杭州电子科技大学

Python与算法设计实验

综合实验报告

**姓名：**

**学号：**

**专业：**

**老师：** 余善恩

**学期：** 2019-2020-2

**日期：** 2020.6

2020年06月

**实验一：数据分析及可视化**

针对提供的name.txt，已知1880个名字中，是按年级排列的，2015-2018级每年级470人。试进行以下进行分析处理：

1. 以饼状图形式（6块）呈现前5大姓氏及其他姓氏（占比1位小数）。

2. 列出名字中前10个最常用字，以柱状图呈现；

3. 以二维格式输出每个年级的前5大姓氏，示例：

2015级： 张 李 王 林 黄

2016级： 李 王 林 黄 张

2017级： 王 林 黄 张 李

2018级： 林 黄 张 李 王

4. 分析第2题里得到的10个常用字在各年级的次数变化，以多折线呈现；

（纵轴次数，横轴年级，每个常用字一根折线）

5．按自己理解，尝试得到其他有用的信息

6. 实验一总结：（问题及解决过程，心得、体会、收获等）

**实验二：GUI设计**

运用TK设计一个简易计算器，要求附源码及最终的运行界面。

界面需要2个图，第1个图显示输入的内容：123\*321-121/11，第2个图显示按等号后的输出结果。

**实验三：爬虫**

选取一个自己感兴趣的内容，对相应网页进行数据抓取并输出。

要求附源码及最终的输出结果。

**实验四：自由实验**

不限主题，用python实现一个小功能

**课程总结：**

对整个课程进行总结，随便写点什么吧

**注意：以上5个模块，每个模块都另起一页**

**实验一：数据分析及可视化**

1. *# coding = utf-8*
2. import os
3. import math
4. import numpy as np
5. import matplotlib.pyplot as plt
6. from mpl\_toolkits.mplot3d import Axes3D
7. from pylab import mpl
8. if \_\_name\_\_ == "\_\_main\_\_" :
9. *# 读入文件*
10. file\_name = open("name.txt", "rt", encoding='UTF-8')
11. lists\_name = [name.rstrip('\n') for name in file\_name.readlines()[1:]]
12. file\_name.close()
13. *# 题1*
14. n = 1880
15. dic = {}
16. for name in lists\_name :
17. *# 狗屎数据有空行*
18. if len(name) > 0 :
19. dic[name[0]] = dic.get(name[0], 0)+1
20. lis = list(dic.items())
21. lis.sort(key=lambda k : k[1], reverse=True)
22. *# 标签*
23. labels = [lis[i][0] for i in range(5)]
24. labels.append('其他')
25. *# 占比*
26. sizes = [lis[i][1]/n\*100 for i in range(5)]
27. sizes.append(100.0)
28. for i in range(5) : sizes[5] -= sizes[i]
29. *# 分离*
30. explode = [0 for i in range(5)]
31. explode.append(0.05)
32. *# 设置中文字体*
33. mpl.rcParams['font.sans-serif']=['SimHei']
34. fig = plt.figure()
35. ax = fig.gca()
36. *# 1位小数*
37. ax.pie(sizes, explode=explode, labels=labels, autopct='%1.1f%%')
38. plt.title('前5大姓氏及其他姓氏')
39. plt.show()
40. *# 题2*
41. dic = {}
42. for name in lists\_name :
43. for i in range(1, len(name)) :
44. dic[name[i]] = dic.get(name[i], 0)+1
45. lis = list(dic.items())
46. lis.sort(key=lambda k : k[1], reverse=True)
47. x\_data = [lis[i][0] for i in range(10)]
48. y\_data = [lis[i][1] for i in range(10)]
49. plt.bar(x=x\_data, height=y\_data)
50. plt.xlabel('字')
51. plt.ylabel('出现次数')
52. plt.title('名字中前10个最常用字')
53. plt.show()
54. *# 题4*
55. dic2 = {}
56. x\_data = [lis[i][0] for i in range(10)]
57. y\_data = []
58. for grade in range(0, 4) :
59. *# 清空字典*
60. for i in range(10) :
61. dic2[lis[i][0]] = 0
62. names = lists\_name[grade\*470:grade\*470+470]
63. for name in names :
64. for i in range(1, len(name)) :
65. if name[i] in dic2 :
66. dic2[name[i]] += 1
67. y\_data = list(dic2.values())
68. plt.plot(x\_data, y\_data, label=str(2015+grade))
69. plt.legend()
70. plt.show()
71. *# 题3*
72. for grade in range(0, 4) :
73. *# represent 2015+grade*
74. names = lists\_name[grade\*470:grade\*470+470]
75. dic = {}
76. for name in names :
77. *# 狗屎数据有空行*
78. if len(name) > 0 :
79. dic[name[0]] = dic.get(name[0], 0)+1
80. lis = list(dic.items())
81. lis.sort(key=lambda k : k[1], reverse=True)
82. print(str(2015+grade)+'级： ', end='')
83. for i in range(5) :
84. print(lis[i][0], end=' ')
85. print()
87. *# 题5*
88. '''
89. 我要这名单有何用?
90. '''
91. cnt = 0
92. for name in lists\_name :
93. if len(name) > 0 :
94. cnt += 1
95. print('已知1880个同学中，共有', 1880-cnt, '个无名氏')
96. *# os.system('pause')*

6．实验一总结

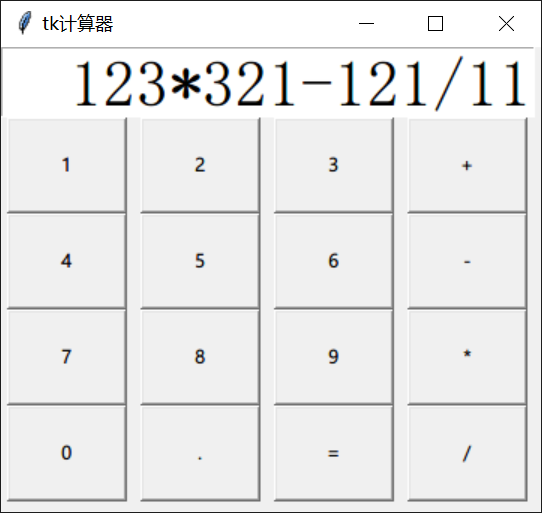
问题:不会画图

解决过程:百度,依样画葫芦

心得:百度真好用

**实验二：GUI设计**

1. *# coding = utf-8*
2. import tkinter as tk
3. import time
4. ls = []
5. def f(x):
6. global ls
7. if x == '=':
8. if '+' in ls or '-' in ls or '\*' in ls or '/' in ls:
9. res = str(eval(enstr.get()))
10. enstr.set(res)
11. ls = [res]
12. else:
13. enstr.set('0')
14. ls = ['0']
15. else:
16. ls.append(x)
17. enstr.set(''.join(ls))
18. if \_\_name\_\_ == "\_\_main\_\_":
19. root = tk.Tk()
20. root.title("tk计算器")
21. root.geometry("360x310")
22. *# root.config(bg='black')*
23. enstr = tk.StringVar()
24. tk.Entry(root, textvariable=enstr, fg='black', font=('仿宋', 32), justify=tk.RIGHT,\
25. width=16).grid(row=0, column=0, rowspan=1, columnspan=4)
26. tk.Button(root, text="1", width=10, height=3, command=lambda:f('1')).grid(row=1, column=0)
27. tk.Button(root, text="2", width=10, height=3, command=lambda:f('2')).grid(row=1, column=1)
28. tk.Button(root, text="3", width=10, height=3, command=lambda:f('3')).grid(row=1, column=2)
29. tk.Button(root, text="+", width=10, height=3, command=lambda:f('+')).grid(row=1, column=3)
30. tk.Button(root, text="4", width=10, height=3, command=lambda:f('4')).grid(row=2, column=0)
31. tk.Button(root, text="5", width=10, height=3, command=lambda:f('5')).grid(row=2, column=1)
32. tk.Button(root, text="6", width=10, height=3, command=lambda:f('6')).grid(row=2, column=2)
33. tk.Button(root, text="-", width=10, height=3, command=lambda:f('-')).grid(row=2, column=3)
34. tk.Button(root, text="7", width=10, height=3, command=lambda:f('7')).grid(row=3, column=0)
35. tk.Button(root, text="8", width=10, height=3, command=lambda:f('8')).grid(row=3, column=1)
36. tk.Button(root, text="9", width=10, height=3, command=lambda:f('9')).grid(row=3, column=2)
37. tk.Button(root, text="\*", width=10, height=3, command=lambda:f('\*')).grid(row=3, column=3)
38. tk.Button(root, text="0", width=10, height=3, command=lambda:f('0')).grid(row=4, column=0)
39. tk.Button(root, text=".", width=10, height=3, command=lambda:f('.')).grid(row=4, column=1)
40. tk.Button(root, text="=", width=10, height=3, command=lambda:f('=')).grid(row=4, column=2)
41. tk.Button(root, text="/", width=10, height=3, command=lambda:f('/')).grid(row=4, column=3)
42. root.mainloop()

**实验三：爬虫**

1. *# coding = utf-8*
2. import time
3. from bs4 import BeautifulSoup
4. if \_\_name\_\_ == "\_\_main\_\_":
5. import requests
6. print('读取杭电党委宣传部(新闻中心)党委教师工作部要闻')
7. print(time.strftime("%Y-%m-%d %H:%M:%S", time.localtime()))
8. url = 'http://www.hdu.edu.cn/news/important'
9. web\_data = requests.get(url)
10. web\_data.encoding = 'utf-8'
11. soup = BeautifulSoup(web\_data.text, 'html.parser')
12. *# print(soup)*
13. for news in soup.select('.title'):
14. if str(news)[20] == '<' : continue
15. print(str(news)[19:-6])

**输出**

读取杭电党委宣传部(新闻中心)党委教师工作部要闻

2020-06-15 22:09:02

我校博士研究生招生网络远程考试圆满结束

我校举行2020届首场线下双选会

学校召开统一战线传达学习全国两会精神暨学校情况通报交流会

学校召开发展党员质量提升工程暨“两个标准”建设推进会

首届“星耀杭电”工程候选人投票环节正式启动

我校与诸暨市教育体育局签订共建研学基地战略合作协议

杭电智库连续得到省领导肯定性批示12件次

省社科联党组书记郭华巍一行来校调研

省委副秘书长、省档案局局长刘芸一行来校调研档案工作

2019年度校十佳大学生评选暨百名优秀大学生云直播举行

学校召开党委会扩大会议传达学习 2020年全国两会精神

我校与6所高职院校专升本联合培养签约仪式举行

我校与西湖区签订全面战略合作协议

第二届“全国创新争先奖”公布 我校徐铭恩教授获奖

**实验四：自由实验**

1. import threading
2. x = []
3. q = []
4. p = []
5. def my\_find(num) :
6. for i in range(len(x)) :
7. if x[i] == num :
8. return i
9. return -1
10. def my\_fun(l, r) :
11. for i in range(l, r) :
12. p[i] = my\_find(q[i])
13. if \_\_name\_\_ == "\_\_main\_\_":
14. thread\_count = 2
15. x = list(map(int, input('输入原始数据:').split()))
16. q = list(map(int, input('输入要查询的数据:').split()))
17. n = len(q)
18. p = [-1 for i in range(n)]
19. t = n//thread\_count
20. threads = []
21. for i in range(thread\_count) :
22. l = i\*t
23. r = max(l+t, n)
24. threads.append(threading.Thread(target=my\_fun, args=[l, r]))
25. threads[i].start()
26. for i in range(thread\_count) :
27. threads[i].join()
28. print('查询的数据的位置分别为:', ' '.join(map(str, p)))

**说明**

用多线程写了个查询程序.对于q中每个数查找在x中的下标,如果不存在则为-1

**测试**

输入原始数据0 1 2 3 4 5

输入要查询的数据2 3 5 9 1

查询的数据的位置分别为 2 3 5 -1 1

**课程总结**

这门神奇的Python课可真是令鄙人受益匪浅，在老师的伟大指引下，我对Python深入浅出，从入门到精通，充分得体会到了Python作为一门时下热门语言的魅力所在，人生苦短，我用Python

通过这门课程我最大的收获还是自学能力，师傅领进门，修行靠个人，仅靠课堂所学终究不够，路漫漫其修远兮，从入门到精通Python还有很长一段路要走，遇到难题要勇于面对，积极在网上搜索相关资料，大数据时代下相信每个人都不难得到海量的学习资料，至于如何筛选和吸收，道阻且长，再接再厉