第7章 文件、数据格式化与Python计算生态

目录

CONTENTS

7.1 文件的使用

目录

CONTENTS

7.1.1 文件的类型

- 文件是一个存储在辅助存储器上的数据序列，可以包含任何数据内容。
- 文件包括两种类型：文本文件和二进制文件。

- 文本文件可以处理各种语言所需的字符，只包含基本文本字符，不包括诸如字体、字号、颜色等信息。它可以在文本编辑器和浏览器中显示。即在任何情况下，文本文件都是可读的。
- 使用其它编码方式的文件即二进制文件，如Word文档、PDF、图像和可执行程序等。

- 二进制文件和文本文件最主要的区别在于是否有统一的字符编码。
- 无论文件创建为文本文件或者二进制文件，都可以用“文本文件方式”和“二进制文件方式”打开，但打开后的操作不同。

```
1  f = open("C:\\Users\\li\\python\\hello.txt","rt")
2  print(f.readline())                          #t表示文本文件方式
3  f.close()
```

```
========= RESTART: C:\Users\li\python\open函数.py =========
你好，河南！
>>>
```

文本文件hello.txt，采用二进制方式打开

```
1  f = open("C:\\Users\\li\\python\\hello.txt","rb")
2  print(f.readline())                    #b表示二进制文件方式
3  f.close()
```

运行结果如下：

```
========= RESTART: C:\Users\li\python\open函数.py =========
b'\xc4\xe3\xba\xc3\xa3\xac\xba\xd3\xc4\xcf\xa3\xa1'
>>>
```

- 结论：
- 采用文本方式读入文件，文件经过编码形成字符串，打印出有含义的字符；
- 采用二进制方式打开文件，文件被解析为字节流。

7.1.2 文件的打开和关闭

- 在Python中对文件的操作通常按照以下三个步骤进行：

7.1.2 文件的打开和关闭

- 打开：使用open()函数打开（或建立）文件，返回一个file对象。
- 操作：使用file对象的读/写方法对文件进行读/写的操作。
- 关闭：使用file对象的close()方法关闭文件。

- open()函数

  用来打开文件。open()函数需要一个字符串路径，表明希望打开文件，并返回一个文件对象。语法如下：

- fileobj=open(filename[,mode[,buffering]])

```
>>> helloFile=open("d:\\python\\hello.txt","r",1)
```

open()函数中mode参数常用值

值	描述
'r'	只读模式，如果文件不存在，则发生异常，默认值
'w'	覆盖写模式，如果文件不存在，则创建文件再打开；如果文件存在，则清空文件内容再打开
'a'	追加模式，如果文件不存在，则创建文件再打开；如果文件存在，打开文件后将新内容追加至原内容之后
'b'	二进制模式，可添加到其他模式中使用
't'	文本文件模式，默认值
'+'	读/写模式，可添加到其他模式中使用

- 打开模式使用字符串方式表示，根据字符串定义，单引号或者双引号均可。上述打开模式中，'r'、'w'可以和'b'、't'、'+'组合使用，形成既表达读写又表达文件模式的方式。
- 如：rb表示二进制只读方式

```
>>> helloFile=open(“d:\\python\\hello.txt”,”rb”)
```

```
>>> helloFile=open("d:\\python\\hello.txt","r",1)
```

open函数的第三个参数buffuring控制缓冲。

- 当参数取0或False时，输入/输出是无缓冲的，所有读写操作直接针对硬盘。

- 当参数取1或True时，输入/输出有缓冲，此时Python使用内存代替硬盘，使程序运行速度更快，只有使用flush或close时才会将数据写入硬盘。
- 当参数大于1时，表示缓冲区的大小，以字节为单位；负数表示使用默认缓冲区的大小。

练　习

1．关于Python文件的'+'打开模式，以下选项中描述正确的是：D

A. 只读模式

B. 覆盖写模式

C. 追加写模式

D. 与r/w/a/x一同使用，在原功能基础上增加同时读写功能

练　习

2．在读写文件之前，需要创建文件对象，采用的方法是 C

A. create　　B. folder　　C. open　　D. File

7.1.3 文件的读写

我们可以调用文件file对象的多种方法读取文件内容。

1．read()方法

不设置参数的read()方法将整个文件的内容读取为一个字符串。

```
1 helloFile=open("C://Users//li//Python//hello1.txt")
2 fileContent=helloFile.read()
3 helloFile.close()
4 print(fileContent)
```

运行结果如下：

```
========= RESTART: C:/Users/li/python/read函数.py =========
关关雎鸠，在河之洲。
窈窕淑女，君子好逑。
```

2．readline()方法

readline()方法从文件中获取一个字符串，每个字符串就是文件中的每一行。

```
1  helloFile=open("C:\\Users\\li\\Python\\hello1.txt")
2  fileContent=""
3  while True:
4      line=helloFile.readline()
5      if line=="":    # 或者 if not line
6          break
7      fileContent+=line
8  helloFile.close()
9  print(fileContent)
```

运行结果如下：

```
========= RESTART: C:/Users/li/python/readline函数.py =========
关关雎鸠，在河之洲。
窈窕淑女，君子好逑。
```

3．readlines()方法

readlines方法返回一个字符串列表，其中的每一项是文件中每一行的字符串。

```
1 helloFile=open("C://Users//li//Python//hello1.txt")
2 fileContent=helloFile.readlines()
3 helloFile.close()
4 print(fileContent)
5 for line in fileContent:
6     print(line)
```

```
====== RESTART: C:\Users\li\python\readlines.py ======
['关关雎鸠，在河之洲。 \n', '窈窕淑女，君子好逑。 ']
关关雎鸠，在河之洲。

窈窕淑女，君子好逑。
```

4. seek()方法

能够移动读取指针的位置，f.seek(0)将读取指针移动到文件开头，f.seek(2)将读取指针移动到文件结尾。

```
>>>f = open("C:\\Users\\li\\Python\\hello1.txt", "r")
>>>s = f.read()
>>>print(s)
======== RESTART: C:/Users/li/python/seek函数.py ========
关关雎鸠，在河之洲。
窈窕淑女，君子好逑。
>>>f.seek(0)   # 将读取指针重置到文件开头
>>>ls = f.readlines()
>>>print(ls)
['关关雎鸠，在河之洲。 \n', '窈窕淑女，君子好逑。 ']
>>>f.close()
```

练　习

3．下列选项中不是Python对文件读操作方法的是：C

A. read()　　B. readline()　　C. readtext()　　D. readlines()

- 写文件与读文件相似，都需要先创建文件对象连接。所不同的是，打开文件时是以“写”模式或“添加”模式打开。如果文件不存在，则创建该文件。
- 与读文件时不能添加或修改数据类似，写文件时也不允许读取数据。“w”写模式打开已有文件时，会覆盖文件原有内容，从头开始，就像我们用一个新值覆写一个变量的值。

1. write ()方法

write方法将字符串参数写入文件。

2. writelines()方法

writelines(sequence)方法向文件写入一个序列字符串列表，如果需要换行，则要自己加入每行的换行符。

```
1  helloFile=open("C:/Users/li/Python/hello1.txt","w")
2  helloFile.write("蒹葭苍苍，白露为霜。\n")
3  helloFile.close()
4  helloFile=open("C:/Users/li/Python/hello1.txt","a")
5  helloFile.write("所谓伊人，在水一方。\n")
6  helloFile.close()
7  helloFile=open("C:/Users/li/Python/hello1.txt")
8  fileContent=helloFile.read()
9  helloFile.close()
10 print(fileContent)
```

运行结果如下：

```
======= RESTART: C:\Users\li\python\write函数.py =======
蒹葭苍苍，白露为霜。
所谓伊人，在水一方。
```

练　习

4.
```
fname=input("请输入要写入的文件：")
fo=open(fname,"w+")
ls=["唐诗","宋词","元曲"]
fo.writelines(ls)
fo.seek(0)
for line in fo:
    print(line)
fo.close
```
上述代码的运行结果是：C

A.唐诗
宋词
元曲

B. “唐诗”
“宋词”
“元曲”

C.唐诗宋词元曲

D. “唐诗宋词元曲”

THANKS