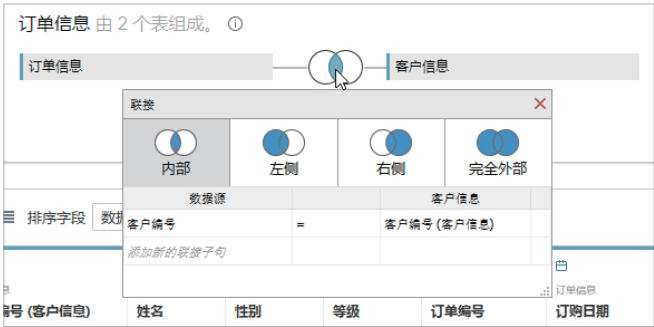


图 3-19 更改联接类型 2

与创建关联时可以更改或添加匹配字段类似，创建联接时也可以设置匹配字段，联接中的匹配字段称为联接子句。在“联接”对话框中可以更改联接的联接子句，也可以添加多组联接子句，如图 3-20 所示。

注意：创建联接时，用于为两个表建立联接的匹配字段必须具有相同的数据类型。如果在创建联接后更改联接字段的数据类型，则会立刻断开两个表的联接。



图 3-20 设置联接子句

3.4 合并表中的记录

通过为表创建并集，可以将多个表中的记录合并到一个表中，它是一种纵向的合并，这种合并方式类似于关系数据库软件（例如 Access）中的追加记录操作。为了获得最好的合并结果，要合并的各个表应该具有相同的字段数，相关字段的名称和数据类型也必须相同。

3.4.1 通过并集合并表中的记录

如图 3-21 所示是要合并的 3 个表，它们具有相同的结构，即每个表都有相同数量、名称和数据类型的字段。

客户编号	姓名	性别	等级	客户编号	姓名	性别	等级
KH001	胥田歌	男	特级	KH003	米寄翠	女	一级
KH002	赵泓轩	男	特级	KH004	姚储	男	一级
				KH005	黄夏萍	女	一级
				KH006	宣丽	女	一级
				KH007	周芹	女	二级
				KH008	余严	男	二级
				KH009	楚诗夏	女	二级
				KH010	张恭	男	二级

图 3-21 要合并的 3 个表

合并 3 个表中的记录的操作步骤如下：

Step01 在“数据源”页面的左侧窗格中双击“新建并集”，如图 3-22 所示。

Step02 打开“并集”对话框，在上方选择“特定（手动）”选项卡，然后将左侧窗格中的一个表拖动到“并集”对话框中，如图 3-23 所示。

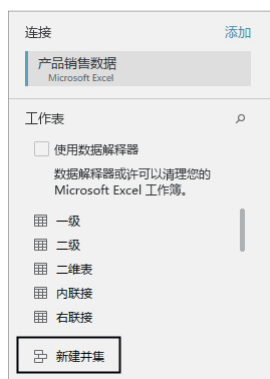


图 3-22 双击“新建并集”



图 3-23 在对话框中添加要合并的第一个表

提示：在左侧窗格中通过按住 Ctrl 键或 Shift 键，并配合鼠标单击，可以同时选择多个表，并将它们一次性添加到“并集”对话框中。

Step03 将要合并的另一个表拖动到“并集”对话框中，并放置到第一个表的下方，如图 3-24 所示。

Step04 使用相同的方法，将第三个表添加到“并集”对话框中，如图 3-25 所示。

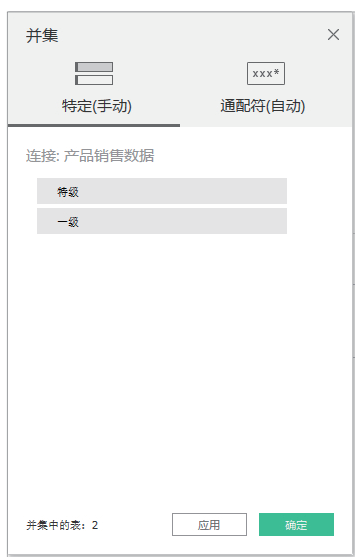


图 3-24 在对话框中添加要合并的第二个表

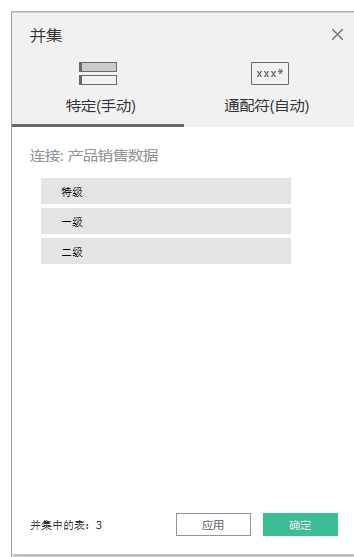


图 3-25 在对话框中添加要合并的第三个表

Step05 单击“确定”按钮，将“并集”对话框的所有表中的记录合并到一个新建的名为“并集”的表中，如图 3-26 所示。

并集

需要更多数据
将表拖到此处以使其相关。

排序字段 数据源顺序

并集 客户编号	并集 姓名	并集 性别	并集 等级	并集 工作表	并集 表名称
KH003	米奇翠	女	一级	一级	一级
KH004	姚陆	男	一级	一级	一级
KH005	黄夏萍	女	一级	一级	一级
KH006	宣丽	女	一级	一级	一级
KH007	周芹	女	二级	二级	二级
KH008	余严	男	二级	二级	二级
KH009	楚诗夏	女	二级	二级	二级
KH010	张赫	男	二级	二级	二级

图 3-26 合并表中的记录

3.4.2 重命名合并记录后的表

如需修改合并记录后由 Tableau 自动创建的逻辑表的名称，可以在画布中右击该表，在弹出的菜单中选择“重命名”命令，然后输入所需的名称并按 Enter 键，如图 3-27 所示。

如需修改合并记录后的逻辑表中的物理表的名称，可以双击画布中的逻辑表，进入物理层后双击其中的物理表，然后输入所需的名称并按 Enter 键，如图 3-28 所示。

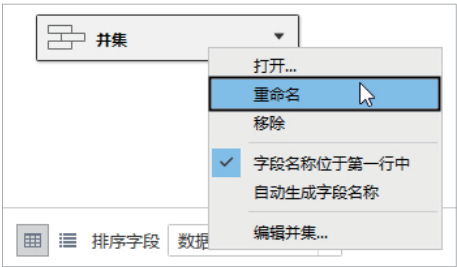


图 3-27 选择“重命名”命令



图 3-28 修改物理表的名称

3.4.3 添加要合并的表或删除已合并的表

如需在当前已合并记录的表中继续合并其他表中的记录，可以在画布中右击合并记录后的逻辑表，在弹出的菜单中选择“编辑并集”命令，打开“并集”对话框，然后从左侧窗格中将要合并的表拖动到该对话框中。



图 3-29 从并集中删除表

如需从当前已合并记录的表中删除一个或多个表中的记录，可以打开“编辑并集”对话框，然后将鼠标指针移动到要删除的表上，当右侧显示图标时，单击该图标即可将表删除，如图 3-29 所示。

3.5 创建计算字段

如果数据源中缺少分析所需的数据，则可以通过创建对现有字段执行计算的公式来获得所需的数据。通过输入计算公式创建的字段称为计算字段。下面列出了需要创建计算字段的一些常见应用场景：

- 为数据分段。
- 聚合数据。
- 筛选结果。
- 计算比率。

可以在“数据源”页面或“工作区”页面中创建计算字段。

3.5.1 创建简单的计算字段

如图 3-30 所示，数据源中有“销售额”和“利润”两个字段，可以通过这两个字段计算出成本。此处介绍在“数据源”页面中创建计算字段的方法，创建“成本”计算字段的操作步骤如下：

Step01 将包含“销售额”和“利润”两个字段的“订单明细”表添加到“数据源”页面的画布中。

Step02 在数据网格中右击“销售额”字段的顶部区域，然后在弹出的菜单中选择“创建计算字段”命令，如图 3-31 所示。

#	#	#
订单明细	订单明细	订单明细
数量	销售额	利润
23	230	55.20
80	400	112.00
43	3,010	662.20
91	273	54.60
71	355	53.25
98	4,900	882.00
44	3,080	616.00
72	1,080	248.40

图 3-30 “销售额”和“利润”两个字段

#	#	#
订单明细	订单明细	订单明细
数量	销售额	利润
23		
80		
43		
91		
71		
98		
44		
72	1,080	248.40

图 3-31 选择“创建计算字段”命令

Step03 打开如图 3-32 所示的计算编辑器，在上方的文本框中输入计算字段的名称，本例为“成本”，然后在下方输入以下公式：

[销售额]-[利润]



图 3-32 输入计算字段的名称和公式

提示：对话框的左下角会显示对输入公式的正确性的判断结果，显示“计算有效”表示用户输入的公式在语法上是正确的，可以正常计算。

Step04 单击“确定”按钮，关闭计算编辑器，新建的“成本”计算字段将显示在数据网格中，如图 3-33 所示。

提示：为了与数据源中的原生字段相区别，Tableau 会在计算字段的数据类型图标上显示一个等号，如图 3-34 所示。

# 订单明细 数量	# 订单明细 销售额	# 订单明细 利润	=# 计算 成本
23	230	55.20	174.80
80	400	112.00	288.00
43	3,010	662.20	2,347.80
91	273	54.60	218.40
71	355	53.25	301.75
98	4,900	882.00	4,018.00
44	3,080	616.00	2,464.00
72	1,080	248.40	831.60

图 3-33 创建的计算字段显示在数据网格中



图 3-34 计算字段的数据类型图标

3.5.2 使用函数创建计算字段

函数提供了丰富的计算方式，只需为函数提供要计算的数据，函数就能计算出结果。创建计算字段时，为了满足不同需求的计算，可以在输入的公式中使用函数，而不只是简单的四则运算。

如图 3-35 所示，客户编号由字母和数字两部分组成，现在要创建一个名为“数字编号”的计算字段，其中只包含客户编号中的后 3 位数字。

此处介绍在“工作区”页面中创建计算字段的方法，创建“数字编号”字段的操作步骤如下：

- Step01** 将包含“客户编号”字段的“客户信息”表添加到“数据源”页面的画布中。
- Step02** 单击 Tableau Desktop 窗口下方的“工作表 1”，切换到“工作区”页面，然后单击菜单栏中的“分析”|“创建计算字段”命令，如图 3-36 所示。

客户信息	客户信息	客户信息	客户信息
客户编号	姓名	性别	等级
KH001	胥田歌	男	特级
KH002	赵泓轩	男	特级
KH003	米奇翠	女	一级
KH004	姚皓	男	一级
KH005	黄夏萍	女	一级
KH006	宣丽	女	一级
KH007	周萍	女	二级
KH008	余严	男	二级

图 3-35 客户编号由字母和数字组成

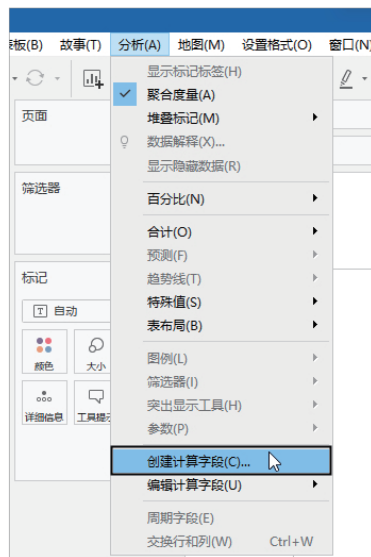


图 3-36 选择“创建计算字段”命令

技巧：由于要在公式中输入“客户编号”字段的名称，所以可以直接在“数据”窗格中右击“客户编号”字段，然后在弹出的菜单中选择“创建”|“计算字段”命令，在打开的计算编辑器中会自动输入“客户编号”字段的名称，如图 3-37 所示。

Step03 打开计算编辑器，在上方的文本框中输入计算字段的名称，本例为“数字编号”，如图 3-38 所示。

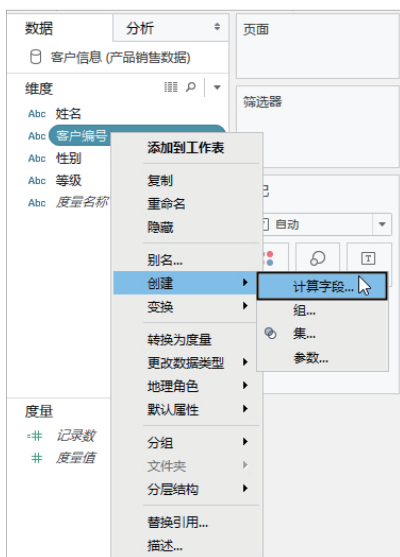


图 3-37 选择“计算字段”命令



图 3-38 输入计算字段的名称

Step04 单击计算编辑器右侧中间位置处的三角按钮，展开函数列表，然后双击其中的 RIGHT，将 RIGHT 函数添加到左侧的公式文本框，如图 3-39 所示。

提示：添加函数后，可以再次单击三角按钮将函数列表折叠起来。

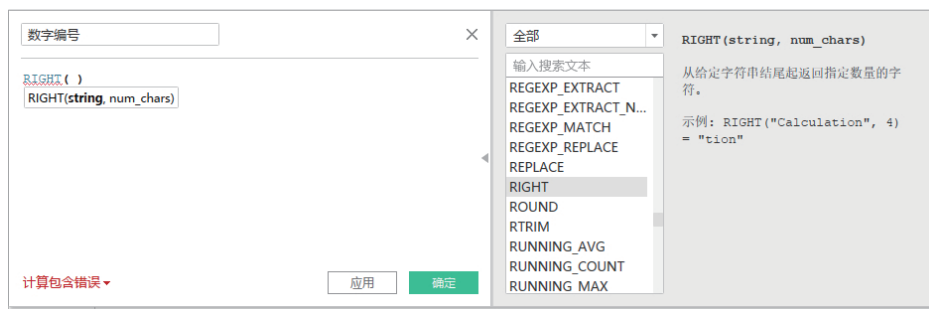


图 3-39 添加 RIGHT 函数

Step05 在左侧的 RIGHT 函数右侧的圆括号中输入 “[客户编号],3”，完整公式如下，如图 3-40 所示。

```
RIGHT([客户编号],3)
```



图 3-40 输入公式

提示：在 Step05 中输入的内容由两部分组成，它们之间以逗号分隔。两个部分都是 RIGHT 函数的参数，第一个部分是 RIGHT 函数要处理的数据，本例为“客户编号”字段中的每个客户编号；第二个部分是要提取的字符个数，本例输入 3，表示从每个客户编号的最右侧提取 3 个字符。

Step06 单击“确定”按钮，关闭计算编辑器，创建的“数字编号”计算字段将显示在“数据”窗格中，如图 3-41 所示。



图 3-41 创建的计算字段显示在“数据”窗格中

有时可能需要使用聚合函数来创建计算字段，例如创建“销售利润率”字段。销售利润率的计算公式如下：

$$\text{利润} \div \text{销售额} \times 100\%$$

在一个包含“销售额”和“利润”的表中创建一个名为“销售利润率”的计算字段，在计算编辑器中输入以下公式，然后单击“确定”按钮，如图 3-42 所示。

```
SUM([利润])/SUM([销售额])
```



图 3-42 创建“销售利润率”计算字段

提示：创建“销售利润率”计算字段时，如果在公式中输入“ $\times 100\%$ ”或“ $\ast 100\%$ ”，则公式会出现错误。如需以百分比格式显示销售利润率，可以在将该计算字段添加到视图后，设置该字段的数字格式，具体方法将在第 4 章中介绍。

3.5.3 编辑计算字段

如需修改计算字段的名称或公式，可以在“数据源”页面的数据网格或“工作区”页面的“数据”窗格中右击计算字段，然后在弹出的菜单中选择“编辑”命令，如图 3-43 所示，在打开的计算编辑器中进行修改，完成后单击“确定”按钮。



图 3-43 选择“编辑”命令

3.6 在数据窗格中管理字段

第1章曾经介绍过“数据”窗格的外观和布局结构，本节将介绍在“数据”窗格中对字段执行的一些常用操作，其中的一些操作与在“数据源”页面管理字段的方法类似。“数据”窗格位于“工作区”页面中，为了在“数据”窗格中显示正确的字段，需要先连接数据并设置数据源，然后单击 Tableau Desktop 窗口下方的“工作表 1”（或其他工作表名称），切换到“工作区”页面。

3.6.1 更改字段名称

如需在“数据”窗格中更改字段的名称，可以使用以下几种方法：

- 在“数据”窗格中右击字段，然后在弹出的菜单中选择“重命名”命令，如图 3-44 所示。
- 在“数据”窗格中单击字段并按住鼠标左键，当字段名称显示为编辑状态时，释放鼠标左键。
- 在“数据”窗格中单击字段，然后按 F2 键。

使用任意一种方法，将进入字段名称的编辑状态，如图 3-45 所示。输入字段的新名称，然后按 Enter 键，即可完成名称的修改。

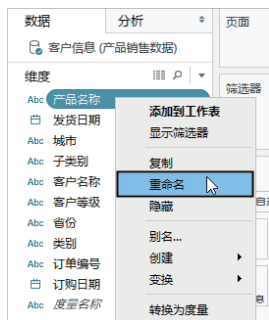


图 3-44 选择“重命名”命令



图 3-45 字段名称显示为编辑状态

如需将字段名称恢复为默认名称，可以进入字段名称的编辑状态，然后单击字段名称右侧的箭头，如图 3-46 所示。

如需一次性恢复多个字段的默认名称，可以在“数据”窗格中选择这些字段，然后右击选中的任意一个字段，在弹出的菜单中选择“重置名称”命令，如图 3-47 所示。



图 3-46 单击右侧箭头恢复字段的默认名称

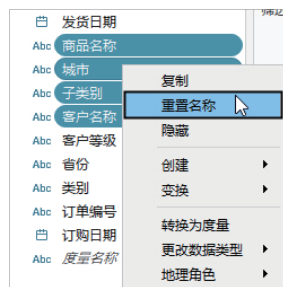


图 3-47 选择“重置名称”命令

3.6.2 为字段中的成员创建别名

通过为字段中的成员创建别名，可以使字段成员在视图中以指定的名称显示。只能为离散维度字段中的成员创建别名，不能为连续维度字段、日期字段和度量字段创建别名。

如需为字段中的成员创建别名，可以在“数据”窗格中右击该字段，然后在弹出的菜单中选择“别名”命令，打开“编辑别名”对话框，如图 3-48 所示。选择要创建别名的字段成员，该成员在“值”列中的内容自动显示为编辑状态，输入所需的名称，然后按 Enter 键，即可为该成员创建别名，如图 3-49 所示。



图 3-48 “编辑别名”对话框



图 3-49 输入别名

提示：如果已经为字段成员创建别名，则会在该字段成员所在行的“有别名”列中显示一个星号，如图 3-50 所示。



图 3-50 使用星号标记已创建的别名

如需删除字段成员的别名，可以在“编辑别名”对话框中单击“清除别名”按钮。

3.6.3 更改字段的数据类型

如需在“数据”窗格中更改字段的数据类型，可以在该窗格中右击字段，在弹出的菜单中选择“更改数据类型”命令，然后在子菜单中选择一种数据类型，如图 3-51 所示。

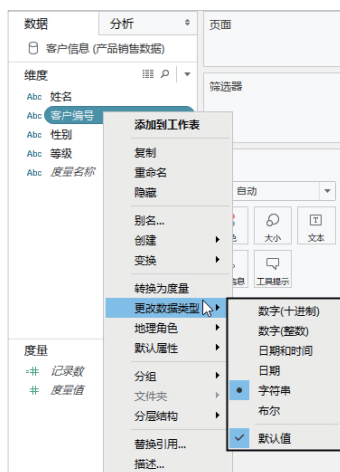


图 3-51 更改字段的数据类型

3.6.4 在维度和度量之间转换

在 Tableau Desktop 中设置数据源后，Tableau 会根据字段的数据类型自动将字段划分为维度或度量，通常将文本类型的字段划分为维度，将数值类型的字段划分为度量。如果字段的分类有误，或者希望字段在视图中发挥特殊的作用，则可以手动将字段在维度和度量之间转换。

如需将字段在维度和度量之间转换，可以在“数据”窗格中右击字段，然后在弹出的菜单中选择“转换为度量”或“转换为维度”命令，如图 3-52 所示。显示哪个命令取决于当前右击的字段是维度还是度量。



图 3-52 将字段在度量和维度之间转换

另一种转换发生在离散字段和连续字段之间，可以将度量字段在离散和连续之间转换。对于维度字段来说，只有日期维度和数值类型维度可以在离散和连续之间转换。在“数据”窗格中右击要转换的字段，然后在弹出的菜单中选择“转换为连续”或“转换为离散”命令。

3.6.5 隐藏字段

为了避免造成干扰，可以将不使用的字段隐藏起来，这样字段就不会显示在“数据”窗格中或者仅显示为灰色。如需隐藏一个或多个字段，可以先在“数据”窗格中选择这些字段，然后右击选中的任意一个字段，在弹出的菜单中选择“隐藏”命令，如图 3-53 所示。

如需查看隐藏的字段，可以单击“数据”窗格中的下拉按钮，在弹出的菜单中选择“显示隐藏字段”命令，如图 3-54 所示。

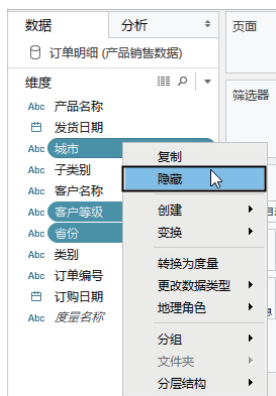


图 3-53 选择“隐藏”命令

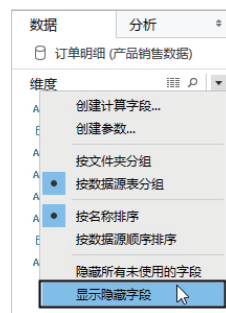


图 3-54 选择“显示隐藏字段”命令

提示：如需隐藏所有未在视图中使用的字段，可以在图 3-54 所示的菜单中选择“隐藏所有未使用的字段”命令。

处于隐藏状态的字段在“数据”窗格中显示为灰色，如图 3-55 所示的“城市”“客户等级”和“省份”3 个字段是处于隐藏状态的字段。



图 3-55 处于隐藏状态的字段显示为灰色

如需使隐藏的字段恢复正常显示，可以选择这些字段，然后右击选中的任意一个字段，在弹出的菜单中选择“取消隐藏”命令。

3.6.6 使用文件夹组织字段

如果“数据”窗格中的字段数量较多，可以使用文件夹组织字段，以便对字段进行分类管理，使“数据”窗格更简洁。为了使用文件夹组织字段，首先需要单击“数据”窗格中的下拉按钮，在弹出的菜单中选择“按文件夹分组”命令，如图 3-56 所示。

现在“数据”窗格中的字段按文件夹分组显示，看不到分组效果是因为在“数据”窗格中还没有创建文件夹。如需创建文件夹，可以右击“数据”窗格中的空白处，在弹出的菜单中选择“创建文件夹”命令，如图 3-57 所示。

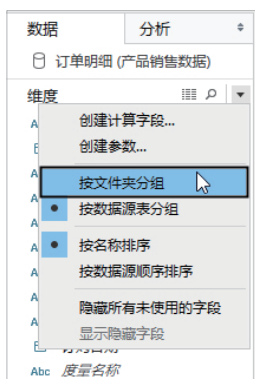


图 3-56 选择“按文件夹分组”命令

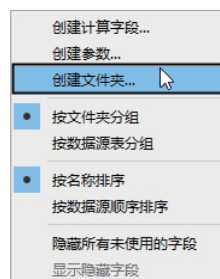


图 3-57 选择“创建文件夹”命令

打开“创建文件夹”对话框，在“名称”文本框中输入文件夹的名称，然后单击“确定”按钮，如图 3-58 所示。创建文件夹后，如需将字段添加到文件夹中，可以直接将字段拖动到文件夹上，当鼠标指针附近出现 + 号时，释放鼠标按键，即可将字段移入文件夹，如图 3-59 所示。

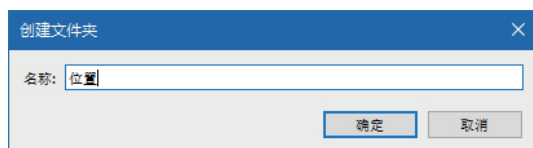


图 3-58 输入文件夹的名称

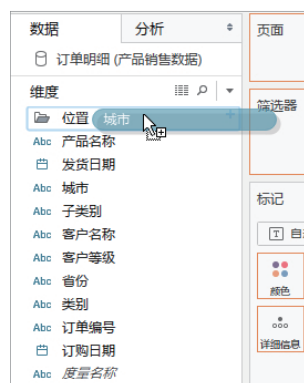


图 3-59 将字段拖动到文件夹上

如果文件夹中包含字段，则可以单击文件夹图标左侧的按钮展开或折叠文件夹，以便显示或隐藏其中的字段，如图 3-60 所示。



图 3-60 展开或折叠文件夹

如果在“数据”窗格中有多个文件夹，则可以将字段从一个文件夹移动到另一个文件夹，只需将字段从一个文件夹拖动到另一个文件夹即可。还可以右击字段，在弹出的菜单中选择“文件夹”|“添加到文件夹”命令，然后在子菜单中选择移动到的目标文件夹，如图 3-61 所示。

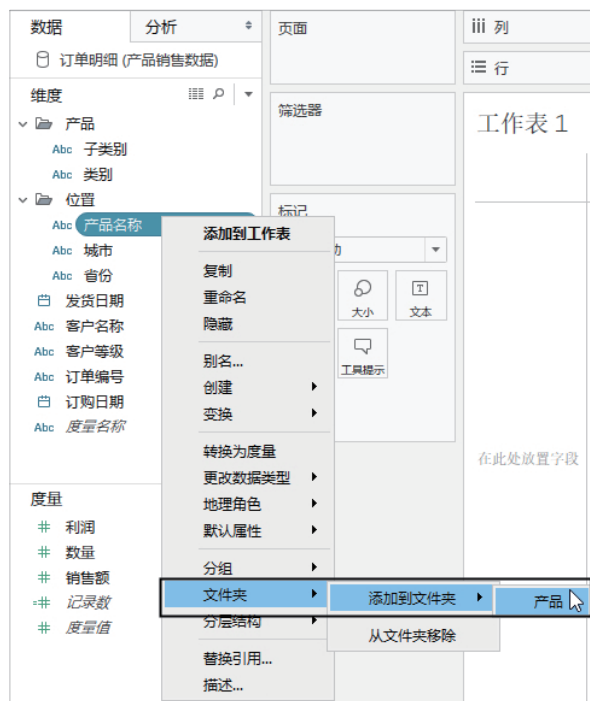


图 3-61 选择移动到的目标文件夹

提示：如需删除文件夹中的字段，可以选择图 36-1 中的“从文件夹移除”命令。

如需删除文件夹，可以右击文件夹，在弹出的菜单中选择“移除文件夹”命令，如图 3-62 所示。

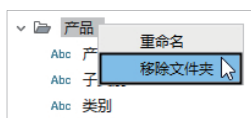


图 3-62 选择“移除文件夹”命令

3.6.7 更改字段的默认值

将字段添加到视图时，字段在视图中最初的显示方式由字段的默认值决定，Tableau 会为每个字段设置默认值，用户可以根据实际需求，更改字段的默认值。如果将字段的默认值设置为日常所需的值，在将字段添加到视图时，就不必每次都要重复更改字段的设置。

如需更改字段的默认值，可以在“工作区”页面的“数据”窗格中右击字段，在弹出的菜单中选择“默认属性”命令，然后在子菜单中选择要更改的默认值类型。如图 3-63 所示是度量字段包含的默认值类型，维度字段包含的默认值类型略有不同。

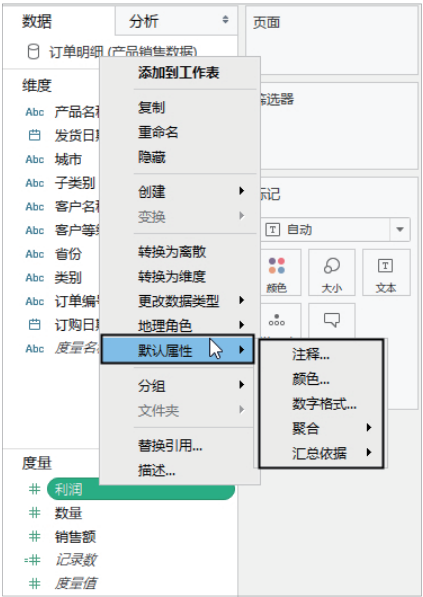


图 3-63 选择默认值的类型

选择一种类型后，再进一步设置默认值。以后将字段添加到视图时，会使用新的默认值来显示字段。