

# Cấu Trúc Rời Rạc

## Ứng dụng thống kê

*Discrete Structures for Computing* on June 25, 2020

Nguyễn Duy Phước 1914768  
Dương Bá Tình 1915530  
Võ Nguyễn Minh Tú 1915828  
Nguyễn Hữu Lợi 1914047  
Đào Quốc Bảo 1927004

Faculty of Computer Science and Engineering  
University of Technology - VNUHCM

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tình 1915530  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9



#### Mô tả dữ liệu

#### Dọc file và Xử lí dữ liệu

Dọc file

Xử lí dữ liệu

#### Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

## ① Mô tả dữ liệu

## ② Đọc file và Xử lí dữ liệu

Dọc file

Xử lí dữ liệu

## ③ Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

## Mô tả dữ liệu

Đính kèm đề bài tập lớn là 24 files filename.xlsx trong đó chứa thông tin về điểm qua các lần nộp các bài Quiz của các sinh viên trên BKEL.(Báo cáo này sử dụng 4 file). Thông tin bài tập lớn:

- 1 tid là mã số bài tập (gồm có 4 file dữ liệu File1: 16, File2: 19, File3: 23, File4: 20)
- 2 Mã số ID ta gọi là uid là mã số định danh sinh viên nộp bài, mỗi sinh viên có một mã số id duy nhất và không trùng với một mã số id của các sinh viên khác
- 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. Trong đó, “e” là ngày (1..31) “B” tên tháng đầy đủ, “Y” là năm (0..9999), “l” là giờ (01..12), “M” là phút (00–59), “p” chỉ định AM/PM
- 5 Thời gian thực hiện: Khoảng thời gian làm bài
- 6 Điểm/10: Tổng số điểm của các quiz cộng lại thấp nhất là 0, tối đa là 10
- 7 Q.i/1 là điểm số của bài quiz chỉ 0 hoặc 1

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

## Đọc file

Giả sử file nằm trong C:/Users/Asus/Desktop/CTRR/vl  
Tên file là 5.1.xlsx

```
library(dplyr)
library(tidyr)
library(ggplot2)
library(PerformanceAnalytics)
setwd("C:/Users/Asus/Desktop/CTRR/vl")
library(readxl)
df <- read_excel("5.1.xlsx", col_names =
c("uid", "Tinh_trang", "Bat_dau", "Hoan_thanh",
"Thuc_hien", "Diem", "Q1", "Q2", "Q3", "Q4", "Q5",
"Q6", "Q7", "Q8", "Q9", "Q10"))
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

## Xử lý dữ liệu

Chuyển cột 1 và cột 6->16 sang kiểu số(numeric)

Thay na bằng 0 ở các cột 6->16

Loại bỏ dòng cuối cùng tính trung bình

```
df<-df[-1,]  
df$uid<-as.numeric(gsub(",", ".", df$uid))  
df$Diem<-as.numeric(gsub(",", ".", df$Diem))  
df$Q1<-as.numeric(gsub(",", ".", df$Q1))  
df$Q2<-as.numeric(gsub(",", ".", df$Q2))  
df$Q3<-as.numeric(gsub(",", ".", df$Q3))  
df$Q4<-as.numeric(gsub(",", ".", df$Q4))  
df$Q5<-as.numeric(gsub(",", ".", df$Q5))  
df$Q6<-as.numeric(gsub(",", ".", df$Q6))  
df$Q7<-as.numeric(gsub(",", ".", df$Q7))  
df$Q8<-as.numeric(gsub(",", ".", df$Q8))  
df$Q9<-as.numeric(gsub(",", ".", df$Q9))  
df$Q10<-as.numeric(gsub(",", ".", df$Q10))  
for(i in 6:16)  
for(j in 1:nrow(df))  
{  
  if (is.na(df[j,i])) df[j,i] =0  
}  
df<-df %>%  
drop_na(uid)
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 1

### Câu hỏi đề bài

Xác định số lượng sinh viên trong tập mẫu

Code R :

```
SoSV<-n_distinct(df[,1])  
print(SoSV)
```

Kết quả:

File 1: 232

File 2: 182

File 3: 99

File 4: 168

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 a

### Câu hỏi đề bài

Xác định điểm số là điểm tổng của các câu hỏi với mỗi câu hỏi đơn vị đều có điểm tối đa là 1 điểm

Code R :

```
Diemtong_full1<-sum(filter(df,df[,6]==10)[,6])
```

Kết quả:

File 1: 1930

File 2: 1790

File 3: 840

File 4: 1400

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 b

### Câu hỏi đề bài

Xác định điểm số thấp nhất (có nộp bài)

Code R :

```
a<-filter(df,df[,2]=="Đã_hoàn_thành")  
diem_min<-min(a[,6])
```

Kết quả:

File 1: 0

File 2: 0

File 3: 2

File 4: 0

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 2 c

### Câu hỏi đề bài

Xác định danh sách các sinh viên có ít nhất một bài nộp có số điểm thấp nhất

Code R :

```
c<-filter(df[,1],df[,6]==diem_min)
dsdiem_min<-distinct(c[,1])
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 c

Kết quả:

File 1:

	uid
1	1914661
2	1911900

File 2:

	uid
1	1914661

File 3:

	uid
1	1812477

File 4:

	uid
1	1914661
2	1914424

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tĩnh 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

[Đọc file và Xử lý dữ liệu](#)

Đọc file

Xử lý dữ liệu

[Giải quyết bài toán](#)

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 d

### Câu hỏi đề bài

Xác định phổ theo số lần nộp bài của các sinh viên có ít nhất một bài nộp có số điểm thấp nhất

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 d

Code R :

```
c<-semi_join(df,c)
max1<-0
n<-table(c[,1])
for(i in 1:n_distinct(c[,1])){
  if(max1<n[[i]][1])max1=n[[i]][1]
}
x<-NULL
y<-NULL
for(j in 1:max1){
  value<-0
  y<-c(y,j)
  for(i in 1:n_distinct(c[,1])){
    if(n[[i]][1]==j)value<-value+1}
  x<-c(x,value)
}
barplot(x,
main="Phổ theo số lần nộp bài các sinh viên có ít nhất một bài
có số điểm thấp nhất",
xlab="Số lần",
ylab="Số sinh viên",
names.arg=y,
col="darkred")
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

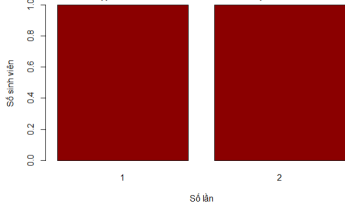
# Giải quyết bài toán

## Bài 2 d

Kết quả:

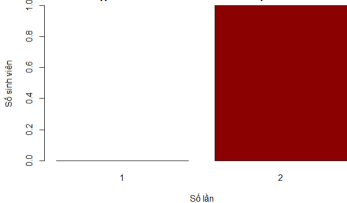
File 1:

Phổ theo số lần nộp bài các sinh viên có ít nhất một bài có số điểm thấp n1



File 2:

Phổ theo số lần nộp bài các sinh viên có ít nhất một bài có số điểm thấp n2



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

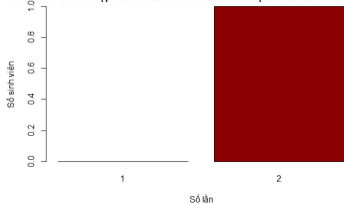
# Giải quyết bài toán

## Bài 2 d

Kết quả:

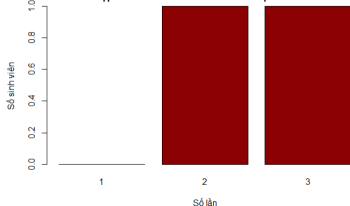
File 3:

Phổ theo số lần nộp bài các sinh viên có ít nhất một bài có số điểm thấp n1



File 4:

Phổ theo số lần nộp bài các sinh viên có ít nhất một bài có số điểm thấp n2



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 e

### Câu hỏi đề bài

Xác định điểm số tổng kết thấp nhất

Code R :

```
dg<-distinct(df[,1])
diem_tk_min<-11
i<-1
for(i in 1:nrow(table(df[,1]))) {
  l<-filter(df,df[,1]==nth(dg[i,1],1))
  if(max(l[,6])<diem_tk_min) diem_tk_min<-max(l[,6])
  diem_tk_min
}
```

Kết quả:

File 1: 0

File 2: 9

File 3: 7

File 4: 7

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 f

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

### Câu hỏi đề bài

Xác định danh sách các sinh viên có điểm số tổng kết thấp nhất

Code R :

```
ds_tk_min<-data.frame()
for(i in 1:nrow(table(df[,1]))){
  l<-filter(df,df[,1]==nth(dg[i,1],1))
  if(max(l[,6])==diem_tk_min)ds_tk_min<-rbind(ds_tk_min,list(dg[[i
,1]]))}
colnames(ds_tk_min)<-c("uid")
ds_tk_min
```



# Giải quyết bài toán

## Bài 2 f

Kết quả:

File 1:

	uid
1	1911900

File 2:

	uid
1	1911900

File 3:

	uid
1	1912579
2	1913386
3	1914864
4	1937024
5	1911569
6	1915486

File 4:

	uid
1	1912041
2	1913268

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

[Dọc file và Xử lý dữ liệu](#)

Dọc file

Xử lý dữ liệu

[Giải quyết bài toán](#)

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 g

### Câu hỏi đề bài

Xác định phổ theo số lần nộp bài của các sinh viên có điểm số tổng kết thấp nhất

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 g

Code R :

```
c<-semi_join(df,ds_tk_min)
max1<-0
n<-table(c[,1])
for(i in 1:n_distinct(c[,1])){
  if(max1<n[[i]][1])max1=n[[i]][1]
}
x<-NULL
y<-NULL
for(j in 1:max1){
  value<-0
  y<-c(y,j)
  for(i in 1:n_distinct(c[,1])){
    if(n[[i]][1]==j)value<-value+1}
  x<-c(x,value)
}
barplot(x,
main="Phổthểsốlầnnộpbàicủacácsinhviêncóđiểm
sốtổngkếtthấpnhất",
      xlab="Sốlần",
      ylab="Sốsinhviên",
      names.arg=y,
      col="darkred")
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

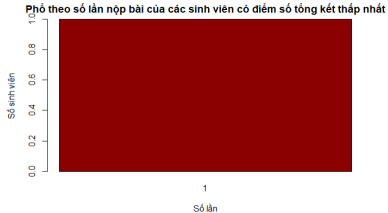
Bài 9

# Giải quyết bài toán

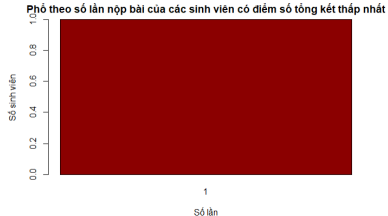
## Bài 2 g

Kết quả:

File 1:



File 2:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 g

Kết quả:

File 3:



File 4:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 h

### Câu hỏi đề bài

Xác định điểm số cao nhất

Code R :

```
diem_max<-max(df[,6])
```

Kết quả:

File 1: 10

File 2: 10

File 3: 10

File 4: 10

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 i

### Câu hỏi đề bài

Xác định danh sách các sinh viên có tối thiểu một bài nộp có số điểm số cao nhất

Code R :

```
c<-filter(df[,1],df[,6]==diem_max)
ds_diemmax<-distinct(c[,1])
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 i

Kết quả:

File 1:

	▲	uid	▼
1		1910666	
2		1914220	
3		1914003	
4		1911314	
5		1913186	
6		1913355	
7		1912084	
8		1912237	

File 2:

	▲	uid	▼
1		1910666	
2		1914220	
3		1914003	
4		1911314	
5		1913186	
6		1913355	
7		1912084	
8		1912237	

File 3:

	▲	uid	▼
1		1915775	
2		1913014	
3		1911591	
4		1912237	
5		1913186	
6		1912817	
7		1912980	
8		1914713	

File 4:

	▲	uid	▼
1		1913014	
2		1912980	
3		1913396	
4		1911314	
5		1912371	
6		1911591	
7		1913355	
8		1915182	

Ứng dụng thông kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 2 j

### Câu hỏi đề bài

Xác định phổ theo số lần nộp bài của các sinh viên có tối thiểu một bài nộp có điểm số cao nhất

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 j

Code R :

```
c<-semi_join(df,c)
max1<-0
n<-table(c[,1])
for(i in 1:n_distinct(c[,1])){
  if(max1<n[[i]][1])max1=n[[i]][1]
}
x<-NULL
y<-NULL
for(j in 1:max1){
  value<-0
  y<-c(y,j)
  for(i in 1:n_distinct(c[,1])){
    if(n[[i]][1]==j)value<-value+1}
  x<-c(x,value)
}
barplot(x,
main="Phổ theo số lần nộp bài các sinh viên có ít nhất
một bài có số điểm cao nhất",
xlab="Số lần",
ylab="Số sinh viên",
names.arg=y,
col="darkred")
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

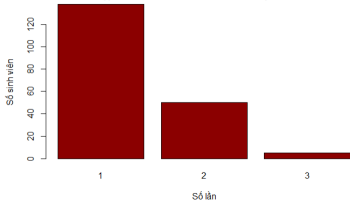
# Giải quyết bài toán

## Bài 2 j

Kết quả:

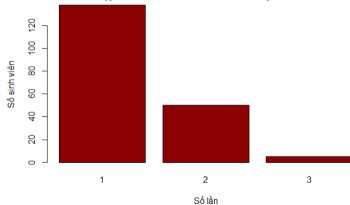
File 1:

Phổ theo số lần nộp bài các sinh viên có ít nhất một bài có số điểm cao nh



File 2:

Phổ theo số lần nộp bài các sinh viên có ít nhất một bài có số điểm cao nh



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

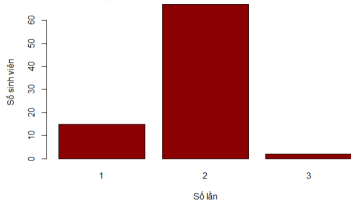
# Giải quyết bài toán

## Bài 2 j

Kết quả:

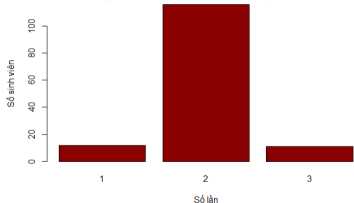
File 3:

Phổ theo số lần nộp bài các sinh viên có ít nhất một bài có số điểm cao nh



File 4:

Phổ theo số lần nộp bài các sinh viên có ít nhất một bài có số điểm cao nh



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 k

### Câu hỏi đề bài

Xác định điểm số tổng kết cao nhất

Code R :

```
diem_tk_max<-0
for(i in 1:nrow(table(df[,1]))) {
  l<-filter(df,df[,1]==nth(dg[i,1],1))
  if(max(l[,6])>diem_tk_max)diem_tk_max<-max(l[,6])}
diem_tk_max
```

Kết quả:

File 1: 10

File 2: 10

File 3: 10

File 4: 10

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 I

### Câu hỏi đề bài

Xác định danh sách các sinh viên có điểm số tổng kết cao nhất

Code R :

```
ds_tk_max<-data.frame()
for(i in 1:nrow(table(df[,1]))){
  l<-filter(df,df[,1]==nth(dg[i,1],1))
  if(max(l[,6])==diem_tk_max)ds_tk_max<-rbind(ds_tk_max,list(dg[[i,1]]))}
colnames(ds_tk_max)<-c("uid")
ds_tk_max
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 I

Kết quả:

File 1:

	▲	uid	▼
1		1910666	
2		1914220	
3		1914003	
4		1911314	
5		1913186	
6		1913355	
7		1912084	
8		1912237	

File 2:

	▲	uid	▼
1		1910666	
2		1914220	
3		1914003	
4		1911314	
5		1913186	
6		1913355	
7		1912084	
8		1912237	

File 3:

	▲	uid	▼
1		1913014	
2		1915775	
3		1911591	
4		1912237	
5		1913186	
6		1912817	
7		1912980	

File 4:

	▲	uid	▼
1		1913014	
2		1912980	
3		1913396	
4		1911314	
5		1912371	
6		1911591	
7		1913355	
8		1915182	

Ứng dụng thông kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 m

### Câu hỏi đề bài

Xác định phổ theo số lần nộp bài của các sinh viên có điểm số tổng kết cao nhất

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 2 m

Code R :

```
c<-filter(df[,1],df[,6]==diem_tk_max)
c<-semi_join(df,c)
max2<-0
n<-table(c[,1])
for(i in 1:n_distinct(c[,1])){
  if(max2<n[[i]][1])max2=n[[i]][1]}
x<-NULL
y<-NULL
for(j in 1:max2){
  value<-0
  y<-c(y,j)
  for(i in 1:n_distinct(c[,1])){
    if(n[[i]][1]==j)value<-value+1}
  x<-c(x,value)}
barplot(x,
main="Phổ theo số lần nộp bài của các sinh viên có điểm",
số_tổng_kết_cao_nhất",
xlab="Số lần",
ylab="Số sinh viên",
names.arg=y,
col="darkred")
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

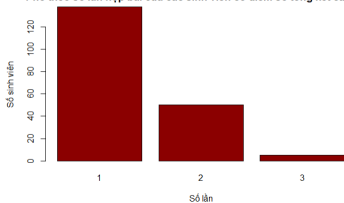
# Giải quyết bài toán

## Bài 2 m

Kết quả:

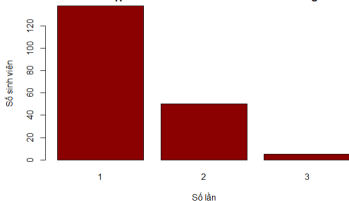
File 1:

Phò theo số lần nộp bài của các sinh viên có điểm số tổng kết cao nhất



File 2:

Phò theo số lần nộp bài của các sinh viên có điểm số tổng kết cao nhất



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

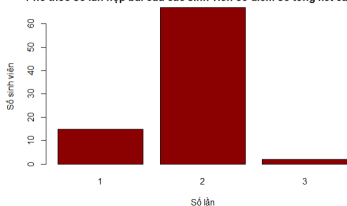
# Giải quyết bài toán

## Bài 2 m

Kết quả:

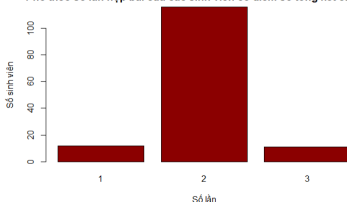
File 3:

Phò theo số lần nộp bài của các sinh viên có điểm số tổng kết cao nhất



File 4:

Phò theo số lần nộp bài của các sinh viên có điểm số tổng kết cao nhất



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 n

### Câu hỏi đề bài

Xác định điểm số trung bình của của các sinh viên trong mẫu (làm tròn còn 0 số lẻ)

Code R :

```
dtb<-round(mean(df$Diem),digits=0)  
dtb
```

Kết quả:

File 1: 9

File 2: 10

File 3: 9

File 4: 9

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 o

### Câu hỏi đề bài

Xác định số lượng sinh viên có điểm số trung bình

Code R :

```
a<-filter(df[,1],df[,6]==dtb)
soSvDiemtb<-n_distinct(a)
```

Kết quả:

File 1: 20

File 2: 176

File 3: 30

File 4: 56

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 p

### Câu hỏi đề bài

Tính trung vị mẫu, cực đại mẫu, cực tiểu mẫu của trên

Code R :

```
Trungvi<-median(df$Diem)
Cucdai<-max(df$Diem)
Cuctieu<-min(df$Diem)
```

Kết quả:

File 1: 10 10 0

File 2: 10 10 0

File 3: 9 10 2

File 4: 9 10 0

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 q

### Câu hỏi đề bài

Hãy đo mức độ phân tán của điểm số (xung quanh giá trị trung bình) của mẫu.

Code R :

```
DLC<-round(sd(df$Diem),1)
```

Kết quả:

File 1: 1.3

File 2: 0.9

File 3: 1.5

File 4: 1.5

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 r

### Câu hỏi đề bài

Tính độ méo lệch (skewness), và độ nhọn (kurtosis) của dữ liệu trong mẫu trên.

Code R :

```
skewnessDiem<-round(skewness(df$Diem),1)  
kurtosisDiem<-round(kurtosis(df$Diem),1)
```

Kết quả:

File 1: -3.9 19.4

File 2: -6.7 58.1

File 3: -1.5 2.2

File 4: -2.4 9.9

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 2 s

### Câu hỏi đề bài

Tính tứ phân vị (quartile) thứ nhất (Q1) và thứ ba (Q3) của mẫu.

Code R :

```
Q1<-quantile(df$Diem,probs=c(0.25))  
Q3<-quantile(df$Diem,probs=c(0.75))
```

Kết quả:

File 1: 9.5 10

File 2: 10 10

File 3: 8 10

File 4: 8 10

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 t

### Câu hỏi đề bài

Xác định số lượng sinh viên có điểm số nằm trong 2 mức điểm cao nhất

Code R :

```
a<-filter(df[,1],df[,6]>=max(df$Diem[df$Diem<max(df$Diem)]))  
SLSV1_2<-n_distinct(a)
```

File 1: 215

File 2: 182

File 3: 91

File 4: 139

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 u

### Câu hỏi đề bài

Xác định phổ theo số lần nộp bài của các sinh viên có điểm số tổng kết ở 2 mức điểm cao nhất

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 u

Code R :

```
x<-NULL
for(i in 1:nrow(table(df[,1]))){
  l<-filter(df,df[,1]==nth(dg[i,1],1))
  if((max(l[,6]) %in%x)==FALSE)
  x<-c(x,max(l[,6]))}
x<-sort(x,decreasing=T)
ds_tk_2<-data.frame()
for(i in 1:nrow(table(df[,1]))){
  l<-filter(df,df[,1]==nth(dg[i,1],1))
  if(max(l[,6])==x[1]|max(l[,6])==x[2])
  ds_tk_2<-rbind(ds_tk_2,list(dg[[i,1]]))}
colnames(ds_tk_2)<-c("uid")
c<-semi_join(df,ds_tk_2)
max2<-0
n<-table(c[,1])
for(i in 1:n_distinct(c[,1])){
  if(max2<n[[i]][1])max2=n[[i]][1]
}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 u

Code R :

```
x<-NULL
y<-NULL
for(j in 1:max2){
  value<-0
  y<-c(y,j)
  for(i in 1:n_distinct(c[,1])){
    if(n[[i]][1]==j)value<-value+1}
  x<-c(x,value)
}
barplot(x,
  main="Phổ theo số lần nộp bài của các sinh viên có điểm
số tổng kết ở 2 mức điểm cao nhất",
  xlab="Số lần", ylab="Số sinh viên", names.arg=y, col="darkred")
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

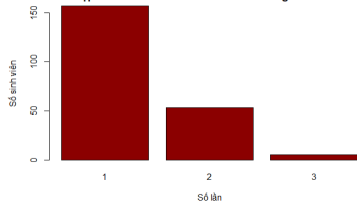
# Giải quyết bài toán

## Bài 2 u

Kết quả:

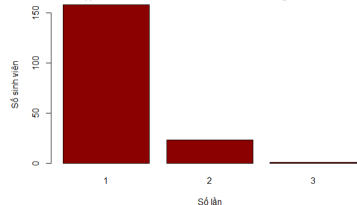
File 1:

ồ theo số lần nộp bài của các sinh viên có điểm số tổng kết ở 2 mức điểm cao



File 2:

ồ theo số lần nộp bài của các sinh viên có điểm số tổng kết ở 2 mức điểm cao



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

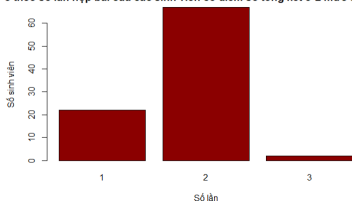
# Giải quyết bài toán

## Bài 2 u

Kết quả:

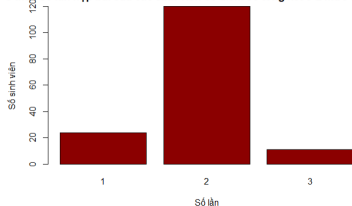
File 3:

ồ theo số lần nộp bài của các sinh viên có điểm số tổng kết ở 2 mức điểm cơ



File 4:

ồ theo số lần nộp bài của các sinh viên có điểm số tổng kết ở 2 mức điểm cơ



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 v

### Câu hỏi đề bài

Xác định số lượng sinh viên có điểm số tổng kết ở mức điểm cao thứ  $k$  với  $k$  cho trước

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 2 v

Code R :

```
readinteger <- function(){
n <- readline(prompt="Enter an integer:")
n <- as.integer(n)
if (is.na(n)){
n <- readinteger()}
return(n)}
k<-readinteger()
x<-NULL
for(i in 1:nrow(table(df[,1]))){
l<-filter(df,df[,1]==nth(dg[i,1],1))
if((max(l[,6]) %in%x)==FALSE)
x<-c(x,max(l[,6]))}
x<-sort(x,decreasing=T)
x
ds_tk_k<-data.frame()
if(!is.na(x[k])){
for(i in 1:nrow(table(df[,1]))){
l<-filter(df,df[,1]==nth(dg[i,1],1))
if(max(l[,6])==x[k])ds_tk_k<-rbind(ds_tk_k,list(dg[[i,1]]))}
colnames(ds_tk_k)<-c("uid")}
n_distinct(ds_tk_k)
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 v

Kết quả:

File 1: 10

File 2: 0

File 3: 6

File 4: 11

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 w

### Câu hỏi đề bài

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

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

**Bài 2**

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 2 w

Code R :

```
if(!is.na(x[k])){  
  c<-semi_join(df,ds_tk_k)  
  max2<-0  
  n<-table(c[,1])  
  for(i in 1:n_distinct(c[,1])){  
    if(max2<n[[i]][1])max2=n[[i]][1]  
    x<-NULL  
    y<-NULL  
    for(j in 1:max2){  
      value<-0  
      y<-c(y,j)  
      for(i in 1:n_distinct(c[,1])){  
        if(n[[i]][1]==j)value<-value+1  
        x<-c(x,value)}  
      barplot(x,  
        main="Phổ theo số lần nộp bài của các sinh viên có  
        điểm số tổng kết cao thứ k",  
        xlab="Số lần",  
        ylab="Số sinh viên",names.arg=y,col="darkred")  
    }else print('Không có điểm số tổng kết cao thứ k nào')
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

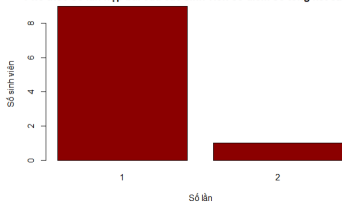
# Giải quyết bài toán

## Bài 2 w

Kết quả:

File 1:

Phổ theo số lần nộp bài của các sinh viên có điểm số tổng kết cao thứ k



File 2:

"không có điểm số tổng kết cao thứ k nào"

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

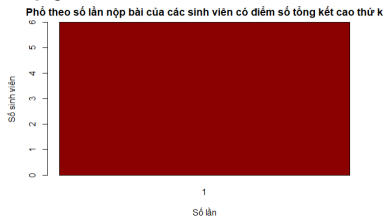
Bài 9

# Giải quyết bài toán

## Bài 2 w

Kết quả:

File 3:



File 4:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 Xử lý số liệu

Tạo dataframe ds\_TongKet là danh sách Sinh Viên với số lần nộp bài của họ

Cột 1: ID(uid)

Cột 2: Số lần nộp(n)

```
ds_TongKet<-count(filter(df,Tinh_trang!="Chưa_bao_giờ_gửi"  
"),filter(df,Tinh_trang!="Chưa_bao_giờ_gửi")[,1])  
dsf<-count(filter(df,Tinh_trang=="Chưa_bao_giờ_gửi")  
,filter(df,Tinh_trang=="Chưa_bao_giờ_gửi")[,1])  
dsf[,2]=0  
ds_TongKet<-union(ds_TongKet,dsf)
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 a

### Câu hỏi đề bài

Xác định số lần nộp bài ít nhất (gồm sv không nộp bài)

Code R :

```
min_SLNB<-min(ds_TongKet[,2])  
print(min_SLNB)
```

Kết quả:

File 1: 0

File 2: 1

File 3: 1

File 4: 0

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 3 b

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

### Câu hỏi đề bài

Xác định danh sách các sinh viên có số lần nộp bài ít nhất

Code R :

```
minSL<-filter(ds_TongKet,ds_TongKet[,2]==min_SLNB)[,1]  
print(minSL)
```

# Giải quyết bài toán

## Bài 3 b

Kết quả:

File 1:

	uid
1	1911900

File 2:

uid	uid	uid	uid	uid	uid	uid	uid	uid
1 1613010	20 1910565	41 1911591	62 1912602	83 1913341	104 1914084	125 1914845	146 1915541	
2 1814096	21 1910620	42 1911704	63 1912683	84 1913355	105 1914093	126 1914864	147 1915557	
3 1814518	22 1910644	43 1911837	64 1912700	85 1913356	106 1914121	127 1914880	148 1915562	
4 1820028	23 1910650	44 1911841	65 1912715	86 1913386	107 1914220	128 1914881	149 1915598	
5 1910006	24 1910663	45 1911881	66 1912749	87 1913419	108 1914227	129 1915040	150 1915822	
6 1910032	25 1910666	46 1911907	67 1912798	88 1913424	109 1914291	130 1915076	151 1915882	
7 1910038	26 1910892	47 1911931	68 1912817	89 1913433	110 1914310	131 1915083	152 1915903	
8 1910060	27 1910916	48 1912041	69 1912912	90 1913446	111 1914316	132 1915130	153 1915919	
9 1910101	28 1910964	49 1912064	70 1912980	91 1913464	112 1914352	133 1915268	154 1915988	
10 1910137	29 1911015	50 1912123	71 1913014	92 1913560	113 1914384	134 1915294	155 1927007	
11 1910202	30 1911110	51 1912164	72 1913026	93 1913678	114 1914405	135 1915323	156 1936024	
12 1910238	31 1911165	52 1912190	73 1913075	94 1913713	115 1914472	136 1915329	157 1937024	
13 1910298	32 1911186	53 1912268	74 1913094	95 1913764	116 1914474	137 1915350	158 1937064	
14 1910339	33 1911262	54 1912371	75 1913102	96 1913775	117 1914641	138 1915351		
15 1910346	34 1911283	55 1912366	76 1913114	97 1913832	118 1914651	139 1915439		
16 1910351	35 1911296	56 1912410	77 1913123	98 1913844	119 1914677	140 1915473		
17 1910402	36 1911314	57 1912463	78 1913241	99 1913949	120 1914685	141 1915474		
18 1910409	37 1911441	58 1912523	79 1913254	100 1914003	121 1914698	142 1915482		
19 1910473	38 1911478	59 1912539	80 1913268	101 1914011	122 1914713	143 1915486		
20 1910565	39 1911565	60 1912579	81 1913334	102 1914054	123 1914738	144 1915520		
21 1910620	40 1911569	61 1912594	82 1913336	103 1914055	124 1914807	145 1915540		

Ứng dụng thông kê

Nguyễn Duy Phức  
1914768  
Dương Bá Tĩnh 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 b

Kết quả:

File 3:

uid	uid	uid
1 1910032	11 1912371	21 1913419
2 1910473	12 1912463	22 1913446
3 1910663	13 1912523	23 1913678
4 1910666	14 1912715	24 1913832
5 1911262	15 1913075	25 1914220
6 1911569	16 1913123	26 1914424
7 1911704	17 1913241	27 1914651
8 1911837	18 1913268	28 1914880
9 1911931	19 1913355	29 1915146
10 1912041	20 1913386	30 1915562

File 4:

uid
1 1914424

Ứng dụng thông kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 c

### Câu hỏi đề bài

Xác định phổ điểm của các sinh viên có số lần nộp bài ít nhất

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 c

Code R :

```
dg<-distinct(ds_TongKet[,1])
y<-slice(df,0:0)
for(i in 1:nrow(dg[,1])){
  l<-filter(df,df[,1]== nth(dg[i,1],1))
  y<-union(y,filter(l,l[,6]==max(l[,6])))
}
y<-distinct(y[,c(1,6)])
Phodiem_minSL<-inner_join(y,filter(ds_TongKet,ds_TongKet[,2]
==min_SLNB),by="uid")
Phodiem_minSL<-count(rename(Phodiem_minSL,u=3),rename(P
hodiem_minSL,u=3)[,2])
ds0<-data.frame(Diem=seq(0, 10, by=0.5),n=rep(0,each=21))
ds0<-anti_join(ds0,Phodiem_minSL, by="Diem")
Phodiem_minSL<-union(ds0,Phodiem_minSL)

BD_Phodiem_minSL<-ggplot(Phodiem_minSL, aes(x = factor(Diem), y = n)
)+ geom_bar(stat = "identity") + theme_minimal() + labs(title
="PhôođiêmosinhoviênocóosôolânoNBoítonhất", x = "Điêm",
y = "Sốosinhoviên") + coord_fixed(ratio = 10/max(Phodiem_minSL[,2]))
print(BD_Phodiem_minSL)
```

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

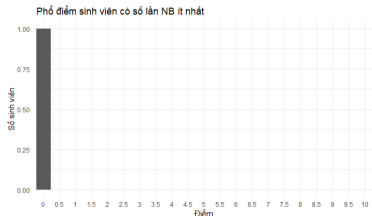
Bài 9

# Giải quyết bài toán

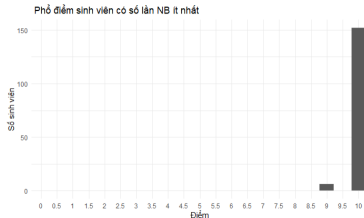
## Bài 3 c

Kết quả:

File 1:



File 2:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

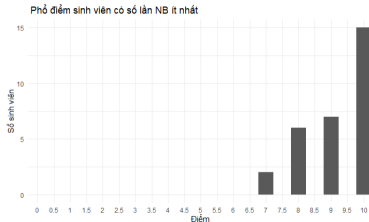
Bài 9

# Giải quyết bài toán

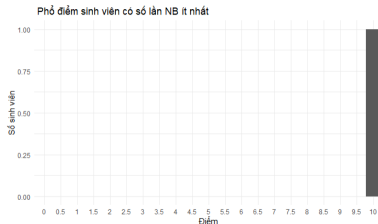
## Bài 3 c

Kết quả:

File 3:



File 4:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tình 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 d

### Câu hỏi đề bài

Xác định số lần nộp bài nhiều nhất

Code R :

```
max_SLNB<-max(ds_TongKet[,2])  
print(max_SLNB)
```

Kết quả:

File 1: 3

File 2: 3

File 3: 3

File 4: 3

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 3 e

### Câu hỏi đề bài

Xác định các sinh viên có số lần nộp bài nhiều nhất

Code R :

```
maxSL<-filter(ds_TongKet,ds_TongKet[,2]==max_SLNB)[,1]  
print(maxSL)
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 e

Kết quả:

File 1:

	uid
1	1910650
2	1911000
3	1913418
4	1915787
5	1916061

File 2:

	uid
1	1914704

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tĩnh 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Dọc file và Xử lý dữ liệu

Dọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 e

Kết quả:

File 3:

	uid
1	1912237
2	1913186

File 4:

	uid
1	1910402
2	1911015
3	1911907
4	1912539
5	1912817
6	1913014
7	1913268
8	1913652
9	1915775
10	1927007

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

### Câu hỏi đề bài

Xác định phổ điểm của các sinh viên có số lần nộp bài nhiều nhất

Code R :

```
Phodiem_maxSL<-inner_join(y,filter(ds_TongKet,ds_TongKet[,  
2]==max_SLNB),by="uid")  
Phodiem_maxSL<-count(rename(Phodiem_maxSL,u  
=3),rename(Phodiem_maxSL,u=3)[,2])  
ds0<-data.frame(Diem=seq(0, 10, by=0.5),n=rep(0,each=21))  
ds0<-anti_join(ds0,Phodiem_maxSL, by="Diem")  
Phodiem_maxSL<-union(ds0,Phodiem_maxSL)  
  
BD_Phodiem_maxSL<-ggplot(Phodiem_maxSL, aes(x = factor(Diem)  
, y = n))+ geom_bar(stat = "identity") + theme_minimal() +  
labs(title = "Phổ điểm sinh viên có số lần nộp bài nhiều nhất", x = "Điểm",  
y = "Số sinh viên") + coord_fixed(ratio = 10/max(Phodiem_maxSL[,2]))  
print(BD_Phodiem_maxSL)
```

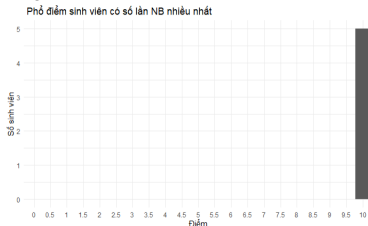


# Giải quyết bài toán

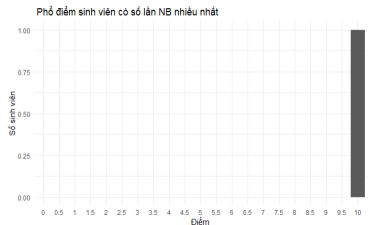
## Bài 3 f

Kết quả:

File 1:



File 2:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

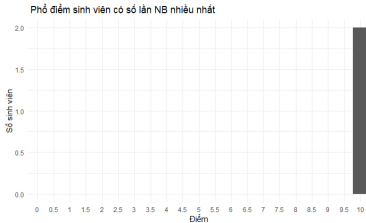
Bài 9

# Giải quyết bài toán

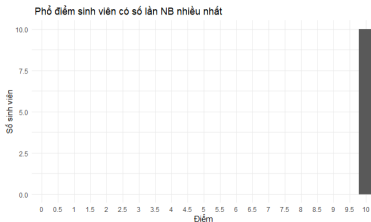
## Bài 3 f

Kết quả:

File 3:



File 4:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 g

### Câu hỏi đề bài

Xác định số lần nộp bài trung bình của của các sinh viên (làm tròn đến số nguyên)

Code R :

```
SL_TB<-round(mean(ds_TongKet$n,0))  
print(SL_TB)
```

Kết quả:

File 1: 1

File 2: 1

File 3: 2

File 4: 2

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 h

### Câu hỏi đề bài

Xác định số lượng sinh viên có số lần nộp trung bình

Code R :

```
SVN_TB<-n_distinct(filter(ds_TongKet,n==SL_TB))  
print(SVN_TB)
```

Kết quả:

File 1: 170

File 2: 158

File 3: 67

File 4: 121

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 3 i

### Câu hỏi đề bài

Xác định phổ theo điểm số của các sinh viên có lần nộp bài trung bình

Code R :

```
Phodiem_tbSL<-inner_join(y,filter(ds_TongKet
,ds_TongKet[,2]==SL_TB),by="uid")
Phodiem_tbSL<-count(rename(Phodiem_tbSL,u=3)
,rename(Phodiem_tbSL,u=3)[,2])
ds0<-data.frame(Diem=seq(0, 10, by=0.5),n=rep(0,each=21))
ds0<-anti_join(ds0,Phodiem_tbSL, by="Diem")
Phodiem_tbSL<-union(ds0,Phodiem_tbSL)

BD_Phodiem_tbSL<-ggplot(Phodiem_tbSL, aes(x = factor(Diem), y = n))
+ geom_bar(stat = "identity") + theme_minimal() + labs(title
="Phổ điểm sinh viên có số lần nộp bài trung bình", x = "Điểm",
y = "Số sinh viên") + coord_fixed(ratio = 10/max(Phodiem_tbSL[,2]))
print(BD_Phodiem_tbSL)
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

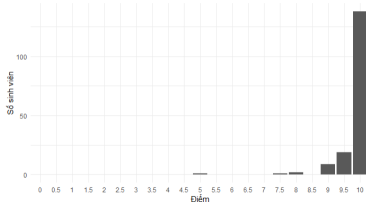
# Giải quyết bài toán

## Bài 3 i

Kết quả:

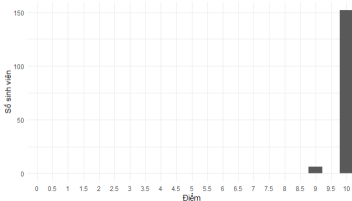
File 1:

Phổ điểm sinh viên có số lần NB trung bình



File 2:

Phổ điểm sinh viên có số lần NB trung bình



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

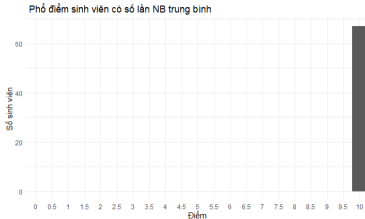
Bài 9

# Giải quyết bài toán

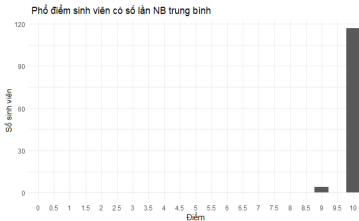
## Bài 3 i

Kết quả:

File 3:



File 4:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tình 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 j

### Câu hỏi đề bài

Tính trung vị mẫu, cực đại mẫu, cực tiểu mẫu của trên.

Code R :

```
Trungvi_SL<-median(Phodiem_tbSL$n)  
Cucdai_SL<-max(Phodiem_tbSL$n)  
Cuctieu_SL<-min(Phodiem_tbSL$n)  
print(Trungvi_SL)  
print(Cucdai_SL)  
print(Cuctieu_SL)
```

Kết quả:

File 1: 0 138 0

File 2: 0 152 0

File 3: 0 67 0

File 4: 0 117 0

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 k

### Câu hỏi đề bài

Hãy đo mức độ phân tán của điểm số (xung quanh giá trị trung bình) của mẫu.

Chính là độ lệch chuẩn:

Code R :

```
DLC_SL<-round(sd(Phodiem_tbSL$n,1))  
print(DLC_SL)
```

Kết quả:

File 1: 30

File 2: 33

File 3: 15

File 4: 26

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 I

### Câu hỏi đề bài

Tính độ méo lệch (skewness), và độ nhọn (kurtosis) của dữ liệu trong mẫu trên.

Code R :

```
skewness_tbSL<-round(skewness(Phodiem_tbSL[,2]),1)
kurtosis_tbSL<-round(kurtosis(Phodiem_tbSL[,2]),1)
print(sknewess_tbSL)
print(kurtosis_tbSL)
```

Kết quả:

File 1: 4.1 15.2

File 2: 4.2 16

File 3: 4.2 16

File 4: 4.2 16

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 m

### Câu hỏi đề bài

Tính tứ phân vị (quartile) thứ nhất (Q1) và thứ ba (Q3) của mẫu.

Code R :

```
Q1_SL<-data.matrix(arrange(Phodiem_tbSL, n))[,2][3]
Q3_SL<-data.matrix(arrange(Phodiem_tbSL, desc(n)))[,2][3]
print(Q1_SL)
print(Q3_SL)
```

Kết quả:

File 1: 0 9

File 2: 0 0

File 3: 0 0

File 4: 0 0

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

### Câu hỏi đề bài

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

Code R :

```
maxSL2<-anti_join(ds_TongKet,maxSL, by = "uid")
maxSL2<-filter(maxSL2,maxSL2[,2]==max(maxSL2[,2]))[,1]
print(maxSL2)
```





# Giải quyết bài toán

## Bài 3 n

Kết quả:

File 1:

uid	uid	uid	uid	uid	uid	uid	uid
1 1812477	9 1910644	17 1912288	25 1913186	33 1913651	41 1914093	49 1914768	
2 1910101	10 1910892	18 1912683	26 1913241	34 1913775	42 1914227	50 1915130	
3 1910202	11 1910916	19 1912700	27 1913306	35 1913828	43 1914472	51 1915146	
4 1910238	12 1911015	20 1912713	28 1913380	36 1913918	44 1914474	52 1915268	
5 1910276	13 1911110	21 1912798	29 1913386	37 1914003	45 1914661	53 1915323	
6 1910346	14 1911441	22 1912817	30 1913396	38 1914022	46 1914674	54 1915442	
7 1910473	15 1912184	23 1913014	31 1913424	39 1914052	47 1914677	55 1915865	
8 1910643	16 1912237	24 1913026	32 1913433	40 1914054	48 1914704	56 1915991	

File 2:

uid	uid
1 1812477	13 1913621
2 1910110	14 1913651
3 1910276	15 1913652
4 1911000	16 1914424
5 1911136	17 1914661
6 1911217	18 1914720
7 1912237	19 1915146
8 1912676	20 1915442
9 1913186	21 1915775
10 1913396	22 1915865
11 1913418	23 1916061
12 1913467	

Ứng dụng thông kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 n

Kết quả:

File 3:

uid	uid	uid	uid	uid	uid	uid
1 1613010	11 1910339	21 1911314	31 1912817	41 1913844	51 1914713	61 1915541
2 1812477	12 1910346	22 1911478	32 1912980	42 1913949	52 1914720	62 1915775
3 1910006	13 1910351	23 1911591	33 1913014	43 1914003	53 1914845	63 1915822
4 1910060	14 1910402	24 1912084	34 1913341	44 1914316	54 1915083	64 1915865
5 1910101	15 1910409	25 1912288	35 1913356	45 1914352	55 1915294	65 1916061
6 1910110	16 1910620	26 1912386	36 1913396	46 1914405	56 1915323	66 1927007
7 1910137	17 1910644	27 1912410	37 1913464	47 1914474	57 1915329	67 1936024
8 1910238	18 1910892	28 1912602	38 1913467	48 1914641	58 1915439	
9 1910276	19 1911015	29 1912700	39 1913560	49 1914685	59 1915473	
10 1910298	20 1911283	30 1912798	40 1913764	50 1914698	60 1915482	

File 4:

uid	uid	uid	uid	uid	uid	uid
1 1812477	19 1910620	37 1912084	55 1913186	73 1914003	91 1914720	109 1915482
2 1814096	20 1910644	38 1912288	56 1913254	74 1914011	92 1914738	110 1915540
3 1820028	21 1910650	39 1912371	57 1913341	75 1914054	93 1914845	111 1915541
4 1910006	22 1910666	40 1912386	58 1913355	76 1914055	94 1914864	112 1915562
5 1910032	23 1910892	41 1912410	59 1913396	77 1914084	95 1914880	113 1915598
6 1910038	24 1910916	42 1912579	60 1913418	78 1914093	96 1914881	114 1915622
7 1910060	25 1911000	43 1912594	61 1913424	79 1914291	97 1915040	115 1915865
8 1910101	26 1911110	44 1912602	62 1913433	80 1914316	98 1915076	116 1915903
9 1910110	27 1911136	45 1912676	63 1913464	81 1914352	99 1915083	117 1915919
10 1910137	28 1911185	46 1912700	64 1913467	82 1914384	100 1915130	118 1915968
11 1910202	29 1911217	47 1912715	65 1913560	83 1914405	101 1915146	119 1916061
12 1910276	30 1911262	48 1912749	66 1913651	84 1914424	102 1915268	120 1936024
13 1910298	31 1911283	49 1912798	67 1913678	85 1914474	103 1915323	121 1937024
14 1910339	32 1911296	50 1912980	68 1913713	86 1914641	104 1915329	
15 1910346	33 1911314	51 1913075	69 1913764	87 1914661	105 1915439	
16 1910409	34 1911441	52 1913094	70 1913832	88 1914685	106 1915442	
17 1910473	35 1911591	53 1913114	71 1913844	89 1914704	107 1915473	
18 1910565	36 1911881	54 1913123	72 1913949	90 1914713	108 1915474	

Ứng dụng thông kê

Nguyễn Duy Phức  
1914768  
Dương Bá Tĩnh 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 o

### Câu hỏi đề bài

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

Code R :

```
maxSL12<-union(maxSL2,maxSL)  
print(maxSL12)
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 o

Kết quả:

File 1:

uid	uid	uid	uid	uid	uid	uid
1 1812477	11 1910916	21 1912798	31 1913424	41 1914093	51 1915146	61 1916061
2 1910101	12 1911015	22 1912817	32 1913433	42 1914227	52 1915268	
3 1910202	13 1911110	23 1913014	33 1913651	43 1914472	53 1915323	
4 1910238	14 1911441	24 1913026	34 1913775	44 1914474	54 1915442	
5 1910276	15 1912184	25 1913186	35 1913828	45 1914661	55 1915865	
6 1910346	16 1912237	26 1913241	36 1913918	46 1914674	56 1915991	
7 1910473	17 1912288	27 1913306	37 1914003	47 1914677	57 1910650	
8 1910643	18 1912683	28 1913380	38 1914022	48 1914704	58 1911000	
9 1910644	19 1912700	29 1913386	39 1914052	49 1914768	59 1913418	
10 1910892	20 1912713	30 1913396	40 1914054	50 1915130	60 1915787	

File 2:

uid	uid
1 1812477	13 1913621
2 1910110	14 1913651
3 1910276	15 1913652
4 1911000	16 1914424
5 1911136	17 1914661
6 1911217	18 1914720
7 1912237	19 1915146
8 1912676	20 1915442
9 1913186	21 1915775
10 1913396	22 1915065
11 1913418	23 1916061
12 1913467	24 1914704

Ứng dụng thông kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 o

Kết quả:

File 3:

uid	uid	uid	uid	uid	uid	uid	uid
1 1613010	11 1910339	21 1911314	31 1912817	41 1913844	51 1914713	61 1915541	
2 1812477	12 1910346	22 1911478	32 1912980	42 1913949	52 1914720	62 1915775	
3 1910006	13 1910351	23 1911591	33 1913014	43 1914003	53 1914845	63 1915822	
4 1910060	14 1910402	24 1912084	34 1913341	44 1914316	54 1915083	64 1915865	
5 1910101	15 1910409	25 1912288	35 1913356	45 1914352	55 1915294	65 1916061	
6 1910110	16 1910620	26 1912386	36 1913396	46 1914405	56 1915323	66 1927007	
7 1910137	17 1910644	27 1912410	37 1913464	47 1914474	57 1915329	67 1936024	
8 1910238	18 1910892	28 1912602	38 1913467	48 1914641	58 1915439	68 1912237	
9 1910276	19 1911015	29 1912700	39 1913560	49 1914685	59 1915473	69 1913186	
10 1910298	20 1911283	30 1912798	40 1913764	50 1914698	60 1915482		

File 4:

uid	uid	uid	uid	uid	uid	uid	uid	uid
1 1812477	19 1910620	37 1912084	55 1913186	73 1914003	91 1914720	109 1915482	128 1913268	
2 1814096	20 1910644	38 1912288	56 1913254	74 1914011	92 1914738	110 1915540	129 1913652	
3 1820028	21 1910650	39 1912371	57 1913341	75 1914054	93 1914845	111 1915541	130 1915775	
4 1910006	22 1910666	40 1912386	58 1913355	76 1914055	94 1914864	112 1915562	131 1927007	
5 1910032	23 1910892	41 1912410	59 1913396	77 1914084	95 1914880	113 1915598		
6 1910038	24 1910916	42 1912579	60 1913418	78 1914093	96 1914881	114 1915822		
7 1910060	25 1911000	43 1912594	61 1913424	79 1914291	97 1915040	115 1915865		
8 1910101	26 1911110	44 1912602	62 1913433	80 1914316	98 1915076	116 1915903		
9 1910110	27 1911136	45 1912676	63 1913464	81 1914352	99 1915083	117 1915919		
10 1910137	28 1911185	46 1912700	64 1913467	82 1914384	100 1915130	118 1915968		
11 1910202	29 1911217	47 1912715	65 1913560	83 1914405	101 1915146	119 1916061		
12 1910276	30 1911262	48 1912749	66 1913651	84 1914424	102 1915268	120 1936024		
13 1910298	31 1911283	49 1912798	67 1913678	85 1914474	103 1915323	121 1937024		
14 1910339	32 1911296	50 1912800	68 1913713	86 1914641	104 1915329	122 1910402		
15 1910346	33 1911314	51 1913075	69 1913764	87 1914661	105 1915439	123 1911015		
16 1910409	34 1911441	52 1913094	70 1913832	88 1914685	106 1915442	124 1911907		
17 1910473	35 1911591	53 1913114	71 1913844	89 1914704	107 1915473	125 1912539		
18 1910565	36 1911681	54 1913123	72 1913949	90 1914713	108 1915474	126 1912817		

Ứng dụng thông kê

Nguyễn Duy Phức  
1914768  
Dương Bá Tĩnh 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 p

### Câu hỏi đề bài

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

Code R :

```
n_maxSL12<-n_distinct(maxSL12)
```

Kết quả:

File 1: 61

File 2: 24

File 3: 69

File 4: 131

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 q

### Câu hỏi đề bài

Xác định phổ theo điểm số của các sinh viên có lần nộp bài nhiều nhất hoặc nhiều nhì

Code R :

```
Phodiem_maxSL12<-count(semi_join(y, maxSL12, by = "uid")
,semi_join(y, maxSL12, by = "uid")[,2])

ds0<-data.frame(Diem=seq(0, 10, by=0.5),n=rep(0,each=21))
ds0<-anti_join(ds0,Phodiem_maxSL12, by="Diem")
Phodiem_maxSL12<-union(ds0,Phodiem_maxSL12)

BD_Phodiem_maxSL12<-ggplot(Phodiem_maxSL12, aes(x =
  factor(Diem), y = n))+ geom_bar(stat = "identity") +
  theme_minimal() + labs(title = "Phổ điểm sinh viên có
  số lần nộp nhiều nhất và nhì", x = "Điểm",
  y = "Số sinh viên") + coord_fixed(ratio = 10/max(Phodiem_maxSL12[,2]))
print(BD_Phodiem_maxSL12)
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

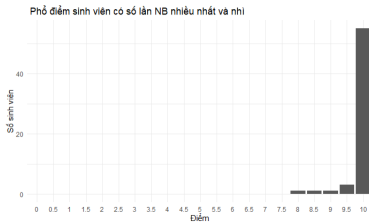
Bài 9

# Giải quyết bài toán

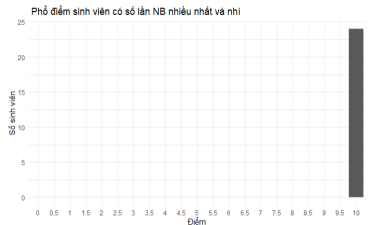
## Bài 3 q

Kết quả:

File 1:



File 2:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9



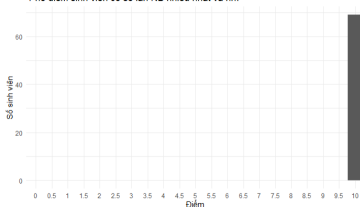
# Giải quyết bài toán

## Bài 3 q

Kết quả:

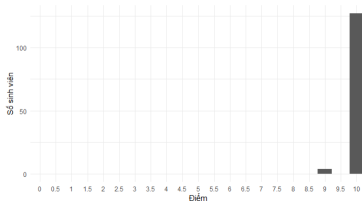
File 3:

Phổ điểm sinh viên có số lần NB nhiều nhất và nhì



File 4:

Phổ điểm sinh viên có số lần NB nhiều nhất và nhì



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 r

### Câu hỏi đề bài

Xác định danh sách các sinh viên nằm trong nhóm một phần ba đầu theo thứ tự số lần nộp bài giảm dần

Code R :

```
ds1_3_SLmax<-arrange(ds_TongKet, desc(n))[1:round  
(nrow(ds_TongKet)/3,0),1]  
print(ds1_3_SLmax)
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 r

Kết quả:

File 1:

uid	uid	uid	uid	uid	uid	uid
1 1910650	12 1910473	23 1912683	34 1913386	45 1914054	56 1915146	67 1910032
2 1911000	13 1910643	24 1912700	35 1913396	46 1914093	57 1915268	68 1910038
3 1913418	14 1910644	25 1912713	36 1913424	47 1914227	58 1915323	69 1910060
4 1915787	15 1910892	26 1912798	37 1913433	48 1914472	59 1915442	70 1910110
5 1916061	16 1910916	27 1912817	38 1913651	49 1914474	60 1915865	71 1910113
6 1812477	17 1911015	28 1913014	39 1913775	50 1914661	61 1915991	72 1910123
7 1910101	18 1911110	29 1913026	40 1913828	51 1914674	62 1613010	73 1910137
8 1910202	19 1911441	30 1913186	41 1913918	52 1914677	63 1814096	74 1910224
9 1910238	20 1912184	31 1913241	42 1914003	53 1914704	64 1814518	75 1910296
10 1910276	21 1912237	32 1913306	43 1914022	54 1914768	65 1820028	76 1910339
11 1910346	22 1912288	33 1913380	44 1914052	55 1915130	66 1910006	77 1910351

File 2:

uid	uid	uid	uid	uid	uid	uid
1 1914704	11 1913396	21 1915442	31 1910038	41 1910402	51 1910916	61 1911441
2 1812477	12 1913418	22 1915775	32 1910060	42 1910409	52 1910984	
3 1910110	13 1913467	23 1915865	33 1910101	43 1910473	53 1911015	
4 1910276	14 1913621	24 1916061	34 1910137	44 1910565	54 1911110	
5 1911000	15 1913651	25 1613010	35 1910202	45 1910620	55 1911185	
6 1911136	16 1913652	26 1814096	36 1910238	46 1910644	56 1911186	
7 1911217	17 1914424	27 1814518	37 1910298	47 1910650	57 1911262	
8 1912237	18 1914661	28 1820028	38 1910339	48 1910663	58 1911283	
9 1912676	19 1914720	29 1910006	39 1910346	49 1910666	59 1911296	
10 1913186	20 1915146	30 1910032	40 1910351	50 1910892	60 1911314	

Ứng dụng thông kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tĩnh 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 r

Kết quả:

File 3:

uid	uid	uid
1 1912237	12 1910298	23 1911314
2 1913186	13 1910339	24 1911478
3 1613010	14 1910346	25 1911591
4 1812477	15 1910351	26 1912084
5 1910006	16 1910402	27 1912288
6 1910060	17 1910409	28 1912386
7 1910101	18 1910620	29 1912410
8 1910110	19 1910644	30 1912602
9 1910137	20 1910892	31 1912700
10 1910238	21 1911015	32 1912798
11 1910276	22 1911283	33 1912817

File 4:

uid	uid	uid	uid	uid	uid	uid
1 1910402	9 1915775	17 1910060	25 1910346	33 1910892	41 1911283	49 1912371
2 1911015	10 1927007	18 1910101	26 1910409	34 1910916	42 1911296	50 1912386
3 1911907	11 1812477	19 1910110	27 1910473	35 1911000	43 1911314	51 1912410
4 1912539	12 1814096	20 1910137	28 1910565	36 1911110	44 1911441	52 1912579
5 1912817	13 1820028	21 1910202	29 1910620	37 1911136	45 1911591	53 1912594
6 1913014	14 1910006	22 1910276	30 1910644	38 1911185	46 1911881	54 1912602
7 1913268	15 1910032	23 1910298	31 1910650	39 1911217	47 1912084	55 1912676
8 1913652	16 1910038	24 1910339	32 1910666	40 1911262	48 1912288	56 1912700

Ứng dụng thông kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tĩnh 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 s

### Câu hỏi đề bài

Xác định số lượng các sinh viên nằm trong nhóm một phần ba đầu theo thứ tự số lần nộp bài giảm dần

Code R :

```
SV_Nopbai1_3<-round(nrow(ds_TongKet)/3,0)  
print(SV_Nopbai1_3)
```

Kết quả:

File 1: 77

File 2: 61

File 3: 33

File 4: 56

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 t

### Câu hỏi đề bài

Xác định phổ theo điểm số của các sinh viên nằm trong nhóm một phần ba đầu theo thứ tự số lần nộp bài giảm dần

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 t

Code R :

```
Phodiem_ds1_3_SLmax<-count(semi_join(y,ds1_3_SLmax , by = "uid")
,semi_join(y,ds1_3_SLmax , by = "uid")[,2])
ds0<-data.frame(Diem=seq(0, 10, by=0.5),n=rep(0,each=21))
ds0<-anti_join(ds0,Phodiem_ds1_3_SLmax, by="Diem")
Phodiem_ds1_3_SLmax<-union(ds0,Phodiem_ds1_3_SLmax)

BD_Phodiem_ds1_3_SLmax<-ggplot(Phodiem_ds1_3_SLmax
, aes(x = factor(Diem), y = n))+ geom_bar(stat = "identity")
+ theme_minimal() + labs(title =
"Phổ điểm sinh viên có số lần NB trong khoảng 1/3 nhiều nhất"
, x = "Điểm",
y = "Số sinh viên") +
coord_fixed(ratio = 10/max(Phodiem_ds1_3_SLmax[,2]))
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

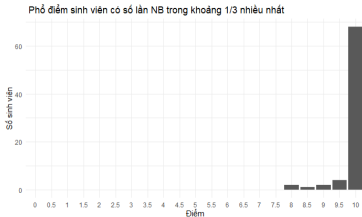
Bài 9

# Giải quyết bài toán

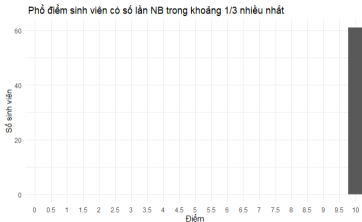
## Bài 3 t

Kết quả:

File 1:



File 2:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9



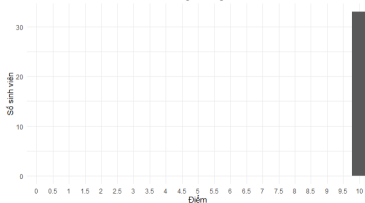
# Giải quyết bài toán

## Bài 3 t

Kết quả:

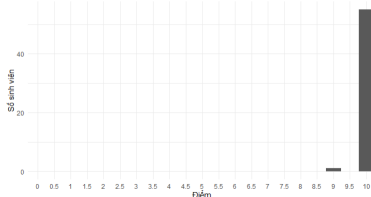
File 3:

Phổ điểm sinh viên có số lần NB trong khoảng 1/3 nhiều nhất



File 4:

Phổ điểm sinh viên có số lần NB trong khoảng 1/3 nhiều nhất



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 u

### Câu hỏi đề bài

Xác định phổ theo điểm số của các sinh viên nằm trong  $k$  nhóm đầu mà mỗi nhóm chứa các sinh viên có cùng số lần nộp bài và các nhóm được sắp xếp theo thứ tự giảm dần của có số lần nộp bài (với  $k$  cho trước).

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

**Bài 3**

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 3 u

Code R :

```
while(TRUE){  
  k<- readline("Nhap gia tri cho k <- ")  
  k<-as.numeric(k)  
  if(!is.na(k)) break  
}
```

```
maxSL_k<-filter(ds_TongKet,ds_TongKet[,2]>=data.matrix  
  (arrange(distinct(ds_TongKet[,2]),desc(n))))[,1][k]][,1]  
Phodiem_maxSL_k<-count(semi_join(y,maxSL_k , by = "uid")  
  ,semi_join(y,maxSL_k , by = "uid")[,2])  
ds0<-data.frame(Diem=seq(0, 10, by=0.5),n=rep(0,each=21))  
ds0<-anti_join(ds0,Phodiem_maxSL_k, by="Diem")  
Phodiem_maxSL_k<-union(ds0,Phodiem_maxSL_k)
```

```
BD_Phodiem_maxSL_k<-ggplot(Phodiem_maxSL_k, aes(x = factor(Diem)  
  ,y = n)) + geom_bar(stat= "identity") + theme_minimal() + labs(title  
  ="Phổ điểm sinh viên có số lần NB trong nhóm đầu", x = "Điểm",  
  y = "Số sinh viên") + coord_fixed(ratio = 10/max(Phodiem_maxSL_k[,2]))
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

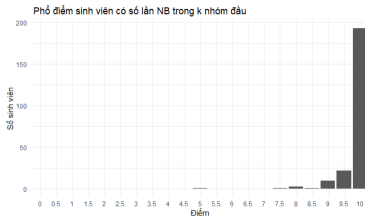
Bài 9

# Giải quyết bài toán

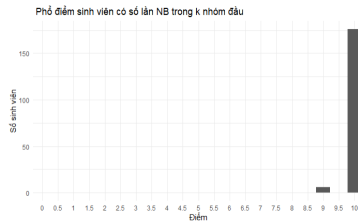
## Bài 3 u

Kết quả(lấy  $k=3$ ):

File 1:



File 2:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tình 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

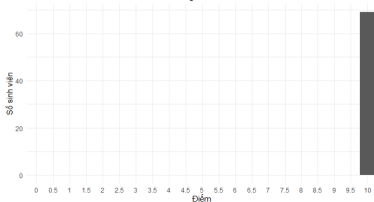
# Giải quyết bài toán

## Bài 3 u

Kết quả(lấy  $k=3$ ):

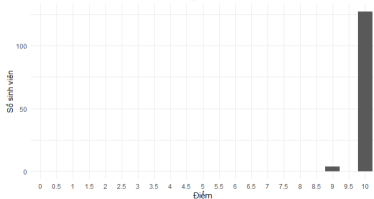
File 3:

Phổ điểm sinh viên có số lần NB trong k nhóm đầu



File 4:

Phổ điểm sinh viên có số lần NB trong k nhóm đầu



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

Chuyển char sang dạng TG ở các cột chỉ TG:

Code R :

```
for (i in 1:nrow(df))
{
  if(!grepl("giờ",df$Thuc_hien[i]))
  {
    if(grepl("phút",df$Thuc_hien[i]))
    {
      df$Thuc_hien[i]<-paste(df$Thuc_hien[i],"0giờ", step =" ")
    }
    else{ df$Thuc_hien[i]<-paste(df$Thuc_hien[i],"0phút_0giờ"
    , step =" ")}}
  df$Bat_dau<-parse_date_time2(df$Bat_dau,orders="dbYIMp")
  df$Hoan_thanh<-parse_date_time2(df$Hoan_thanh,orders="dbYIMp")
  df$Thuc_hien<-difftime(parse_date_time(df$Thuc_hien,order=c(
    "MS", "S", "HMS"))
    ,parse_date_time("0",orders = "S"),units="secs")
  df<-df %>%
  drop_na(Hoan_thanh,Bat_dau)
```



# Giải quyết bài toán

## Bài 4 a

### Câu hỏi đề bài

Với mỗi sinh viên, xác định thời gian dài nhất tính từ lần nộp bài đầu tiên đến lần nộp cuối.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 a

Code R :

```
cau_a <- function(needReturn = FALSE) {  
  if (needReturn == FALSE) cat("Câu a: ", '\n', '\n')  
  studentID <- df[[1]]  
  count <- 1  
  scope <- c(1:(nrow(df)-1))  
  uid <- character()  
  average_vector <- numeric()  
  nSubmit <- 0  
  for (n in df[[5]]) if (n != 0 && n != "") nSubmit <- nSubmit + 1  
  frequence_vector <- numeric()  
  frequency_vector <- numeric()  
  cumulativefre <- 0  
  cumfre_vector <- numeric()  
  for (i in (1:nrow(df))) {  
    maxTime <- 0  
    times <- 0  
    average <- 0  
    frequence <- 0  
    frequency <- 0  
    if (studentID[i] != FALSE && studentID[i] != "") {
```

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 4 a

Code R :

```
if (df[[5]][i] != 0) {
  maxTime <- as.numeric(df[[5]][i])
  prev <- as.numeric(df[[3]][i]) + as.numeric(df[[5]][i])
  prevInTheory <- as.numeric(df[[4]][i])
  totalTime <- 0
  times <- 1
  frequence <- 1
}
count <- count + 1
for (j in (scope + 1)) {
  if (i != nrow(df)) {
if(studentID[i] == studentID[j] && i != j && studentID[i] != FALSE &&
studentID[j] != FALSE && studentID[i] != "" && studentID[j] != "") {
  times <- times + 1
  if (df[[5]][j] != 0) {
    if (maxTime < df[[5]][j]) maxTime <- as.numeric(df[[5]][j])
    if (as.numeric(df[[3]][j]) == prevInTheory) {
      totalTime <- totalTime + as.numeric(df[[5]][j])
      current <- prev + as.numeric(df[[5]][j])
    } else {
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 a

Code R :

```
    } else {  
current <- as.numeric(df[[3]][j]) + as.numeric(df[[5]][j])  
totalTime <- totalTime + (current - prev)  
}  
prev <- current  
prevInTheory <- as.numeric(df[[4]][j])  
frequence <- frequence + 1  
}  
if (i < j) studentID[j] <- FALSE  
}  
}  
}  
if (studentID[i] != FALSE && studentID[i] != "") {  
if (times == 1) totalTime <- as.numeric(df[[5]][i])  
average <- totalTime / times  
frequency <- frequence / nSubmit * 100  
frequency_vector <- c(frequency_vector, frequency)  
uid <- c(uid, studentID[i])  
average_vector <- c(average_vector, average)
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 a

Code R :

```
current <- as.numeric(df[[3]][j]) + as.numeric(df[[5]][j])
totalTime <- totalTime + (current - prev)
}
prev <- current
prevInTheory <- as.numeric(df[[4]][j])
frequency <- frequency + 1
}
if (i < j) studentID[j] <- FALSE
}
}
}
if (studentID[i] != FALSE && studentID[i] != "") {
if (times == 1) totalTime <- as.numeric(df[[5]][i])
average <- totalTime / times
frequency <- frequency / nSubmit * 100
frequency_vector <- c(frequency_vector, frequency)
uid <- c(uid, studentID[i])
average_vector <- c(average_vector, average)
frequency_vector <- c(frequency_vector, frequency)
frequency <- frequency / nSubmit * 100
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 a

Code R :

```
frequency_vector <- c(frequency_vector, frequency)
cumulativefre <- cumulativefre + frequency
cumfre_vector <- c(cumfre_vector, cumulativefre)
}
if (needReturn == FALSE) {
  if (studentID[i] != FALSE && studentID[i] != "" && maxTime != 0) {
    maxTime <- seconds_to_period(maxTime)
  }
  if (maxTime != 0) {
    cat("Thời gian nộp bài dài nhất của sinh viên có mã số",
      , studentID[i], " là",
      sprintf('%02d giờ, %02d phút, %02d giây', hour(maxTime)
      , minute(maxTime), second(maxTime)), '\n')
  } else cat("Sinh viên có mã số", studentID[i], " chưa bao giờ
    nộp bài", '\n')
}
if (studentID[i] != "") studentID[i] <- FALSE
scope <- c(count:(nrow(df)-1))
}
}
if (needReturn == TRUE) return(list(uid,
  average_vector, frequency_vector,
  frequency_vector, cumfre_vector))
```

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 a

Kết quả:

File 1: Danh sách:

1913014

1910666

1914220

1914003

1911314

File 2: Danh sách

1911591

1814518

1911136

1912980

1915562

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Dọc file và Xử lý dữ liệu

Dọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 a

Kết quả:

File 3:Danh sách

1913014

1912980

1913396

1911314

1912371

File 4:Danh sách

1913014

1915775

1914220

1911591

1912237

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Dọc file và Xử lý dữ  
liệu

Dọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 b

### Câu hỏi đề bài

Xác định phổ thời gian làm việc (được tính từ lần nộp bài đầu tiên đến lần nộp cuối) của các sinh viên.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 b

Code R :

```
cau_b <- function() {  
  cat("Câu b:", '\n', '\n')  
  studentID <- df[[1]]  
  count <- 1  
  scope <- c(1:(nrow(df)-1))  
  for (i in (1:nrow(df))) {  
    if (studentID[i] != FALSE && studentID[i] != "") {  
      if (df[[5]][i] != 0) {  
        before <- df[[4]][i]  
        after <- df[[4]][i]  
      }  
    }  
    count <- count + 1  
    for (j in (scope + 1)) {  
      if (i != nrow(df)) {  
        if (studentID[i] == studentID[j] && i != j && studentID[i] != FALSE &&  
            studentID[j] != FALSE && studentID[i] != "" && studentID[j] != "") {  
          if (df[[5]][j] != 0) {  
            after <- df[[4]][j]  
            if (i < j) studentID[j] <- FALSE  
          }  
        }  
      }  
    }  
  }  
}
```

Ứng dụng thông kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 4 b

Code R :

```
    }}}
if (studentID[i] != FALSE && studentID[i] != "") {
  if (df[[5]][i] != 0) {
    cat("Phổ thời gian làm việc của sinh viên có mã số", studentID[i]
      , " là từ ",
      hour(as.POSIXct(before, origin = "1970-01-01")), " giờ",
      minute(as.POSIXct(before, origin = "1970-01-01")), " phút", " đến ",
      hour(as.POSIXct(after, origin = "1970-01-01")), " giờ",
      minute(as.POSIXct(after, origin = "1970-01-01")), " phút", '\n')
  } else cat("Sinh viên có mã số", studentID[i]
    , " chưa báo giờ nộp bài", '\n')
  }
if (studentID[i] != "") studentID[i] <- FALSE
scope <- c(count:(nrow(df)-1))
}
hour <- numeric()
for (i in df[[4]]) if (!is.na(i) && i != "") hour <- c(hour
  , hour(as.POSIXct(i, origin = "1970-01-01")))
pie(table(factor(hour)), main = "PHỔ THỜI GIAN LÀM VIỆC", edges = 1000,
  radius = 1, clockwise = TRUE, init.angle = 90)}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

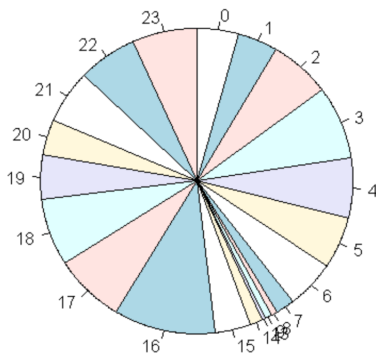
# Giải quyết bài toán

## Bài 4 b

Kết quả:

File 1:

**PHỔ THỜI GIAN LÀM VIỆC**



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

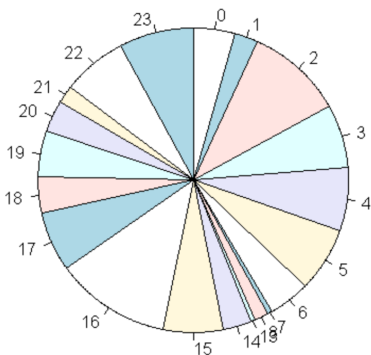
# Giải quyết bài toán

## Bài 4 b

Kết quả:

File 2:

**PHỔ THỜI GIAN LÀM VIỆC**



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

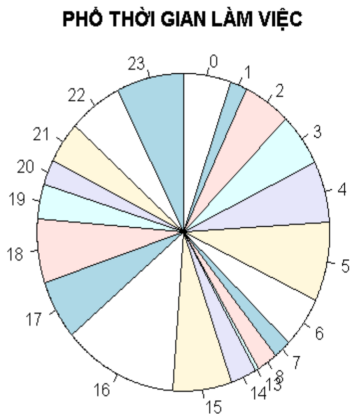
Bài 9

# Giải quyết bài toán

## Bài 4 b

Kết quả:

File 3:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

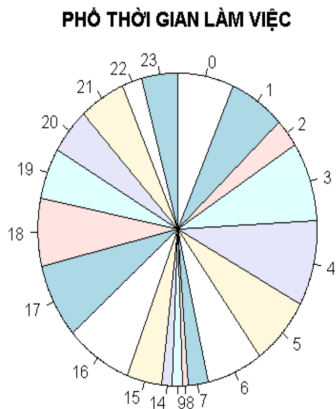
Bài 9

# Giải quyết bài toán

## Bài 4 b

Kết quả:

File 4:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 c

### Câu hỏi đề bài

Tần suất nộp bài được tính bằng phân số giữa khoảng thời gian tính từ lần nộp bài đầu tiên đến lần nộp cuối và số lần nộp bài.

Ứng dụng thông kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 c

Code R :

```
cau_c <- function(needReturn = FALSE) {  
  if (needReturn == FALSE) cat("Câu c:", '\n', '\n')  
  studentID <- df[[1]]  
  count <- 1  
  scope <- c(1:(nrow(df)-1))  
  submitFrequency_vector <- numeric()  
  uid <- character()  
  scoreSpectrums <- list()  
  totalScores <- df$Diem  
  for (i in (1:nrow(df))) {  
    submitFrequency <- 0  
    subtraction <- 0  
    times <- 0  
    first <- 0  
    last <- 0  
    if (studentID[i] != FALSE && studentID[i] != "") {  
      if (df[[5]][i] != 0) {  
        first <- as.numeric(df[[3]][i]) + as.numeric(df[[5]][i])  
        last <- as.numeric(df[[3]][i]) + as.numeric(df[[5]][i])  
        score_vector <- totalScores[i]
```

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 c

Code R :

```
    }  
    count <- count + 1  
    for (j in (scope + 1)) {  
      if (i != nrow(df)) {  
        if (studentID[i] == studentID[j] && i != j && studentID[i] != FALSE &&  
            studentID[j] != FALSE && studentID[i] != "" && studentID[j] != "") {  
          if (df[[5]][j] != 0) {  
            times <- times + 1  
            score_vector <- c(score_vector, totalScores[j])  
            if (df[[4]][i] == df[[3]][j]) subtraction <- as.numeric(df[[5]][j])  
            if (first == 0 && last == 0) first <- as.  
numeric(df[[3]][j]) + as.numeric(df[[5]][j])  
            last <- as.numeric(df[[3]][j]) + as.numeric(df[[5]][j])  
          }  
          if (i < j) studentID[j] <- FALSE  
        }  
      }  
      if (studentID[i] != FALSE && studentID[i] != "") {  
        if (times != 1) { if (as.numeric(subtraction) == 0  
        ) subtraction <- last - first  
        } else if (df[[5]][i] != 0) subtraction  
        <- as.numeric(df[[5]][i])  
      }  
    }  
  }  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 4 c

Code R :

```
times <- 1
}
count <- count + 1
for (j in (scope + 1)) {
  if (i != nrow(df)) {
    if (studentID[i] == studentID[j] && i != j && studentID[i] != FALSE &&
        studentID[j] != FALSE && studentID[i] != "" && studentID[j] != "") {
      if (df[[5]][j] != 0) {
        times <- times + 1
        score_vector <- c(score_vector, totalScores[j])
        if (df[[4]][i] == df[[3]][j]) subtraction <- as.numeric(df[[5]][j])
        if (first == 0 && last == 0) first <- as.numeric(df[[3]][j]) + as.numeric(df[[5]][j])
        last <- as.numeric(df[[3]][j]) + as.numeric(df[[5]][j])
      }
      if (i < j) studentID[j] <- FALSE
    }
  }
}
if (studentID[i] != FALSE && studentID[i] != "") {
  if (times != 1) { if (as.numeric(subtraction) == 0) subtraction <- last - first
```

Ứng dụng thông kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 c

Code R :

```
} else if (df[[5]][i] != 0) subtraction <- as.numeric(df[[5]][i])
}
submitFrequency <- subtraction %/% times
if (df[i, 5] != 0) {
  uid <- c(uid, studentID[i])
  submitFrequency_vector <- c(submitFrequency_vector, submitFrequency)
  scoreSpectrums <- c(scoreSpectrums, list
    (c(min(score_vector), max(score_vector))))
}
if (needReturn == FALSE) {
  submitFrequency <- seconds_to_period(submitFrequency)
  if (times != 0) {
    cat("Tần suất nộp bài của sinh viên có mã số ",
      studentID[i], " là ",
      sprintf('%02d tháng, %02d ngày, %02d giờ, %02d
phút, %02d giây',
month(submitFrequency), day(submitFrequency),
  hour(submitFrequency),
minute(submitFrequency), second(submitFrequency)), '\n')
  } else cat("Sinh viên có mã số ", studentID[i]
, " chưa bao giờ nộp bài", '\n')
}
if (studentID[i] != "") studentID[i] <- FALSE
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 d

### Câu hỏi đề bài

Xác định danh sách các sinh viên có tần suất nộp bài ít nhất

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tình 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 d

Code R :

```
cau_d <- function() {  
  cat("Câu_d: ", '\n', '\n')  
  data <- cau_c(TRUE)  
  minFrequency <- min(data[[2]])  
  indexStudent <- which(data[[2]] %in% minFrequency)  
  minFrequency <- seconds_to_period(minFrequency)  
  cat("Sinh_viên_có_tần_suất_nộp_bài_ít_nhất_có_mã_số: ",  
    , data[[1]][indexStudent], '\n')  
  cat("Với_tần_suất_nộp_bài_là: ", sprintf('%02d_tháng, %02d_ngày',  
    , %02d_giờ, %02d_phút, %02d_giây',  
    month(minFrequency), day(minFrequency), hour(minFrequency)  
    , minute(minFrequency), second(minFrequency)), '\n')  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 d

kết quả:

File 1: 1912817

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tĩnh 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Dọc file và Xử lý dữ  
liệu

Dọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 e

### Câu hỏi đề bài

Xác định phổ điểm của các sinh viên có tần suất nộp bài ít nhất

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tĩnh 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 e

Code R :

```
cau_e <- function() {  
  cat("Câu_e: ", '\n', '\n')  
  data <- cau_c(TRUE)  
  minFrequency <- min(data[[2]])  
  indexStudent <- which(data[[2]] %in% minFrequency)  
  for (i in indexStudent) {  
    cat("Sinh_viên_có_tần_suất_nộp_bài_ít_nhất_có_mã_số: ",  
      , data[[1]][i], '\n')  
    cat('\t', "Phổ_điểm_của_sinh_viên_này_là_từ"  
      , data[[3]][[i]][1], "điểm",  
      "đến", data[[3]][[i]][2], "điểm", '\n')  
  }  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 e

Kết quả:

File 1: 1912817 từ 9,5-10

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 4 f

### Câu hỏi đề bài

Xác định số lượng sinh viên có tần suất nộp bài nhiều nhất

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 f

Code R :

```
cau_f <- function() {  
  cat("Câu f: ", '\n', '\n')  
  data <- cau_c(TRUE)  
  maxFrequency <- max(data[[2]])  
  indexStudent <- which(data[[2]] %in% maxFrequency)  
  maxFrequency <- seconds_to_period(maxFrequency)  
  cat("Số lượng sinh viên có tần suất nộp bài nhiều nhất là: ",  
    , length(indexStudent), '\n')  
  cat("Với tần suất nộp bài là: ", sprintf('%02d tháng, %02d ngày',  
    , %02d giờ, %02d phút, %02d giây',  
    month(maxFrequency), day(maxFrequency), hour(maxFrequency)  
    , minute(maxFrequency), second(maxFrequency)), '\n')  
}
```

Kết quả:

File 1: 1

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 g

### Câu hỏi đề bài

Xác định các sinh viên có tần suất nộp bài nhiều nhất.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tình 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 g

Code R :

```
cau_g <- function() {  
  cat("Câu g: ", '\n', '\n')  
  data <- cau_c(TRUE)  
  maxFrequency <- max(data[[2]])  
  indexStudent <- which(data[[2]] %in% maxFrequency)  
  maxFrequency <- seconds_to_period(maxFrequency)  
  cat("Sinh viên có tần suất nộp bài nhiều nhất có mã số: ",  
    data[[1]][indexStudent], '\n')  
  cat("Với tần suất nộp bài là: ", sprintf('%02d tháng, %02d ngày',  
    %02d giờ, %02d phút, %02d giây',  
    month(maxFrequency), day(maxFrequency), hour(maxFrequency)  
    , minute(maxFrequency), second(maxFrequency)), '\n')  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 g

Kết quả:

File 1: 1914022

Ứng dụng thông kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tình 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Dọc file và Xử lý dữ  
liệu

Dọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 h

### Câu hỏi đề bài

Xác định phổ điểm của các sinh viên có tần suất nộp bài nhiều nhất.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 h

Code R :

```
cau_h <- function() {  
  cat("Câu_h: ", '\n', '\n')  
  data <- cau_c(TRUE)  
  maxFrequency <- max(data[[2]])  
  indexStudent <- which(data[[2]] %in% maxFrequency)  
  for (i in indexStudent) {  
    cat("Sinh_vien_có_tần_suất_nộp_bài_nhiều_nhất  
    có_mã_số: ", data[[1]][i], '\n')  
    cat('\t', "Phổ_điểm_của_sinh_vien_này_là_từ"  
    , data[[3]][[i]][1], "điểm",  
    "đến", data[[3]][[i]][2], "điểm", '\n')  
  }  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 h

File 1: 1914022 phổ điểm 8 -10 điểm

Ứng dụng thông kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tình 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 4 i

### Câu hỏi đề bài

Xác định các sinh viên nằm trong nhóm có tần suất nộp bài nhiều nhì.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 i

Code R :

```
cau_i <- function() {  
  cat("Câu i: ", '\n', '\n')  
  data <- cau_c(TRUE)  
  maxFrequency <- max(data[[2]])  
  indexStudent <- which(data[[2]] %in% maxFrequency)  
  data[[2]][indexStudent] <- FALSE  
  secondFrequency <- max(data[[2]])  
  indexStudent <- which(data[[2]] %in% secondFrequency)  
  secondFrequency <- seconds_to_period(secondFrequency)  
  cat("Sinh viên có tần suất nộp bài nhiều nhất có mã số: ",  
    , data[[1]][indexStudent], '\n')  
  cat("Với tần suất nộp bài là: ", sprintf('%02d tháng  
    , %02d ngày, %02d giờ, %02d phút, %02d giây',  
    month(secondFrequency), day(secondFrequency)  
    , hour(secondFrequency),  
    minute(secondFrequency), second(secondFrequency)), '\n')  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 i

Kết quả:  
File 1:1914093

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tình 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 j

### Câu hỏi đề bài

Xác định các sinh viên nằm trong nhóm có tần suất nộp bài nhiều nhất hoặc nhiều nhì.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 j

Code R :

```
cau_j <- function() {  
  cat("Câu_j: ", '\n', '\n')  
  data <- cau_c(TRUE)  
  maxFrequency <- max(data[[2]])  
  indexStudent <- which(data[[2]] %in% maxFrequency)  
  data[[2]][indexStudent] <- FALSE  
  secondFrequency <- max(data[[2]])  
  indexStudent <- c(indexStudent, which(data[[2]]  
    %in% secondFrequency))  
  cat("Các sinh viên có tần suất nộp bài nhiều  
nhất hoặc nhiều nhất có mã số: ", data[[1]]  
    [indexStudent], '\n')  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 j

Kết quả:

File 1:1914022 1914093

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 k

### Câu hỏi đề bài

Hãy tính thời gian trung bình (tính bằng giây) giữa hai lần nộp bài liên nhau sau cùng của cùng một sinh viên trong mẫu đã chọn.

Ứng dụng thông kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 k

Code R :

```
cau_k <- function() {  
  cat("Câu_k:", '\n', '\n')  
  data <- cau_a(TRUE)  
  for (i in (1:length(data[[1]]))) {  
    if (data[[2]][i] != 0) {  
      cat("Thời_gian_trung_bình_giữa_hai_lần_nộp_bài_liên_nhau  
_của_sinh_viên_có_mã_số",  
data[[1]][i], "là", round(data[[2]][i], 2), "giờ", '\n')  
    } else cat("Sinh_viên_có_mã_số", data[[1]][i],  
"chưa_bao_giờ_nộp_bài", '\n')  
  }  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 4 I

### Câu hỏi đề bài

Tính tần số, tần suất và tần suất tích lũy của mẫu trên.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 I

Code R :

```
cau_1 <- function() {  
  cat("Câu 1: ", '\n', '\n')  
  data <- cau_a(TRUE)  
  for (i in (1:length(data[[1]]))) {  
    if (data[[3]][i] != 0) {  
      cat("Sinh viên có mã số", data[[1]][i], ": ", '\n')  
      cat('\t', "Tên số nộp bài là", data[[3]][i], "  
(lần nộp)", '\n')  
      cat('\t', "Tần suất nộp bài là", round(data[[4]][i]  
      , 2), '%', '\n')  
      cat('\t', "Tần suất tích lũy nộp bài là", round(data  
      [[5]][i], 2), '%', '\n')  
    } else cat("Sinh viên có mã số", data[[1]][i], " chưa  
    nộp bài", '\n')  
  }  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 m

### Câu hỏi đề bài

Vẽ biểu đồ tần số của mẫu trên. Hãy nhận xét về biểu đồ.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 m

Code R :

```
cau_m <- function() {  
  cat("Câu_m: ", '\n', '\n')  
  data <- cau_a(TRUE)  
  countfrequency <- count(data[[3]])  
  maxfrequency <- max(countfrequency[[2]])  
  hist(data[[3]], main = "TẦN SỐ", xlab =  
    "Tần số nộp bài", ylab = "Số sinh viên",  
    xlim = c(0, max(data[[3]]) + (max(data[[3]]) -  
    min(data[[3]])) / length(countfrequency[[1]])),  
    ylim = c(0, maxfrequency + 25), col="blue")  
  cat("Nhận xét: Tần số (số lần) nộp bài của các sinh viên là:  
    ", countfrequency[[1]], ", trong đó:", '\n')  
  n <- 0  
  for (i in countfrequency[[1]]) {  
    n <- n + 1  
    cat('\t', "Nộp", i, " lần: ", countfrequency[[2]][n]  
    , " sinh viên", '\n')  
  }  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

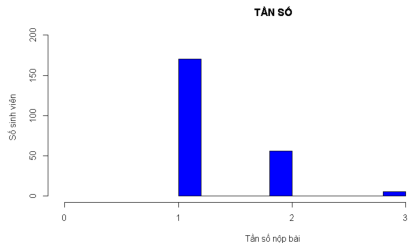
Bài 9

# Giải quyết bài toán

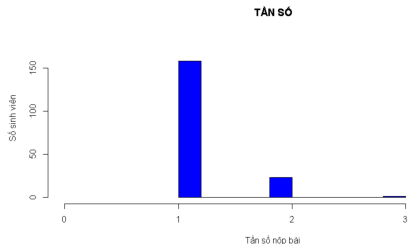
## Bài 4 m

Kết quả:

File 1:



File 2:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

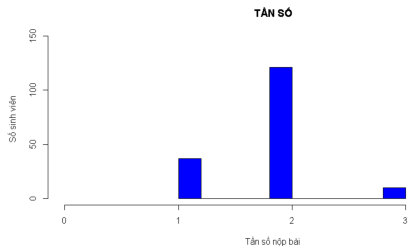
Bài 9

# Giải quyết bài toán

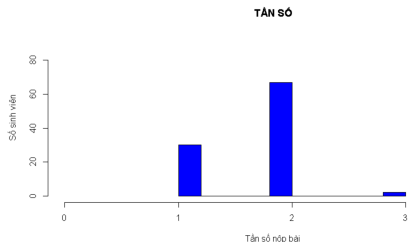
## Bài 4 m

Kết quả:

File 3:



File 4:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 n

### Câu hỏi đề bài

Vẽ biểu đồ tần suất của mẫu trên. Hãy nhận xét về biểu đồ.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 n

Code R :

```
cau_n <- function() {  
  cat("Câu_n: ", '\n', '\n')  
  data <- cau_a(TRUE)  
  countfrequency <- count(data[[4]])  
  maxfrequency <- max(countfrequency[[2]])  
  hist(data[[4]], main = "TẦN SUẤT", xlab =  
    "Tần suất nộp bài (%)", ylab = "Số sinh viên",  
    xlim = c(0, max(data[[4]]) + (max(data[[4]]) -  
      min(data[[4]])) / length(countfrequency[[1]])),  
    ylim = c(0, maxfrequency + 25), col="blue")  
  cat("Nhận xét: Tần suất nộp bài của các sinh viên là: ",  
    round(countfrequency[[1]], 2), ", trong đó: ", '\n')  
  n <- 0  
  for (i in countfrequency[[1]]) {  
    n <- n + 1  
    cat('\t', "Tần suất", round(i, 2), "%: ",  
      countfrequency[[2]][n], "sinh viên", '\n')  
  }  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

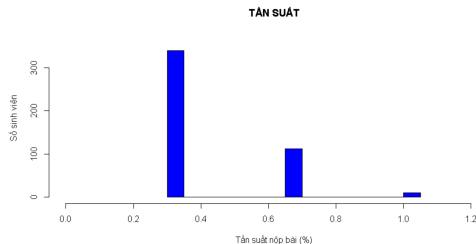


# Giải quyết bài toán

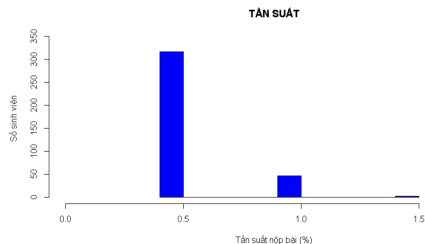
## Bài 4 n

Kết quả:

File 1:



File 2:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

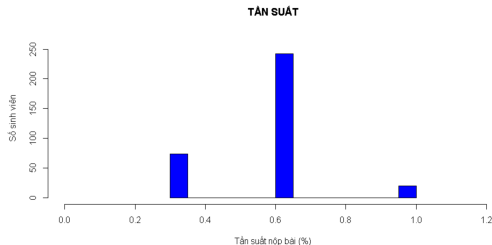
Bài 9

# Giải quyết bài toán

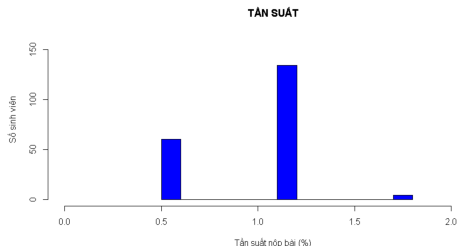
## Bài 4 n

Kết quả:

File 3:



File 4:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 o

### Câu hỏi đề bài

Vẽ biểu đồ tần suất tích lũy của mẫu trên. Hãy nhận xét về biểu đồ.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 o

Code R :

```
cau_o <- function() {  
  cat("Câu_o:", '\n', '\n')  
  data <- cau_a(TRUE)  
  
  plot(data[[5]], main = "TẦN_SUẤT_TÍCH_LŨY", xlab = "Sinh_viên",  
        ylab = "Số_tần_suất_tích_lũy", type = "l", col = "red")  
  if ((max(data[[4]]) - min(data[[4]])) <= 2) {  
    cat("Nhận_xét: Có_thể_xem_đồ_thị_như_là_một  
đường_thẳng, suy_ra_tần_suất_nộp_bài_của_các  
sinh_viên_tương_đối_bằng_nhau", '\n')  
  } else if ((max(data[[4]]) - min(data[[4]])) <= 4) {  
    cat("Nhận_xét: Đồ_thị_hơi_méo_mó, suy_ra_tần_suất  
nộp_bài_của_các_sinh_viên_có_sự_chênh_lệch  
nhưng_không_nhiều_lắm", '\n')  
  } else cat("Nhận_xét: Đồ_thị_có_sự_gập_khúc, suy  
ra_tần_suất_nộp_bài_của_các_sinh_viên_chênh_lệch_khá_nhiều", '\n')  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

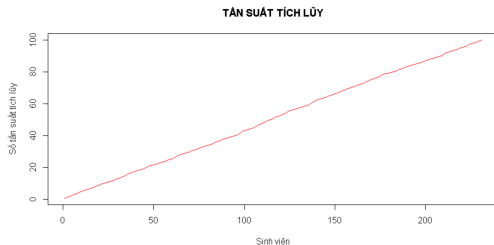
Bài 9

# Giải quyết bài toán

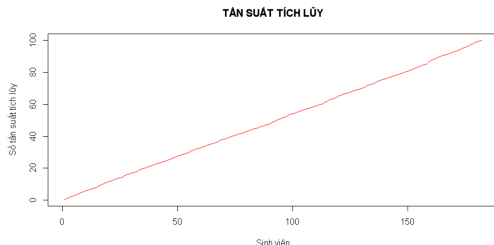
## Bài 4 o

Kết quả:

File 1:



File 2:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

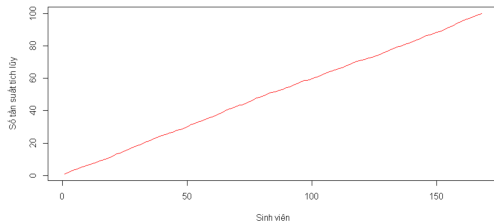
# Giải quyết bài toán

## Bài 4 o

Kết quả:

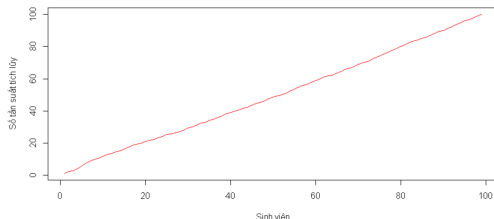
File 3:

TẦN SUẤT TÍCH LŨY



File 4:

TẦN SUẤT TÍCH LŨY



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 p

### Câu hỏi đề bài

Tính trung vị mẫu, cực đại mẫu, cực tiểu mẫu của trên.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 p

Code R :

```
cau_p <- function() {  
  cat("Câu p: ", '\n', '\n')  
  data <- cau_a(TRUE)  
  cat("Đối với tần số: ", '\n')  
  cat('\t', "Trung vị mẫu là: ", median(data[[3]]), '\n')  
  cat('\t', "Cực đại mẫu là: ", max(data[[3]]), '\n')  
  cat('\t', "Cực tiểu mẫu là: ", min(data[[3]]), '\n')  
  cat("Đối với tần suất: ", '\n')  
  cat('\t', "Trung vị mẫu là: ",  
    round(median(data[[4]]), 2), '\n')  
  cat('\t', "Cực đại mẫu là: ",  
    round(max(data[[4]]), 2), '\n')  
  cat('\t', "Cực tiểu mẫu là: ",  
    round(min(data[[4]]), 2), '\n')  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 4 q

### Câu hỏi đề bài

Hãy đo mức độ phân tán của điểm số (xung quanh giá trị trung bình) của mẫu.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 q

Code R :

```
cau_q <- function() {  
  cat("Câu q: ", '\n', '\n')  
  cat("Mức độ phân tán của điểm số xung quanh  
    giá trị trung bình: ", '\n')  
  cat('\t', "Phương sai bằng: ",  
    round(var(df$Diem), 2), '\n')  
  cat('\t', "Độ lệch chuẩn bằng: ",  
    round(sd(df$Diem), 2), '\n')  
  cat('\t', "Hệ số biến thiên bằng: ", round(  
    sd(df$Diem) / mean(df$Diem) * 100, 2 ), "%", '\n')  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 r

### Câu hỏi đề bài

Tính độ méo lệch (skewness), và độ nhọn (kurtosis) của dữ liệu trong mẫu trên.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 r

Code R :

```
cau_r <- function() {  
  cat("Câu r: ", '\n', '\n')  
  cat("Độ méo lệch (skewness) của điểm số là: "  
    , round(skewness(df$Diem), 2), '\n')  
  cat("Độ nhọn (kurtosis) của điểm số là: "  
    , round(kurtosis(df$Diem), 2), '\n')  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 s

### Câu hỏi đề bài

Tính tứ phân vị (quartile) thứ nhất (Q1) và thứ ba (Q3) của mẫu.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 4 s

Code R :

```
cau_s <- function() {  
  cat("Câu s: ", '\n', '\n')  
  quantile <- quantile(df$Diem)  
  cat("Tứ phân vị thứ nhất (Q1) của mẫu là: "  
    , quantile[[2]], '\n')  
  cat("Tứ phân vị thứ ba (Q3) của mẫu là: "  
    , quantile[[4]], '\n')  
}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

**Bài 4**

Bài 5

Bài 7

Bài 9

Code R:

```
data2 <-df[,c(1,6)]
df1 <-data_frame(data2)
dt<- NULL
for(i in 1:nrow(df)){
    dt <-c(dt,df1[i,1])
}
dt2 <- NULL
for(i in 1 :nrow(df)){
    dt2 <-c(dt2,df1[i,2])
}
dt2 <-unlist(dt2)
dt <-unlist(dt)
dt1 <-dt[!duplicated(dt)]
dt1 <-unlist(dt1)
```



# Giải quyết bài toán

## Bài 5 a

### Câu hỏi đề bài

Hãy tính và vẽ biểu đồ sự phân bố về điểm đạt được của sinh viên sau  $k = 6$  lần nộp bài.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

**Bài 5**

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 5 a

Code R :

```
cau5<-NULL
for(i in 1:length(dt1)){
  x <- seq(0,0,length.out = 6)
  k<-1
  for(j in 1:length(dt)){
    if(dt1[i]==dt[j]&& k<=6){
      x[k]=dt2[j]
      k<-k+1
    }
  }
  cau5 <- c(cau5,max(x))
}

hist(cau5,breaks=10,col="red",border="white",xlab="điểm số đạt được",
,ylab="số lượng sinh viên")
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

**Bài 5**

Bài 7

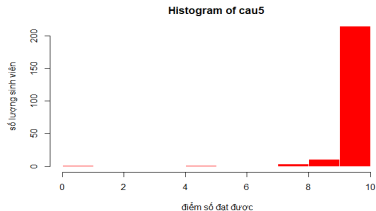
Bài 9

# Giải quyết bài toán

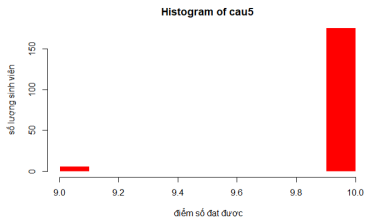
## Bài 5 a

Kết quả:

File 1:



File 2:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tình 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

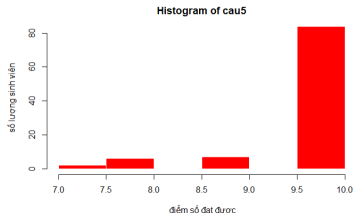
Bài 9

# Giải quyết bài toán

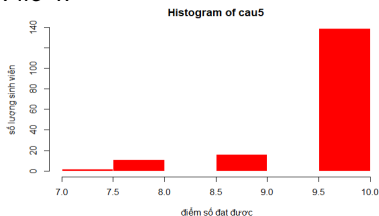
## Bài 5 a

Kết quả:

File 3:



File 4:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

**Bài 5**

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 5 b

### Câu hỏi đề bài

Áp dụng câu 5a với k được tính theo công thức sau:  $MD \bmod 3 + 1$

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

**Bài 5**

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 5 b

=> k=1

Code R :

```
cau5<-NULL
for(i in 1:length(dt1)){
  x <- seq(0,0,length.out = 1)
  k<-1
  for(j in 1:length(dt)){
    if(dt1[i]==dt[j]&& k<=1){
      x[k]=dt2[j]
      k<-k+1
    }
  }
  cau5 <- c(cau5,max(x))
}
hist(cau5,breaks=10,col="red",border="white"
,xlab="điểm số đạt được",ylab="số lượng sinh viên")
```

Ứng dụng thống kê

Nguyễn Duy Phước

1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú

1915828

Nguyễn Hữu Lợi

1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

**Bài 5**

Bài 7

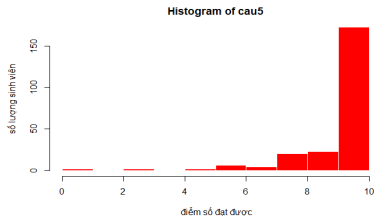
Bài 9

# Giải quyết bài toán

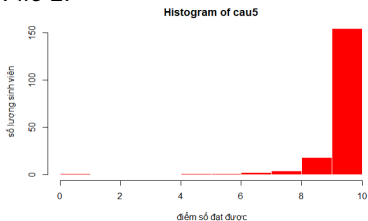
## Bài 5 b

Kết quả:

File 1:



File 2:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

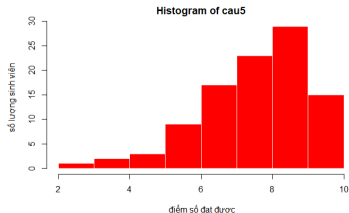
Bài 9

# Giải quyết bài toán

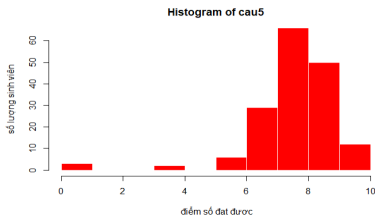
## Bài 5 b

Kết quả:

File 3:



File 4:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 5 c

### Câu hỏi đề bài

Hãy tính các giá trị  $T_{Bk}$  và vẽ biểu đồ thể hiện sự thay đổi của các giá trị trung bình này với sự thay đổi của  $k$ . Hãy nhận xét về biểu đồ mà các em vừa vẽ được.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

**Bài 5**

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 5 c

Code R :

```
cau5c <- NULL
for(h in 1:8){
  cau5 <- NULL
  for(i in 1:length(dt1)){
    x <- seq(0,0,length.out = h)
    k<-1
    for(j in 1:length(dt)){
      if(dt1[i]==dt[j]&& k<=h){
        x[k]=dt2[j]
        k<-k+1
      }
    }
    cau5 <- c(cau5,max(x))
  }
  cau5c <-c(cau5c,mean(cau5))
}
stt <-seq(1,8,1)
hinhcau5c <-data.frame(stt,cau5c)
ggplot(data=hinhcau5c,aes(x=stt,y=cau5c))+ geom_line()
+ xlab("k")+ylab("TBk")
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

**Bài 5**

Bài 7

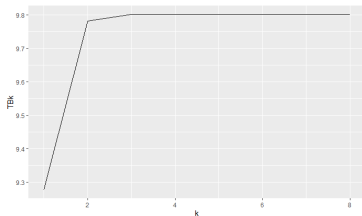
Bài 9

# Giải quyết bài toán

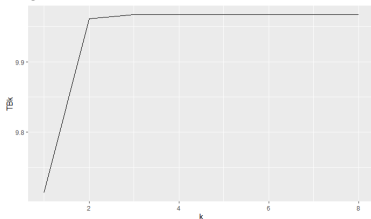
## Bài 5 c

Kết quả:

File 1:



File 2:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

**Bài 5**

Bài 7

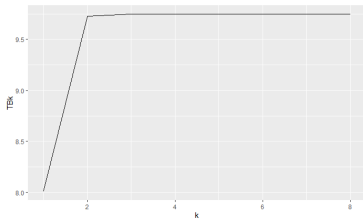
Bài 9

# Giải quyết bài toán

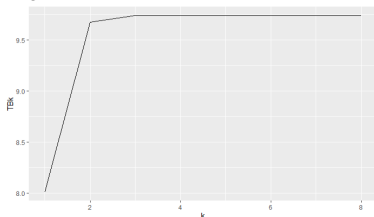
## Bài 5 c

Kết quả:

File 3:



File 4:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

**Bài 5**

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 5 c

Nhận xét: Dựa vào đồ thị ta có thể thấy rằng khi  $k$  tăng thì điểm  $TB_k$  cũng tăng đến giá trị  $k=3$ , sau đó không tăng nữa và cũng không giảm  
 $\Rightarrow$  số lần làm bài cho phép là 3

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

**Bài 5**

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 5 d

### Câu hỏi đề bài

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

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

**Bài 5**

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 5 d

Theo câu 5c thì số lần làm bài cho phép của sinh viên là 3 nên trung bình điểm số mà các sinh viên đạt được qua bài tập tin này là:

cau5c[3]

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

**Bài 5**

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 7 a

### Câu hỏi đề bài

Hãy xác định thời điểm  $t_2$  phù hợp.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 7 a

Tiêu chí chọn thời điểm  $t_2$  = thời gian hoàn thành bài làm lần cuối của sinh viên nộp bài cuối cùng - 3 ngày Ví dụ: trong file đang xử lý có thời gian nộp bài lần cuối của sinh viên nộp bài cuối cùng là "18 May 2020 10:47 PM" Nên ta có  $t_2$  = ("18 May 2020 10:47 PM") - 3 ngày = "15 May 2020 10:47 PM" Theo tiêu chí trên, nếu sinh viên nào có nộp bài lần đầu sau thời điểm  $t_2$  = "15 May 2020 10:47 PM" sẽ bị xem như học đối phó

Code R :

```
cau7a <- function(df7a)
{
  df7a.input.data <- df7a
  list.submit.7a <- df7a.input.data
  list.submit.7a <- list.submit.7a[order(strptime(list
    .submit.7a$Hoan_thanh, "%d_%B_%Y_%I:%M_%p"), na.last = TRUE), ]
  list.submit.7a <- list.submit.7a[list.submit.7a$Hoan_thanh != "-", ]
  t2 <- strptime(list.submit.7a[nrow(list.submit.7a), "Hoan_thanh"]
    , "%d_%B_%Y_%I:%M_%p") - ddays(3)
  return(t2)
}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9



Code R :

```
ds.sv.doi.pho <- function(df7dp)
{
  df7dp.input.data <- df7dp
  t2dp <- cau7a(df7dp.input.data)
  first.submit <- df7dp.input.data

  for(i in 1:nrow(first.submit))
  {
    if (first.submit[i,"Hoan_thanh"] == "-")
    {
      first.submit[i,"Hoan_thanh"] <- first.submit[i,"Bat_dau"]
    }
  }

  first.submit <- first.submit %>% group_by(.,uid) %>% arrange(.,
  , strptime(Hoan_thanh, "%d_%B_%Y_%I:%M_%p"), .by_group = TRUE)
  %>% distinct(., uid, .keep_all = TRUE) %>% ungroup()
  first.submit.after.t2 <- first.submit %>% filter(.,
  , strptime(Hoan_thanh, "%d_%B_%Y_%I:%M_%p") > t2dp)
  %>% arrange(., strptime(Hoan_thanh, "%d_%B_%Y_%I:%M_%p"))
  return(first.submit.after.t2)
}
```



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 7 b

### Câu hỏi đề bài

Xác định số lượng sinh viên học đôi phó.

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 7 b

Code R :

```
cau7b <- function(df7b)
{
df7b.input.data <- df7b
first.submit.after.t2 <- ds.sv.doi.pho(df7b.input.data)
num.SVDP <- nrow(first.submit.after.t2)
return(num.SVDP)
}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

Code R :

```
tao.bang.tan.so.raw <- function(df7craww)
{
  df7c.input.data.raww <- df7craww
  submit.after.t2 <- ds.sv.doi.pho(df7c.input.data.raww)
  frequency.table.7c.raww <- submit.after.t2[, "uid"]
  frequency.table.7c.raww <- frequency.table.7c.raww %>% left_join
  (., df7c.input.data.raww, by = "uid") %>% select(
  ,c("uid", "Diem")) %>% count(., Diem)
  names(frequency.table.7c.raww)[2] <- "SL"
  return(frequency.table.7c.raww)
}

tao.bang.tan.so.full <- function(input.from.tao.bang.tan.so.raw)
{
  frequency.table.7c.full <- tao.bang.tan.so.raw(
  input.from.tao.bang.tan.so.raw)
  h1 <- data.frame("Diem"=seq(0,10,by=0.5), stringsAsFactors = FALSE)
  frequency.table.7c.full <- left_join(
  h1, frequency.table.7c.full, by = "Diem")
  frequency.table.7c.full[is.na(frequency.table.7c.full)] <- 0
  return(frequency.table.7c.full)
}
```



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

Code R :

```
total.score.num <- function(input.from.tao.bang.tan.so.full)
{
  frequency.table.7c.total <- tao.bang.tan.so.full(
    input.from.tao.bang.tan.so.full)
  total.score.num <- sum(frequency.table.7c.total$SL)
  return(total.score.num)
}

bang.tan.so.ctr <- function(df7c.ctr)
{
  frequency.table.7c.ctr <- tao.bang.tan.so.full(df7c.ctr)
  %>% spread(., Diem, SL)
  row.names(frequency.table.7c.ctr) <- NULL
  frequency.table.7c.ctr$Diem <- c("SL")
  frequency.table.7c.ctr$Tong <- total.score.num(df7c.ctr)
  frequency.table.7c.ctr <- frequency.table.7c.ctr[ , c(22,1:21,23)]
  return(frequency.table.7c.ctr)
}

get.name.data.frame <- function(file)
{
  data.frame.name <- deparse(substitute(file))
  return(data.frame.name)
}
```



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 7 c

### Câu hỏi đề bài

Xác định phổ điểm của các sinh viên học đối phó.

Code R :

```
cau7c.func <- function(input.data, name.of.data.frame)
{
  pho.diem.data <- tao.bang.tan.so.full(input.data)
  data.frame.input.name <- name.of.data.frame
  pho.diem <- ggplot(pho.diem.data, aes(x=factor(Diem), y=SL))
  + geom_bar(stat="identity", fill = "#274CBF", colour = "#17202A"
  , aes(col = Diem)) + geom_text(aes(label=SL), vjust=-0.3
  , size=3.5, color="#000000") + labs(title = "PHỔ ĐIỂM", subtitle
= "(SINH VIÊN HỌC ĐỐI PHÓ)", caption = paste("input:", as.character(
data.frame.input.name))) + xlab("Điểm") + ylab("Số lượng") +
  theme_bw() + theme(axis.text = element_text(colour = "#000080"
  , size = rel(1), face = "bold"), plot.title = element_text(
size = rel(2)), panel.border = element_rect(linetype = "solid"
  , fill = NA)) + theme(axis.title = element_text(family = "sans"
  , face = "bold"))+scale_y_continuous(labels=label_comma(accuracy = 1))
return(pho.diem)
}
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 7 c

Kết quả:

File 1:

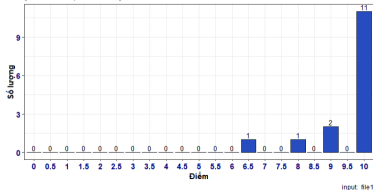
a> 2020-05-15 23:35:00

b> 12

c>

PHỔ ĐIỂM

(SINH VIÊN HỌC ĐỐI PHỎ)



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 7 c

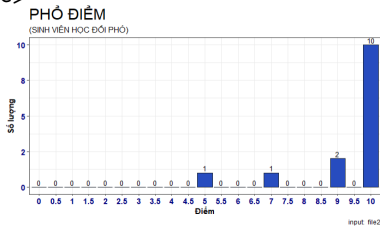
Kết quả:

File 2:

a> 2020-05-15 22:47:00

b> 11

c>



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9



# Giải quyết bài toán

## Bài 7 c

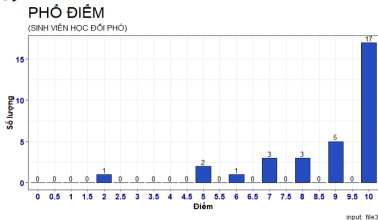
Kết quả:

File 3:

a> 2020-05-15 22:14:00

b> 18

c>



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 7 c

Kết quả:

File 4:

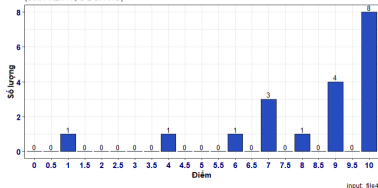
a> 2020-05-16 00:40:00

b> 10

c>

PHỔ ĐIỂM

(SINH VIÊN HỌC ĐỐI PHÓ)



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

Chuyển dạng char về dạng thời gian ở các cột chỉ TG:

Code R :

```
for (i in 1:nrow(df))
{
  if(!grepl("giờ",df$Thuc_hien[i]))
  {
    if(grepl("phút",df$Thuc_hien[i]))
    {
      df$Thuc_hien[i]<-paste(df$Thuc_hien[i],"0giờ", step =" ")
    }
    else{ df$Thuc_hien[i]<-paste(df$Thuc_hien[i],"0phút_0giờ"
    , step =" ")}}
  }
df$Bat_dau<-parse_date_time2(df$Bat_dau,orders="dbYIMp")
df$Hoan_thanh<-parse_date_time2(df$Hoan_thanh,orders="dbYIMp")
df$Thuc_hien<-difftime(parse_date_time(df$Thuc_hien,order=
c("MS","S","HMS")),parse_date_time("0",orders = "S"),units="secs")
df<-df %>%
drop_na(Hoan_thanh,Bat_dau)
```



# Giải quyết bài toán

## Bài 9 a

### Câu hỏi đề bài

Hãy xác định giá trị  $k$  và  $n$  phù hợp.

Code R :

```
k<-9.5
```

```
n<-2
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Dọc file và Xử lí dữ  
liệu

Dọc file

Xử lí dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 9 b

### Câu hỏi đề bài

Xác định số lượng sinh viên thông minh.

Code R :

```
for(j in 1:nrow(df))

{
  t<-arrange(filter(df,df[1]==df$uid[j]),Bat_dau)
  if (nrow(t)<n){
    for(i in 1:(n-nrow(t)))
    {
      do<-t[,nrow(t)]
      t<-bind_rows(t,do)
    }
    if(j==1){
      SV<-data.frame(uid=c(t$uid[1]),Diem=c(max(t$Diem[1:n])))
    } else{
      SV<-union(data.frame(uid=c(t$uid[1]),Diem=c(max(t$Diem[1:n]))),SV)
    }
  }
  SV_ThongMinh<-filter(SV,SV$Diem>=k)
  SVTM<-n_distinct(SV_ThongMinh$uid)
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 9 b

Kết quả:

File 1: 56

File 2: 23

File 3: 67

File 4: 120

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tĩnh 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

# Giải quyết bài toán

## Bài 9 c

### Câu hỏi đề bài

Xác định phổ điểm của các sinh viên thông minh.

Code R :

```
BD_Phodiem_SVTM<-ggplot(Phodiem_SVTM, aes(x = factor  
(Diem), y = n))+ geom_bar(stat = "identity") + theme_minimal()  
+ labs(title = "Phổ điểm sinh viên thông minh", x = "Điểm",  
y = "Số sinh viên") + coord_fixed(ratio = 10/max(Phodiem_SVTM[,2]))
```

Ứng dụng thống kê

Nguyễn Duy Phước  
1914768

Dương Bá Tình 191553

Võ Nguyễn Minh Tú  
1915828

Nguyễn Hữu Lợi  
1914047

Đào Quốc Bảo 192700



Mô tả dữ liệu

Đọc file và Xử lý dữ  
liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

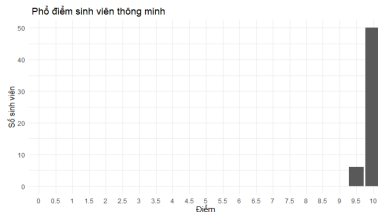
Bài 9

# Giải quyết bài toán

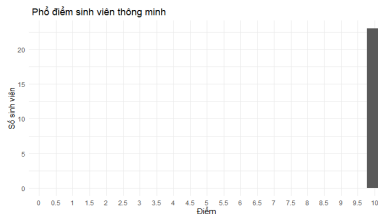
## Bài 9 c

Kết quả:

File 1:



File 2:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tình 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9

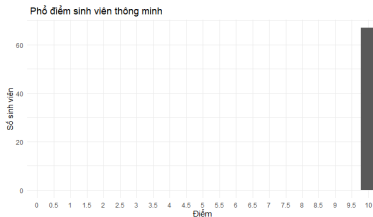


# Giải quyết bài toán

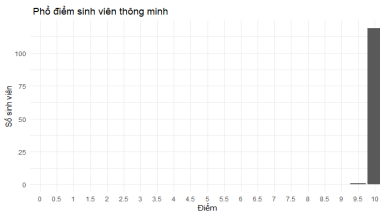
## Bài 9 c

Kết quả:

File 3:



File 4:



Ứng dụng thống kê

Nguyễn Duy Phước  
1914768  
Dương Bá Tình 191553  
Võ Nguyễn Minh Tú  
1915828  
Nguyễn Hữu Lợi  
1914047  
Đào Quốc Bảo 1927004



Mô tả dữ liệu

Đọc file và Xử lý dữ liệu

Đọc file

Xử lý dữ liệu

Giải quyết bài toán

Bài 1

Bài 2

Bài 3

Bài 4

Bài 5

Bài 7

Bài 9