

# 跨数据库联接

---

欢迎观看“跨数据库联接”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。

## 来自不同数据源的数据

许多组织在不同的系统中保存不同的数据。SQL Server 数据库中可能有财务数据，而 Amazon Redshift 中可能有产品数据。数据可能位于不同的环境中，但我们希望将所有东西放在一起分析。

在该视频中，我们使用基于以下两种不同文件的数据源：Excel 和 CSV。为方便您跟着操作，我们仅使用平面文件。

在此处所用的数据结构中，销售数据包含产品 ID，但没有关于所售产品的其他信息。CSV 文件中的产品数据是包含产品 ID、类别、子类别和产品名称的简单表格。

这种情况非常适合跨数据库联接。如果根据产品 ID，以左联接的方式将产品数据联接 to 销售数据，我们就可以将产品信息添加到每笔相关销售交易。要详细了解联接，请观看关于联接类型的视频。

## 跨数据库联接

为使用跨数据库联接创建该集成数据源，

- 从“销售”开始 - 单击 Excel 并打开“Sales 2016 (2016 年销售)”文件
  - 我们双击并将此表命名为“销售”，以便理清头绪
- 接下来，我们添加一个与该数据源的数据连接
  - 在“连接”区域旁单击“添加”按钮，而非工具栏中的“新建数据源”按钮
  - 依次单击“添加”和“文本文件”
  - 打开“Products 2016 (2016 年产品)”
- 工作表已添加到画布中
  - 将其重命名为“产品”
- 然后，将数据源命名为更具描述性的“2016 年销售和产品”
- 与同一数据源中的所有联接一样，也可单击图标编辑联接，我们将选择左联接。

- 可看到数据已基于产品 ID 链接到行级别

在网格中，可看到产品信息（黄色）已添加到“销售”数据（蓝色）的每一行。

## 使用集成数据源

我们现在得到一个新的单一数据源，让我们测试一下。

- 可在数据窗格中看到单一数据源，但字段已按表格分开 - 如同相同数据源中的联接。
  - 必要时可单击右键并选择“按文件夹分组”将其关闭，而不选择“按数据源表分组”
- 为子类别和类别创建分层结构
  - 添加产品 ID 和产品名称
  - 可在 CSV 表中隐藏产品 ID，因为它仅用于联接子句。
- 我们将“类别”移至行并将其展开
- 将“销售”移至列
- 按订单日期分开

集成数据源的功能与任何其他单一来源数据源相似，甚至还可以发布和保存，这与混合不同。

## 结语

感谢您观看“跨数据库联接”视频。诚邀您继续观看免费培训视频，进一步了解如何使用 Tableau。