

Chapter 1

Combining Pandas Objects

1.1 引言

有多种选项可用于将两个或多个 DataFrame 或 Series 组合在一起。追加方法是最不灵活的，只允许将新行追加到 DataFrame 中。concat 方法非常通用，可以在任一轴上组合任意数量的 DataFrame 或 Series。join 方法通过将一个 DataFrame 的列与其他 DataFrame 的索引对齐来提供快速查找。merge 方法提供了类似 SQL 的功能来将两个 DataFrame 连接在一起。

1.2 Understanding the differences between concat, join, and merge

The .merge and .join DataFrame (and not Series) methods and the concat function all provide very similar functionality to combine multiple pandas objects together.

- concat
 - A pandas function
 - Combines two or more pandas objects vertically or horizontally
 - Aligns only on the index
 - Errors whenever a duplicate appears in the index
 - Defaults to outer join with the option for inner join
- .join
 - A DataFrame method
 - Combines two or more pandas objects horizontally
 - Aligns the calling DataFrame's column(s) or index with the other object's index (and not

the columns)

- Handles duplicate values on the joining columns/index by performing a Cartesian product
 - Defaults to left join with options for inner, outer, and right
- `.merge`
 - A DataFrame method
 - Combines exactly two DataFrames horizontally
 - Aligns the calling DataFrame's column(s) or index with the other DataFrame's column(s) or index
 - Handles duplicate values on the joining columns or index by performing a cartesian product
 - Defaults to inner join with options for left, outer, and right