

# 一：创建交叉表

## 1、crosstab的常用参数及其说明

交叉表是一种特殊的透视表，主要用于计算分组频率。利用pandas提供的crosstab函数可以制作交叉表，crosstab函数的常用参数和使用格式如下。

由于交叉表是透视表的一种，其参数基本保持一致，不同之处在于crosstab函数中的index，columns，values填入的都是对应的从Dataframe中取出的某一列。

```
pandas.crosstab(index, columns, values=None ,  
aggfunc=None, margins=False, dropna=True, normalize=False)
```

参数名称	说明
index	接收string或list。表示行索引键。无默认。
columns	接收string或list。表示列索引键。无默认。
values	接收array。表示聚合数据。默认为None。
aggfunc	接收function。表示聚合函数
dropna	接收boolean。表示是否删掉全为NaN的。默认为False。
margins	接收boolean。默认为True。汇总（Total）功能的开关，设为True后结果集中会出现名为“ALL”的行和列。
normalize	接收boolean。表示是否对值进行标准化。默认为False。

## 2、练习代码

我们想要直观的看到此样本数据中，按照性别分组后统计他们用手习惯的次数，如下图所示

	Sample	Gender	Handedness
0	1	Female	Right-handed
1	2	Male	Left-handed
2	3	Female	Right-handed
3	4	Male	Right-handed
4	5	Male	Left-handed
5	6	Male	Right-handed
6	7	Female	Right-handed
7	8	Female	Left-handed
8	9	Male	Right-handed
9	10	Female	Right-handed

Handedness	Left-handed	Right-handed	All
Gender			
Female	1	4	5
Male	2	3	5
All	3	7	10

```
import pandas as pd
import numpy as np
```

```
data = pd.DataFrame({'Sample': range(1, 11),
'Gender': ['Female', 'Male', 'Female', 'Male',
'Male', 'Male', 'Female', 'Female', 'Male',
'Female'],
                    'Handedness': ['Right-handed',
'Left-handed', 'Right-handed', 'Right-handed',
'Left-handed', 'Right-handed', 'Right-handed',
'Left-handed', 'Right-handed', 'Right-handed']})
print(data)
# 方法一：用pivot_table
data1=data.pivot_table(index=['Gender'],
                        columns='Handedness'
                        ,aggfunc=len,margins=True)

data1
# 方法二：用crosstab
# crosstab()的前两个参数可以是数组、Series或数组列表
data2=pd.crosstab(index=data['Gender'],columns=data[
'Handedness'],margins=True)
data2
```