

基操

模式	说明
r	读模式（默认模式，可省略），如果文件不存在，抛出异常
w	写模式，如果文件已存在，先清空原有内容；如果文件不存在，创建新文件
x	写模式，创建新文件，如果文件已存在则抛出异常
a	追加模式，不覆盖文件中原有内容
b	二进制模式（可与r、w、x或a模式组合使用）
t	文本模式（默认模式，可省略）
+	读、写模式（可与其他模式组合使用）

方法	功能说明
close()	把缓冲区的内容写入文件，同时关闭文件，释放文件对象
read([size])	从文本文件中读取并返回size个字符，或从二进制文件中读取并返回size个字节，省略size参数表示读取文件中全部内容
readline()	从文本文件中读取并返回一行内容
readlines()	返回包含文本文件中每行内容的列表
seek(cookie, whence=0, /)	定位文件指针，把文件指针移动到相对于whence的偏移量为cookie的位置。其中whence为0表示文件头，1表示当前位置，2表示文件尾。对于文本文件，whence=2时cookie必须为0；对于二进制文件，whence=2时cookie可以为负数
write(s)	把s的内容写入文件，如果写入文本文件则s应该是字符串，如果写入二进制文件则s应该是字节串
writelines(s)	把列表s中的所有字符串写入文本文件，并不在s中每个字符串后面自动增加换行符。也就是说，如果确实想让s中的每个字符串写入文本文件之后各占一行，应由程序员保证每个字符串以换行符结束

语句with

在实际开发中，读写文件应优先考虑使用上下文管理语句with。关键字with可以自动管理资源，不论因为什么原因跳出with块，总能保证文件被正确关闭。除了用于文件操作，with关键字还可以用于数据库连接、网络连接或类似场合。用于文件内容读写时，with语句的语法形式如下：with open(filename, mode, encoding) as fp: # 这里写通过文件对象fp读写文件内容的语句块

In [2]:

```
with open(filename, mode, encoding) as fp:  
# 这里写通过文件对象fp读写文件内容的语句块
```

```
File "C:\Users\23113\AppData\Local\Temp\ipykernel_11704\3075165646.py",  
line 2  
# 这里写通过文件对象fp读写文件内容的语句块  
^
```

IndentationError: expected an indented block

合并两个.txt文件的内容，两个文件的多行内容交替写入结果文件，如果一个文件内容较少，则把另一个文件的剩余内容写入结果文件尾部。

json文件

In [3]:

```
import json  
information = [  
    {'小区名称': '小区A', '均价': 8000, '月交易量': 20},  
    {'小区名称': '小区B', '均价': 8500, '月交易量': 35},  
    {'小区名称': '小区C', '均价': 7800, '月交易量': 50},  
    {'小区名称': '小区D', '均价': 12000, '月交易量': 18}]  
  
with open('房屋信息.json', 'w') as fp:  
    json.dump(information, fp, indent=4, separators=(',', ':'))  
  
with open('房屋信息.json') as fp:  
    information = json.load(fp)  
    for info in information:  
        print(info)
```

```
{'小区名称': '小区A', '均价': 8000, '月交易量': 20}  
{'小区名称': '小区B', '均价': 8500, '月交易量': 35}  
{'小区名称': '小区C', '均价': 7800, '月交易量': 50}  
{'小区名称': '小区D', '均价': 12000, '月交易量': 18}
```

CSV文件

In [4]:

```
from csv import reader, writer
from random import randrange
from datetime import date, timedelta
fn = 'data.csv'
with open(fn, 'w') as fp:
    wr = writer(fp)                                # 创建csv文件写对象
    wr.writerow(['日期', '销量'])                 # 写入表头
    # 第一天的日期, 2020年1月1日
    startDate = date(2020, 1, 1)
    for i in range(100):                          # 生成100个模拟数据
        amount = 500 + i*5 + randrange(5, 50)     # 生成一个模拟数据, 写入csv文件
        wr.writerow([str(startDate), amount])
        startDate = startDate + timedelta(days=1) # 下一天
# 读取并显示上面代码生成的csv文件内容
with open(fn) as fp:
    for line in reader(fp):
        if line:
            print(*line)
```

```
2020-02-22 780
2020-02-23 780
2020-02-24 789
2020-02-25 792
2020-02-26 795
2020-02-27 815
2020-02-28 800
2020-02-29 839
2020-03-01 846
2020-03-02 836
2020-03-03 857
2020-03-04 852
2020-03-05 850
2020-03-06 869
2020-03-07 852
2020-03-08 858
2020-03-09 858
2020-03-10 857
2020-03-11 884
2020-03-12 900
2020-03-13 888
```

文件的打开与关闭

In [19]:

```
file1 = open(r'C:\Users\23113\A.txt')           # 以只读方式打开D盘的文本文件a.txt
file2 = open('b.txt', 'w')                       # 以只写方式打开当前目录的文本文件b.txt
file3 = open('c.txt', 'w+')                      # 以读/写方式打开文本文件c.txt
file4 = open('d.txt', 'wb+')                    # 以读/写方式打开二进制文件d.txt
```

In [22]:

```
file1.close()
```

In [27]:

```
with open(r'C:\Users\23113\A.txt') as f:
    pass
```

文件的读写

In [36]:

```
with open(r'C:\Users\23113\file.txt', mode='r') as f:
    print(f.read(2))    # 读取两个字节的数据
    print(f.read())
```

```
i
like you all time
maybe you don't konw
but i always online
```

In [35]:

```
with open(r'C:\Users\23113\file.txt', mode='r', encoding='utf-8') as f:
    print(f.readline())
    print(f.readline())
```

```
i like you all time

maybe you don't konw
```

In [38]:

```
with open(r'C:\Users\23113\file.txt', mode='r', encoding='utf-8') as f:
    print(f.readlines())    # 使用readlines()方法读取数据
```

```
['i like you all time\n', 'maybe you don't konw\n', 'but i always online']
```

In [53]:

```
string = "Here we are all, by day; by night." # 字符串
with open(r'C:\Users\23113\file.txt', mode='w', encoding='utf-8') as f:
    size = f.write(string)    # 写入字符串
    print(size)              # 打印字节数
```

34

In [64]:

```
string = "Here we are all, by day;\nby night we're hurl'd By dreams, " \
        "each one into a several world."
with open(r'C:\Users\23113\file.txt', mode='w', encoding='utf-8') as f:
    f.writelines(string)
    size = f.write(string)    # 写入字符串
    print(size)
```

88

```
Here we are all, by day;  
by night we're hurl'd By dreams, each one into a several world.  
Here we are all, by day;  
by night we're hurl'd By dreams, each one into a several world.
```

文件的定位读写

In [65]:

```
with open(r'C:\Users\23113\file.txt') as f:  
    print(f.tell()) # 获取文件读写位置  
    print(f.read(5)) # 利用read()方法移动文件读写位置  
    print(f.tell())
```

```
0  
Here  
5
```

In [66]:

```
with open(r'C:\Users\23113\file.txt') as f:  
    f.tell() # 获取文件读写位置  
    loc = f.seek(5,0) # 相对文件首部移动5字节  
    print(loc) # 打印当前文件读写位置
```

```
5
```

In [70]:

```
with open(r'C:\Users\23113\file.txt') as f: # 打开文本文件  
    f.seek(5,0) # 相对文件首部移动5字节  
    f.seek(3,1)
```

```
-----  
-  
UnsupportedOperation                                Traceback (most recent call last)  
~\AppData\Local\Temp\ipykernel_7372\4282129401.py in <module>  
      1 with open(r'C:\Users\23113\file.txt') as f: # 打开文本文件  
      2     f.seek(5,0) # 相对文件首部移动5字节  
----> 3     f.seek(3,1)
```

UnsupportedOperation: can't do nonzero cur-relative seeks

In [71]:

```
with open(r'C:\Users\23113\file.txt', 'rb') as f:  
    f.seek(5,0)  
    f.seek(3,1)
```

文件与目录管理

In [2]:

```
import os
os.remove(r'C:\Users\23113\A.txt')
# A文件没有了
```

```
-----
FileNotFoundError                                Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_14972\3856628903.py in <module>
      1 import os
----> 2 os.remove(r'C:\Users\23113\A.txt')
      3 # A文件没有了
```

FileNotFoundError: [WinError 2] 系统找不到指定的文件。: 'C:\\Users\\23113\\A.txt'

A		
	.condarc	2022/4/17 17:42
	.wakatime	2022/4/19 23:57
	.wakatime-internal	2022/4/19 23:15
	3_原始数据(1)	2022/4/21 14:45
	2017dianying	2022/4/22 17:51
	2018dianying	2022/4/22 17:56
	dianying	2022/4/21 17:31
	file	2022/4/22 18:37
	insurance	2022/1/13 3:16
	insurence 数据可视化实战项目	2022/4/21 14:48
	Untitled	2022/4/21 18:11

In [10]:

```
os.rename(r'C:\Users\23113\A.txt', r'C:\Users\23113\test.txt')
```

```
-----  
-  
AttributeError                                Traceback (most recent call last)  
~\AppData\Local\Temp\ipykernel_14972\2896109231.py in <module>  
----> 1 os.rename(r'C:\Users\23113\A.txt', r'C:\Users\23113\test.txt')
```

AttributeError: module 'os' has no attribute 'rename'

In [9]:

```
os.mkdir('dir')
```

In []:

```
os.rmdir('dir')
```

In []:

```
print(os.getcwd())
```

In []:

```
os.chdir('E:\\')                                # 更改默认目录为E:\  
print(os.getcwd())                             # 获取当前工作目录
```

数据维度与数据格式化

In [20]:

```
csv_file = open(r'C:\Users\23113\2018dianying.csv' )  
lines = []  
for line in csv_file:  
    line = line.replace('\n', '')  
    lines.append(line.split(','))  
print(lines)  
csv_file.close()
```

```
-----  
-  
UnicodeDecodeError                            Traceback (most recent call last)  
~\AppData\Local\Temp\ipykernel_14972\1155775050.py in <module>  
      1 csv_file = open(r'C:\Users\23113\2018dianying.csv' )  
      2 lines = []  
----> 3 for line in csv_file:  
      4     line = line.replace('\n', '')  
      5     lines.append(line.split(','))  
  
UnicodeDecodeError: 'gbk' codec can't decode byte 0xb0 in position 14: illegal mul  
tibyte sequence
```

数据格式问题，可以查找相关数据

A57																			
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
60	《妖铃铃》喜剧	2D		12/29		寒假档	36292万	艺恩数据	吴君如	0	沈腾	6286636	岳云鹏	15563042	吴君如	20885597	张栢	84317	4.3
61	《二代妖》奇幻/喜剧	2D/3D		12/29	续集	寒假档	29207万	艺恩数据	肖洋	8	李光洁	5413566	冯绍峰	19687698	刘亦菲	55039518	郭京飞	4791293	5.7
62	《解忧杂货店》奇幻/剧情	2D		12/29	改编	寒假档	22298万	艺恩数据	韩杰	6	董子健	436426	王俊凯	40681076	迪丽热巴	40172001	李鸿其	337442	5.1
63	《妖猫传》奇幻/爱情	2D/IMAX		12/22	改编	寒假档	52998万	艺恩数据	陈凯歌	34	染谷将太	64169	黄轩	8628861	李淳	214667	张雨绮	6692391	7
64	《奇门遁甲》动作/奇幻	3D/IMAX		12/14		寒假档	29964万	艺恩数据	袁和平	23	李治廷	10630981	大鹏	7875356	倪妮	19190910	周冬雨	19329506	4.5
65	《英雄本色》动作/剧情	2D		1/18		寒假档	6298万	艺恩数据	丁晟	4	马天宇	25609768	王凯	15325837	李雪	219044	吴健	1415999	4.9
66	《无问西东》剧情/爱情	2D		1/12		寒假档	75394万	艺恩数据	李芳芳	3	吴尊	8206548	黄晓明	56091135	王力宏	59051775	张译	6098535	7.7
67	《谜巢》动作/惊悚	3D		1/12		寒假档	4943万	艺恩数据	金波·兰道	5	凯尔希·格兰莫		凯南·鲁兹	3257	李冰冰	40024240	章子怡	129293	3.6
68	《红海行动》动作/剧情	2D/3D/IM		2/16		寒假档	363232.07	艺恩数据	林超贤	23	张译	6049742	黄景瑜	5580108	海清	11157338	杜江	10554729	8.5
69	《唐人街》喜剧/剧情	2D/IMAX		2/16	续集	寒假档	339604.07	艺恩数据	陈思诚	18	王宝强	29053398	王迅	3176639	刘昊然	15852625	尚语贤	574931	7
70	《捉妖记2》奇幻/动作	2D/IMAX		2/16	续集	寒假档	232660.87	艺恩数据	许诚毅	8	梁朝伟	2075575	井柏然	21669809	白百合	24270562	李宇春	9983825	5.1
71	《祖宗十九代》喜剧/剧情	2D		2/16		寒假档	16979万	艺恩数据	郭德纲	0	岳云鹏	15563042	吴京	11561245	吴秀波	9137313	王宝强	29053398	4.4
72	《西游记》动作	3D/IMAX		2/16	改编	寒假档	72721万	艺恩数据	郑保瑞	7	冯绍峰	19687698	郭富城	13060605	小沈阳	16169031	赵丽颖	60390077	4.5
73	《南极之恋》爱情/剧情	2D/3D		2/2	改编	寒假档	23435万	艺恩数据	吴有音	0	赵又廷	13024484	杨子姗	5436475					6.8
74	《泡芙小姐》动作	2D		2/9	改编	寒假档	1929万	艺恩数据	张歆艺	0	王栎鑫	5714434	张歆艺	18533461	于莎莎	1791158	蒋欣	24492622	3.6
75	《遇见你》爱情	2D		2/16	改编	寒假档	4949万	艺恩数据	顾长卫	14	张海宇	2144479	白客	7649883	蓝盈莹	3459592	王玉雯	201158	4.8
76	《坏爸爸》喜剧/情感	2D		2/2		寒假档	2445万	艺恩数据	兰城序	0	邵兵	3037667	孙绍龙	516816	孔琳	127056	赵圆圆	318713	3.6
77	《玩命》动作	2D		3/30		寒假档	24万	艺恩数据	李炳渊	2	李炳雷	13618	黄译萱	99719	王洪强	1312	徐少强	1839	3.8
78	《闺蜜2》剧情/喜剧	2D		3/2	续集	寒假档	6427万	艺恩数据	黄真真	7	陈意涵	18485455	范湉湉	2240002	薛凯琪	6893329	张钧甯	11739639	2.7
79	《我说的都是真的》喜剧/爱情	2D		3/30		寒假档	1721万	艺恩数据	刘仪伟	0	小沈阳	16169031	连凯	454275	罗海琼	1048176	陈意涵	18485455	4.8
80	《宇宙有爱》爱情	2D		3/2		寒假档	154万	艺恩数据	陈思行	0	张亮	24518485	王梓轩	256724	郭碧婷	15882531	白冰	3973050	4.5
81	《爱的帕拉梅拉》励志	2D		3/30		寒假档	306万	艺恩数据	刘海	0	王栎鑫	5714434	周知	1198971	奚美娟	1490	许还山	165	2.4
82	《卧底巨星》喜剧/动作	2D		1/12		寒假档	3925.5万	艺恩数据	谷德昭	2	陈奕迅	20396297	李荣浩	13613663	李一桐	4301502	陈国坤	402346	3.8
83	《妈妈咪呀！》喜剧	2D		3/9		寒假档	3711.7万	艺恩数据	赵锐	0									7
84	《大世界》动画	2D		1/12		寒假档	261.8万	艺恩数据	刘健	4									7
85	《熊出没》喜剧	2D/3D		2/16	续集	寒假档	60531.5万	艺恩数据	丁亮/林汇川	7									6.8

In [13]:

```
csv_file = open('score.csv')
file_new = open('count.csv', 'w')
lines = []
for line in csv_file:
    line = line.replace('\n', '')
    lines.append(line.split(','))
# 添加表头字段
lines[0].append('总分')
# 添加总分
for i in range(len(lines)-1):
    idx = i+1
    sun_score = 0
    for j in range(len(lines[idx])):
        if lines[idx][j].isnumeric():
            sun_score += int(lines[idx][j])
    lines[idx].append(str(sun_score))
for line in lines:
    print(line)
    file_new.write(','.join(line)+'\n')
csv_file.close()
file_new.close()
```

FileNotFoundError

Traceback (most recent call last)

~\AppData\Local\Temp\ipykernel_14972\511618113.py in <module>

----> 1 csv_file = open('score.csv')

2 file_new = open('count.csv', 'w')

3 lines = []

4 for line in csv_file:

5 line = line.replace('\n', '')

FileNotFoundError: [Errno 2] No such file or directory: 'score.csv'

In []: