Python
File (Simple)

郭忠義

jykuo@ntut.edu.tw

臺北科技大學資訊工程系

開啟檔案

- <file> = open(<filename>, <mode>)
  - <file>：開啟檔案物件
  - <filename>：檔名
  - <mode>：開檔模式
    - r: 唯讀
    - w (write):覆蓋寫入內容
    - a (append): 在後加入新內容

```
infile = open('hello.txt', 'r') # 以讀取模式開檔
outfile = open('hello.txt', 'w') # 以寫入模式開檔
```

r	**開啟檔案只讀取**
rb	開啟檔案只讀取格式為 binary
r+	開啟檔案可讀取寫入
rb+	開啟檔案可讀取寫入binary
w	開啟檔案只寫入
w+	可讀取可寫入模式
wb	開啟檔案只寫入二進位格式
a	只允許新增
a+	允許新增與讀取

關閉檔案

- <file>.close() #將緩衝區內容寫入檔案，關閉檔案
  - 檔案讀寫可能產生IOError，而不會呼叫fp.close()造成關檔錯誤 。
  - 為保證是否出錯都正確關閉檔案，可使用try ... finally

```
try:
  fp = open('hello.txt', 'r')
  print(fp.read())
  fp.write('test')
except FileNotFoundError:
  print("file not found")
except:
  print("Something wrong")
finally:
  fp.close()
```

- 用with 開檔，若錯誤，檔案仍會自動關閉，
  - 執行 return, continue, break跳出 with 指令區塊
  - 發生例外 (Exception)

```
with open('hello.txt', 'r') as fp:
```

讀取檔案

- <file>.read()
  - 讀取全部或剩餘資料，回傳長字串
- <file>.readline()
  - 讀取下一行資料，回傳字串
  - 利用迴圈一次讀一行資料

```
with open('filename.txt', 'r') as infile:
    while True:
        data = infile.readline()     # 一次讀一行資料
        print(type(data))
        print(data)
        if not data:                 # 所有資料讀取完畢
            break
        print(line, end='')          # end='': 不要自動加斷行
```

讀取檔案

- <file>.readlines()
  - 讀取全部或剩餘資料，
  - 回傳串列，每個元素都是一行資料

```
with open('filename.txt', 'r') as fp:
    data2 = fp.readlines()
    print(type(data2))
    print(data2)
```

```
with open('filename.txt', 'r') as infile:
    for line in infile.readlines():   # 一次讀取所有資料，再一行一行處理
        print(line, end='')
```

```
# Python 讀檔將每行資料存到串列中的元素，上述程式可簡化
with open('filename.txt', 'r') as infile:
    for line in infile:
        print(line, end='')
```

寫檔 4 個方法

- write

```
# 開啟檔案
fp = open("filename.txt", "a")
# 寫入 This is a testing! 到檔案緩衝區
fp.write("This is a testing!")
#將緩衝區寫入檔案，關閉檔案
fp.close()
```

- writelines

```
# 開啟檔案
fp = open("filename.txt", "w")
# 將 lines 所有內容寫入到緩衝區
lines = ["One\n", "Two\n", "Three\n"]
fp.writelines(lines)
#將緩衝區寫入檔案，關閉檔案
fp.close()
```

- print

```
# 開啟檔案
fp = open("filename.txt", "a")
# 寫入 This is a testing! 到檔案緩衝區
print("This is a testing!\n", file=fp)
#將緩衝區寫入檔案，關閉檔案
fp.close()
```

- with open

```
with opefilename.txt", 'r') as in_file:
    with open("filename.txt", 'a') as out_file:
        for line in in_n("file:
            out_file.write(line)
```

Exercise

- 利用迴圈一次讀一行資料，將偶數行資料印出

```
with open('filename.txt', 'r') as infile:
    line_num=0
    for line in i
        line_num+=1
        if line_num%
            print(line, end='')
```

- 一次讀取、印出多行資料，將全部資料的第一個字與最後一個字印出

```
fp = open('hello.txt', 'w')
fp.write("First line\n#Second line\n#Third")
fp.close()
with open('hello.txt', 'r') as infile:
    data = infile.r
    print(data)
    print(len(data))
    print(data[0], d
```

讀寫檔案

- 顯示檔案所有行，忽略以#開頭的行

```
with open("./hello.txt") as f:
    for line in f:
        if line.strip()[0] !=  "#":
            print(line)
```

- 把/passwd檔案中'root'字串用'west'替換，另存tmp檔案

```
with open("passwd.txt") as f1:
    # 遍歷檔案的每一行內容；
    for line in f1:
        # 字串替換
        bline = line.replace("root", "west")
        with open("tmp", "a+") as f2:
            # 寫入新檔案
            f2.write(bline)
```

```
root        word
user        pass
```

```
west        word
user        pass
```

Exercise

- 把./passwd檔案中 xi 字串用 yi 替換，另存tmp檔案
  - X =['x1', 'x2', 'x3'], Y=['y1', 'y2', 'y3']

```
import copy
x=["x1","x2","x3"]
y=["y1","y2","y3"]

with open('passwd') as f1:
   # 遍歷檔案的每一行內容 ;
   for line in f1:
      bline=copy.copy(line)
      # 字串替換
      for i in range(    ):
#        print(x[i],y[i])
        bline = bline.replace(      )
      with open("tmp", "a+") as f2:
         # 寫入新檔案
         f2.write(bline)
```

讀取CSV檔案

- 一列一列的讀取出csv資料

```
import csv
#f= open('data.csv', encoding='utf-8')
f= open('data.csv')
readFile = csv.reader(f)
for row in readFile:
    print(row)
f.close()
```

- 使用with開啟csv檔案
  - 加上 newline=''，為讓資料中包含的換行字元可正確解析

```
import csv
with open('data.csv', newline='') as csvfile:
  readFile = csv.reader(csvfile)
  for row in readFile:
    print(row)
```

讀取CSV檔案

- 指定分隔字元
  - 資料欄位分隔字元非使用預設逗號，而是其他字元，讀取時要指定分隔字元

```
import csv
with open('data.csv', newline='') as csvfile:
  readFile = csv.reader(csvfile, delimiter=':')
  for row in readFile:
    print(row)
```

- 讀取成 Dictionary
  - 讀取csv 檔案內容後，轉為dictionary 格式
  - csv.DictReader()自動把第一列(row)當作欄位名稱，第二列後的每一列轉為 dictionary，如此可使用欄位名稱存取資料

```
import csv
with open('D:\\Courses\Python\data.csv', newline='') as csvfile:
  readFile = csv.DictReader(csvfile)
  for row in readFile:
    print(row['班級'], row['學號'], row['期中考成績'])
```

Exercise

- 讀取檔案中 xi 字串被 yi 替換，另存tmp檔案
  - X =[xi], Y=[yi]，X是英文，Y是中文翻譯
  - X, Y 分別存在 data.csv檔案的第一 row和第三row，第二和第四row是備註

```
import csv
def getHeader():  #讀取檔案第0行和第2行
    i=0
    f= open('data.csv')
    readFile = csv.
    for row in readFile:
        if i==0:
            eng=row
        if i==2:
            chi=row
        i=i+1
    f.
    print(eng)
    print(chi)
    return eng, chi
getHeader()
```

Exercise Solution

```
def convert(aFile, bFile, eng, chi):
    f1 = open(aFile)
    f2 = open(bFile,'w', encoding='utf-8-sig')
    data = f1.read()
    print(type(data))
    print(data)
    bline=data
    #zip 可將eng, chi 對應的元素打包成一個個 tuple，回傳 tuples 組成的 list
    for e,c in zip(          ):
        bline = bline.          )
#       if e in data:
#           bline = bline.replace(e, c)
    print(bline, '####')
    f2.write(bline)
    f1.close()
    f2.close()


eng, chi = getHeader()
convert('d:\\Courses\Python\data.csv', 'data.csv', eng, chi)
```

寫入CSV檔案

- 一次寫入二維表格
  - 若資料是已整理好二維表格，可一次把整張表格寫進 csv 檔案

```
import csv
# 二維表格
table = [['班級', '學號', '成績'],
         ['資工一', '109590001', 90],
         ['資工一', '109590002', 85]]
with open('output.csv', 'w', newline='') as csvfile:
  writer = csv.writer(csvfile)
  writer.writerows(table)   # 寫入二維表格
```

- 寫入 Dictionary
  - 資料格式是 dictionary，可使用 csv.DictWriter()寫入 csv 檔案中

```
import csv
with open('output.csv', 'w', newline='') as csvfile:
  columns = ['班級', '學號', '成績']
  # 將 dictionary 寫入 CSV 檔
  writer = csv.DictWriter(csvfile, fieldnames=columns, delimiter=':')
  writer.writeheader()   # 寫入第一列的欄位名稱
  writer.writerow({'班級': '資工一', '學號': '109590003', '成績': 95})   # 寫入資料
  writer.writerow({'班級': '資工一', '學號': '109590004', '成績': 88})   # 寫入資料
```

delimiter=':' is optional

Exercise

- 一行一行讀檔案 score.txt
  - 計算平均，將平均寫到最下面

Input file: score.txt

```
班級,學號,期中考成績,
資工一,109590001,88,
資工一,109590002,90,
資工一,109590003,92,
資工一,109590004,85,
資工一,109590005,87,
資工一,109590006,95,
資工一,109590007,80,
資工一,109590008,84,
資工一,109590009,86,
資工一,109590010,83,
```

Output file: avg_score.txt

```
Class,Student ID,Score,
資工一,109590001,88,
資工一,109590002,90,
資工一,109590003,92,
資工一,109590004,85,
資工一,109590005,87,
資工一,109590006,95,
資工一,109590007,80,
資工一,109590008,84,
資工一,109590009,86,
資工一,109590010,83,
平均,,87.0,
```

Exercise

- 製作一個csv檔 score.csv
  - 一行一行讀檔案 score.csv，製作成字典
  - 計算每位學生平均，寫在學生資料最後，計算全班平均，寫到最下面

Input file: score.csv

```
班級,學號,國文,數學,英文
資工一,109590001,80,80,80
資工一,109590002,90,90,90
資工一,109590003,70,70,70
資工一,109590004,60,60,60,
```

Output file: output.csv

```
Class,Student ID,average,
資工一,109590001,80,
資工一,109590002,90,
資工一,109590003,70,
資工一,109590004,60,
75,75,75,75
```

Exercise

- 製作一個csv檔 score.csv
  - 一行一行讀檔案 score.csv，製作成字典
  - 計算每位學生平均，寫在學生資料最後，計算全班平均，寫到最下面
- 輸出成 output.csv

score.csv
班級,學號,國文,數學,英文
資工一,109590001,80,80,80
資工一,109590002,90,90,90
資工一,109590003,70,70,70
資工一,109590004,60,60,60,

Output.csv
Class,Student ID,average,
資工一,109590001,80,
資工一,109590002,90,
資工一,109590003,70,
資工一,109590004,60,
75,75,75,75

Exercise

```
import csv
def trans(row):
    data = {}
    score = 0
    subject = ['國文','英文','數學']
    for key, value in row.items():
        print('=>', key, value)
        if key in subject:
            score = score + int(value)
    for key, value in row.items():
        
    data['average'] = score//3
    return data

with open('D:\\x.csv', newline='') as csvfile:
  readFile = csv.DictReader(csvfile)
  #print(readFile)
  inData = []
  for row in readFile:
      print(row)
      inData
print(inData)

with open('y.csv', 'w', newline='') as csvfile:
  #columns = ['班級', '學號','國文','數學','英文']
  columns = ['Class', 'Student Id','average']
  # 將 dictionary 寫入 CSV 檔
  writer = csv.DictWriter(csvfile, fieldnames=columns, delimiter=',')
  writer.writeheader()  # 寫入第一列的欄位名稱
  for data in inData:
      # 寫入資料
```