

玩转数据处理120题重制发布与PDF版本下载

刘早起早起 早起Python 今天

点击上方“早起Python”，关注并星标公众号

和我一起玩Python

大家好，在十天前发布Pandas进阶修炼120题后，收到来自微信、CSDN、知乎、GitHub等热爱数据分析的读者的关注，并且收到了很多鼓励与建议，因此我再次对该专题进行整理，修改了一些错误(有些错误将错就错给出新的回答)，同时联合中山大学博士@陈熹制作了R语言版本今天再次发布，在这个全新的版本中**不仅每一题同时给出了Pandas与R两种解法，同时对部分习题给出了多种解法与注解**👉

2

数据提取

题目：提取含有字符串"Python"的行

难度：☆☆

期望结果

```
grammar  score
0 Python    1.0
7 Python   10.0
```

Python解法：

```
#> 1
df[df['grammar'] == 'Python']
#> 2
results = df['grammar'].str.contains("Python")
results.fillna(value=False,inplace = True)
df[results]
```

R语言解法

```
df[which(df$grammar == 'Python'),]
```

题目：将DataFrame保存为EXCEL

难度：☆☆

Python解法

```
df.to_excel('filename.xlsx')
```

R解法

```
#R对EXCEL文件不太友好  
#第一种方法：利用readr包转为csv再用EXCEL打开  
#文件本质依然是csv  
library(readr)  
write_excel_csv(df, 'filename.csv')  
  
#第二种方法：利用openxlsx包  
openxlsx::write.xlsx(df, 'filename.xlsx')  
  
#也可以用xlsx包，但需要先配置JAVA环境  
#确保JAVA配置到环境变量中并命名为JAVA_HOME  
Sys.getenv("JAVA_HOME")  
install.packages('rJava')  
install.packages("xlsxjars")  
library(rJava)  
library(xlsxjars)  
xlsx::write.xlsx(df, 'filename.xlsx')
```

相信经过升级后的120题能让各位读者更好的学习相关操作，同时还有B站视频讲解，现在我已经将该版本整理为综合版、Pandas单独版、R语言单独版发布在公众号，当然也将PDF版本上传至后台，**回复pandas**即可获取下载链接，不过建议还是在公众号内阅读会有最好的体验。最后来听听作者怎么说👉

@**刘早起**(*Pandas进阶修炼120题*): 『**Pandas进阶修炼120题**』系列现已完结，我们对Pandas中常用的操作以习题的形式发布。从读取数据到高级操作全部包含，希望可以**通过刷题的方式来完整学习pandas中数据处理的各种方法**，当然如果你是高手，也欢迎尝试给出与答案不同的解法。

@**陈熹**(*R语言数据处理120题*): R语言和Pandas都是数据处理的重要工具，而二者的高下争论时有发生，我相信对于数据而言没有绝对的孰优孰劣，需要做的应该是在必要时权衡最合适的办法。我将用R语言实现Pandas进阶修炼120题来帮助我们**比较了两种语言处理不同问题的共性**。当然也发现了各自的灵活和缺陷，它们覆盖多数数据分析初期

可能遇到的问题，无论是对R语言还是对Python技能的提升相信都有很大帮助。

@井钰([Pandas系列B站视频讲解](#)):Pandas是一种实用的数据处理工具，相比较理论，在实践中才能更好的掌握它。幸运的是[pandas进阶修炼120题]提供这样一个实际操作的机会，当你做完这120题后，你已经对pandas中常见的操作有所了解，**能够运用这个工具去解决一些实际问题**。当然，由于能力有限，视频制作的不够完善，欢迎大家批评指正，我们一起共同进步。

好了，以上就是Pandas进阶修炼120升级为玩转数据处理120题的说明，整个系列均由我校对，整理下来共近5万字符，如果在使用过程中发现任何错误可以联系我！另外NumPy与Matplotlib专题正在更新中，欢迎关注与转发，谢谢！

