|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Python程序设计作业评分表  作业：去除灌水评论；统计出现最多的英文单词；作者鉴定 | | | | | | |
| 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 | 作者鉴定 | | 25% |  |
| 2017339960012 | 17电信实验班 | 张文韬 | 去除灌水评论；统计出现最多的英文单词；作者鉴定 | | 25% |  |
| 2017339960025 | 17计科(3)班 | 鲍晨翔 | 统计出现最多的英文单词 | | 25% |  |
| 2017339960024 | 17计科(2)班 | 王旭 | 去除灌水评论 | | 25% |  |
| 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% |  |

**实验/作业目的：**

建立Python环境  
掌握在Jupyter下进行Python初步编程的知识和能力  
掌握列表，元组，集合，字典的概念

#### 实验/作业 内容：

**题1. 以下是一系列书评,但是很多是灌水的,请你写一段代码,把灌水的书评尽可能去掉。灌水书评一般会有一个特点：重复的字比较多，利用这个特点，把灌水书评去掉。**

|  |
| --- |
| 代码详见附件  张文韬\_2017339960012 |
| comments=['这是一本非常好的书','作者用心了','作者大大辛苦了','好书，感谢作者提供了这么多的好案例','书在运输的路上破损了，我好悲伤。。。','为啥我买的书上有菜汤。。。。','啊啊啊啊啊啊，我怎么才发现这么好的书啊，相见恨晚','书的质量有问题啊，怎么会开胶呢？？？？？？','好好好好好好好好好好好','好难啊看不懂好难啊看不懂好难啊看不懂','书的内容很充实','你的书上好多代码啊，不过想想也是，编程的书嘛，肯定代码多一些','书很不错!!一级棒!!买书就上当当，正版，价格又实惠，让人放心!!!','无意中来到你小铺就淘到心意的宝贝，心情不错!','送给朋友的、很不错','这是一本好书，讲解内容深入浅出又清晰明了，推荐给所有喜欢阅读的朋友同好们。','好好好，下一个看看','渣渣渣渣渣渣渣渣','没用没用没用没用','很好的书，五颗星']  rule=lambda s:len(set(s))/len(s)>0.5  result=filter(rule,comments)  print("原始书评：")  for comment in comments:  print (comment),    print("="\*30)  print ("过滤后的书评：")  for comment in result :  print(comment)  王旭\_2017339960024 |
| 图 1刘颜铭\_2017339960040  【源码】：  comments=['这是一本非常好的书,作者用心了','作者大大辛苦了','好书，感谢作者提供了这么多的好案例','书在运输的路上破损了，我好悲伤。。。','为啥我买的书上有菜汤。。。。','啊啊啊啊啊啊，我怎么才发现这么好的书啊，相见恨晚','书的质量有问题啊，怎么会开胶呢？？？？？？','好好好好好好好好好好好','好难啊看不懂好难啊看不懂好难啊看不懂','书的内容很充实','你的书上好多代码啊，不过想想也是，编程的书嘛，肯定代码多一些','书很不错!!一级棒!!买书就上当当，正版，价格又实惠，让人放心!!!','无意中来到你小铺就淘到心意的宝贝，心情不错!','送给朋友的、很不错','这是一本好书，讲解内容深入浅出又清晰明了，推荐给所有喜欢阅读的朋友同好们。','好好好，下一个看看','渣渣渣渣渣渣渣渣','没用没用没用没用','很好的书，五颗星']  for comment in comments:  new\_comment=''  a=list(comment)  b=a.copy()  for i in range(len(a)):  if(b.count(a[i])>1 and a.count(a[i])>2 ):  b.remove(a[i])  print(new\_comment.join(b)) |

**题2. 一段文字，统计出现最多的英文单词。**

|  |
| --- |
| 代码详见附件  张文韬\_2017339960012 |
| table = str.maketrans(',.!"?:[]-',' ')  f = open(r'C:\Users\lenovo\Desktop\样本.txt')  file1 = f.read()  f.close()  f1 = file1.translate(table).lower()  wordlist = f1.split(None)  d1 = {}  for word in wordlist:  d1[word] = d1.get(word,0) +1  itemli =sorted(d1.items(), key=lambda x:x[1],reverse=True)  print (itemli[0:3])    鲍晨翔\_2017339960025 |
|  |

**题3. 作者鉴定。一直有人认为《水浒传》的作者不是施耐庵，而是罗贯中。现在就请你对《水浒传》来做一个判定，看看和《三国演义》到底风格是否相近。鉴定方法由你自己来设计，当然，必须用Python来实现哦。看看最后你的结论是什么？**

|  |
| --- |
| 代码详见附件  张文韬\_2017339960012 |
|  |
|  |