**py 4 read and write files**

**一、文件操作：open**

# 关于open 模式：

w 以写方式打开，

a 以追加模式打开

(从 EOF 开始, 必要时创建新文件)

r+ 以读写模式打开

w+ 以读写模式打开 (参见 w )

a+ 以读写模式打开 (参见 a )

rb 以二进制读模式打开

wb 以二进制写模式打开 (参见 w )

ab 以二进制追加模式打开 (参见 a )

rb+ 以二进制读写模式打开 (参见 r+ )

wb+ 以二进制读写模式打开 (参见 w+ )

ab+ 以二进制读写模式打开 (参见 a+ )

fid = open('D:\py35\scikit\_exercise.py','w+')

**# 读一行**

**# 如果定义了size，**

**# 有可能返回的只是一行的一部分**

fid.readline()

**# 一次读取所有行**

**# 文件指针指向文件尾部把文件每一行作为**

**# 一个list的一个成员，并返回这个list。**

**# 其实它的内部是通过循环**

**# 调用readline()来实现的。**

**# 如果提供size参数，size是表示读取内容的**

**# 总长，也就是说可能只读到文件的一部分。**

fid.readlines()

**#把str写到文件中，**

**# write()并不会在str后加上一个换行符**

fid.write('abcdefghijklmnopqrstuvw;mn')

**# fid.writelines(seq)**

**#把seq的内容全部写到文件中**

**# (多行一次性写入)。**

**# 返回下一行，并将文件操作标记位移到下一行。**

fid.next()

**# 把一个file用于for … in file这样的语句时，**

**# 就是调用next()函数来实现遍历的。**

**#　返回文件指针位置，为字节数**

fid.tell()

**# 0 移动到开头**

**# 0 从文件头偏移0**

fid.seek(0)

* 以上为方法（后加()），以下是属性（不加()）

**# 当前文件是否关闭 False**

fid.closed

**# 当前文件编码方式 'none'**

fid.encoding

**# 当前文件名字**

fid.name

**# 当前文件打开模式 'r'**

fid.mode

**# 当前行的换行方式 'none'**

print fid.newlines

**二、TXT文件**

* 打开以Tab分隔分割字符串的文档

fr=**open**('lenses.txt')

lenses=[inst.**strip**().**split**('\t') **for** inst infr.**readlines**()]

*# w时，没有文件则新建*

fid1 = open('ex1.txt','w+')

*# 写入文件内容*

fid1.write('new line. \n')

*# 关闭文件*

fid1.close()

**三、csv 文件**

**1 pandas 读写**

* 默认填写路径即可，如果不想添加索引（就是one two three这个）
* 如果只想输出year和state这两个字段的数据，那么就如代码所示操作。

df=pd.read\_csv('data/xxx.csv');

df.to\_csv('result.csv', index=**False**, header=[‘year’,'state'])