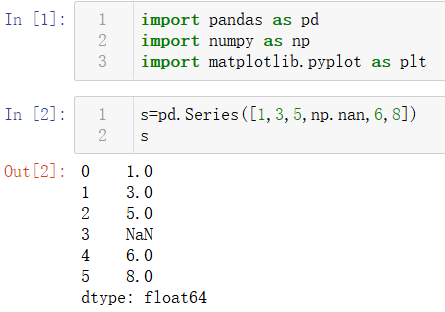
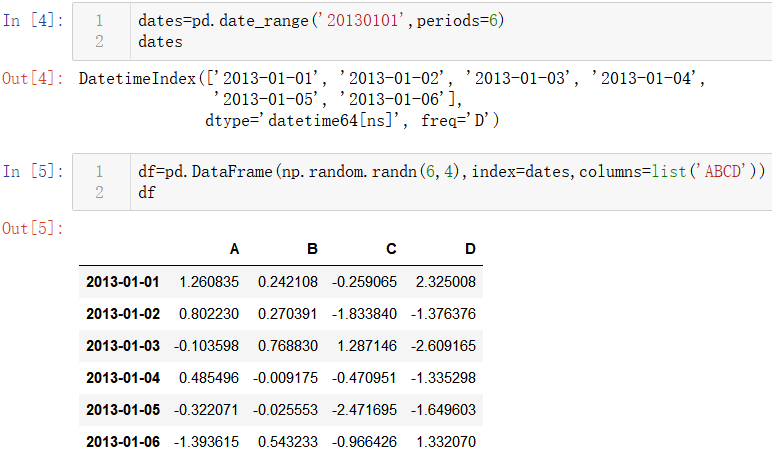
1.创建对象

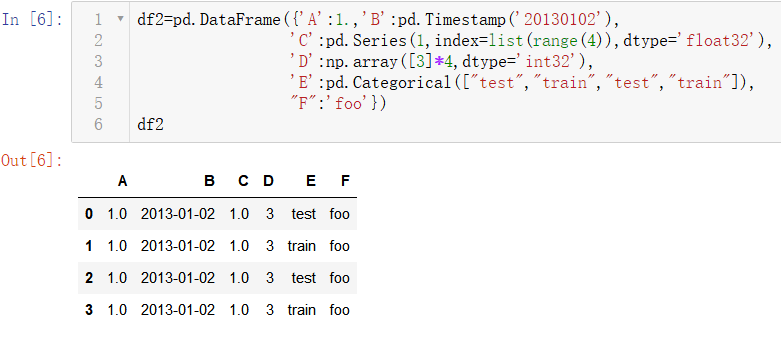
通过传递值列表来创建Series，让pandas创建一个默认的整数索引：



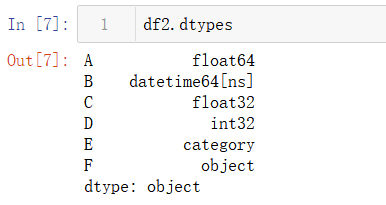
通过传递带有日期时间索引和标记列的NumPy数组来创建DataFrame：



通过传递可以转换为类似series的对象的dict来创建DataFrame。

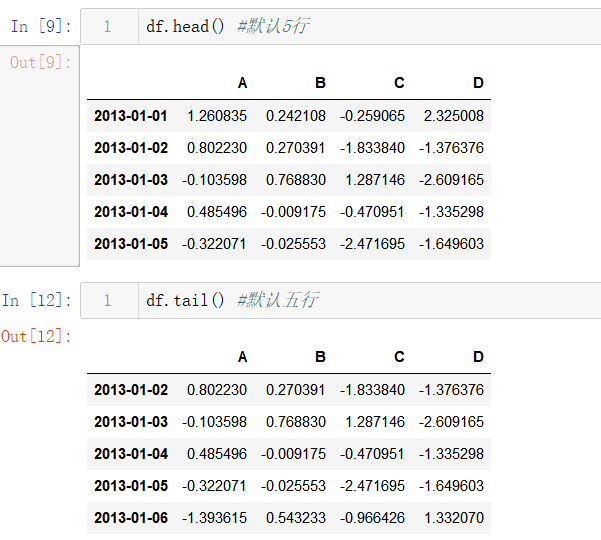


生成的DataFrame的列具有不同的dtypes。

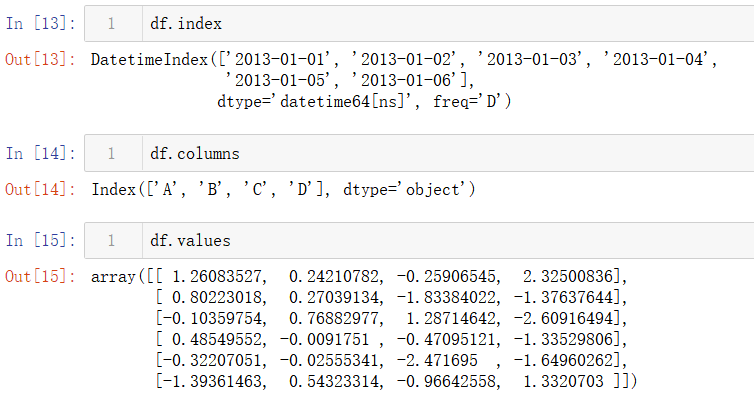


2.查看数据

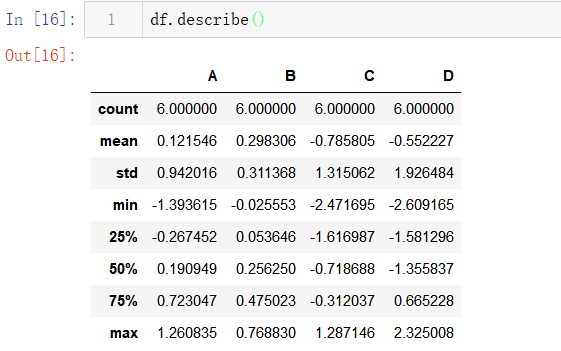
以下是查看frame的顶部和底部行的方法：



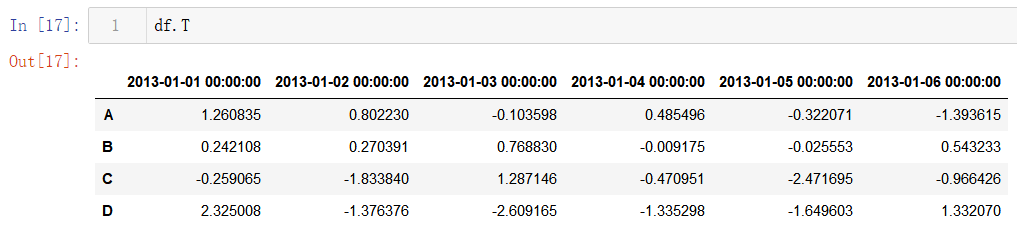
显示索引，列和基础NumPy数据：



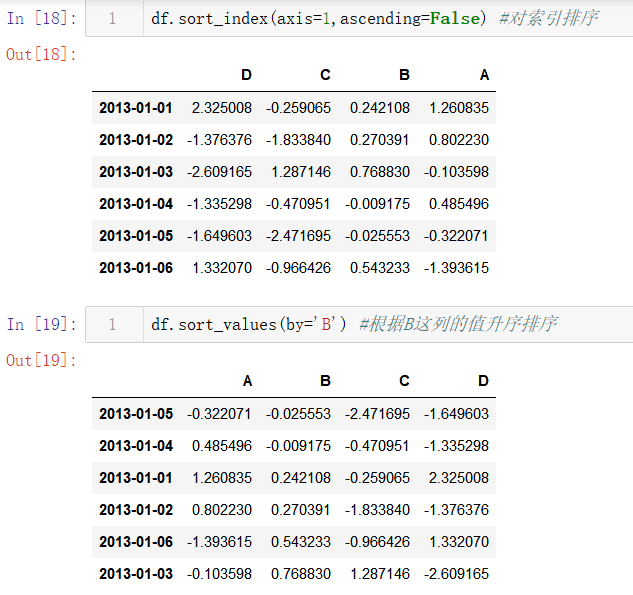
Describe()方法显示了数据的快速统计摘要：



数据转置:

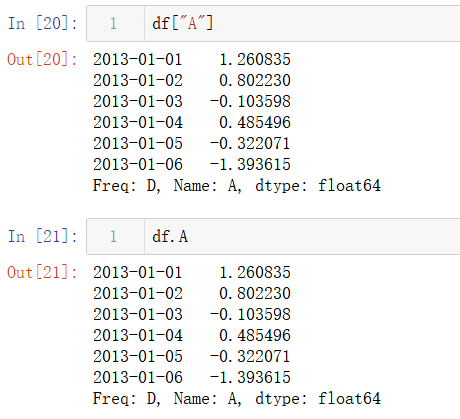


根据某个轴排序

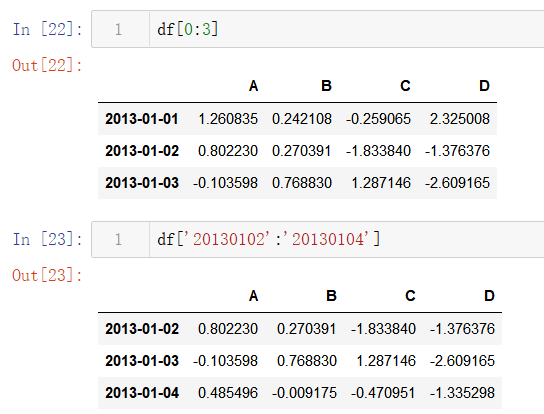


3.数据选择

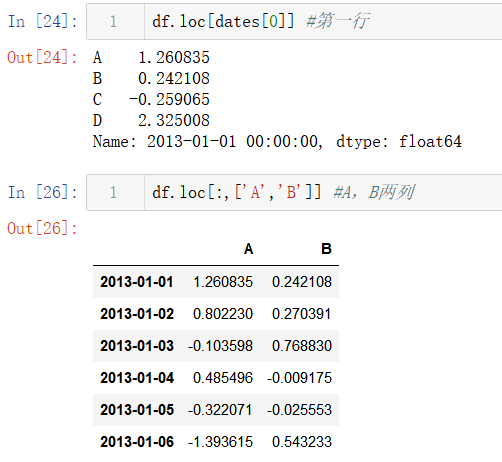
选择一列产生一个Series，相当于df.A.

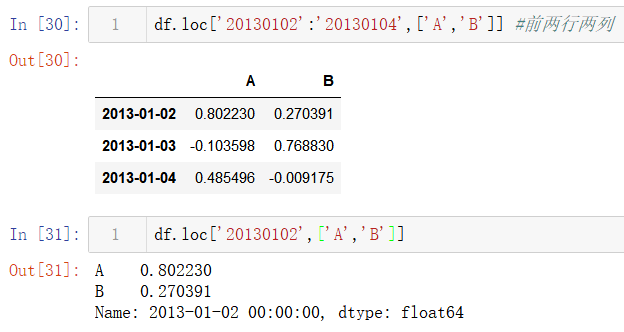


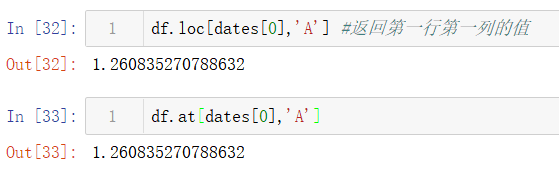
通过[]选择，对行进行切片。



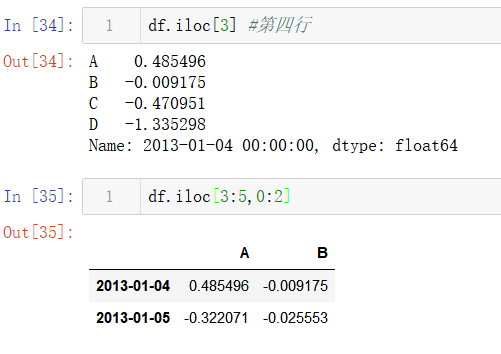
根据标签选择

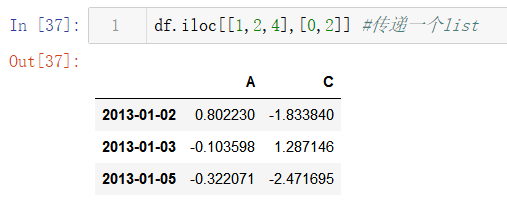


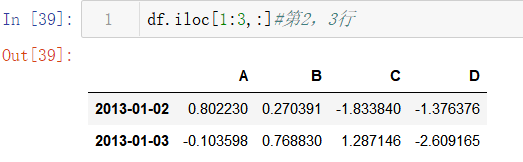


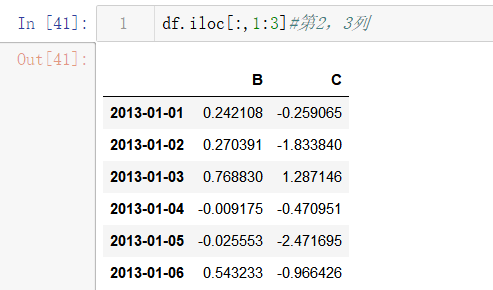


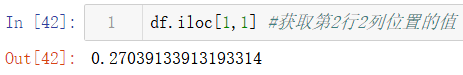
根据位置选择：



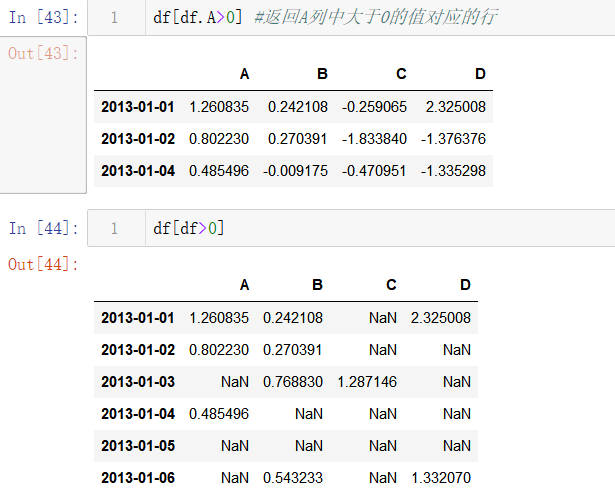




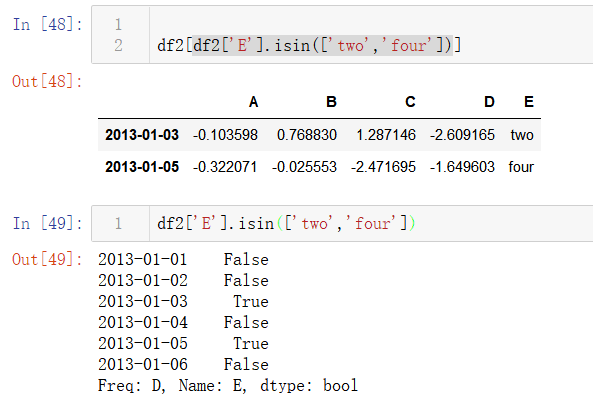




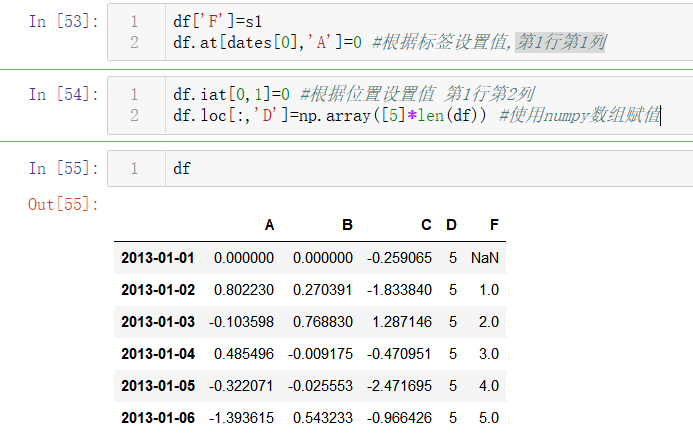
布尔类型索引



使用isin()方法过滤



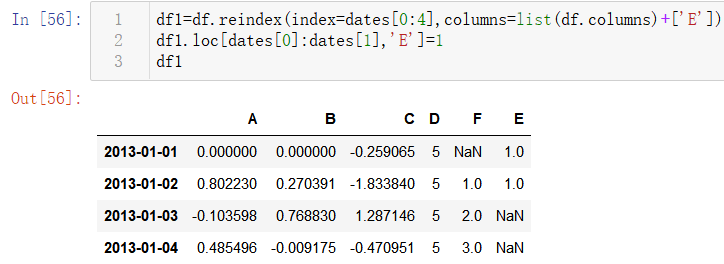
设置值



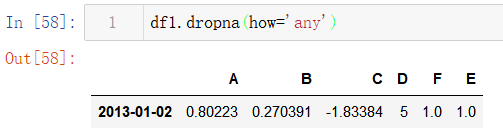
4.缺失值

pandas主要使用值np.nan来表示缺失的数据。 它默认不包含在计算中。

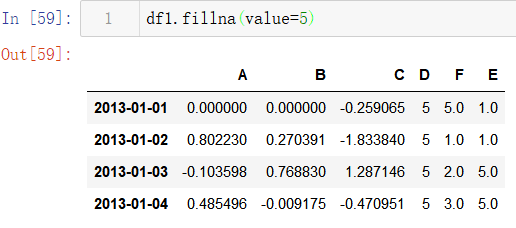
重建索引允许您更改/添加/删除指定轴上的索引。 这将返回数据的副本。



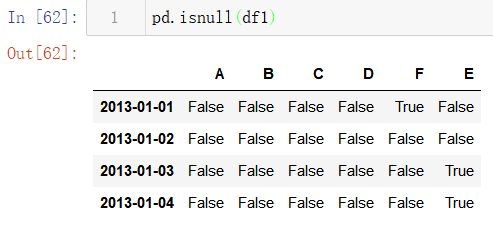
删除含有缺失值的行



填充缺失值



获取值为nan的布尔掩码。

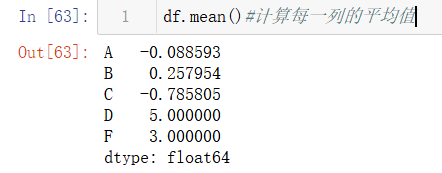


5.操作

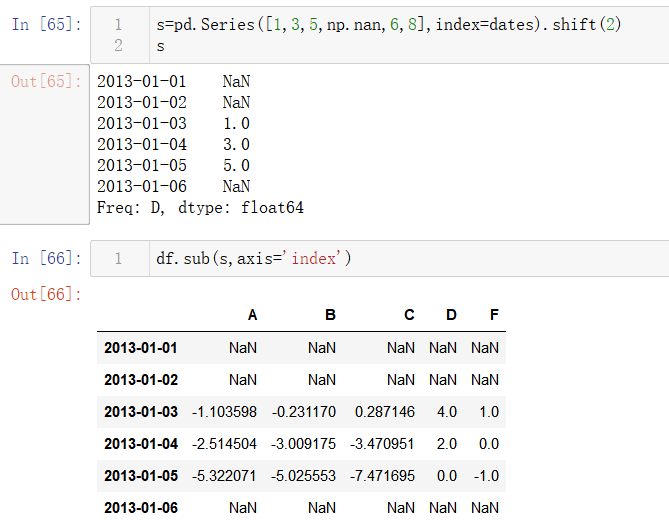
5.1 统计

操作通常排除丢失的数据。

执行描述性统计：

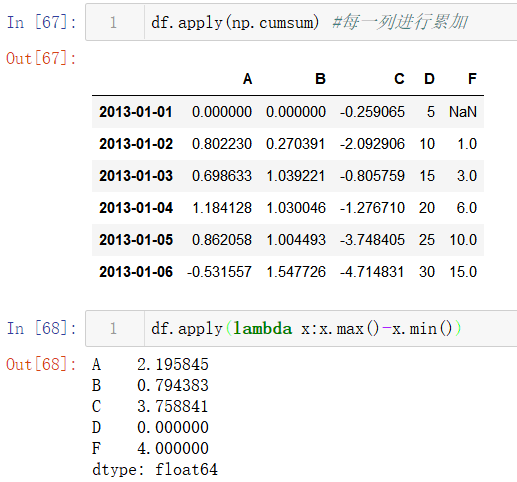


使用具有不同维度的对象进行操作并需要对齐。 此外，pandas会自动沿指定维度进行广播。

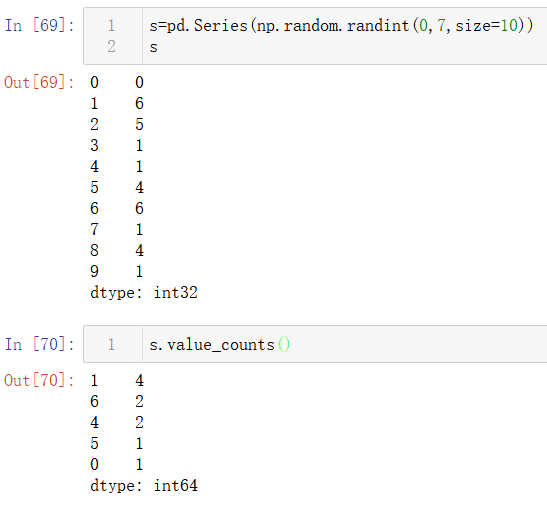


5.2 apply函数

将函数应用于数据：

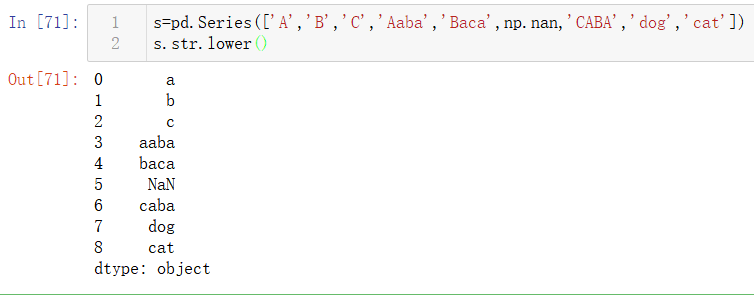


5.3 Histogramming(直方图化)



5.4 字符串方法

Series在str属性中配备了一组字符串处理方法，可以轻松地对数组的每个元素进行操作，如下面的代码片段所示。请注意，str中的模式匹配通常默认使用正则表达式（在某些情况下总是使用它们）。

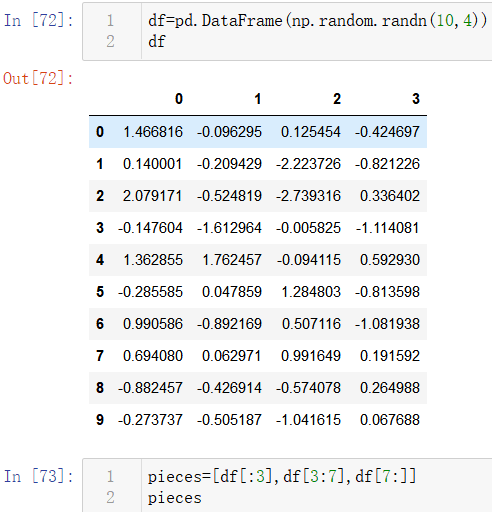


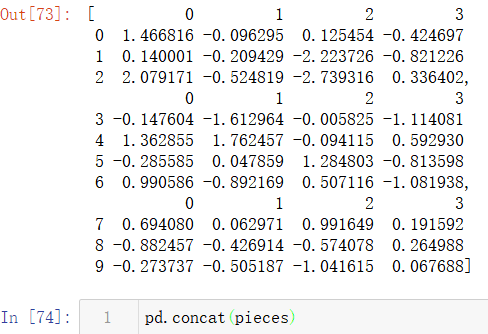
6.合并(merge)

6.1 concat

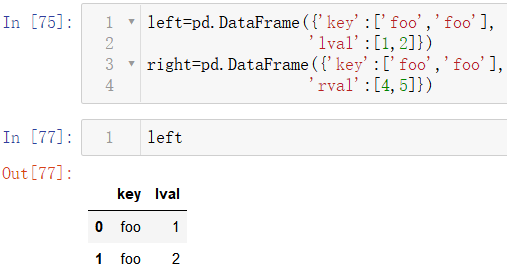
pandas提供了各种工具，可以在连接/合并类型操作的情况下，轻松地将Series，DataFrame和Panel对象与索引和关系代数功能的各种设置逻辑组合在一起。

使用concat()连接pandas对象：





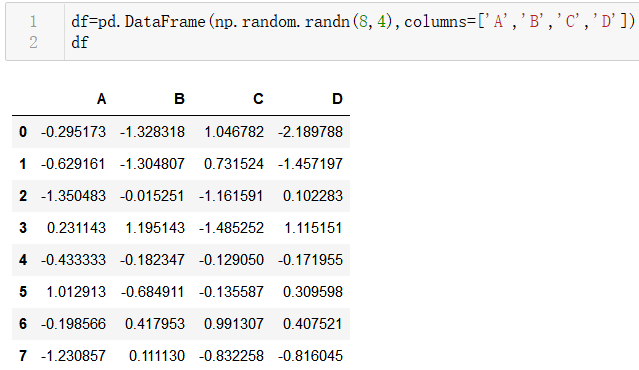
6.2 Join

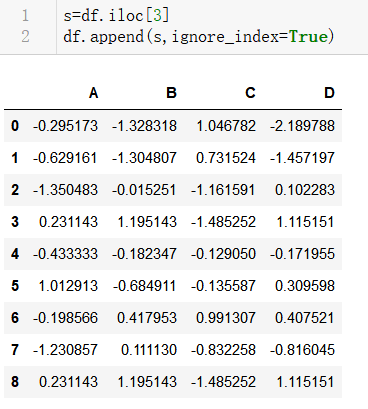




6.3 Append

添加行到dataframe

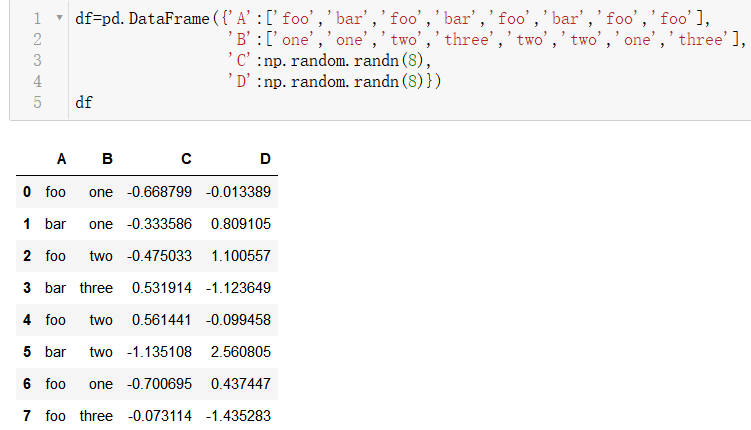


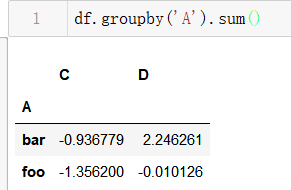


7.Grouping

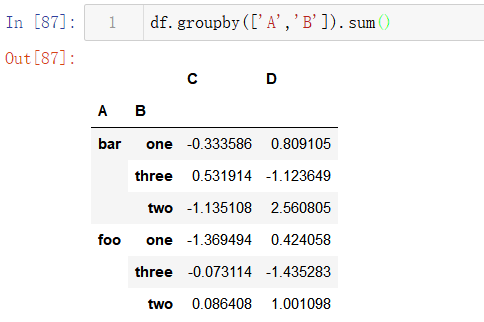
通过“group by”，我们指的是涉及以下一个或多个步骤的过程：

* 根据某些标准将数据拆分为组
* 将函数独立应用于每个组
* 将结果组合到数据结构中





按多列分组形成分层索引，我们再次应用sum函数。



8.Reshaping

8.1 Stack