**《Python程序设计》第一次作业（基本数据类型）**

1. 某百科上“Python”词条的部分内容可以字符数形式存储为：

text = **"Python is a widely used general purpose, high level programming language. It was created by Guido van Rossum in 1991 and further developed by the Python Software Foundation. It was designed with an emphasis on code readability, and it allows programmers to express their concepts in fewer lines of code.\nPython is a programming language that lets you work quickly and integrate systems more efficiently.\nThere are two major Python versions: Python 2 and Python 3. Both are quite different.\nBefore we start Python programming, we need to have an interpreter to interpret and run our programs."**

其中不同段落间用换行符（**\n**）分隔。请编写Python程序完成以下任务：

1. 计算该文本由几个段落构成。
2. 提取出第一段的内容。
3. 统计第一段中包括哪些单词（不重复）。
4. 计算第一段中Python出现的次数，并利用字符串格式化操作打印输出如下：

**第一段中Python出现了x次**

其中**x**为计算所得的Python出现次数。

答案：

*# a.*1）num\_para = len(text.split(**"\n"**)) *# 方法1*2）num\_para = text.count(**"\n"**) + 1 *# 方法2  
  
# b.*1）para\_1 = text.split(**"\n"**)[0] *# 方法1*2）para\_1 = text[:text.find(**"\n"**)] *# 方法2  
  
# c.*1）all\_words = para\_1.lower().replace(**"."**,**""**).replace(**","**,**""**).split(**" "**) unique\_words = list(set(all\_words)) *# 方法1*

2）p1w=set(filter(None,re.split("[\n |\.|\,]",para\_1.lower()))) *# 方法2*

3）first\_paragraph\_cleaned = ''.join(c for c in para\_1 if c.isalpha() or c.isspace()).lower() *# 方法3（此方法有一点点小问题，课上讨论）*

words = first\_paragraph\_cleaned.split()

4）words=text1.split()

dict\_words=[]

for i in words:

if i not in dict\_words:

dict\_words.append(i) *# 方法*

*# d.*

1）freq = all\_words.count(**"Python"**.lower())*# 方法1*

print(**"第一段中Python出现了%d次"** % freq)

2）x = all\_words.count("Python")*# 方法2*

print(f"第一段Python出现了{x}次。")

3）H=all\_words.count('Python') *# 方法3*

I='第一段中Pyhon出现了%d次'

print(I % H)

4)fre = 0

for i in paragraph\_1:

if i == 'Python':

fre += 1

print('第一段中Python出现了{}次'.format(fre)) *# 方法4*

1. 某外卖平台部分饮品类商家的信息（包括名称、评分、配送费）如下所示：

* 名称：瑞幸咖啡，评分：4.9，配送费：3元
* 名称：CoCo都可，评分：4.8，配送费：0元
* 名称：快乐柠檬，评分：4.9，配送费：0元

请编写Python程序完成以下任务：

1. 创建一个字典info，用于存储以上商家信息。
2. 将以下热销饮品信息添加到info中。

* 瑞幸咖啡：生椰拿铁（289）、厚乳拿铁（173）、拿铁（126）
* CoCo都可：鲜百香双响炮（616）、红豆奶茶（493）
* 快乐柠檬：柠檬菠萝饮（194）

1. 将“瑞幸咖啡”的配送费改为6元。
2. 删除“快乐柠檬”的信息。
3. 对“瑞信咖啡”热销饮品按月销由低到高排序。

答案：

*# a.*info = {**"瑞幸咖啡"**:{**"评分"**:4.9,**"配送费"**:3},**"CoCo都可"**:{**"评分"**:4.8,**"配送费"**:0},**"快乐柠檬"**:{**"评分"**:4.9,**"配送费"**:0}}  
  
*# b.*1）info[**"瑞幸咖啡"**][**"热销饮品"**] = {**"生椰拿铁"**:289,**"厚乳拿铁"**:173,**"拿铁"**:126} *# 方法1*  
info[**"CoCo都可"**][**"热销饮品"**] = {**"鲜百香双响炮"**:616,**"红豆奶茶"**:493}  
info[**"快乐柠檬"**][**"热销饮品"**] = {**"柠檬菠萝饮"**:194}  
2）x={'热销饮品': {"生椰拿铁":289,"厚乳拿铁":173,"拿铁":126}} *# 方法2*

info['瑞幸咖啡'].update(x)

*# c.*info[**"瑞幸咖啡"**][**"配送费"**] = 6  
  
*# d.***1）del** info[**"快乐柠檬"**] *# 方法1*  
2）info.pop("快乐柠檬")*# 方法2*

*# e.*sale\_sort = sorted(info[**"瑞幸咖啡"**][**"热销饮品"**].items(),key=**lambda** x:x[1])

1. 某网页中呈现的“Python常用第三方库”列表可用如下字符串表示：

lib = **"<ul><li>matplotlib</li><li>numpy</li><li>pandas</li><li>bs4</li><li>requests</li><li>selenium</li></ul>"**

其中**<ul>**和**</ul>**分别是关于列表开始和结束的标记，**<li>**和**</li>**之间是列表中每一项的内容。请编写Python程序完成以下任务：

1. 将以上字符串格式的列表转为列表格式的变量lib\_list，即：

lib\_list = [**'matplotlib'**, **'numpy'**, **'pandas'**, **'bs4'**, **'requests'**, **'selenium'**]

1. 将scipy和jieba插入requests之后。
2. 删除bs4。
3. 按字符长度对lib\_list中各项进行降序排序。

答案：

*# a.*1）lib\_list = lib.strip(**"<ul><li></li></ul>"**).split(**"</li><li>"**) *# 方法1*

2）A=lib.replace('<ul><li>','</li><li>') *# 方法2*

B=A.replace('</li></ul>','</li><li>')

C=B.split('</li><li>')

C.remove('')

C.remove('')

3）a = lib.strip('<ul>i/') *# 方法3*

b = a.replace('</li>'," ")

c = b.replace('<li>'," ")

lib\_list = c.split(' ')

4）lib\_list = re.findall(r"<li>(.\*?)</li>", lib) *# 方法4*

5）[i.strip("</li>") for i in lib.split("<li>")[1:-1]] *# 方法5*

*# b.*1）i = lib\_list.index(**"requests"**)*# 方法1*  
lib\_list[i+1:i+1] = [**"scipy"**,**"jieba"**]

2）insert(i+1,"scipy")*# 方法2*

3）lib\_list.insert(-1,'scipy') *# 方法3*

lib\_list.insert(-1,'jieba')

4）lib\_list.insert(lib\_list.index('requests')+1,'spcipy') *# 方法4*

lib\_list.insert(lib\_list.index('requests')+2,'jieba')

5）lib\_list.insert(lib\_list.index('requests')+1, 'spicy')

lib\_list.insert(lib\_list.index('spicy')+1, 'jieba') *# 方法5*

5）lib\_list = lib\_list[:idx + 1] + ["scipy", "jieba"] + lib\_list[idx + 1:] *# 方法6*

*# c.*1）lib\_list.remove(**"bs4"**)*# 方法1*  
2）del lib\_list[lib\_list.index("bs4")]*# 方法2*

*# d.*1）lib\_list.sort(key=**lambda** x:len(x),reverse=**True**) *# 方法1*

2）lib\_list.sort(key=len,reverse=True) *# 方法2*

3）lib\_list.sort(key=len) *# 方法3*

lib\_list.reverse()

4）c=sorted(b,key=len) *# 方法4*

c.reverse()

5）b=sorted( lib\_list,key=lambda x: -len(x)) *# 方法5*

6）lib\_list.sort(key=lambda x:-len(x)) *# 方法6*

7）sorted\_list=sorted(lib\_list,key=len,reverse=True) *# 方法7*

作业提交截止时间：2023年3月17日晚24:00前（迟交会酌情扣分）

作业提交方式：邮件提交。发送邮件标题命名为：Python第一次作业+姓名+学号。

每道题目的程序单独建一个python文件（.py或.ipynb），例如本次作业建3个文件，分别命名为1.py，2.py，3.py或1.ipynb，2.ipynb，3.ipynb。

本次作业统一打包成一个压缩文件(.zip或.rar)，以邮件附件的形式提交给助教，压缩文件命名方式为“Python第一次作业+姓名+学号.zip，或者第一次作业+姓名+学号.rar”