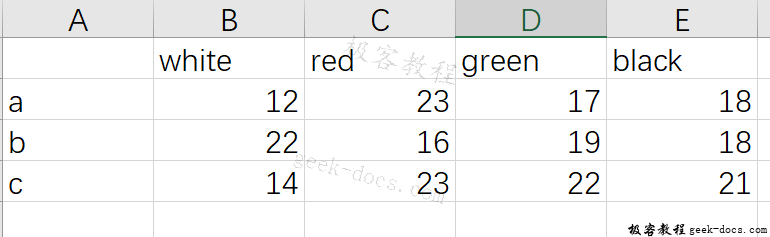
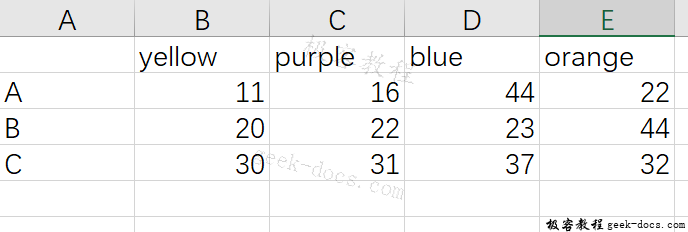
[Pandas 读写excel](https://geek-docs.com/pandas/pandas-read-write/pandas-to-read-and-write-excel.html)

**Pandas 读写excel**，除了CSV文件，使用Excel工作表存放列表形式的数据也很常见，Pandas定义了两个API函数来专门处理Excel文件：read\_excel()和to\_excel()。read\_excel()函数能够读取Excel 2003(.xls)和Excel 2007(.xlsx)两个类型的文件，该函数之所以能够读取Excel，是因为它整合了xlrd模块。

首先打开一个Excel文件，在sheet1和sheet2中输入数据，然后将文件保存为data.xlsx。  
  


**读取Excel文件**

要读取excel文件中的数据，并将其转换为DataFrame对象，只需要使用read\_excel()函数即可，如下所示:

import pandas as pd

frame = pd.read\_excel("data.xlsx")

print(frame)

输出结果如下:

white red green black

a 12 23 17 18

b 22 16 19 18

c 14 23 22 21

如上所见，读取excel时，默认返回的DataFrame对象包含第一个工作表中的数据。若要读取第二个工作表中的数据，需要用第二个参数指定工作表的名称或工作表的序号或索引，如下所示:

import pandas as pd

frame = pd.read\_excel("data.xlsx",'Sheet2')

print(frame)

输出结果如下:

yellow purple blue orange

A 11 16 44 22

B 20 22 23 44

C 30 31 37 32

也可以如下方式读取数据:

import pandas as pd

frame = pd.read\_excel("data.xlsx",1)

print(frame)

输出结果如下:

yellow purple blue orange

A 11 16 44 22

B 20 22 23 44

C 30 31 37 32

**写入Excel文件**

上述操作也适用于Excel写操作，将DataFrame对象转换为Excel，如下所示:

import pandas as pd

import numpy as np

frame = pd.DataFrame(np.random.random((4,4)),

index=['exp1','exp2','exp3','exp4'],

columns=['jan2015','Fab2015','Mar2015','Apr2005'])

print(frame)

frame.to\_excel("data2.xlsx")

输出结果如下:

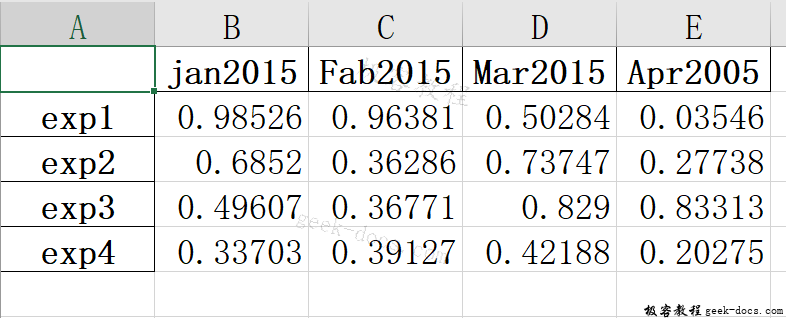
jan2015 Fab2015 Mar2015 Apr2005

exp1 0.985263 0.963805 0.502845 0.035461

exp2 0.685198 0.362865 0.737473 0.277377

exp3 0.496066 0.367708 0.828998 0.833133

exp4 0.337033 0.391268 0.421885 0.202747

工作目录中会生成一个包含数据的新Excel文件，存放数据如下所示:  


**操作Excel行列**

1. 读取指定的单行，数据会存在列表里面

import pandas as pd

df=pd.read\_excel('data.xlsx') #这个会直接默认读取到这个Excel的第一个表单

data=df.ix[0].values #0表示第一行 这里读取数据并不包含表头，要注意哦！

print(data)

输出结果如下所示:

[12 23 17 18]

1. 读取指定的多行，数据会存在嵌套的列表里面

import pandas as pd

df=pd.read\_excel('data.xlsx') #这个会直接默认读取到这个Excel的第一个表单

data=df.ix[[1,2]].values #读取指定多行的话，就要在ix[]里面嵌套列表指定行数

print(data)

输出结果如下所示:

[[22 16 19 18]

[14 23 22 21]]

1. 读取指定的行列

import pandas as pd

df=pd.read\_excel('data.xlsx') #这个会直接默认读取到这个Excel的第一个表单

data=df.ix[1,2] #读取第一行第二列的值，这里不需要嵌套列表

print(data)

输出结果如下所示:

19

1. 读取指定的多行多列值

import pandas as pd

df=pd.read\_excel('data.xlsx') #这个会直接默认读取到这个Excel的第一个表单

data=df.ix[[1,2],['red','green']].values#读取第一行第二行的title以及data列的值，这里需要嵌套列表

print(data)

输出结果如下所示:

[[16 19]

[23 22]]

1. 获取所有行的指定列

import pandas as pd

df=pd.read\_excel('data.xlsx') #这个会直接默认读取到这个Excel的第一个表单

data=df.ix[:,['red','green']].values#读取第一行第二行的title以及data列的值，这里需要嵌套列表

print(data)

输出结果如下所示:

[[23 17]

[16 19]

[23 22]]

1. 获取行号并打印输出

import pandas as pd

df=pd.read\_excel('data.xlsx') #这个会直接默认读取到这个Excel的第一个表单

print(df.index.values)

输出结果如下所示:

['a' 'b' 'c']

1. 获取列名并打印输出

import pandas as pd

df=pd.read\_excel('data.xlsx') #这个会直接默认读取到这个Excel的第一个表单

print(df.columns.values)

输出结果如下所示:

['white' 'red' 'green' 'black']

1. 获取指定行数的值

import pandas as pd

df=pd.read\_excel('data.xlsx') #这个会直接默认读取到这个Excel的第一个表单

print(df.sample(3).values)#这个方法类似于head()方法以及df.values方法

输出结果如下所示:

[[12 23 17 18]

[22 16 19 18]

[14 23 22 21]]

1. 获取指定列的值

import pandas as pd

df=pd.read\_excel('data.xlsx') #这个会直接默认读取到这个Excel的第一个表单

print(df['white'].values)

输出结果如下所示:

[12 22 14]