# Python程序设计实验报告9

班级：17应统一班

学号：117060400124

姓名：王贵珍

指导老师：林卫中

实验题目： 文本字符分析

实验要求：编写程序接收字符串，按字符出现的频率的将需打印字母。分别尝试录入一些中英文文章片段，比较不同语言之间字符频率的差别

实验代码：

from operator import itemgetter

def analyseText(text):

D = {}

tlist = list(text)

for e in tlist:

D[e] = D.get(e,0) + 1

return D

text = input("请输入一段文本：")

D = analyseText(text)

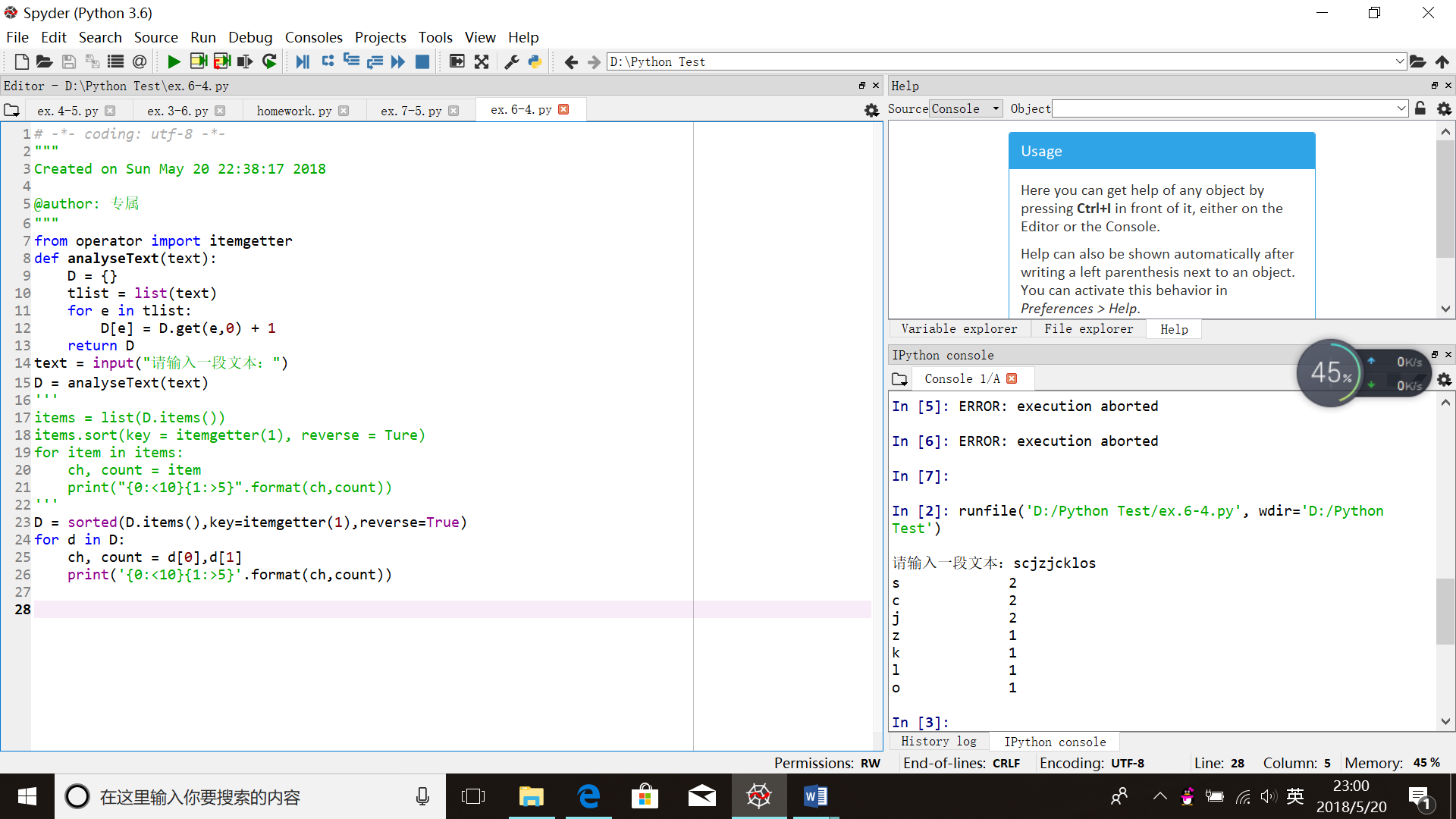
D = sorted(D.items(),key=itemgetter(1),reverse=True)

for d in D:

ch, count = d[0],d[1]

print('{0:<10}{1:>5}'.format(ch,count))

实验结果：



实验题目：《红楼梦》人物统计

实验要求：编写程序统计《红楼梦》中前20位出场最多的人物

实验代码：

import jieba

txt = open('红楼梦.txt', encoding = 'GB18030').read()

excludes = {'什么','一个','我们','你们','如今','老太太','知道','姑娘','这里','出来','一时','没有',

'怎么','那里','两个','说道','起来','众人','自己','太太','一面','只见','不是','不知','这个'}

words = jieba.lcut(txt)

counts = {}

for word in words:

if len(word) == 1:

continue

else:

counts[word] = counts.get(word,0) + 1

for word in excludes:

del(counts[word])

items = list(counts.items())

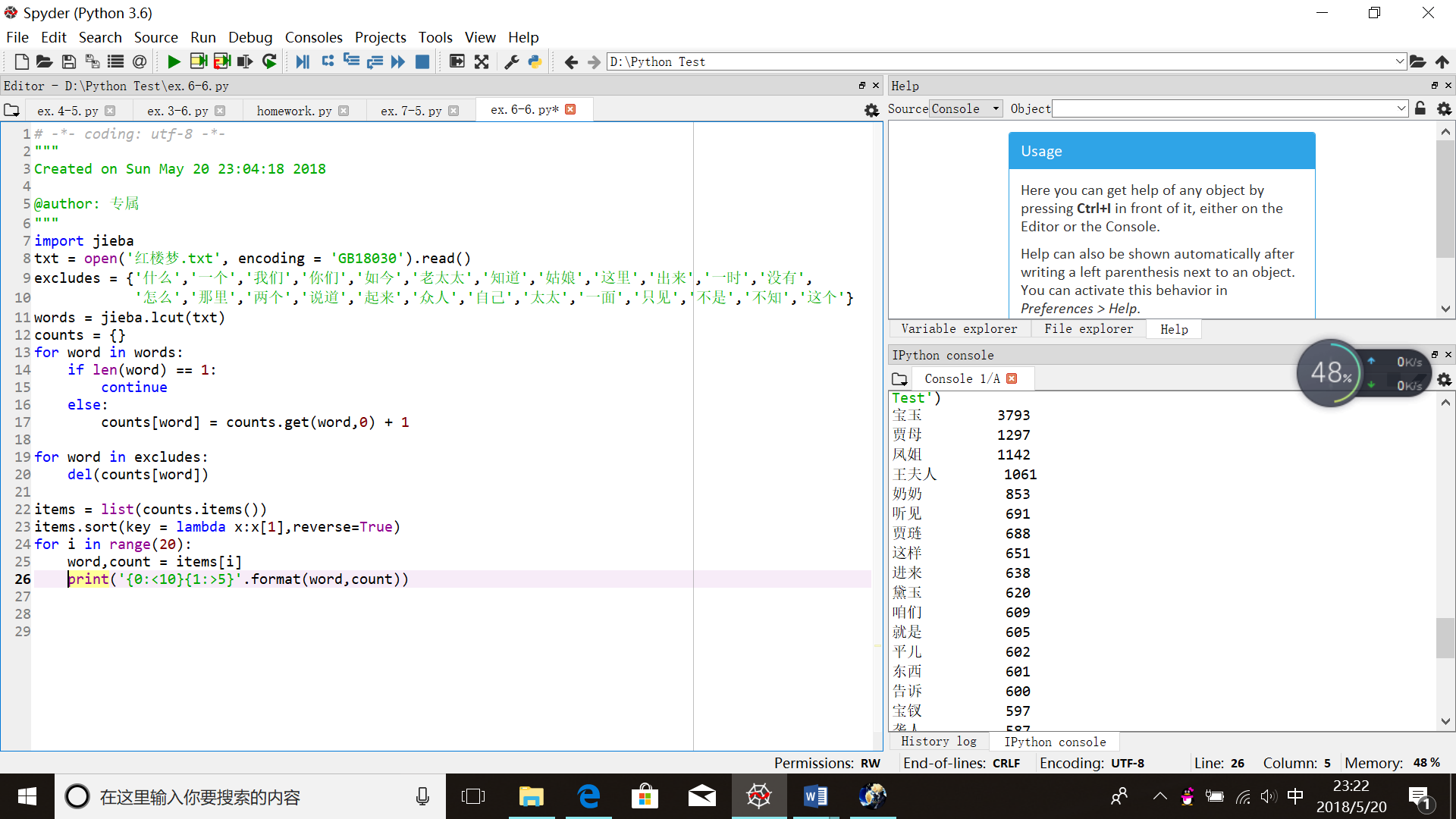
items.sort(key = lambda x:x[1],reverse=True)

for i in range(20):

word,count = items[i]

print('{0:<10}{1:>5}'.format(word,count))

实验结果：



实验小结：1.掌握字典的构建以及合理使用

2.根据value的值进行降序排列（D = sorted(D.items(),key=itemgetter(1),reverse=True)）

3.若字典中没有，则加进去（counts[word] = counts.get(word,0) + 1）