

数据混合

欢迎观看这段介绍数据混合的视频。您可以下载练习工作簿，以便在自己的 Tableau 软件副本中跟着操作。

什么是数据混合？

假设您有两家零售连锁店：Office City 和 Coffee Chain。您将这两家连锁店的数据分别记录在两个不同的系统中。数据混合就是将来自多个数据源的数据合并到单个视图中的一种方式。它不是像跨数据库联接那样在行一级联接数据，而是向不同的数据源分别发送查询，然后在 Tableau 中将返回的查询结果聚合到同一级别。

混合关系

在这里的数据窗格中，有两个数据源：Office City 和 Coffee Chain。我们可以在这两个数据源间轻松切换，切换时请注意观察维度和度量的变化。要进行数据混合，这两个数据源之间需要有至少一个公共字段，比如“州/省”。如果字段名称不同，但是成员相同，我们可以手动定义它们的关系。

我们知道 Office City 的“地区”字段和 Coffee Chain 的“区域”字段都包含值“中部”、“东部”、“南部”和“西部”，所以我们就定义这两个字段是等同的。为此，需打开“数据”菜单，选择“编辑关系”。我们可以将主数据源更改为 Office City，选中“自定义”，然后添加一层关系：两个字段分别选择“地区”和“区域”。

这时您会注意到，Tableau 已经在这两个字段间建立了一层关系，此外还列出了使用公共字段“州/省”自动建立的关系。另一种方案是对一个数据源的字段进行重命名，使其与另一个数据源的相应字段匹配。那我们就缓慢双击 Coffee Chain 的“区域”字段，将其重命名为“地区”。如果两个字段同名，Tableau 可以自动在两个数据源之间建立关系。

至此已经定义好了公共字段，接下来就可以开始混合了。

如何混合

首先，选择 Office City 数据源。将 Office City 的“销售额”字段拖到列功能区，将其“州/省”字段拖到行功能区。请注意，在数据窗格中 Office City 旁边现在有了一个蓝色的复选标记。每当在 Tableau 中连接到多个数据源时，第一个添加到视图中的数据源就会成为主数据源，主数据源就以这个蓝色的复选标记指示。现在，我们来切换到第二个数据源。请注意“州/省”字段旁边的橙色链接图标。由于我们已经将“州/省”拖到了视图中，所以 Tableau 会自动基于该字段混合数据，并以橙色链接指示该字段。

在“地区”字段旁边还有一个灰色的断开状链接。由于“地区”是两个数据源的公共字段，因此也是潜在的链接字段，只是现在未在当前视图使用。如果还要基于该字段混合数据，点击此灰色链接即可将其激活并建立关系。那我们就将 Coffee Chain 的“销售额”字段拖到列功能区来完成数据混合。现在，可以看到在 Coffee Chain 数据源旁边有一个橙色的复选标记，在视图中的胶囊装字段上也有一个。此标记表示该字段来自辅助数据源。

这个视图具体是如何构建的呢？Tableau 向每个数据源查询各个州/省的销售额总计，然后将查询到的这些值一起显示在这个视图中。可以看到，Office City 在每个州/省的销售额信息都显示了出来，这是因为 Tableau 会返回主数据源中所有字段成员（即州/省）的信息，而不管辅助数据源中是否有数据。

交换主数据源和辅助数据源

务必要注意的是，具体哪个是主数据源、哪个是辅助数据源要根据工作表而定，而不是在整个工作簿中一成不变。当我们打开一个新的工作表时，数据窗格中的数据源并没有相应的橙色和蓝色复选标记来指示它们是主数据源还是辅助数据源。这里并不会沿用我们在之前的工作表中建立的关系。

如果我们交换数据源，将 Coffee Chain 用作主数据源，而将 Office City 用作辅助数据源，将会怎样？我们先把 Coffee Chain 的“销售额”字段拖到视图中，再将“州/省”字段拖过来。与之前的示例相比，现在视图中的州/省数量只有原来的一半左右。这是因为在 Coffee Chain 的数据集中，州/省的数量比 Office City 少。Tableau 会先显示 Coffee Chain 中的所有州/省，然后从 Office City 中只提取相关的信息。这意味着，在 Office City 中有，而在 Coffee Chain 中没有的任何州/省都不会在此视图中显示。

结语

感谢您观看这段介绍数据混合的视频。我们邀请您观看“其他数据混合主题”视频以及我们的其他免费培训视频，详细了解如何使用 Tableau。