西南民族大学

**实验报告**

2019 ------2020 学年第 2 学期

课程名称：软件工程课程设计

学 院：计算机科学与技术

专业：计算机科学与技术

年级：2017级

班级：1702班

姓名：毛梦竹

|  |
| --- |
| 西南民族大学学生实验报告  教学单位：计算机科学与技术 实验室名称：BS-329 实验时间： 2020年3月12日  专业：计算机科学与技术 班级：1702班 |
| 实验项目名称：第三次实验 实验成绩： 教师签名：周绪川 |
| 1. 实验目的与背景   目的：编写程序处理数据文件：yq\_in.txt并按要求输出yq\_out.txt文件  要求输入三个参数yq\_in.txt yq\_out.txt 省份，读入文件yq\_in.txt 文件，输出的txt文件命名为yq\_out.txt。输出内容为第三个参数省份的所有城市以及城市患病人数。  要求：  1、生成可执行程序input.exe;  2、读取yq\_in.txt文件，并输出yq\_out.txt文件  **二、材料与方法**  材料：pc机、pycharm编译器  方法：python的pandas、Numpy  GitHub：**<https://github.com/MinusXx/-.git>**  三、实验主要过程与结果  1、设计思路（伪代码）  读取三个外界输入的yq\_in.txt yq\_out.txt 省份三个参数  分别为三个参数命名为StringA、StringB、StringC  省份==String  In= 如果StringA==yq\_in.txt，读取yq\_in.txt文件作为In表格  Group=按照In表格第一列对省份进行分类  Pro=每个省份的索引值  构建City、CityNum两个空列表以及构建Out这个多行两列的空表格  进行循环：  如果Pro中的元素等于省份：  搜集所有相同的元素的下标作为索引ProIndexList  如果k在索引ProIndexList的长度内：  将（k，对应In第二列的城市）元素添加到City中  将（k，对应In第三列的城市患病人数）元素添加到CityNum中  将City、CityNum转换为Serise数组格式  Out=将City\CityNmu拼接起来  转换Out为DataFrame表格格式  输出表格Out为yq\_out.txt文件  2、代码实现  *# -\*- coding:utf-8-\*-* **import** pandas **as** pd **import** numpy **as** np **import** argparse ap=argparse.ArgumentParser() ap.add\_argument(**'StringA'**,type=str,help=**"Para A"**) ap.add\_argument(**'StringB'**,type=str,help=**"Para B"**) ap.add\_argument(**'StringC'**,type=str,help=**"ParaC"**) args=ap.parse\_args() name=args.StringC **if** \_\_name\_\_ == **'\_\_main\_\_'**:  In = pd.read\_table(**'yq\_in.txt'**, header=None)  Group = In.groupby(0).mean() *#* Pro = Group.index  Pro = np.array(Pro)  City=[]  CityNum=[]  Out2 = pd.DataFrame(columns=[0, 1])  **for** i **in** range(len(Pro)):  **if** Pro[i]==name:  ProIndexList=In[In.loc[:,0]==Pro[i]].index.to\_list()  ProIndex=len(ProIndexList)  **for** k **in** range(ProIndex):  City.append(In.loc[ProIndexList[k],1])  CityNum.append(In.loc[ProIndexList[k],2])  City2=pd.Series(City)  CityNum2=pd.Series(CityNum)  out\_put = pd.concat([City2, CityNum2], axis=1)  Out=Out.append(out\_put)  Out = pd.concat([Out, pd.Series([**' '**])])  *# print (Out)* Out.to\_csv(**'yq\_out1.txt'**, sep=**'\t'**, index=False, header=None)  3、将input.py文件转换为input.exe  (1)安装下载pyinstaller    （2）在pycharm中用ALT+F12打开控制台，输入pyinstaller -F -w input.py命令，将input.py文件编译为input.exe文件    （3）成功制成input.exe文件      （4）输入参数    （5）得到浙江省的yq\_out.exe文件     1. 讨论分析：   在本次实验中，我一直遇到的一个问题就是索引不正确，修改了很久都没修改正确，例如在收集表格中与输入省份相匹配的省份下标的时候就需要注意要将收集出来的下标整理合成一个列表就要在后面添加一个to\_list()函数，不然后面在找出城市与城市患病人数时就会出现错误。在添加找到的城市到City数组时也要注意，横坐标是在In表格的第一列。最后的到City与CityNum后也要将其转换为pandas里面的Serise类型，不然在最后拼接时也会出现格式不对不能拼接的错误。整个实验中，要时刻注意每一个函数要处理的类型，我也是在查找了很多资料，从如何输入参数，到如何对每一列相同的元素进行聚合，再到定位索引该使用哪种方法是切分一片数据，再到最后的将得到的列拼接在一起都需要精准，由于python语言的特性是分析数据时的高效所以在使用python语言的时候就要更加了解每一次处理的数据类型，因为这些都是封装好了的所以需要更加小心。  三、教师评阅 |

注：实验报告的内容及格式可由学院根据学科专业特点确定；全校各专业必须使用学校统一封面。