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：于晴 学号：201731102256

同组人：无

|  |
| --- |
| 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-226 实验时间：2020年3月6日  专业：计算机科学与技术 班级：1702班 姓名：于晴 学号：201731102256 |
| 实验项目名称：第二次实验 实验成绩： 教师签名：周绪川 |
| 1. 实验要求 2. 语言要求：C/C++，Java等 3. 实验报告内容：解决思路，流程，代码及结果测试； 4. 源码上传至个人git 5. 实验过程及结果   编写程序处理数据文件：yq\_in.txt  要求：   1. 生成可执行程序yq.exe; 2. 命令行执行：yq yq\_in.txt,得到结果如yq\_out.txt所示   实验过程：  解决思路：首先要完成文档的读取和写入，其次就是处理数据的问题  代码：  **import** java.io.\*;  **public class** Test {  **public static void** main(String[] args) **throws** IOException {  *//按行读中文文档* InputStreamReader reader = **new** InputStreamReader(**new** FileInputStream(**"D:\\SEProject\\yq\_in.txt"**),**"GBK"**);  BufferedReader br = **new** BufferedReader(reader);   *//按行写入中文文档* OutputStreamWriter writer = **new** OutputStreamWriter(**new** FileOutputStream(**"D:\\SEProject\\yq\_out.txt"**),**"GBK"**);  BufferedWriter bw = **new** BufferedWriter(writer);   *//设置空字符串，对格式进行调整* String title=**""**;  String tmp=**""**;  **while**((tmp = br.readLine())!= **null**) {  **if** (tmp.contains(**"待明确地区"**))**continue**;*//可观察到in.txt文件中去除了“待明确地区”的数据，这里进行剔除* String head = tmp.substring(0,3);*//用substring方法分离省份* String msg = tmp.substring(3);*//分离省份之后的详细数据* **if** (title.equals(head)!=**true**){  **if** (title!=**""**){  bw.write(**"\n"**);  }  bw.write(head.substring(0,3)+**"\n"**);  title = head;  }  bw.write(msg.trim()+**"\n"**);  bw.flush();  }  reader.close();  br.close();  writer.close();  bw.close();  } }  实验结果：  yq\_in.txt:    yq\_out.txt:    在这里生成了jar包：  1  个人git链接：https://github.com/YuQ23333/SE-project   1. 分析讨论   在本次的实验中，我使用Java语言来完成本次的作业，我对java中的读取和写入文件的方法又有了更加深入的认识，如何读取带有中文文本的文档和写入带有中文文本的文档，这就需要我们我们用到InputStreamReader（）和OutputStreamWriter（），这两种方法可以成功避免读写中文文档出现乱码的问题。在实验过程中我面对的第二个问题就是如何对数据进行相应的处理，一开始我在输出部分进行调试，但是走进了死胡同，只是单单实现了省份和后面市级及数据的分离，没用办法做到相同省份的去重。之后我又想到是不是会用到大数据处理中的文本分类方法，类似于KNN方法或者朴素贝叶斯算法，但是在网络上搜寻了几篇相关文献后，感觉方向也不对。最后在我不懈的努力下，终于找到了一种简单又有效的方法，直接对字符串进行处理就可以实现目标。这让我也明白了一点，解决问题的方法不止一种，要学会变通的思维，也要学会有耐心和有恒心，在今后完成软件的开发时也要秉承这样的信念最终才能得到让人满意的结果。   1. 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。