

要合并多个键, 则把多个键赋给 on 选项。

```
>>> pd.merge(frame1, frame2, on=['id', 'brand'], how='outer')
   brand  color      id
0  OMG  white    ball
1  ABC   red   pencil
2  ABC   red     pen
3  POD  black     mug
4  POD  green  ashtray
5  OMG   NaN    pencil
6  POD   NaN    pencil
7  ABC   NaN    ball
8  POD   NaN     pen
```

根据索引合并

有时, 合并操作不用 DataFrame 的列而用索引作为键。把 left_index 和 right_index 选项的值置为 True, 将其激活, 就可将其作为合并 DataFrame 的基准。

```
>>> pd.merge(frame1, frame2, right_index=True, left_index=True)
   brand_x  color  id_x brand_y  id_y
0  OMG  white  ball  OMG  pencil
1  ABC   red  pencil  POD  pencil
2  ABC   red    pen  ABC   ball
3  POD  black    mug  POD    pen
```

但是 DataFrame 对象的 join() 函数更适合根据索引进行合并。还可以用它合并多个索引相同或索引相同但列却不同的 DataFrame 对象。

输入以下代码:

```
>>> frame1.join(frame2)
```

pandas 将会给出错误信息, 因为 frame1 的列名称与 frame2 有重合。因此在使用 join() 函数之前, 要重命名 frame2 的列。

```
>>> frame2.columns = ['brand2', 'id2']
>>> frame1.join(frame2)
   brand  color      id brand2  id2
0  OMG  white    ball  OMG  pencil
1  ABC   red   pencil  POD  pencil
2  ABC   red     pen  ABC   ball
3  POD  black     mug  POD    pen
4  POD  green  ashtray  NaN    NaN
```

上述合并操作是以索引而不是列为基准。合并后得到的 DataFrame 对象包含只存在于 frame1 中的索引 4, 但是整合自 frame2、索引号为 4 的各元素均为 NaN。

6.2 拼接

另外一种数据整合操作叫作拼接 (concatenation)。NumPy 的 concatenate() 函数就是用于数组的拼接操作。