

使用文本和 Excel 文件准备数据

欢迎观看本次有关数据准备的视频。

您可以下载数据集，以便在自己的 Tableau 软件副本中跟着操作。本视频中讲到的一些功能，例如“拆分”，也可用于其他数据源。但是，有一些功能仅适用于文本或 Excel 文件。

这里有一个 Excel 格式的报告，其中显示了每月每位员工解决的事件数。根据以往的经验，像这样的格式化报告加载到 Tableau 中的效果并不好。“理想”选项卡中显示了类似数据库表格式的数据，也就是原始数据。要在 Tableau 中实现最优化分析，最好就选择这种格式。

不过，现在 Tableau Desktop 中包含了多项功能，可用于自动将文本和 Excel 文件转换为这种格式，从而在 Tableau 中获得更好的分析效果。

首先，我们连接到该 Excel 文件。

- 打开 Tableau Desktop 并从左侧选择 Excel。
- 导航到已下载文件并单击打开。
- 此时将显示“已解决的事件”工作表。

尽管 Tableau 可以连接到该报告，但我们在预览这里可以看出，这种格式的显示效果并不是很好。没有列名称、Excel 中的许多标题都是显示为 Null 等。Tableau 也发现了这一问题，因此会建议使用数据解释器。

- 让我们单击“启用”
- 现在可以看到，那些标题和 Null 值都已去除，我们可以正确辨识每一列。
- 我们可以继续，但如果想要获得有关数据解释器作用的更多细节，可以单击右侧的“查看结果”。这将打开一个 Excel 文件，其中说明了所有更改。
- 如果单击我们使用的“已解决的事件”标签，可以看出哪些字段用作标题，显示为红色，还有哪些字段被视为数据，显示为绿色。

在返回 Tableau 和数据连接之前，让我们再仔细看一下这个“理想”选项卡。请注意在这种格式下，不是每个月对应一个包含数据的列，而是有一个“日期”列，其下每行中包含日期和员工的每个独特组合所对应的已解决事件数。这样的数据更长，有更多行，而不是更多列显得更宽。

返回 Tableau，我们想要将这种每月一列的布局格式更改为单个日期列和单个“已解决的事件”列。

- 这很简单，只需选择所有日期列。单击第一列，必要时可滚动鼠标，然后按住 Shift 键的同时

单击最后一列。此时将打开菜单并选择“数据透视表”

- 从本质上说，这种数据透视表功能就是将原来各列和各行中的信息合并成两个新列，分别是数据透视表字段名称和数据透视表字段值。
- 可以看出，“数据透视表字段名称”实际就是我们的“日期”，单击可打开菜单并选择“重命名”。
- “数据透视表字段值”可重命名为“已解决的事件”

如果单击第一列上方的这个图标，将转到元数据网格。这个视图非常有用。垂直布局可能更容易导航，尤其在包含大量字段时更是如此，而且当表格已连接时，使用这种布局更有帮助。而现在，我们就后退到标准视图。

我们的数据几乎无所不在，但我们还可以做一件事。在这里，这个“员工”字段实际是两部分信息 – A、B、C、D 或 E 代表位置代码，后接员工 ID 号。

- 我们可以基于该连字符的共享分隔符对该列进行拆分。
- 单击打开菜单并选择“拆分”
- 现在有两个新字段：“员工” - 拆分 1 和拆分 2
- 我们将使用元数据网格对拆分字段进行重命名
- 只需单击名称即可按顺序进行编辑，拆分 1 应该是“位置”，我们单击选项卡，拆分 2 应该是“员工 ID”
- “日期”字段旁边的 Abc 表示该列被视为字符串。但是我们知道这实际上是日期，因此可以单击 Abc 并选择“日期”以更新数据类型。

现在如果单击工作表 1，就可以看到布局十分整洁的数据可供分析了！

我们将“员工 ID”拖到视图中并将“已解决的事件”拖至“列”，然后进行排序

- 从该视图中可以明显看出，实际上有两组员工 – 一组员工解决的事件数明显多于另一组员工解决的事件数。
- 资深员工通常能够解决更多问题，将他们指定为第 II 层。
- 如果查看 Excel 中的原始数据集，可以发现有一个名为“层”的选项卡。
- 在此报告中，第 II 层员工的员工 ID 末尾都加上了“-II”
- 并非所有行中都包含这个“-II”，所以标准拆分不起作用
- 请注意，“拆分”和“自定义拆分”必须使用一致的分隔符。如果我们的数据中包含不规则分隔符，就像此工作表中一样，则 Tableau 将无法使用这些选项拆分数据。

让我们后退到 Tableau 并添加另一个数据连接

- 单击“保存”按钮旁边的“添加数据”图标

- 选择 Excel
- 和之前一样导航到相同的文件
- 但是这次，我们选择“层”
- 我们将打开数据解释器，但现在不用担心数据透视表的问题
- 在“预览”视图中，单击“员工”列并选择“自定义拆分”
 - 我们可以选择自己的分隔符，我们将使用连字符
 - 现在可以说，我们希望有 3 列
 - 这时 Tableau 会使用第 II 层指示符断开第 3 列

感谢您观看本数据准备培训视频。我们邀请您继续观看“点播式培训”视频，进一步了解如何使用 Tableau。