|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Python程序设计作业评分表  作业：验证字符串的作用；测试+、join()；去除html标签 | | | | | | |
| 组名：正好一桌麻将队 | | 得分： |  |  |  | |
| 学号 | 班级 | 姓名 | 承担内容（答辩人在姓名后加\*） | | 比例 | 得分 |
| 2017339960040 | 17电信实验班 | 刘颜铭 | 验证字符串方法的作用；测试+、join()；去除html标签 | |  |  |
| 2017339960012 | 17电信实验班 | 张文韬 | 验证字符串方法的作用；测试+、join()；去除html标签 | |  |  |
| 2017339960025 | 17计科(3)班 | 鲍晨翔 |  | |  |  |
| 2017339960024 | 17计科(2)班 | 王旭 |  | |  |  |
| 打分项 | 说明 | | | | 分值 | 得分 |
| 程序质量（A） | 从以下几个方面综合评价： 1.代码量 2.题目难度 3.使用的技术 4.最终实现的效果 5.团队的协作性 | | | | 50% |  |
| 报告质量（B） | 从以下几个方面综合评价（A项不合格，本项没分）： 1.文本的格式 2.文本的语言 3.文本的结构 4.文本的内容 | | | | 30% |  |
| 评述的质量（C） | 从以下几个方面综合评价（A项不合格，本项没分）： 1.对工作的熟悉程度 2.对题目的理解 3.对技术的掌握 4.对工作成果的分析和评价 | | | | 20% |  |

作业/实验报告附加信息表

|  |  |
| --- | --- |
| 作业/实验名称 | 验证字符串的作用；测试+、join()；去除html标签 |
| 分工情况(较详细的分工，如果单人完成，可空)： | 张文韬：验证字符串的作用；测试+、join()；去除html标签  刘颜铭：验证字符串的作用；测试+、join()；去除html标签  鲍晨翔：  王旭： |
| 关键技术(本次使用到的关键技术和总结) | 张文韬：第二题生成了大量的数据用于测试用+快，还是使用join快；第三题主要用正则表达式<[^<]+>匹配html标签。  刘颜铭：第1题依据等价类的测试思想设定了不同种类的值进行测试，通过对True和False值的打印结果来验证；第2题定义了一个函数random\_str（）来生成指定长度的字符数组，然后分别对不同数量级的同一数组进行运算，根据时间来比较处理速度；第3题用正则表达式<[^<]+>匹配html标签。  鲍晨翔：  王旭： |
| 参考文献(查阅的资料、网站和参考文献) | 张文韬：  https://blog.csdn.net/qq\_42209411/article/details/8315225  刘颜铭：  http://outofmemory.cn/code-snippet/12195/usage-python-produce-holding-length-random-charaeter  鲍晨翔：  王旭： |
| 遇到的问题和解决方法 | 张文韬：对正则表达式的使用有些不熟悉，仔细学习了上课的课件后初步了解了正则表达的规则。  刘颜铭：刚开始想到通过处理不同数量级字符数组的时间来比较两者的运算速度，但是不知道如何生成指定长度如此大的字符数组，通过查阅资料才写出了random\_str（）函数。  鲍晨翔：  王旭： |
| 自我评价(对工作过程、结果和团队协作的评价) | 张文韬：小组成员在完成作业过程中，查阅了相关资料，对查到的资料消化学习后形成了自己的方法，完成了本次作业。在此过程中，组内每位成员都积极参与，尽到了一个组员的义务。  刘颜铭：通过查阅网上的资源学会了很多简便方法，很实用。  鲍晨翔：  王旭： |
| 改进策略(针对前面的评价，打算在下一次实验中如何改进) | 张文韬：匹配html标签可以有很多种表达方法，可以尝试使用其他的。  刘颜铭：可以使用等价类、边界值、因果图、正交排列等测试思想来进行比较。  鲍晨翔：  王旭： |

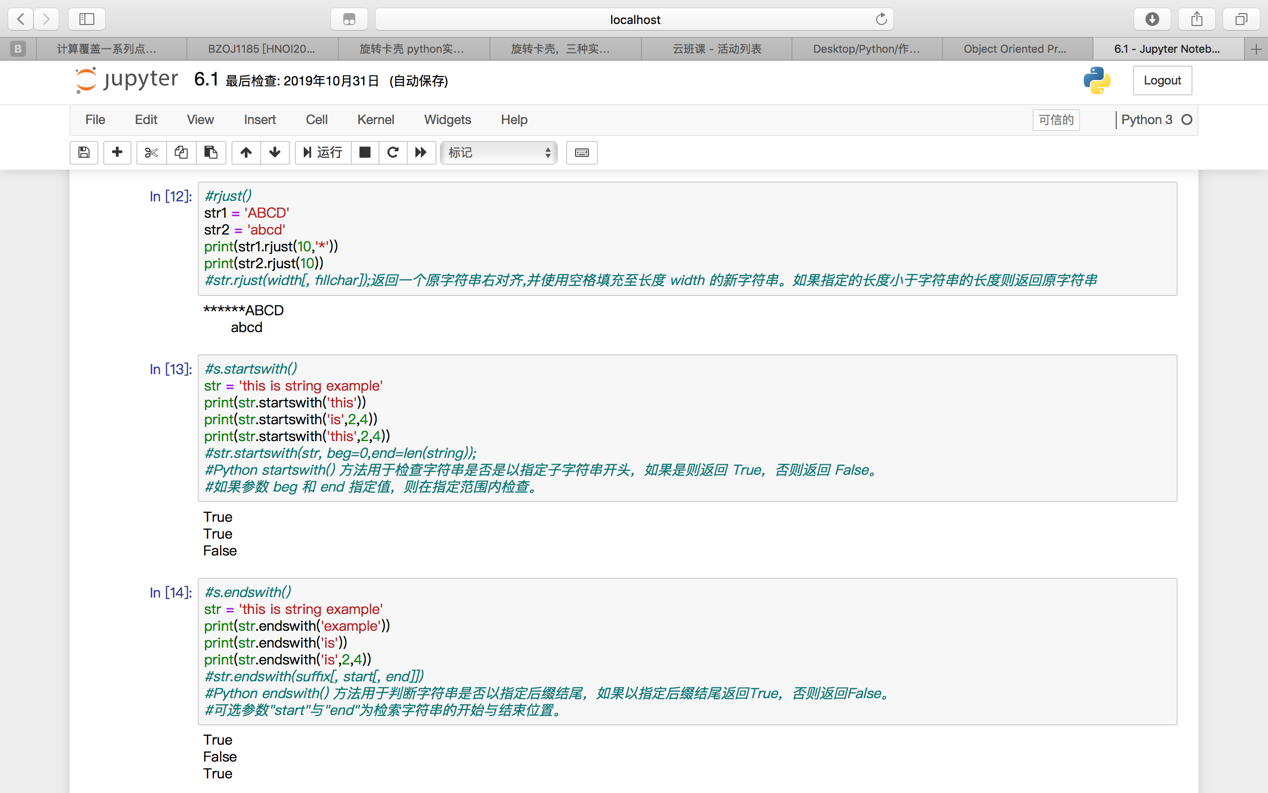
#### 实验/作业 内容：

**题1、验证这些字符串方法的作用**

**isalnum()、isalpha()、isdigit()、isdecimal()、isnumeric()、isspace()、isupper()、islower()**

**zfill()、center()、ljust()、rjust()、s.startswith()、s.endswith()**

-----------------------张文韬\_2017339960012------------------------



具体的验证过程详见附件

-----------------------刘颜铭\_2017339960040------------------------



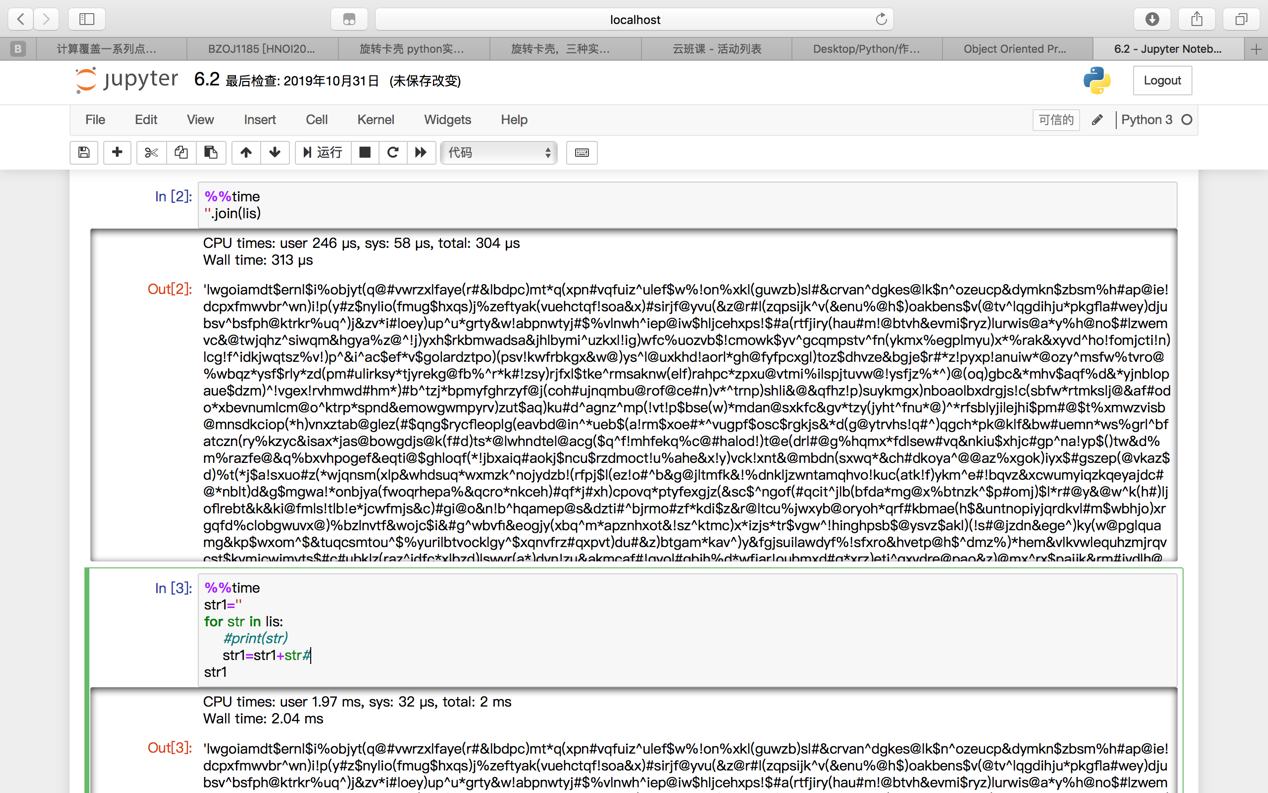
根据各个方法之间能够识别的字符种类的不同，将这几个方法分成了六大类，并在每一类中针对细节问题依据等价类的测试思想设定了不同种类的值进行测试，通过对True和False值的打印结果来验证。

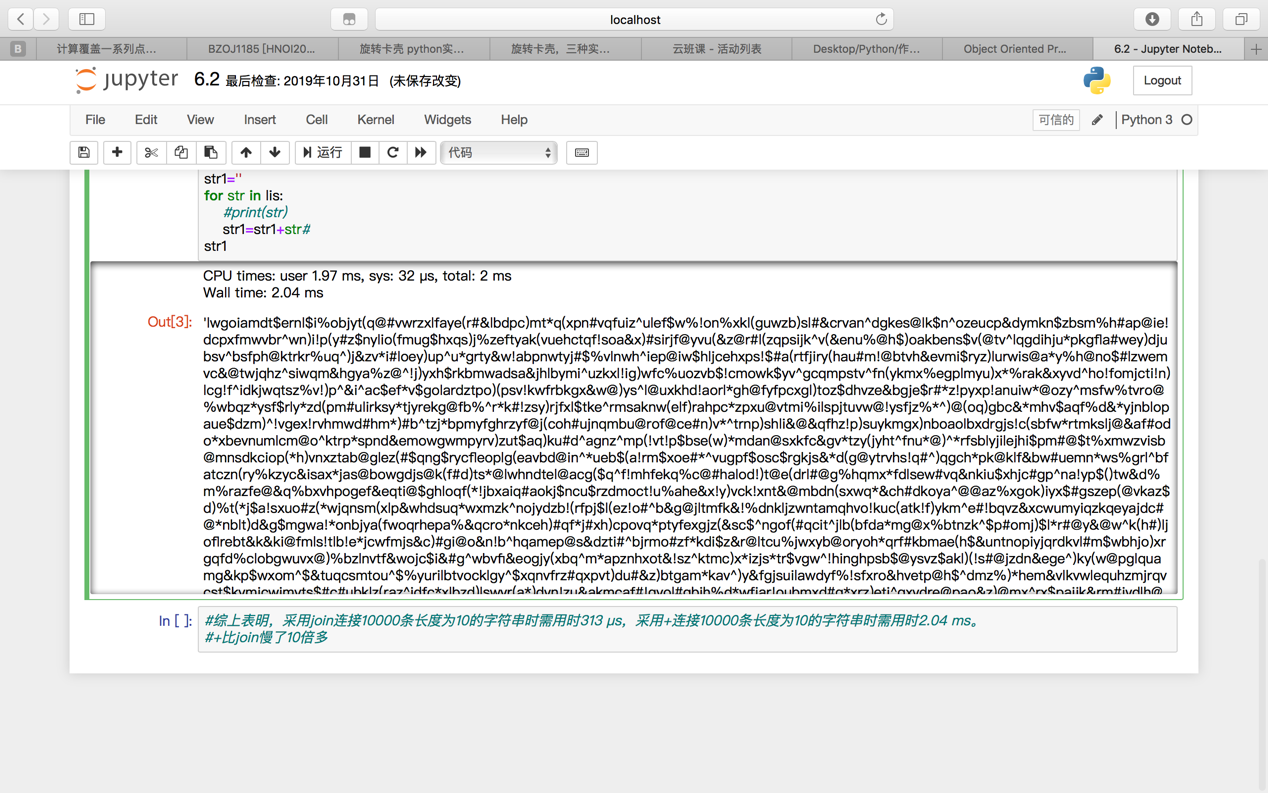
-----------------------鲍晨翔\_2017339960025------------------------

-----------------------王旭\_2017339960024------------------------

**题2、测试字符串连接是用+快，还是使用join快。**

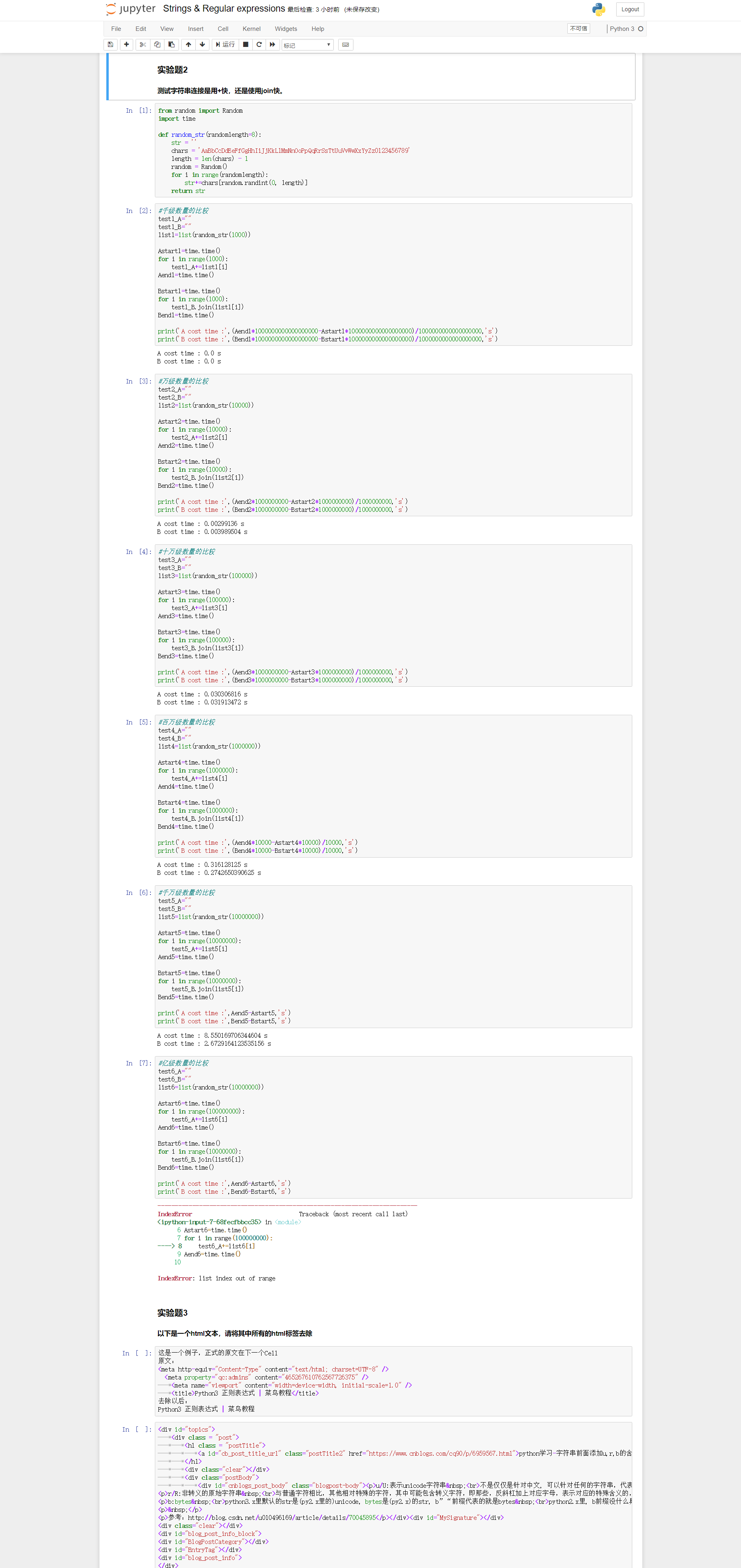
-----------------------张文韬\_2017339960012------------------------





具体的验证过程详见附件

-----------------------刘颜铭\_2017339960040------------------------



由于是要比较两种方法的运算速度，为了得到更具有普适性的结论，我将两种方法处理的数据量设为变量，进行了6个递增数量级的测试。依据30次运行结果的数据统计显示，当处理的数据量在“千”、“万”、“十万”级时，+方法的运算速度始终快于join()方法，但是两者之间处理单位数据的速度差距在不断缩小，并于“百万”级数据量时join()方法完成了反超，处理速度快于+方法，在此之后两者处理速度的差距越来越大。

另外为了完成不同数据量的测试，我编写了一个random\_str（）方法用于生成指定个数的字符数组。

-----------------------鲍晨翔\_2017339960025------------------------

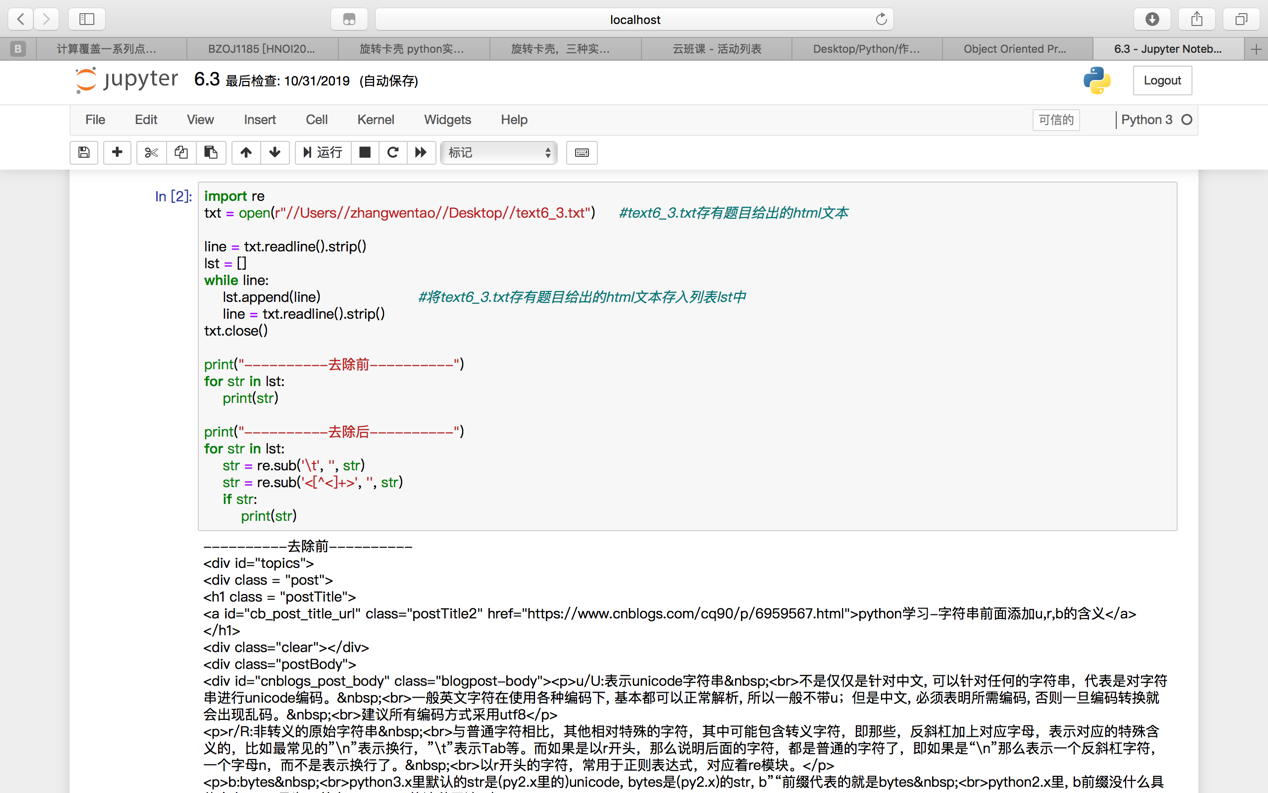
-----------------------王旭\_2017339960024------------------------

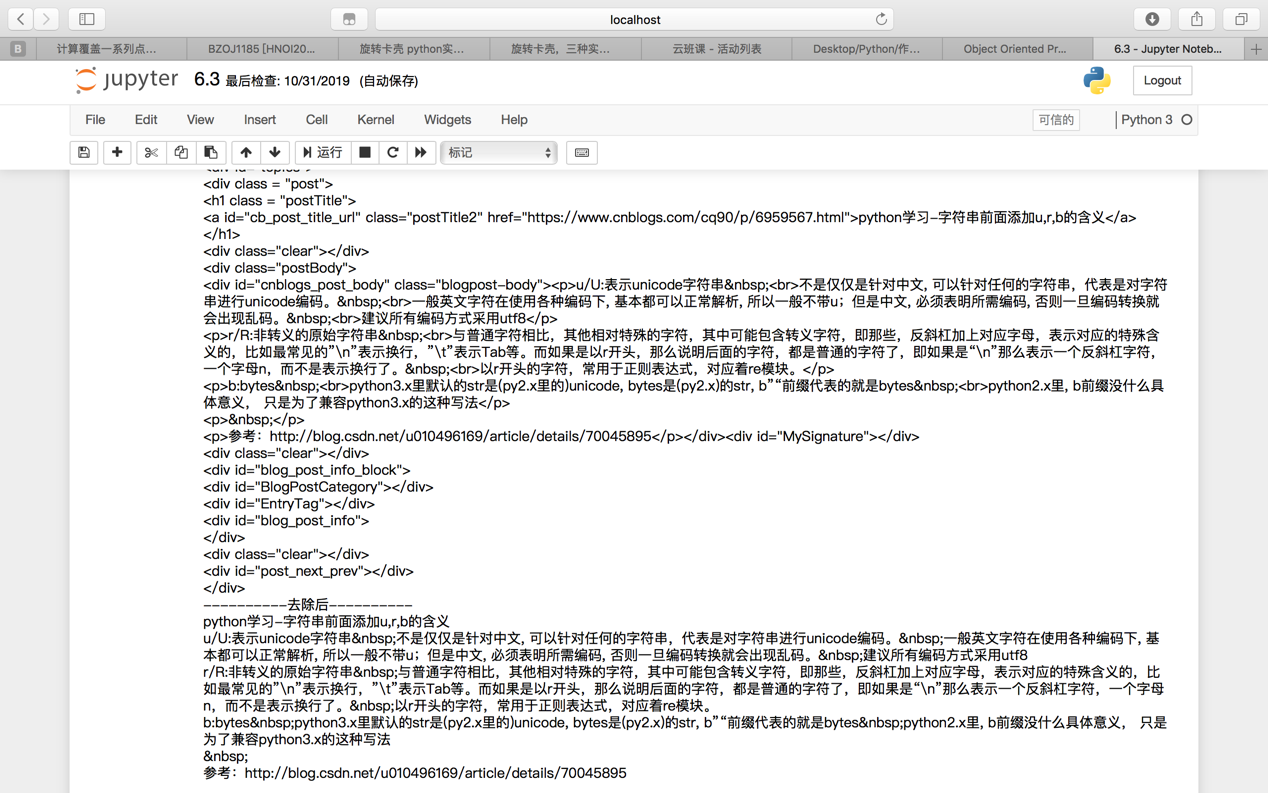
**题3、以下是一个html文本，请将其中所有的html标签去除**

**这是一个例子，正式的原文在下一个Cell 原文：**

**去除以后： Python3 正则表达式 | 菜鸟教程**

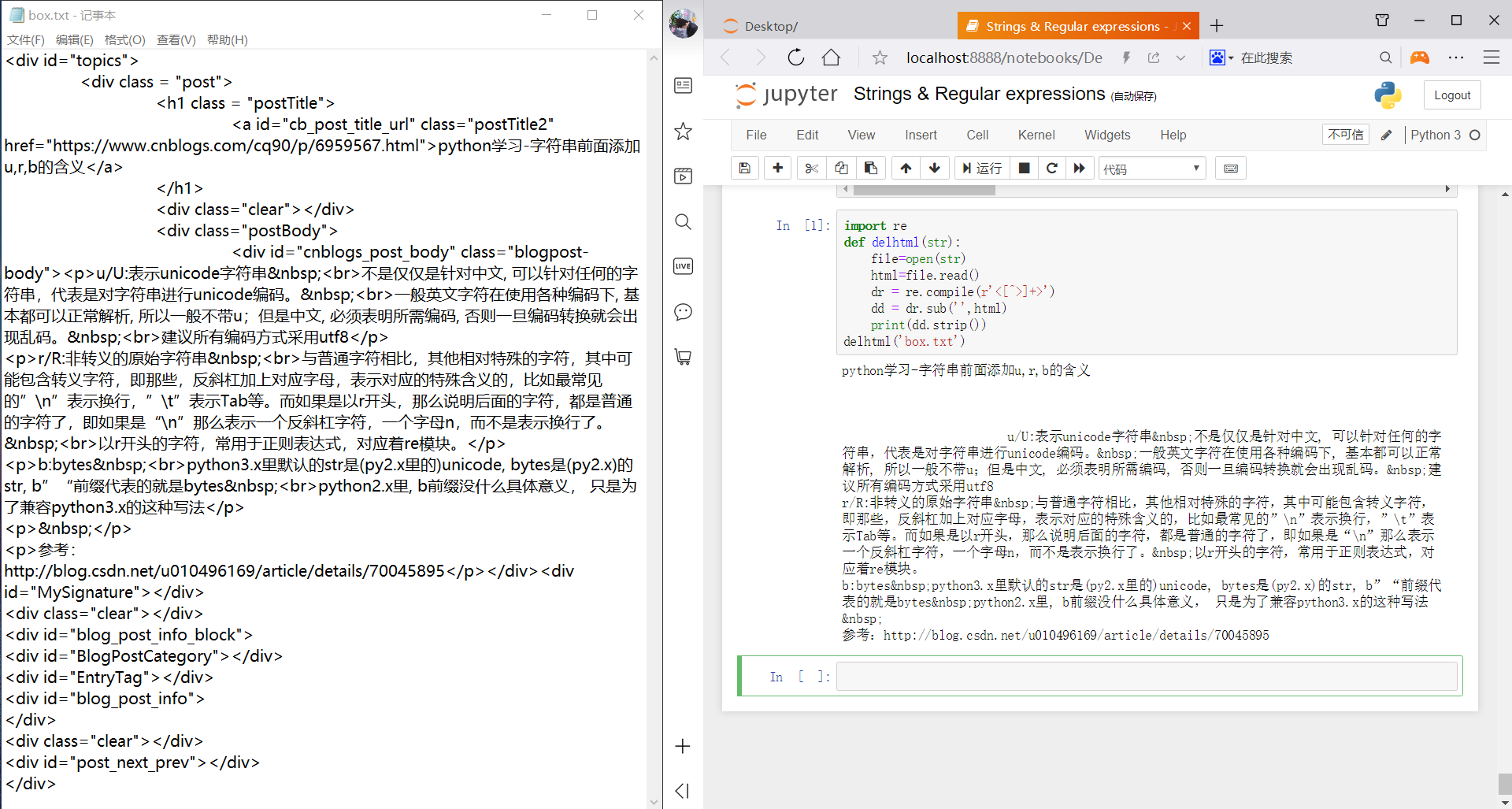
-----------------------张文韬\_2017339960012------------------------





具体代码实现详见附件

-----------------------刘颜铭\_2017339960040------------------------



-----------------------鲍晨翔\_2017339960025------------------------

-----------------------王旭\_2017339960024------------------------