



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Chapter

Bài tập lớn Môn học Cấu trúc Rời rạc dành cho Khoa học Máy tính

on June 24, 2020

Nhóm Sinh viên thực hiện: Nhóm 18
Mã đề 7325

Giảng viên hướng dẫn: Huỳnh Tường Nguyên
Trần Tuấn Anh, Nguyễn Ngọc Lễ
Đại học Bách Khoa, ĐHQG TpHCM



Danh sách thành viên nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu
Mục tiêu
Mô tả dữ liệu
Nội dung bài tập
Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ số thống kê
Tìm hiểu về R

Kết quả thực nghiệm, đánh giá số liệu

Các bước xử lý dữ liệu
Nhóm các câu hỏi kế thừa
Câu hỏi tổng kết

Đánh giá quá trình làm việc

1 Danh sách thành viên nhóm

2 Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu
Mục tiêu
Mô tả dữ liệu
Nội dung bài tập
Nhiệm vụ

3 Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ số thống kê
Tìm hiểu về R

4 Kết quả thực nghiệm, đánh giá số liệu

Các bước xử lý dữ liệu
Nhóm các câu hỏi kế thừa
Câu hỏi tổng kết

5 Đánh giá quá trình làm việc

Nhóm 18 Mã đề 7325

Họ và tên MSSV

Võ Hồng Phúc	1911881
Đinh Lệ Trân	1912267
Dương Ngọc Yến	1912498
Trương Hồng Hoa	1911185
Nguyễn Duy Uyên	1912410
Lương Thị Quỳnh Hương	1911314

Bài tập lớn Môn học
Cấu trúc Rời rạc dành
cho Khoa học Máy tính

Nhóm Sinh viên thực
hiện: Nhóm 18
Mã đề 7325



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu
Mục tiêu
Mô tả dữ liệu
Nội dung bài tập
Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê
Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu
Nhóm các câu hỏi kế thừa
Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Động cơ nghiên cứu

Phân tích & thống kê dữ liệu qua các lần nộp bài của sinh viên nhằm giúp giáo viên đánh giá toàn diện hiệu quả dạy và học online của môn học trong thời gian qua:

- : Đánh giá điểm mạnh, điểm yếu trong hệ thống kiến thức của sinh viên.
- : Xây dựng kế hoạch cải thiện môn học trong thời gian còn lại của môn học.
- : Định hướng và điều chỉnh phương pháp giảng dạy.
- : Bổ sung học liệu cho đề cương trong tương lai.



Mục tiêu

- a Khai phá dữ liệu từ hệ thống kiểm tra online.
- b Bắt đầu với các bài toán thống kê đơn giản từ những dữ liệu được cung cấp.
- c Tìm hiểu ý nghĩa đối với các dữ liệu thực tế.
- d Nâng cao kỹ năng lập trình và giải quyết vấn đề, tích hợp với các hệ thống quản lý và cải thiện chất lượng dạy và học.

Bài tập lớn Môn học
Cấu trúc Rời rạc dành
cho Khoa học Máy tính

Nhóm Sinh viên thực
hiện: Nhóm 18
Mã đề 7325



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Mô tả dữ liệu

Mỗi nhóm nhận được 1 MD là mã đề riêng cho mỗi nhóm (gồm 4 ký số) không trùng nhau, và 24 files chứa thông tin điểm số qua các lần nộp. Các sinh viên sẽ thực hiện yêu cầu đề bài dựa trên các thông tin được cung cấp: Nội dung file:

- 1 tid là mã số bài tập.
- 2 Mã số ID ta gọi là uid là mã số định danh sinh viên nộp bài.
- 3 **Tình trạng**: Đã hoàn thành hoặc chưa bao giờ gửi.
- 4 **Đã bắt đầu vào lúc, Đã hoàn thành**: Thời gian theo dạng “d B Y l:M p” là thời gian bắt đầu và kết thúc làm bài.
- 5 **Thời gian thực hiện**: Khoảng thời gian làm bài
- 6 **Điểm/10**: Tổng số điểm của các quiz.
- 7 $Q.i/1$ là điểm số của bài quiz chỉ 0 hoặc 1.



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Nội dung bài tập

- Xác định số lượng sinh viên trong tập mẫu
- Nhóm câu hỏi liên quan đến điểm số của các sinh viên
- Nhóm câu hỏi liên quan đến số lần nộp bài
- Nhóm câu hỏi liên quan đến thời gian, tần suất nộp bài của các sinh viên.
- Nhóm câu hỏi phân loại sinh viên.
 - Sinh viên **siêng năng**
 - Sinh viên học **đổi phò**
 - Sinh viên **giỏi**
 - Sinh viên **thông minh**
 - Sinh viên **chủ động**
- Xác định phần giao của các loại sinh viên đánh giá ở trên và vẽ biểu đồ thống kê minh họa.



Yêu cầu

- ① Cài đặt R, RStudio
- ② Soạn thảo Latex
- ③ Tạo tài nguyên dùng chung
- ④ Lên kế hoạch nhóm
- ⑤ Quản lí tiến độ
- ① Báo cáo Bài tập lớn
 - ① Lời giải
 - ② Kết quả thực nghiệm
 - ③ Source code R, kết quả tính toán
- ② File logs đánh giá tiến độ, trao đổi

Bài tập lớn Môn học
Cấu trúc Rời rạc dành
cho Khoa học Máy tính

Nhóm Sinh viên thực
hiện: Nhóm 18
Mã đề 7325



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lí dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Một số cơ sở lý thuyết

- 1 Số trung bình (mean)
- 2 Trung bình mẫu (median)
- 3 Tần số (frequency)
- 4 Tần suất và tần suất tích lũy
- 5 Phương sai (variance)
- 6 Độ lệch chuẩn (standard deviation)
- 7 Độ méo lệch (skewness)
- 8 Độ nhọn (kurtosis)
- 9 Tứ phân vị

Bài tập lớn Môn học
Cấu trúc Rời rạc dành
cho Khoa học Máy tính

Nhóm Sinh viên thực
hiện: Nhóm 18
Mã đề 7325



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

R là một ngôn ngữ lập trình và môi trường phần mềm dành cho tính toán và đồ họa thống kê.

Ngôn ngữ R được sử dụng rộng rãi để phát triển phần mềm thống kê và phân tích dữ liệu.

RStudio là IDE (Integrated Development Environment), một phần mềm cho phép truy cập trực tiếp vào phần mềm R, đồng thời hỗ trợ rất nhiều tính năng nâng cao, giúp cho việc quản lý và phân tích dữ liệu với R được hiệu quả hơn.



Danh sách thành viên nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu
Mục tiêu
Mô tả dữ liệu
Nội dung bài tập
Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm, đánh giá số liệu

Các bước xử lý dữ liệu
Nhóm các câu hỏi kế thừa
Câu hỏi tổng kết

Đánh giá quá trình làm việc

Câu lệnh và hàm cơ bản trong R

```
library(xlsx)
library(readxl)
library(dplyr)
library(lubridate)
library(utf8)
library(e1071)
library(ggplot2)
c <- read_xlsx("C01007_TV_HK192-Quiz
2.1-diem.xlsx")
ggplot(df_2l,aes(NoS)) +
geom_histogram(binwidth = 0.5, fill = "blue",
col = "black") +
xlab("Số lần nộp bài") +
ylab("Số lượng sinh viên") +
ggtitle("Phổ theo số lần nộp bài của các sinh
viên")
```

Bài tập lớn Môn học
Cấu trúc Rời rạc dành
cho Khoa học Máy tính

Nhóm Sinh viên thực
hiện: Nhóm 18
Mã đề 7325



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Các bước xử lý dữ liệu

Bài tập lớn Môn học
Cấu trúc Rời rạc dành
cho Khoa học Máy tính

Nhóm Sinh viên thực
hiện: Nhóm 18
Mã đề 7325



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

1 Thực hiện khai báo các thư viện, đính kèm các file

```
(xlsx)
library(readxl)
library(dplyr)
library(lubridate)
library(utf8)
library(e1071)
library(ggplot2)
c <- read_xlsx("C01007_TV_HK192-Quiz
2.1-diem.xlsx")
```

Các bước xử lý dữ liệu

- 1 Thực hiện khai báo các thư viện, đính kèm các file
- 2 Xử lý số liệu của bảng: Đổi tên cột, chuyển đổi cú pháp

```
colnames(c) <- c("ID", "Status", "Stime", "Etime",  
"Time", "Diem", "Q1", "Q2", "Q3", "Q4", "Q5", "Q6",  
"Q7", "Q8", "Q9", "Q10")  
c$Diem=as.numeric(gsub(",", ".", gsub("\\.", "",  
c$Diem)))  
c$Q1=as.numeric(gsub(",", ".", gsub("\\.", "",  
c$Q1)))  
c$Q2=as.numeric(gsub(",", ".", gsub("\\.", "",  
c$Q2)))  
c$Q3=as.numeric(gsub(",", ".", gsub("\\.", "",  
c$Q3)))  
...  
)
```



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Các bước xử lý dữ liệu

- 1 Thực hiện khai báo các thư viện, đính kèm các file
- 2 Xử lý số liệu của bảng: Đổi tên cột, chuyển đổi cú pháp
- 3 Tạo dataframe để chọn lọc, sắp xếp dữ liệu các sinh viên.

```
sapxep_ID=arrange(c, ID) #Tạo một dataframe sắp xếp dữ liệu
ID theo thứ tự tăng dần
n=nrow(c) #số hàng = số lần nộp + 1 IDlist=c()
IDlist <- c(IDlist, sapxep_ID[["ID"]][1])
for( i in 1:(n-2)){
  if(sapxep_ID[["ID"]][i]!=sapxep_ID[["ID"]][i+1])
  IDlist <- c(IDlist, sapxep_ID[["ID"]][i+1])
}
```

*Thông thường, ta sử dụng vòng lặp for, toán tử so sánh ==
... để sắp xếp thông tin*



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kể thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Các bước xử lý dữ liệu

- 1 Thực hiện khai báo các thư viện, đính kèm các file
- 2 Xử lý số liệu của bảng: Đổi tên cột, chuyển đổi cú pháp
- 3 Tạo dataframe để chọn lọc, sắp xếp dữ liệu các sinh viên.
- 4 Trích xuất và sử dụng dữ liệu đã qua xử lý

```
View(IDlist)
n_sv=length(IDlist) #Số SV nộp bài
n_sv #Xuất ra kết quả trên màn hình
```

ID	Status	Stime	Etime	Time	Diem	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
1	Đã hoàn thành	29 March 2020 9:03 AM	29 March 2020 9:05 AM	2 phút 15 giây	9.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	0
2	Đã hoàn thành	29 March 2020 9:25 AM	29 March 2020 9:26 AM	2 phút 52 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
3	Đã hoàn thành	29 March 2020 9:35 AM	29 March 2020 9:40 AM	4 phút 53 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
4	Đã hoàn thành	29 March 2020 10:10 AM	29 March 2020 10:13 AM	3 phút	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
5	Đã hoàn thành	29 March 2020 11:02 AM	29 March 2020 11:17 AM	14 phút 57 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
6	Đã hoàn thành	29 March 2020 11:26 AM	29 March 2020 11:29 AM	3 phút 10 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
7	Đã hoàn thành	29 March 2020 11:51 AM	29 March 2020 11:54 AM	3 phút 4 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
8	Đã hoàn thành	29 March 2020 2:14 PM	29 March 2020 2:18 PM	3 phút 32 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
9	Đã hoàn thành	29 March 2020 3:39 PM	29 March 2020 3:44 PM	5 phút 29 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
10	Đã hoàn thành	29 March 2020 5:07 PM	29 March 2020 5:15 PM	8 phút 22 giây	9.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	0
11	Đã hoàn thành	29 March 2020 5:24 PM	29 March 2020 5:24 PM	42 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
12	Đã hoàn thành	29 March 2020 5:36 PM	29 March 2020 5:39 PM	2 phút 43 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
13	Đã hoàn thành	29 March 2020 7:39 PM	29 March 2020 7:45 PM	5 phút 12 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
14	Đã hoàn thành	29 March 2020 11:18 PM	29 March 2020 11:23 PM	5 phút 20 giây	9.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	0
15	Đã hoàn thành	30 March 2020 8:23 AM	30 March 2020 8:30 AM	7 phút 15 giây	8.5	0	1	1.0	1.0	1.0	1	0.5	1.0	1	1



Thông tin về các file dữ liệu

Câu 1: Tạo một dataframe sắp xếp dữ liệu ID theo thứ tự tăng dần

ID	Status	Stime	Etime	Time	Diem	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
1 1913014	Đã hoàn thành	29 March 2020 9:03 AM	29 March 2020 9:05 AM	2 phút 15 giây	9.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	0
2 1914807	Đã hoàn thành	29 March 2020 9:25 AM	29 March 2020 9:28 AM	2 phút 52 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
3 1015540	Đã hoàn thành	29 March 2020 9:35 AM	29 March 2020 9:40 AM	4 phút 53 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
4 1915541	Đã hoàn thành	29 March 2020 10:10 AM	29 March 2020 10:13 AM	3 phút	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
5 1913186	Đã hoàn thành	29 March 2020 11:02 AM	29 March 2020 11:17 AM	14 phút 57 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
6 1911591	Đã hoàn thành	29 March 2020 11:26 AM	29 March 2020 11:29 AM	3 phút 10 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
7 1912190	Đã hoàn thành	29 March 2020 11:51 AM	29 March 2020 11:54 AM	3 phút 4 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
8 1911704	Đã hoàn thành	29 March 2020 2:14 PM	29 March 2020 2:18 PM	3 phút 32 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
9 1915329	Đã hoàn thành	29 March 2020 3:39 PM	29 March 2020 3:44 PM	5 phút 29 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
10 1911186	Đã hoàn thành	29 March 2020 5:07 PM	29 March 2020 5:15 PM	8 phút 22 giây	9.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	0
11 1911186	Đã hoàn thành	29 March 2020 5:24 PM	29 March 2020 5:24 PM	42 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
12 1912041	Đã hoàn thành	29 March 2020 5:36 PM	29 March 2020 5:39 PM	2 phút 43 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
13 1910666	Đã hoàn thành	29 March 2020 7:39 PM	29 March 2020 7:45 PM	5 phút 12 giây	10.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	1
14 1914003	Đã hoàn thành	29 March 2020 11:18 PM	29 March 2020 11:23 PM	5 phút 20 giây	9.0	1	1	1.0	1.0	1.0	1	1.0	1.0	1	0
15 1914802	Đã hoàn thành	30 March 2020 8:23 AM	30 March 2020 8:30 AM	7 phút 15 giây	8.5	0	1	1.0	1.0	1.0	1	0.5	1.0	1	1

```
sapxep_ID=arrange(c, ID)
n=nrow(c) #số hàng = số lần nộp + 1 IDlist=c()
IDlist <- c(IDlist, sapxep_ID[["ID"]][1])
for( i in 1:(n-2)){
  if(sapxep_ID[["ID"]][i]!=sapxep_ID[["ID"]][i+1])
}
```



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Thông tin về các file dữ liệu

Bài tập lớn Môn học
Cấu trúc Rời rạc dành
cho Khoa học Máy tính

Nhóm Sinh viên thực
hiện: Nhóm 18
Mã đề 7325



Câu 2: Tạo một dataframe sắp xếp dữ liệu điểm số theo thứ tự tăng dần

```
[1] 9.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 9.0 10.0 10.0 10.0 9.0 8.5 9.5 10.0 10.0
[19] 10.0 10.0 10.0 10.0 10.0 9.5 10.0 9.5 10.0 9.0 10.0 9.0 10.0 9.5 10.0 10.0 10.0 10.0 9.0
[37] 10.0 10.0 10.0 10.0 10.0 10.0 10.0 3.0 10.0 9.5 10.0 10.0 10.0 9.0 10.0 9.5 10.0 10.0 10.0
[55] 9.5 10.0 10.0 9.0 10.0 9.0 10.0 10.0 7.5 9.0 10.0 10.0 10.0 10.0 10.0 9.5 10.0 8.0 10.0
[73] 10.0 10.0 10.0 10.0 9.0 10.0 9.0 9.5 10.0 9.5 10.0 9.5 10.0 10.0 9.5 9.5 10.0 9.0
[91] 9.0 10.0 10.0 7.0 10.0 8.5 9.5 10.0 9.0 10.0 7.0 9.0 9.5 9.5 10.0 10.0 9.0 10.0
[109] 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 9.0 10.0 10.0 9.5 10.0 7.0 8.5 10.0
[127] 9.0 9.5 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 9.5
[145] 9.5 10.0 9.5 9.5 10.0 8.5 9.5 9.5 9.5 10.0 10.0 7.5 9.5 9.0 10.0 10.0 10.0 10.0
[163] 10.0 9.5 10.0 10.0 10.0 10.0 9.5 7.5 10.0 10.0 9.5 10.0 9.5 9.0 10.0 10.0 8.5 9.0
[181] 10.0 5.5 10.0 9.5 10.0 10.0 10.0 9.5 10.0 8.5 9.5 10.0 8.5 10.0 10.0 10.0 5.0 8.0
[199] 9.0 10.0 7.5 10.0 9.0 9.5 10.0 9.0 10.0 10.0 10.0 10.0 9.0 10.0 9.5 10.0 10.0 10.0
[217] 10.0 10.0 10.0 9.0 10.0 10.0 8.5 10.0 9.5 8.5 10.0 10.0 10.0 9.0 8.0 10.0 9.0 10.0
[235] 9.0 10.0 10.0 10.0 10.0 10.0 10.0 9.5 9.5 10.0 10.0 10.0 9.5 9.5 10.0 9.5 10.0 9.5
[253] 10.0 10.0 9.5 10.0 10.0 9.5 10.0 10.0 9.5 10.0 9.5 10.0 9.5 10.0 10.0 7.5 9.5 10.0
[271] 9.5 9.5 10.0 10.0 8.5 10.0 10.0 9.0 10.0 8.5 10.0 10.0 10.0 10.0 9.0 10.0 10.0 10.0
[289] 10.0 10.0 8.5 10.0 10.0 10.0 7.5 9.0 9.0 10.0 10.0 10.0 8.5 10.0 10.0 10.0 9.0 10.0
[307] 10.0 8.0 10.0 10.0 10.0 10.0 10.0 9.0 10.0 10.0 9.5 10.0 10.0 8.5 10.0 9.5 10.0 9.5
[325] 10.0 9.5 10.0 10.0 10.0 10.0 10.0 9.0 10.0 9.0 10.0 9.5 9.0 10.0 10.0 10.0 10.0 10.0
[343] 10.0 10.0 10.0 10.0 10.0 10.0 7.5 9.0 10.0 10.0 10.0 10.0 9.5 10.0 10.0 10.0 10.0 9.5
[361] 10.0 8.0 10.0 10.0 6.5 9.5 10.0 9.5 10.0 10.0 10.0 10.0 10.0 9.5 10.0 10.0 0.0 10.0
[379] 10.0 10.0 10.0 10.0 10.0 10.0 9.5 10.0 10.0 10.0 9.5 10.0 10.0 10.0 10.0 10.0 9.0 10.0
[397] 10.0 9.0 10.0 8.5 10.0 9.0 10.0 9.5 9.5 10.0 10.0 10.0 3.0 10.0 10.0 10.0 9.5 9.5
[415] 10.0 10.0 10.0 10.0 8.5 9.5 10.0 4.5 9.0 9.0 9.5 10.0 9.0 10.0 10.0 10.0 10.0 10.0
[433] 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 10.0 0.0
```

Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Thông tin về các file dữ liệu

Câu 3: Thu thập dữ liệu từ hai vector *IDlist* và *numOfSubs* rồi tạo một dataframe sắp xếp dữ liệu ID theo thứ tự tăng dần

```
numOfSubs=c() #Vector đếm số lần nộp bài
dem=1
f=1 #Biến đếm số lần nộp bài
for( i in 1:(n-2)){
  if(sapxep_ID[["ID"]][i]!=sapxep_ID[["ID"]][i+1]){
    numOfSubs <- c(numOfSubs, f)
    f=1
    if(i==n-2) {
      numOfSubs <- c(numOfSubs, f)
    }
  }
  else{
    f=f+1
    if(i==n-2) {
      numOfSubs <- c(numOfSubs, f)
    }
  }
}
#Tạo một dataframe chứa ID và số lần nộp bài của mỗi sinh viên
New_list <- data.frame(IDlist, numOfSubs)
# Sắp xếp New_list theo thứ tự số lần nộp tăng dần
New_list=arrange(New_list, numOfSubs)
numOfSubs_min=New_list[["numOfSubs"]][1]
numOfSubs_min
```

Bài tập lớn Môn học
Cấu trúc Rời rạc dành
cho Khoa học Máy tính

Nhóm Sinh viên thực
hiện: Nhóm 18
Mã đề 7325



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Thông tin về các file dữ liệu

Câu 4: Tạo một dataframe *New_list_4* chứa các dữ liệu thời gian ở dạng *POSIXct* bằng hàm *dmy_hm()* và 2 vector *time1*, *time2*

```
time1=c() # Vector thời gian nọp lan đầu
time2=c() # Vector thời gian lan cuối
time1<-c(time1, sapply_ID[["Etime"]][1])
for( i in 1:(n-2)){
  if(sapply_ID[["ID"]][i]!=sapply_ID[["ID"]][i+1]){
    time1 <- c(time1, sapply_ID[["Etime"]][i+1])
    time2 <- c(time2, sapply_ID[["Etime"]][i])
    if(i==n-2) {
      time2 <- c(time2, sapply_ID[["Etime"]][i+1])
    }
  }
  else{
    if(i==n-2) {
      time2 <- c(time2, sapply_ID[["Etime"]][i+1])
    }
  }
}
# chuyển đổi dạng thời gian của time1 và time2
time1 <- dmy_hm(time1)
time2 <- dmy_hm(time2)
T=c() # Vector tính thời gian chênh lệch từ nọp đầu và lan cuối
# Hàm difftime thuộc thư viện "lubridate", đo chênh lệch tính bằng giây
T=difftime(time2, time1, units="secs")
New_list_4<-data.frame(IDlist, numOfSubs, time1, time2, T)
View(New_list_4)
```



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Thông tin về các file dữ liệu

Câu 5: Nhóm câu hỏi liên quan điểm tổng hợp, điểm trung bình của sinh viên.

Gọi điểm số lần nộp bài thứ k của sinh viên i với i là uid và $k \in (1, 2, 3, \dots)$. Điểm tổng hợp của sinh viên tính tới lần nộp thứ k là điểm lớn nhất cho bài tập đó mà sinh viên đạt được cho tới lần nộp thứ k , tức là:

$$score_{ik} = \max(s_{i1}, s_{i2}, \dots, s_{ik})$$

Đối với sinh viên nộp ít hơn k lần thì vẫn tính theo công thức với giá trị khuyết xem như là 0. Ta sẽ tạo một ma trận để lưu điểm mỗi lần nộp bài của sinh viên. Tìm điểm tổng hợp của mỗi sinh viên bằng hàm max.

```
k=6
LanNop=matrix('NULL', nrow=n_sv, ncol=k)
i=0
for(l in 1:n_sv){
  for(j in 1:numOfSubs[l]){
    i=i+1
    if(is.na(sapxep_ID[["Diem"]][i])== TRUE{ LanNop[l,j]=0 }
    else{ LanNop[l,j]<-sapxep_ID[["Diem"]][i] }
  }
  for(j in (numOfSubs[l]+1):k){ LanNop[l,j]=0 }
}
colnames(LanNop)<-c("I1" "I2" "I3" "I4" "I5" "I6")
```





Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Câu 1: Xác định số lượng sinh viên trong tập mẫu

- 1 Kế thừa dataframe từ khâu xử lý dữ liệu, ta tiến hành tạo 1 vector để lưu ID của sinh viên.

ID	Name	Matr	Class	Time	Row	SR	GP	Q1	Q2	Q3	Q4	SR	GP	Q1	Q2	Q3	Q4
1	20204	BK-HCM	2019-2020	2019-2020	2019-2020	1	1	1	1	1	1	1	1	1	1	1	1
2	20204	BK-HCM	2019-2020	2019-2020	2019-2020	2	1	1	1	1	1	1	1	1	1	1	1
3	20204	BK-HCM	2019-2020	2019-2020	2019-2020	3	1	1	1	1	1	1	1	1	1	1	1
4	20204	BK-HCM	2019-2020	2019-2020	2019-2020	4	1	1	1	1	1	1	1	1	1	1	1
5	20204	BK-HCM	2019-2020	2019-2020	2019-2020	5	1	1	1	1	1	1	1	1	1	1	1
6	20204	BK-HCM	2019-2020	2019-2020	2019-2020	6	1	1	1	1	1	1	1	1	1	1	1
7	20204	BK-HCM	2019-2020	2019-2020	2019-2020	7	1	1	1	1	1	1	1	1	1	1	1
8	20204	BK-HCM	2019-2020	2019-2020	2019-2020	8	1	1	1	1	1	1	1	1	1	1	1
9	20204	BK-HCM	2019-2020	2019-2020	2019-2020	9	1	1	1	1	1	1	1	1	1	1	1
10	20204	BK-HCM	2019-2020	2019-2020	2019-2020	10	1	1	1	1	1	1	1	1	1	1	1
11	20204	BK-HCM	2019-2020	2019-2020	2019-2020	11	1	1	1	1	1	1	1	1	1	1	1
12	20204	BK-HCM	2019-2020	2019-2020	2019-2020	12	1	1	1	1	1	1	1	1	1	1	1
13	20204	BK-HCM	2019-2020	2019-2020	2019-2020	13	1	1	1	1	1	1	1	1	1	1	1
14	20204	BK-HCM	2019-2020	2019-2020	2019-2020	14	1	1	1	1	1	1	1	1	1	1	1
15	20204	BK-HCM	2019-2020	2019-2020	2019-2020	15	1	1	1	1	1	1	1	1	1	1	1
16	20204	BK-HCM	2019-2020	2019-2020	2019-2020	16	1	1	1	1	1	1	1	1	1	1	1
17	20204	BK-HCM	2019-2020	2019-2020	2019-2020	17	1	1	1	1	1	1	1	1	1	1	1
18	20204	BK-HCM	2019-2020	2019-2020	2019-2020	18	1	1	1	1	1	1	1	1	1	1	1
19	20204	BK-HCM	2019-2020	2019-2020	2019-2020	19	1	1	1	1	1	1	1	1	1	1	1
20	20204	BK-HCM	2019-2020	2019-2020	2019-2020	20	1	1	1	1	1	1	1	1	1	1	1
21	20204	BK-HCM	2019-2020	2019-2020	2019-2020	21	1	1	1	1	1	1	1	1	1	1	1
22	20204	BK-HCM	2019-2020	2019-2020	2019-2020	22	1	1	1	1	1	1	1	1	1	1	1
23	20204	BK-HCM	2019-2020	2019-2020	2019-2020	23	1	1	1	1	1	1	1	1	1	1	1
24	20204	BK-HCM	2019-2020	2019-2020	2019-2020	24	1	1	1	1	1	1	1	1	1	1	1
25	20204	BK-HCM	2019-2020	2019-2020	2019-2020	25	1	1	1	1	1	1	1	1	1	1	1

- 2 Độ dài của vector là số lượng sinh viên trong mẫu
- 3 Kết quả từ R

[1] 322

Kế thừa

Câu 2: Nhóm câu hỏi liên quan đến số lần nộp bài

Kế thừa dataframe từ khâu xử lý dữ liệu, ta tiến hành các phép tính toán như sau

a) Xác định điểm số là điểm tổng của các bài làm với mỗi câu hỏi đơn vị đều có điểm tối đa là 1 điểm. *Sử dụng vòng lặp for và dùng hàm arrange để xác định số lần nộp bài ít nhất*
Điểm số là các phần tử trong cột điểm của New.

```
print(New[["Diem"]])
```

Output file 6:





Câu 2:

b) Xác định điểm số thấp nhất Điểm số thấp nhất là phần tử có giá trị nhỏ nhất của vector chứa điểm - Marklist những sinh viên đã nộp bài. Ta tạo vector Marklist và tìm giá trị nhỏ nhất của Marklist như dưới đây.

```
print(min(Marklist))
```

Output file 6:

```
[1] 3
```

Câu 2:

c) Xác định danh sách các sinh viên có ít nhất một bài có số điểm thấp nhất Ta tạo dataframe df_2c là tập con của c chứa phần tử thỏa điều kiện có điểm bằng điểm số thấp nhất như dưới đây. Danh sách ID của df_2c là danh sách cần xác định.

```
df_2c=subset(c,Diem==min(Marklist)) print(df_2c[["ID"]])
```

Output file 6:

```
[1] "1937019" "1914661"
```



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Câu 2:

d) Xác định phổ theo số lần nộp bài của các sinh viên có điểm số tổng kết ở mức điểm cao thứ k với k cho trước

Ta vẽ phổ theo số lần nộp bài của các sinh viên có điểm số tổng kết ở mức điểm cao thứ k với k cho trước với trục dọc là số lượng sinh viên, trục ngang là số lần làm bài.

Lấy ví dụ với $k=3$.

```
k=3
Submissions % > %
filter(Mark==arrangedMark[nMark-k+1]) % > %
ggplot(aes(NoS)) +
geom_histogram(binwidth = 0.5, fill = "blue", col = "black") +
xlab("Số lần nộp bài") +
ylab("Số lượng sinh viên") +
ggtitle("Phổ theo số lần nộp bài của các sinh viên có điểm số tổng
kết ở mức điểm cao thứ k với k cho trước")
```



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

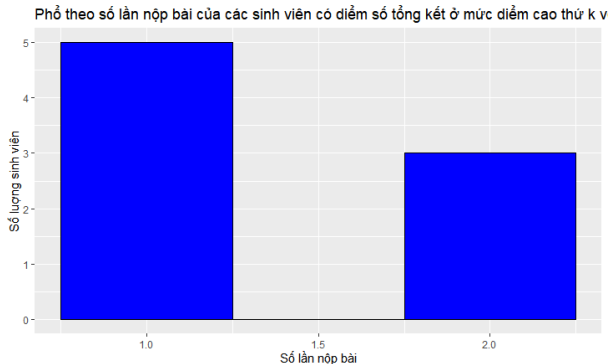
Các bước xử lý dữ liệu

Nhóm các câu hỏi kể thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Câu 2: Output file 6:



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kể thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc



Câu 3: Nhóm câu hỏi liên quan đến số lần nộp bài

Các câu hỏi n, o, p, q

Dựa vào số lượng sinh viên có số lần nộp bài lớn nhất để xác định vị trí của sinh viên đầu tiên có tần suất nộp bài cao thứ nhì (chính là $nrow(df_4f) + 1$)

Xác định danh sách các sinh viên nằm trong nhóm có số lần nộp bài nhiều nhất hoặc nhiều nhì

```
first = subtime[length(subtime)]
first_stus = c()
first_stus = subset(New_list, New_list$numOfSubs == first)
firstandsecond_stus = c()
firstandsecond_stus = rbind(second_stus, first_stus)
View(firstandsecond_stus)
```

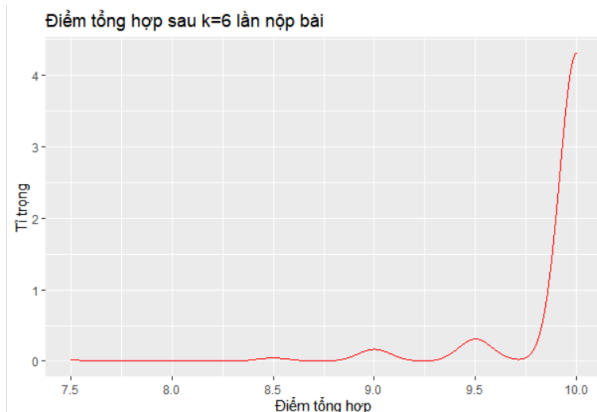
Kể thừa

Câu 5: Nhóm câu hỏi chọn lọc số liệu

a) Ta dùng hàm ggplot để vẽ hàm phân bố điểm

```
p=ggplot(LanNop,aes(x=DiemCuoiCung)) +  
geom_density(alpha=0.5,col="red") + labs(tittle="Điểm tổng hợp  
sau k=6 lần nộp bài", x="Điểm tổng hợp", y="Tỉ trọng")
```

Output của file 6:



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kể thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc



Câu 5: Nhóm câu hỏi chọn lọc số liệu

b) Áp dụng câu a) với k được tính theo công thức sau:

$$MD \bmod 3 + 1$$

Với $k = 3$, ta sẽ lấy điểm từng phần của ba lần nộp bài đầu tiên.
Ta dùng hàm ggplot để vẽ hàm phân bố điểm

```
k=3
LanNop2=data.frame(LanNop2$L1,LanNop2$L2,LanNop2$L3)
DiemCuoiCung2=c()
for(i in 1:n_sv){
  DiemCuoiCung2 <- c(DiemCuoiCung2, max(LanNop2[i,]))
}
LanNop2=data.frame(LanNop2,DiemCuoiCung2) p=ggplot(LanNop2,
aes(x=DiemCuoiCung2)) + geom_density(alpha=0.5,col="red") +
labs(title="Điểm tổng hợp sau k=3 lần nộp bài", x="Điểm tổng
hợp", y="Tỉ trọng")
```

Câu 5: Nhóm câu hỏi chọn lọc số liệu

c) Tính các giá trị TB_k vẽ biểu đồ thể hiện sự thay đổi của các giá trị trung bình này với sự thay đổi của k .

Chọn $k = 6$, ta lưu điểm trung bình sau mỗi lần nộp của tất cả sinh viên trên ma trận tbk . Dùng hàm `mean` để tính trung bình điểm mà các sinh viên đạt được

```
L=c(1,2,3)
tb1=mean(LaNop$L1)
tb2=mean(LaNop$L2)
tb3=mean(LaNop$L3)
tb4=mean(LaNop$L4)
tb5=mean(LaNop$L5)
tb6=mean(LaNop$L6)
tbk=c(tb1,tb2,tb3,tb4,tb5,tb6)
tbk=data.frame(L,tbk)
```



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

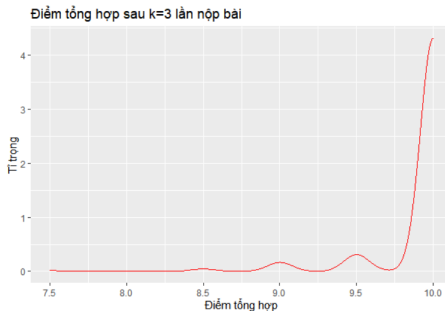
Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Kế thừa

Câu 5:

Output của file 6:



Output file 6:

	L	tb
1	1	9.4708385
2	2	2.9161491
3	3	0.6785714
4	4	0.0310559
5	5	0.0000000
6	6	0.0000000



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc



Câu 5:

d) Hãy cho biết trung bình điểm số mà các sinh viên đạt được qua bài tập tid_n này là bao nhiêu.

```
TrungBinhDiem=mean(LaNop$DiemCuoiCung)
```

Output file 6:

```
> TrungBinhDiem  
[1] 9.880776
```


Nhóm câu hỏi đánh giá số liệu

Câu 7: Đánh giá sinh viên đối phó

a) Thời gian t_2 phù hợp để xác định sinh viên đối phó là sinh viên có thời gian nộp lần đầu sau $2/3$ khoảng thời gian giữa thời điểm sinh viên đầu tiên và thời điểm sinh viên nộp lần cuối

b) Ta tạo vector ID, solannop, landaunop để lưu lại thông tin của tất cả các sinh viên có trong c:

Kết quả thu được như sau:



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Kể thừa

Câu 5: File 6:

```
> ID
[1] 1511191 1613010 1812257 1812477 1812478 1813096 1813528 1813681 1814096 1814518 1814611 1820028 1852443 1910006
[15] 1910032 1910038 1910060 1910076 1910094 1910101 1910110 1910113 1910123 1910137 1910198 1910202 1910224 1910238
[29] 1910265 1910276 1910298 1910339 1910346 1910347 1910351 1910402 1910409 1910473 1910563 1910565 1910620 1910643
[43] 1910644 1910650 1910663 1910666 1910735 1910865 1910892 1910916 1910984 1911000 1911015 1911044 1911056 1911058
[57] 1911066 1911105 1911110 1911136 1911185 1911186 1911207 1911217 1911262 1911283 1911285 1911296 1911314 1911363
[71] 1911441 1911456 1911478 1911520 1911530 1911561 1911565 1911569 1911591 1911594 1911650 1911704 1911736 1911796
[85] 1911837 1911841 1911878 1911881 1911900 1911907 1911931 1911975 1912041 1912046 1912056 1912084 1912123 1912184
[99] 1912190 1912237 1912267 1912288 1912371 1912384 1912386 1912410 1912457 1912463 1912522 1912523 1912526 1912539
[113] 1912579 1912594 1912602 1912675 1912676 1912677 1912683 1912700 1912705 1912713 1912715 1912749 1912761 1912798
[127] 1912811 1912817 1912912 1912916 1912954 1912958 1912966 1912980 1913014 1913021 1913026 1913032 1913040 1913045
[141] 1913048 1913075 1913094 1913102 1913114 1913123 1913167 1913186 1913218 1913228 1913241 1913254 1913260 1913261
[155] 1913268 1913306 1913334 1913336 1913341 1913354 1913355 1913356 1913380 1913386 1913396 1913418 1913419 1913424
[169] 1913430 1913433 1913446 1913457 1913464 1913467 1913560 1913566 1913599 1913609 1913621 1913629 1913651 1913652
[183] 1913678 1913695 1913713 1913729 1913756 1913758 1913763 1913764 1913775 1913817 1913828 1913832 1913844 1913917
[197] 1913918 1913944 1913949 1913990 1914003 1914011 1914022 1914038 1914047 1914052 1914054 1914055 1914064 1914078
[211] 1914079 1914084 1914093 1914121 1914126 1914210 1914220 1914227 1914232 1914291 1914310 1914316 1914352 1914384
[225] 1914405 1914424 1914472 1914474 1914477 1914641 1914651 1914659 1914661 1914674 1914677 1914685 1914697 1914698
[239] 1914704 1914713 1914720 1914738 1914763 1914768 1914802 1914806 1914807 1914830 1914845 1914864 1914878 1914880
[253] 1914881 1914900 1914914 1914979 1915016 1915040 1915063 1915071 1915076 1915083 1915094 1915130 1915133 1915146
[267] 1915161 1915251 1915268 1915275 1915294 1915323 1915329 1915350 1915351 1915378 1915439 1915442 1915470 1915473
[281] 1915474 1915482 1915486 1915520 1915544 1915551 1915557 1915562 1915570 1915598 1915650 1915651 1915676
[295] 1915667 1915745 1915775 1915787 1915795 1915822 1915865 1915866 1915882 1915903 1915905 1915919 1915928 1915931
[309] 1915939 1915940 1915953 1915982 1915983 1915988 1915991 1916022 1916061 1916067 1917007 1936024 1937019 1937024 1937064
```

	ID	solannop
1	1511191	1
2	1613010	2
3	1812257	1
4	1812477	2
5	1812478	1
6	1813096	2
7	1813528	1
8	1813681	1
9	1814096	2

Showing 1 to 10 of 322 entries, 2 total columns

Bài tập lớn Môn học
Cấu trúc Rời rạc dành
cho Khoa học Máy tính

Nhóm Sinh viên thực
hiện: Nhóm 18
Mã đề 7325



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kể thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Kế thừa

Câu 5:

Tạo data7 là một data frame con của data chứa dữ liệu của những ID có thời gian lần đầu nộp sau t2. Thêm vào data7 cột điểm của từng ID:

```
data7=subset(data,landauop > t2)
for(...)
data7 = cbind(data7, diem)
data7 = filter(data7, !is.na(diem))
```

Kết quả thu được data7 như sau:

File 6:



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Câu 11: Xác định phần giao của các loại sinh viên đánh giá (từ câu 6 đến câu 10) và vẽ biểu đồ thống kê minh họa.

Tạo g79 chứa ID của các sinh viên vừa thông minh vừa đôi phó và n79 là số lượng sinh viên vừa thông minh vừa đôi phó:

```
g79=c()  
k=0  
for(i in 1:length(data7$ID7)){  
  for(j in 1:length(ID9)){  
    if(data7$ID7[i]==ID9[j]){  
      k=k+1  
      g79[k]=ID9[j]  
    }  
  }  
}  
n79=length(g79)
```



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Câu 5: Kết quả được như sau:
File 6:

```
> g79
[1] 1813528 1813681 1820028 1910006 1910060 1910198 1910265 1910563 1910663 1910735 1911110 1911285 1911530 1911565
[15] 1911594 1911975 1912384 1912526 1912705 1912713 1913040 1913045 1913167 1913419 1913430 1913446 1913457 1913560
[29] 1913629 1913678 1913990 1914011 1914232 1914477 1914651 1914661 1914685 1914713 1914830 1915161 1915650 1915667
[43] 1915787 1915866 1915931 1915983
> n79
[1] 46
```



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kể thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc



Câu 5:

Tạo g710 chứa ID của các sinh viên vừa chủ động vừa đối phó và n710 là số lượng sinh viên vừa chủ động vừa đối phó.

Tạo g910 chứa ID của các sinh viên vừa chủ động vừa thông minh và n910 là số lượng sinh viên vừa chủ động vừa thông minh.

Tạo g7910 chứa ID của các sinh viên vừa chủ động vừa thông minh vừa đối phó và n7910 là số lượng sinh viên vừa chủ động vừa thông minh vừa đối phó.

Kể thừa

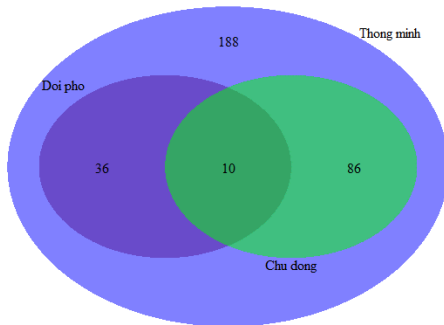
Câu 5:

Vẽ biểu đồ ven minh họa phần giao của các loại sinh viên như sau:

```
grid.newpage()
draw.triple.venn(area1=sosvdoipho,area2=sosinhvienthongminh,area3=sosinhvienchudong,n12=n7
pho","Thong minh","Chu dong"))
```

Ta được các biểu đồ:

File 6:



Bảng phân chia công việc

Bài tập lớn Môn học
Cấu trúc Rời rạc dành
cho Khoa học Máy tính

Nhóm Sinh viên thực
hiện: Nhóm 18
Mã đề 7325



Kết quả phân chia công việc		
Thành viên	Công việc	Ghi chú
Hoa	Bài 1,5,báo cáo	
Trân	Bài 2	
Yến	Bài 3	
Uyên	Bài 4	
Hương	Bài 7,9	
Phúc	Thuyết trình, slide, file logs	
	Bài tập bonus 10,11,12	

Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Quá trình làm việc

Thời gian	Thành viên	Công việc	Ghi chú
14/6	Uyên Hoa	Gửi lên code R tính toán số lần nộp bài Tạo form và lên nội dung Latex	
15/6	Phúc	Gửi link slide thuyết trình	
16/6	Phúc Uyên Hoa Yến Trân Hương	Hoàn thành file logs Gửi lên demo R Bổ sung các khái niệm cho bản Báo cáo, Điều chỉnh bố cục Hoàn thành sơ bộ nhiệm vụ	
17/6	_____ Uyên Yến Trân	Điều chỉnh nội dung Chỉnh sửa phổ điểm Hoàn thiện thuật toán Điều chỉnh thuật toán sau cập nhật	



Danh sách thành viên
nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu
Mục tiêu
Mô tả dữ liệu
Nội dung bài tập
Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê
Tìm hiểu về R

Kết quả thực nghiệm,
đánh giá số liệu

Các bước xử lý dữ liệu
Nhóm các câu hỏi kế thừa
Câu hỏi tổng kết

Đánh giá quá trình làm
việc

Quá trình làm việc			
Thời gian	Thành viên	Công việc	Ghi chú
18/6	_____	Tổng hợp, kiểm tra lần 1	
	Hoa	Chỉnh lỗi báo cáo Gia hạn deadline	
21/6	Hương	Hoàn thành phần code R 7,9,10,11	
23/6	_____	Tập trung toàn nhóm thống nhất bài báo cáo	
	Yên	Cập nhật hình ảnh	
	Hoa	Chỉnh sửa các câu hỏi sử dụng công thức	
	Hương	Thống nhất chọn file code R theo báo cáo	
	Phúc	Thay đổi nội dung slide thuyết trình	
24/6	Trân, Hoa, Hương	Hoàn thành, và sửa lỗi cú pháp các file Latex	

- File hướng dẫn chi tiết, rõ ràng
- Có thêm thời gian tìm hiểu và làm bài tập sau khi thi giữa kì
- Công đoạn thảo luận, bàn bạc được thực hiện online, tận dụng thời gian rảnh của các thành viên
- Các thành viên được lựa chọn công việc
- Nhóm hoạt động tốt, hỗ trợ lẫn nhau trong từng công đoạn
- ...



Danh sách thành viên nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu
Mục tiêu
Mô tả dữ liệu
Nội dung bài tập
Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê
Tìm hiểu về R

Kết quả thực nghiệm, đánh giá số liệu

Các bước xử lý dữ liệu
Nhóm các câu hỏi kế thừa
Câu hỏi tổng kết

Đánh giá quá trình làm việc



Danh sách thành viên nhóm

Tóm tắt yêu cầu đề bài

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nội dung bài tập

Nhiệm vụ

Kiến thức cần thiết

Hiểu biết cơ bản về các chỉ
số thống kê

Tìm hiểu về R

Kết quả thực nghiệm, đánh giá số liệu

Các bước xử lý dữ liệu

Nhóm các câu hỏi kế thừa

Câu hỏi tổng kết

Đánh giá quá trình làm việc

- Các thông tin cần thiết chưa được cập nhật sớm
- Không có nhiều thời gian gặp mặt trực tiếp
- Chưa kịp hoàn thiện toàn bộ bài tập đề ra
- Chưa định hướng rõ các công việc cần làm đối với ngôn ngữ R trong quá trình đề chưa cập nhật
- Bị động trước các thông tin được cung cấp