# Python实验报告6

实验六： 组合数据类型

实验目的：了解列表、字典、集合三种数据类型

## 实验要求：

1.能运用集合的特性去重

2.掌握运用字典概念处理复杂数据信息及字典的排序

3.能运用列表构建数据结构。

4.综合运用组合数据类型进行文本词频统计

## **实验过程：**

1.程序练习题6.1 随机密码生成。编写程序，在26个字母大小写和9个数字组成的列表中随机生成10个8位密码。

编写代码：

import random

ch = []

c = 'A'

while c <= 'Z':

ch.append(c)

c = chr(ord(c) + 1)

c = 'a'

while c <= 'z':

ch.append(c)

c = chr(ord(c) + 1)

c = '1'

while c <= '9':

ch.append(c)

c = chr(ord(c) + 1)

def generatePSW(n:int, chList:list):

psw = []

size = len(chList)

for i in range(n):

k = random.randint(0, size-1)

psw.append( chList[k])

return "".join(psw)

pswlist=[]

for i in range(10):

print("第{0:2} 个随机密码： {1}".format(i+1, generatePSW(8,ch)))

2.程序练习题6.2 重复元素判断。编写一个函数，接受列表作为参数，如果一个元素在列表中出现了不止一次，则返回True,但不要改变原来列表的值。同时编写调用这个函数和测试结果。

3.程序练习题6.3 重复元素续。利用集合的无重复性改编程序练习题6.2的程序，获得一个更快更简洁的版本。

编写代码：

#重复元素续（集合）

def isrepeat(vlist:list):

V = set(vlist)

if len(V) < len(vlist):

return True

return False

ls = [1,2,3,33,2,6]

if isrepeat(ls):

print("有重复元素")

else:

print("没有重复元素")

4.程序练习题6.4 文本字符分析。编写程序接受字符串，按字符出现频率的降序打印字母。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。

编写代码：

#文本字符分析

text="E:/Repoes/pythonProgram/TextbookExercise/exercise6\_2重复元素判定.py"

words = list(text)

counts = {}

for word in words:

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

items = list(counts.items())

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

for d in items:

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

print("{0:<5}{1}".format(ch,count))

5.程序练习题6.5 生日悖论分析。生日悖论指如果一个房间里有23人或以上那么至少有两个人生日相同的概率大于50%。编写程序，输出在不同随机样本数量下，23个人中至少两个人生日相同的概率。

6.程序练习题6.6 《红楼梦》人物统计。编写程序统计《红楼梦》中前20位出场最多的人物。

拓展：统计《红楼梦》中金陵十二钗的出场次数。金陵十二钗是指：宝钗、黛玉、凤姐、元春、探春、湘云、妙玉、迎春、惜春、巧姐、李纨、可卿。

编写代码：

import jieba

text = open("红楼梦.txt","r",encoding = 'utf-8').read()

excludes = {"我们","如今","说道","如今","众人"}

words = jieba.lcut(text)

counts = {} #创立一个集合

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))

## 实验思考：

。