

上面的代码展示了如何将两个 DataFrame 根据某一个列来进行合并。这里提醒读者注意以下几点。

一是传入的两个 DataFrame 必须有相同的列名。上面的例子中，“ind\_name”这个 DataFrame 中行业代码的列名为“code”，所以在上面的代码中使用 rename 方法将其列名改为“ind\_code”，与 trading\_data\_2019 保持一致。读者可以单独使用 rename 这一方法进行观察。

二是通过参数 on 来设置用于合并的列名，合并的列我们称之为合并轴。在上述的例子中，合并轴就是两个 DataFrame 的 ind\_code 列。

三是通过 how 来设置整合过程中的模式。how 共有 4 种参数可选：left、right、outer、inner。其中 left 表示合并过程以左侧 DataFrame 为主（即第 1 个传入的 DataFrame），采用左侧的合并轴来合并，最后整合完成的数据集数据和左侧的数据集数据长度一致；right 同理，以右侧的 DataFrame 为主；outer 表示最终合并的结果是一个并集，是两个合并轴并集的数目；inner 则表示交集。

在上面使用 merge 的例子中，我们对 DataFrame 进行了横向合并。还有一种 DataFrame 的合并是纵向合并，也就是将表征同一件事情的多个 DataFrame 首尾衔接起来。

为简单起见，我们将“trading\_data\_2019”这一 DataFrame 分割为 3 个部分作为示例。

```
In:
part1 = trading_data_2019.iloc[:9]
part2 = trading_data_2019.iloc[9:-10]
part3 = trading_data_2019.iloc[-10:]
total = pd.concat([part1, part2, part3])
```

读者可以自行验证，最后使用“concat”组合起来的 DataFrame 是不是和原来的一样。Pandas 中的 concat 就是将一系列具有相同列名的分裂数据集合在一起的函数，我们只需要将分裂的数据集整理到一个列表当中就可以了。

Pandas 的入门知识就介绍到这里，在本书后续的代码中，遇到新的 Pandas 使用方法时，笔者会加以提示。读者朋友们也可以通过互联网等渠道继续学习 Pandas 的技巧和功能，笔者强烈推荐 Pandas 官网的学习文档，读者可自行查看。