# 实验六：文本处理与正则表达式

## 实验目的

1.掌握Vim编辑器的使用，能够使用Vim修改配置文件和编写代码。

2. 掌握常见的文件查看和文本处理命令。

3. 能够使用正则表达式命令grep查找文件中的内容。

## 二. 实验设备及环境

主机：Windows 11

虚拟机：Oracle VM Virtualbox 7.0、Ubuntu 20.04

## 三、实验内容

## 四、实验步骤（实验过程、步骤及实验结果）

### 1. 实验内容一步骤

1. 掌握vim编辑器的用法。

set number # 显示行号

11G # 移动到第11行

yy # 复制当前行

gg # 移动到第一行

0 # 移动到行首

p 或 P # 粘贴（选择 p 或 P）

u # 撤销上一步操作

G # 移动到最后一行

$ # 移动到行尾

22G # 移动到第22行

dd # 删除当前行

:w! # 强制存盘

gg # 移动到首行

i # 进入插入模式，输入“Hello, this is VIM world!”

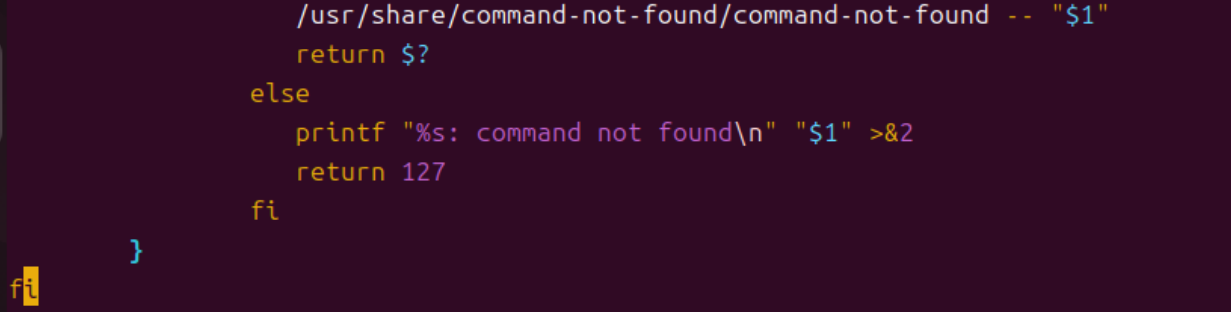
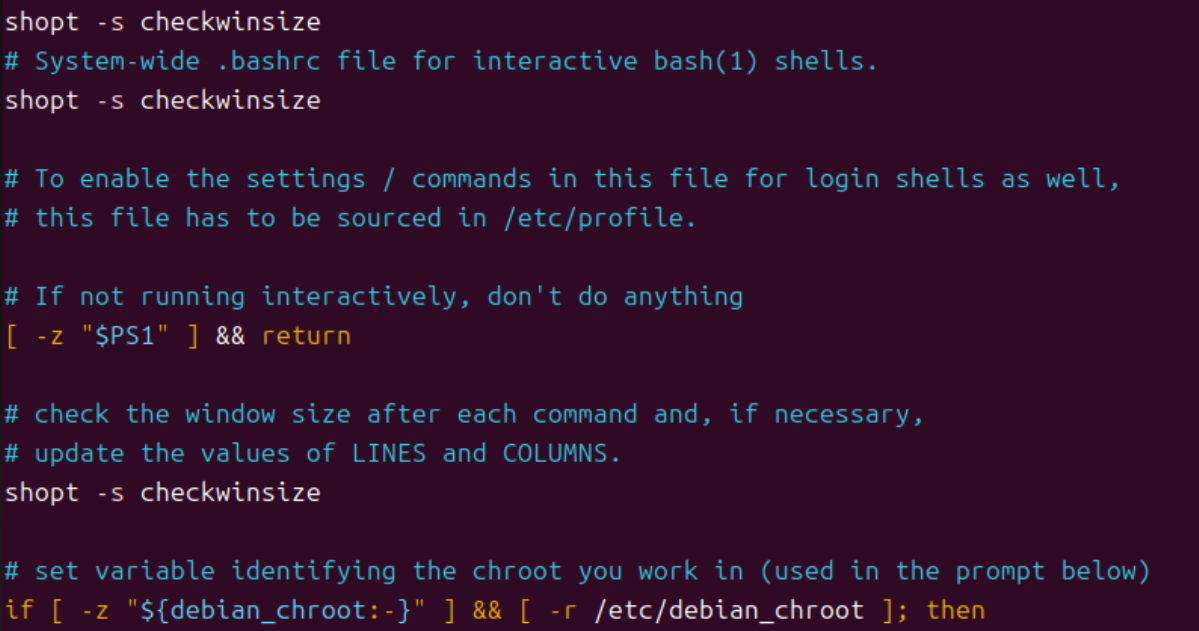
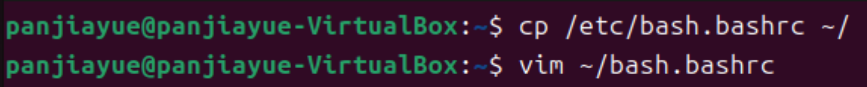
Esc # 返回命令行模式

/shell # 查找字符串“shell”

n/N # 切换到下一个/上一个匹配项

:%s/shell/Shell/g # 全局替换“shell”为“Shell”

:wq 或 :x # 保存并退出



2. 用Vim把课上的综合练习写成一个Shell程序，运行并查看生成的文件。



3. 将以下内容保存到文件 data.txt 中：

13633873456

18723457654

tomcat@139.com

apache@sina.com.cn

tom123@hotmail.com

2022-10-24

98-3-2

2023.11.11

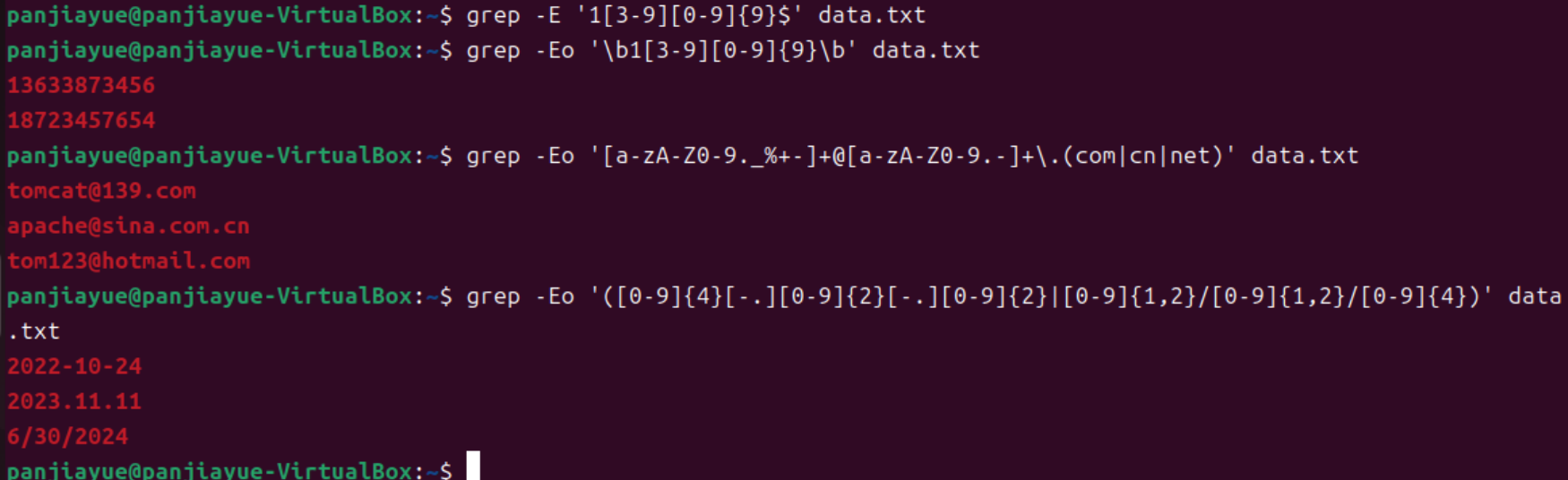
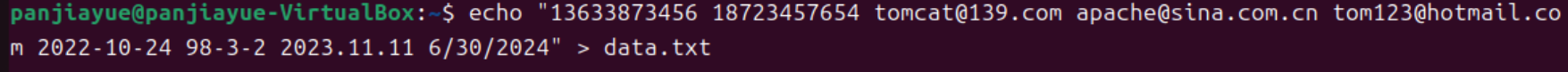
6/30/2024

使用 grep -E 或 egrep 命令分别匹配出：

手机号：要求为11位数字，以1开头的合法手机号。

Email地址：支持常见邮箱域名，如 .com、.cn、.net 等。

日期格式：支持以下三种格式：YYYY-MM-DD（如：2022-10-24）YYYY.MM.DD（如：2023.11.11）MM/DD/YYYY（如：6/30/2024）



4. 将以下数据保存为文件 students.csv：

ID,Name,Age,Grade,Major

101,Tom,20,85,Computer Science

102,Alice,22,90,Mathematics

103,Bob,21,75,Physics

104,Charlie,23,88,Computer Science

105,Daisy,22,82,Mathematics

106,Eve,21,95,Physics

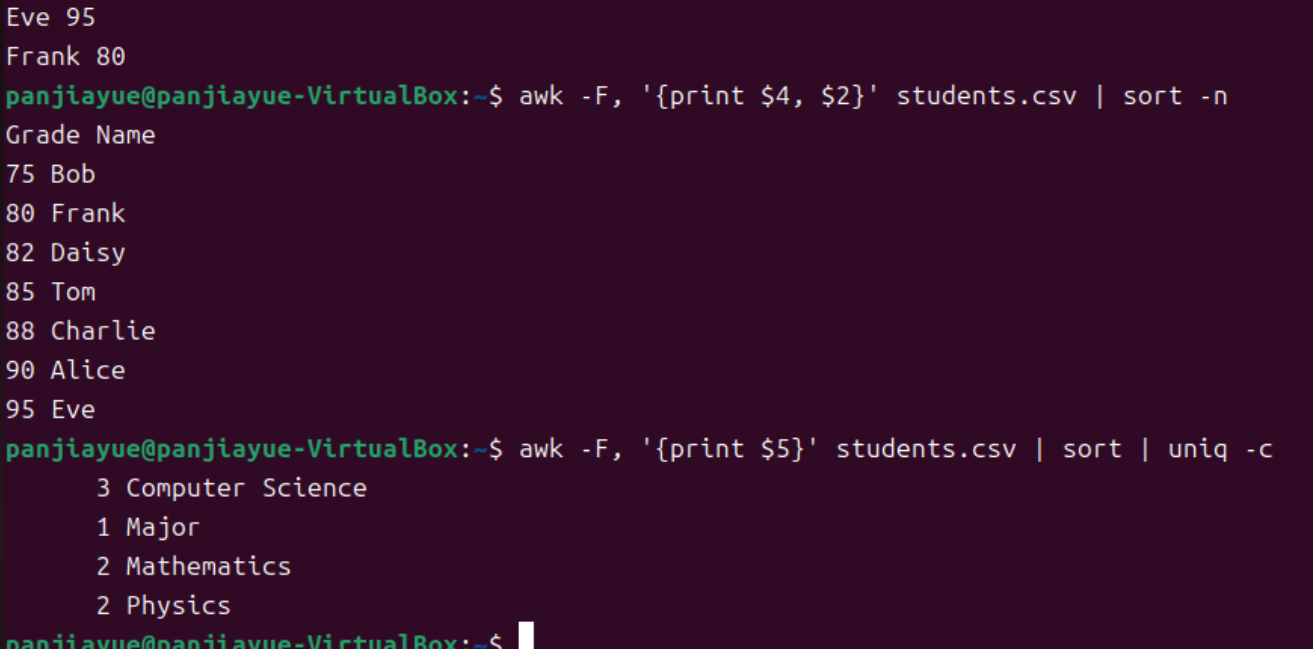
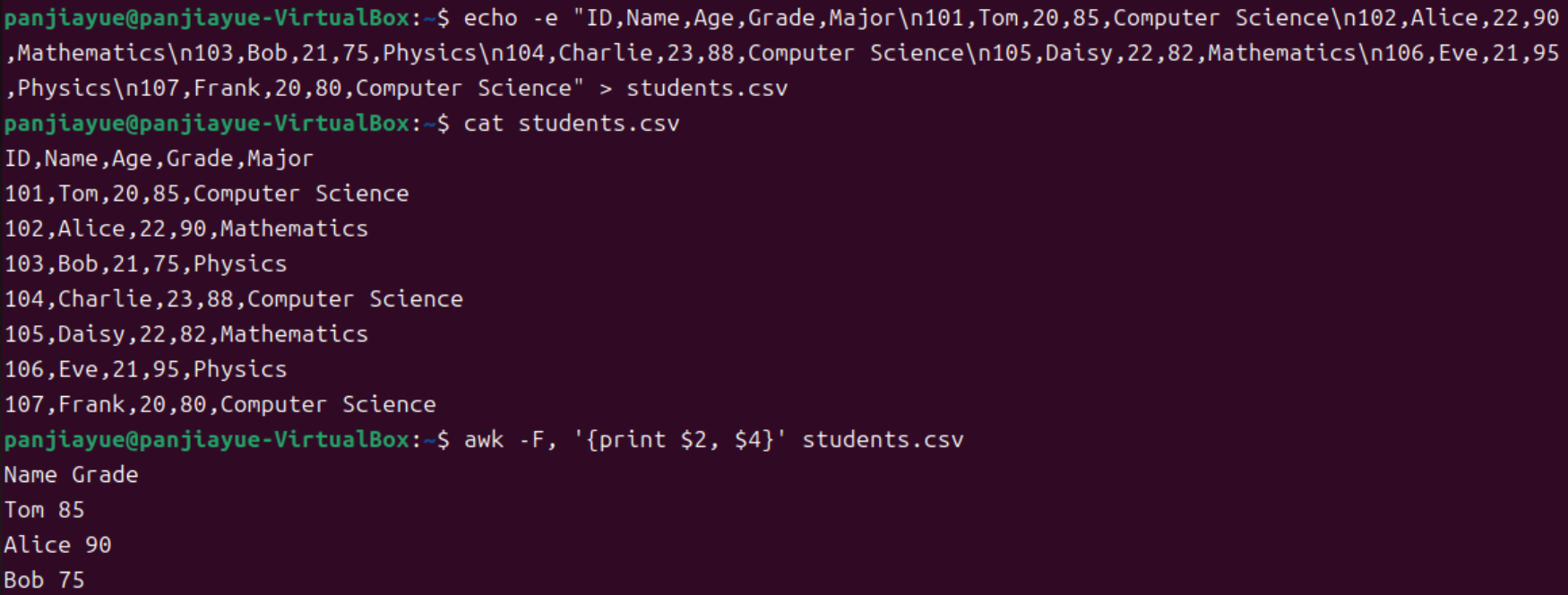
107,Frank,20,80,Computer Science

任务 1：查看文件内容

任务 2：提取学生的姓名和成绩

任务 3：按成绩进行升序排序

任务 4：统计各专业学生人数



## 五、实验中遇到的问题及解决方法

1.第二题文件转换问题