

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC BÁCH KHOA
KHOA KHOA HỌC & KỸ THUẬT MÁY TÍNH



CẤU TRÚC RỜI RẠC CHO KHMT (CO1007)

Ứng dụng thống kê
phân tích kết quả đánh giá việc thực tập của sinh viên

GVHD: Huỳnh Tường Nguyên
Nguyễn Ngọc Lễ
SV thực hiện: Trương Ngọc Trung Anh – 2020004
Huỳnh Bảo Minh – 2020047
Võ Văn Hiền – 2020023
Lê Nam Tiến Thành – 2020088



Mục lục

1	Động cơ nghiên cứu	2
2	Mục tiêu	2
3	Mô tả dữ liệu	2
4	Nhiệm vụ	3
5	Code R và kết quả tính toán	5
6	Kết luận	49
	Tài liệu	49

1 Động cơ nghiên cứu

Khi bạn đang đi học, có những quyết định khó khăn để đưa ra; thật khó để chọn con đường sự nghiệp phù hợp. Bạn có thể không nhận thức được tất cả các lựa chọn của mình, nơi bạn sẽ thành công và cuối cùng, bạn sẽ thích làm gì. Từ đó Trường Đại học Bách Khoa, ĐHQG-HCM kết hợp cùng các công ty trong cũng như ngoài nước để tạo ra môi trường công nghiệp thực sự để giúp các bạn sinh viên thực tập. Bạn sẽ có cơ hội khám phá những gì bạn yêu thích, kết nối với đồng nghiệp, người cố vấn và chuyên gia trong ngành. Đó là cách tốt nhất để có được một chỗ ngồi tại công ty hay thực hiện bước đầu tiên để bắt đầu một sự nghiệp tuyệt vời. Sau khi chương trình kết thúc công ty sẽ cung cấp bảng nhận xét đánh giá các sinh viên qua các tiêu chí cụ thể.

Phân tích & thống kê dữ liệu qua bảng đánh giá sinh viên của công ty không những giúp khoa máy tính có những hướng đúng trong việc phát hiện ra những kiến thức mà sinh viên chưa chắc chắn hay thiếu sót trong đào tạo, cũng như có hướng để cải thiện bổ sung phần học liệu, đào tạo trong tương lai để phù hợp với hơn người học.

2 Mục tiêu

Với tập dữ liệu được thu thập, thực hiện phân tích dữ liệu một cách hợp lý và diễn giải cũng như đưa ra kết luận từ những phân tích đó. Trong bài tập lớn này, mục tiêu trọng tâm là trang bị cho sinh viên những kỹ năng định lượng cần thiết mà họ có thể sử dụng và xây dựng dựa trên những cách thức linh hoạt. Học tập, tìm hiểu và sử dụng công cụ để làm việc với dữ liệu như ngôn ngữ R.

Tìm hiểu [1] các nguyên tắc cơ bản của lý thuyết xác suất, [2] lý luận thống kê và phương pháp suy luận, [3] tính toán thống kê, [4] phân tích mô tả, diễn giải và khám phá dữ liệu bằng đồ họa và các phương tiện khác; [5] hoàn tất bài tập cũng phải học cách trình bày hiệu quả. Học được cách làm việc nhóm và quản lý lượng công việc.

3 Mô tả dữ liệu

Dữ liệu mẫu là 10 sheet dữ liệu từ "E1" đến "E10", mỗi sheet là một bảng tính gồm nhiều dòng, nhiều cột. Mỗi dòng mô tả một tiêu chí đánh giá sinh viên gồm có 26 tiêu chí. Mỗi cột thể hiện kết quả của một sinh viên tương ứng với các tiêu chí. Mỗi sinh viên sẽ được đánh giá bởi 5 mức điểm từ 0 đến 5 thể hiện ở bên dưới (**Bảng 1**).

Mức chọn: 0. Không có thông tin để đánh giá; 1. Rất yếu; 2. Yếu; 3. Trung bình; 4. Tốt		
TIÊU CHÍ	P.1	P.2
01. Khả năng ứng dụng kiến thức tin học trong công việc	3	2
02. Khả năng suy nghĩ logic và/hoặc vận dụng kiến thức toán học trong công việc	3	3
03. Khả năng sử dụng công cụ và công nghệ thích hợp tại đơn vị	3	3
04. Khả năng thiết kế ứng dụng dựa trên yêu cầu của khách hàng	0	0
05. Khả năng kiểm tra một ứng dụng thông qua việc xây dựng một tập kiểm tra mẫu	0	0
06. Khả năng trao đổi ý tưởng và chia sẻ công việc	3	3
07. Nhận thức về trách nhiệm của sinh viên thực tập	2	4
08. Khả năng viết báo cáo	3	3
09. Khả năng chuẩn bị một bài thuyết trình	4	4
10. Kỹ năng nói khi thuyết trình	2	4
11. Khả năng ước lượng chi phí của một dự án phần mềm	0	0
12. Khả năng tự học	3	3
13. Các chứng chỉ chuyên môn mà sinh viên tự trang bị	3	3
14. Khả năng sử dụng Internet để phục vụ công việc	4	4
15. Khả năng sử dụng các ứng dụng xử lý dữ liệu đặc thù trong công việc	0	0
16. Sinh viên làm chủ ít nhất một ngôn ngữ lập trình	0	0
17. Sinh viên thấy được sự trả giá giữa thời gian thực thi và không gian lưu trữ của một giải thuật	0	0
18. Sinh viên có khả năng phân tích ưu và khuyết điểm của một giải pháp kỹ thuật	2	3
19. Sinh viên có khả năng đọc hiểu một phần thiết kế của một dự án thực tế	2	2
20. Khả năng xác định được vai trò và trách nhiệm của mình khi làm việc nhóm	4	4
21. Giúp đỡ các bạn cùng nhóm trong công việc	0	4
22. Đóng góp kỹ thuật vào dự án chung của nhóm	2	2
23. Kỹ năng thu thập và tổng hợp thông tin	4	4
24. Kỹ năng tổ chức thực hiện	3	3
25. Tính kỷ luật trong công việc	4	3
26. Tham gia vào các hoạt động sinh hoạt chung của đơn vị	4	4

Bảng 1: Bảng mẫu đánh giá

4 Nhiệm vụ

Gọi **MD** là mã đề riêng cho mỗi nhóm (gồm 6 ký số) không trùng nhau, nhóm sinh viên sẽ thực hiện các yêu cầu dưới đây với các giá trị xác định như sau:

- Mỗi nhóm sẽ dùng R để thao tác trên 6 sheet dữ liệu được chọn theo cách:
 - **Sheet i:** { "ký số thứ j của MD" + 1}, $i \in \{1, 2, 3, 4, 5, 6\}$

Câu hỏi

- Xác định số lượng sinh viên trong tập mẫu
- Tính điểm trung bình cho mỗi tiêu chí, làm tròn 2 số thập phân
- Tính điểm trung bình cho mỗi SV, làm tròn 2 số thập phân
- Liệt kê danh sách sinh viên có tổng điểm các tiêu chí là nhỏ nhất
- Liệt kê danh sách sinh viên có tổng điểm các tiêu chí là lớn nhất
- Liệt kê danh sách tiêu chí mà tổng điểm của các sinh viên là nhỏ nhất
- Liệt kê danh sách tiêu chí mà tổng điểm của các sinh viên là lớn nhất

8. Liệt kê tổng điểm các tiêu chí và số lượng sinh viên mà tổng điểm các tiêu chí là bằng nhau

Mỗi nhóm sẽ lấy 2 tiêu chí bằng cách sau đây:

- **Tiêu chí 1:** $tc1 = MD \bmod 26 + 1$
 - **Tiêu chí 2:** $tc2 = (MD + tc1) \bmod 26 + 1$
9. Tính minimum, maximum, median, tứ phân vị (quartile) thứ nhất (Q1) và thứ ba (Q3) cho 2 tiêu chí trên của mẫu.
10. Tính tổng điểm trung bình của tập các tiêu chí về kỹ thuật, làm tròn 2 số thập phân
11. Tính tổng điểm trung bình của tập các tiêu chí về kỹ năng mềm, làm tròn 2 số thập phân
12. Tính tổng điểm trung bình cho các tiêu chí về tính kỷ luật, làm tròn 2 số thập phân
13. Liệt kê danh sách sinh viên mà có tổng điểm các tiêu chí về kỹ thuật là nhỏ nhất
14. Liệt kê danh sách sinh viên mà có tổng điểm các tiêu chí về kỹ thuật là lớn nhất
15. Liệt kê danh sách sinh viên mà có tổng điểm các tiêu chí về kỹ năng mềm là nhỏ nhất
16. Liệt kê danh sách sinh viên mà có tổng điểm các tiêu chí về kỹ năng mềm là lớn nhất
17. Liệt kê danh sách sinh viên mà có tổng điểm các tiêu chí về tính kỷ luật là nhỏ nhất
18. Liệt kê danh sách sinh viên mà có tổng điểm các tiêu chí về tính kỷ luật là lớn nhất
19. Kiểm tra các cặp tiêu chí sau xem liệu các công ty đã đánh giá tương thích chưa. Bằng cách liệt kê danh sách các sinh viên mà có sai lệch lớn về đánh giá các tiêu chí có liên quan với nhau cho từng sheets riêng lẻ:
- Khả năng ứng dụng kiến thức tin học trong công việc so với sử dụng Internet để phục vụ công việc
 - Nhận thức về trách nhiệm của sinh viên thực tập so với tính kỷ luật trong công việc
 - Khả năng trao đổi ý tưởng và chia sẻ công việc so với Giúp đỡ các bạn cùng nhóm trong công việc
 - Khả năng trao đổi ý tưởng và chia sẻ công việc so với đóng góp kỹ thuật vào dự án chung của nhóm
 - Khả năng sử dụng công cụ và công nghệ thích hợp tại đơn vị so với khả năng tự học
 - Khả năng thiết kế ứng dụng dựa trên yêu cầu của khách hàng so với làm chủ ít nhất một ngôn ngữ lập trình
 - Khả năng suy nghĩ logic và/hoặc vận dụng kiến thức toán học trong công việc so với sinh viên có thể đọc hiểu một phần thiết kế của một dự án thực tế

Biểu đồ

20. Hãy vẽ biểu đồ hiển thị phân phối tổng điểm các tiêu chí của sinh viên
21. Vẽ phổ tổng điểm các tiêu chí theo sinh viên
22. Vẽ phổ tổng điểm các sinh viên theo tiêu chí
23. Hãy vẽ đồ thị dotplot cho tiêu chí '01'. Khả năng ứng dụng kiến thức tin học trong công việc" và cho giải thích về hình dạng (shape), tâm (center), độ biến thiên (variability), outlier. của đồ thị cho từng sheets riêng lẻ,
- (a) shape: symmetric, skewed

- (b) center: median
 - (c) variability: dùng giá trị nhỏ nhất và lớn nhất khi mô tả sự thay đổi
 - (d) outlier: quan sát nằm phía ngoài tổng thể mẫu
24. Thử so sánh phân phối của 2 biến tiêu chí bằng cách vẽ đồ thị dotplot và thảo luận shape, outliers, center, và variability. Từ 2 tiêu chí được tính ra cho nhóm cho từng sheets riêng lẻ.
 25. Vẽ box-and-whisker cho 2 tiêu chí ở câu trên của nhóm
 26. Vẽ scatterplot cho tổng điểm trung bình gồm các sheets của tập các tiêu chí về kỹ thuật, làm tròn 2 số thập phân
 27. Vẽ scatterplot cho tổng điểm trung bình gồm các sheets của tập các tiêu chí về kỹ năng mềm, làm tròn 2 số thập phân
 28. Vẽ scatterplot cho tổng điểm trung bình gồm các sheets của tập các tiêu chí về tính kỷ luật, làm tròn 2 số thập phân.
 29. Vẽ biểu đồ tần số tương đối tích lũy của điểm của sinh viên trên tập các tiêu chí về kỹ thuật
 30. Vẽ biểu đồ tần số tương đối tích lũy của điểm của sinh viên trên tập các tiêu chí về kỹ năng mềm
 31. Vẽ biểu đồ tần số tương đối tích lũy của điểm của sinh viên trên tập các tiêu chí về tính kỷ luật
 32. Vẽ biểu đồ cho thấy số lượng sinh viên bị sai lệnh lớn trong đánh giá các cặp tiêu chí tương thích cho từng trường hợp ở câu hỏi trên cho 6 sheets

5 Code R và kết quả tính toán

Dữ liệu trong bài

Sinh viên trong mỗi sheet dữ liệu sẽ được phân biệt bằng tên kèm theo với sheet dữ liệu. Ví dụ: Sinh viên thứ nhất của sheet E1 sẽ là P1.1, sinh viên thứ 5 của sheet 2 sẽ là P2.5.

Các tiêu chí kỹ thuật là : 1, 2, 3, 4, 5, 11, 13, 14, 15, 16,17,19, 22.

Các tiêu chí kỹ năng mềm: 6, 8, 9, 10, 12, 18, 21, 23, 24, 26.

Các tiêu chí về kỷ luật: 7, 15 ,20.

Trả lời

1. Xác định số lượng sinh viên trong tập mẫu

```
>cau1<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet ="TONG HOP")
>soLuong= nrow(cau1)
>soLuong
>196
```

2. Tính điểm trung bình cho mỗi tiêu chí, làm tròn 2 số thập phân

Để tính trung bình cộng các tiêu chí của sinh viên

Ta lấy tổng điểm đánh giá sinh viên theo từng sheet rồi chia cho tổng sinh viên của sheet đó

Ví dụ: Ta trung bình cộng điểm sinh viên theo TC1=Tổng điểm của các sinh viên TC1/ 196sv

```
>data2<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "TONG HOP")
>i<-{1:196}
>k<-{2:27}
>cau2<-data.frame(data2[i,k])
>a2=array(dim=26)
>i<-1:196
>for(m in 1:26) a2[m]<-signif(sum(cau2[i,m])/196,digits = 3)
```

```
>q2<-data.frame(a2)
```

```
>q2
```

```
1 2.82
```

```
2 2.88
```

```
3 2.60
```

```
4 2.28
```

```
5 2.37
```

```
6 2.68
```

```
7 2.68
```

```
8 2.62
```

```
9 2.43
```

```
10 2.37
```

```
11 2.03
```

```
12 2.90
```

```
13 2.21
```

```
14 2.90
```

```
15 2.13
```

```
16 2.59
```

```
17 2.30
```

```
18 2.29
```

```
19 2.56
```

```
20 2.71
```

```
21 2.69
```

```
22 2.51
```

```
23 2.46
```

```
24 2.62
```

```
25 2.64
```

```
26 2.62
```

3. Tính điểm trung bình cho mỗi SV, làm tròn 2 số thập phân

```
cau3<-read_excel("C:/BTL/file/BTLCTRR.xlsx",sheet = "TONG HOP")
```

```
>dulieu <- data.frame(cau3[1:196,2:27])
```

```
>sum= c(rowSums(dulieu))
```

```
>SV = data.frame(cau3[,1])
```

```
>colnames(SV) <- c("SV")
```

```
>Trung_Binh_SV = data.frame(SV, GTTB = round(sum/26,digits = 2))
```

```
>Trung_Binh_SV
```

```
SV GTTB
```

```
P1.1 2,15
```

```
P1.2 2,5
```

```
P1.3 2,38
```

```
P1.4 2,58
```

```
P1.5 2,54
```

```
P1.6 2,69
```

```
P1.7 2,73
```

P1.8 2
P1.9 2,31
P1.10 1,88
P1.11 2,58
P1.12 2,23
P1.13 2,42
P1.14 2,77
P1.15 3,35
P1.16 1,96
P1.17 2,54
P1.18 1,92
P1.19 2,12
P1.20 2,96
P1.21 1,77
P1.22 2,62
P1.23 2,35
P1.24 2,31
P1.25 2,46
P1.26 2,42
P1.27 2,08
P1.28 2,35
P1.29 3,31
P1.30 2,27
P2.1 2,81
P2.2 2,08
P2.3 2
P2.4 2
P2.5 2,65
P2.6 1,62
P2.7 2,42
P2.8 2,31
P2.9 2,35
P2.10 2,31
P2.11 2,96
P2.12 2,54
P2.13 2,38
P2.14 2,38
P2.15 2,54
P2.16 2,12
P2.17 2,38
P2.18 2,38
P2.19 2,65
P2.20 2,62

P2.21 2,73
P2.22 3,19
P2.23 2,81
P2.24 2,23
P2.25 2,92
P2.26 2,46
P5.1 3,42
P5.2 3,73
P5.3 3,62
P5.4 3,54
P5.5 3,65
P5.6 3,92
P5.7 3,81
P5.8 3,73
P5.9 2,35
P5.10 2,31
P5.11 2,23
P5.12 2,73
P5.13 2,5
P5.14 2,42
P5.15 2,38
P5.16 2,27
P5.17 2,08
P5.18 2,31
P5.19 2,85
P5.20 2,88
P5.21 2,23
P5.22 2,27
P5.23 2,69
P5.24 2,54
P5.25 2,19
P5.26 1,88
P5.27 2,31
P5.28 2,42
P5.29 2,27
P5.30 2,31
P5.31 2,92
P5.32 1,62
P5.33 2,27
P5.34 2,19
P5.35 2,27
P6.1 2,58
P6.2 0,38

P6.3 2,58
P6.4 2,58
P6.5 2,58
P6.6 2,77
P6.7 2,19
P6.8 2,92
P6.9 2,58
P6.10 2,54
P6.11 2,65
P6.12 2,77
P6.13 1,88
P6.14 2,27
P6.15 2,38
P6.16 2,08
P6.17 3,35
P6.18 2,27
P6.19 2,58
P6.20 2,12
P6.21 2,73
P6.22 2,96
P6.23 2,62
P6.24 2,31
P6.25 2,65
P6.26 2,85
P6.27 2,5
P7.1 4,04
P7.2 3,73
P7.3 2,73
P7.4 2,85
P7.5 2,85
P7.6 2,88
P7.7 2,31
P7.8 2,19
P7.9 2,62
P7.10 2,69
P7.11 2,5
P7.12 2,38
P7.13 2,46
P7.14 2,88
P7.15 2,42
P7.16 2,38
P7.17 2,5
P7.18 2,08

P7.19 2,85
P7.20 2
P7.21 3,27
P7.22 2,54
P7.23 2,23
P7.24 2,5
P7.25 3,23
P7.26 2,08
P7.27 2,5
P7.28 2,65
P7.29 2,62
P7.30 2,96
P7.31 2,23
P7.32 2,92
P7.33 2,65
P7.34 2,69
P7.35 2,85
P7.36 3,23
P7.37 2,65
P7.38 2,69
P7.39 2,08
P7.40 2,46
P8.1 2,73
P8.2 2,77
P8.3 2,58
P8.4 1,65
P8.5 2,27
P8.6 2,58
P8.7 1,65
P8.8 2,5
P8.9 1,42
P8.10 2,81
P8.11 2,65
P8.12 2,27
P8.13 2,38
P8.14 3
P8.15 2,12
P8.16 2,81
P8.17 2,73
P8.18 3,04
P8.19 2,38
P8.20 2,5
P8.21 2,23

P8.22 2,88
P8.23 2,27
P8.24 2,58
P8.25 2,19
P8.26 2,65
P8.27 3,19
P8.28 2,38
P8.29 1,92
P8.30 2,46
P8.31 2,77
P8.32 2,81
P8.33 2,5
P8.34 2,58
P8.35 2,23
P8.36 2,08
P8.37 2,65
P8.38 2,92

4. Liệt kê danh sách sinh viên có tổng điểm các tiêu chí là nhỏ nhất

```
>cau4<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "TONG HOP")
>dulieu <- data.frame(cau4[1:196,2:27])
>tongdiemTC_SV = (rowSums(dulieu))
>minimum = min(tongdiemTC_SV)
>svNameMin = c()
>findMin = c()
>k = 1
>for( i in 1:196){
>if(tongdiemTC_SV[i] == minimum){
>findMin[k] = tongdiemTC_SV[i]
>svNameMin[k] = Trung_Binh_SV[i,1]
>k = k+1
}
}
>SV_Co_tong_diem_nho_nhat = data.frame(Ten_SV = svNameMin , Tong_diem = findMin)
>SV_Co_tong_diem_nho_nhat
>Ten_SV Tong_diem
P6.2 0.38
```

5. Liệt kê danh sách sinh viên có tổng điểm các tiêu chí là lớn nhất

```
>tongdiemTC_SV = (rowSums(dulieu))
>maximum = max(tongdiemTC_SV)
>svNameMax = c()
>findMax = c()
>k = 1
>for( i in 1:196){
```

```
>if(tongdiemTC_SV[i] == maximum){  
>findMax[k] = tongdiemTC_SV[i]  
>svNameMax[k] = Trung_Binh_SV[i,1]  
>k = k+1  
}  
}  
SV_Co_tong_diem_lon_nhat = data.frame(Ten_SVmax = svNameMax , Tong_diemmax= findMax)  
>SV_Co_tong_diem_lon_nhat  
>Ten_SVmax Tong_diemmax  
>P7.1 105
```

6. Liệt kê danh sách tiêu chí mà tổng điểm của các sinh viên là nhỏ nhất

```
>data6<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "TONG HOP")  
>i<-1:196  
>k<-2:27  
>cau6<-data.frame(data6[i,k])  
>b6=array(dim=26)  
>for(m in 1:26) b6[m]<-sum(cau6[i,m])  
>min=600;  
>for(i in 1:26) if(b6[i]>min){min<-min}else{min<-b6[i]}  
>for(i in 1:26) if(b6[i]==min){print(i)}  
[1] 11
```

7. Liệt kê danh sách tiêu chí mà tổng điểm của các sinh viên là lớn nhất

```
>data7<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "TONG HOP")  
>i<-1:196  
>k<-2:27  
>cau7<-data.frame(data7[i,k])  
>c7=array(dim=26)  
>for(m in 1:26) c7[m]<-sum(cau7[i,m])  
>max=0;  
>for(i in 1:26) if(c7[i]>max){max<-c7[i]}else{max<-max}  
>for(i in 1:26) if(c7[i]==max){print(i)}  
[1] 12  
[1] 14
```

8. Liệt kê tổng điểm các tiêu chí và số lượng sinh viên mà tổng điểm các tiêu chí là bằng nhau

```
>ctrr_c8 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet ="TONG HOP")  
>tongdiem_c8 <- rowSums(ctrr_c8 [,2:27])  
>sv_c8 <- ctrr_c8 [,1]  
>thongke_c8 <- data.frame(sv_c8,tongdiem_c8)  
>tongdiem_tc_c8 <- unique(thongke_c8$tongdiem_c8)  
>tk_c8 <- list();  
>soluong_sv_c8 <- c();  
>len_c8 <- length(tongdiem_tc_c8);  
>for (i in 1:len_c8) {  
> tk_c8[[i]] <- thongke_c8[thongke_c8$tongdiem_c8 == >tongdiem_tc_c8[i],];  
>soluong_sv_c8[i] <- length(tk_c8[[i]]$sv_c8);
```

```
}
>c8 <- data.frame(tongdiem_tc_c8, soluong_sv_c8)
Kết quả:
```

	tongdiem_tc	soluong_sv
1	56	1
2	65	10
3	62	12
4	67	12
5	66	7
6	70	5
7	71	7
8	52	4
9	60	10
10	49	3
11	58	8
12	63	6
13	72	5
14	87	2
15	51	1
16	50	2
17	55	4
18	77	4
19	46	1
20	68	5
21	61	4
22	64	5
23	61	5

Showing 1 to 23 of 48 entries, 2 total columns

Mỗi nhóm sẽ lấy 2 tiêu chí bằng cách sau đây:

- **Tiêu chí 1:** $tc1 = MD \bmod 26 + 1$
- **Tiêu chí 2:** $tc2 = (MD + tc1) \bmod 26 + 1$

9. Tính minimum, maximum, median, tứ phân vị (quartile) thứ nhất (Q1) và thứ ba (Q3) cho 2 tiêu chí trên của mẫu.

MD nhóm 8: 506417

Ta tính được: $tc1 = 16$, $tc2 = 6$

```
>ctrr_c9 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet = "TONG HOP")
```

```
>min(ctrr_c9[,7])
```

```
>max(ctrr_c9[,7])
```

```
>median(ctr_c9[,7])
>quantile(ctr_c9[,7])
>min(ctr_c9[,17])
>max(ctr_c9[,17])
>median(ctr_c9[,17])
>quantile(ctr_c9[,17])
>Kết quả:
>Tiêu chí 6
>Min = 0
>Max = 5
>Median = 3
Phân vị thứ nhất: 1
Phân vị thứ ba: 4
0% 25% 50% 75% 100%
0 1 3 4 5
Tương tự cho tiêu chí 16
Min = 0
Max = 5
Trung vị = 3
Phân vị thứ nhất = 1
Phân vị thứ ba = 4
```

10. Tính tổng điểm trung bình của tập các tiêu chí về kỹ thuật, làm tròn 2 số thập phân

```
>data10<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "TONG HOP")
>i<-1:196
>k<-2:27
>cau10<-data.frame(data10[i,k])
>d10<-data.frame(cau10[,2],cau10[,3],cau10[,4],cau10[,5],
+cau10[,6],cau10[,12],cau10[,14],cau10[,15],
cau10[,16],cau10[,17],cau10[,18],cau10[,20],cau10[,23])
> k10=array(dim=13)
> for(m in 1:13) k10[m]<-signif(sum(d10[i,m])/196,digits = 3)
>r10<-c("TC.1", "TC.2", "TC.3", "TC.4", "TC.5", "TC.11", "TC.13", "TC.14", "TC.15", "TC.16",
"TC.17", "TC.19", "TC.22")
>w10<-data.frame(TC=r10,TrungBinhCong=k10)
>w10
TC TrungBinhCong
1 TC.1 2.82
2 TC.2 2.88
3 TC.3 2.60
4 TC.4 2.28
5 TC.5 2.37
6 TC.11 2.03
7 TC.13 2.21
8 TC.14 2.90
9 TC.15 2.13
10 TC.16 2.59
11 TC.17 2.30
```

12 TC.19 2.56

13 TC.22 2.51

11. Tính tổng điểm trung bình của tập các tiêu chí về kỹ năng mềm, làm tròn 2 số thập phân

```
>data11<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "TONG HOP")
>i<-{1:196}
>k<-{2:27}
>cau11<-data.frame(data11[i,k])
>h11<-data.frame(cau11[,6],cau11[,8],cau11[,9],cau11[,10],cau11[,12],cau11[,18],cau11[,21],
cau11[,23],cau11[,24],cau11[,26])
>p11<-array(dim=10)
>for(m in 1:10) p11[m]<-signif(sum(h11[i,m])/196,digits = 3)
>e11<-c("TC.6","TC.8","TC.9","TC.10","TC.12","TC.18","TC.21","TC.23","TC.24","TC.26")
>t11<-data.frame(TC=e11,TrungBinhCong=p11)
>t11
```

TC TrungBinhCong

1 TC.6 2.68

2 TC.8 2.62

3 TC.9 2.43

4 TC.10 2.37

5 TC.12 2.90

6 TC.18 2.29

7 TC.21 2.69

8 TC.23 2.46

9 TC.24 2.62

10 TC.26 2.62

12. Tính tổng điểm trung bình cho các tiêu chí về tính kỷ luật, làm tròn 2 số thập phân

```
> cau12 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "KL",colNames = TRUE)
> hang <- {1:196}
> ky luat12 <- data.frame(cau12[hang, 2:4])
> a12=array(dim=3)
> for (i in 1:3) a12[i]<- signif(sum(ky luat12[hang,i])/196, digits = 3)
> TC <- c("TC 7", "TC 15", "TC 20")
> TB <- c(a12[1],a12[2],a12[3])
> tbkl12 <- data.frame(TC,TB)
> View(tbkl12)
```

TC	TB
TC 7	2.68
TC 15	2.13
TC 20	2.71

13. Liệt kê danh sách sinh viên mà có tổng điểm các tiêu chí về kỹ thuật là nhỏ nhất
- ```
> cau13 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet =8, colNames = TRUE)
> kythuat13 <- data.frame(cau13[hang,1:14])
> b13 = array(dim = 196)
> for(i in 1:196) b13[i] <- sum (kythuat13[i,2:14])
> min13 = 100
> for (j in 1:196) if (b13 < min13) {min13<-b13[j]}
> for (i in 1:196) if (b13[i]== min13) {print(kythuat13[i,1])}
"P1.1"
```

Kết quả in ra màn hình là sinh viên thứ nhất trong sheet E1.

14. Liệt kê danh sách sinh viên mà có tổng điểm các tiêu chí về kỹ thuật là lớn nhất  
Sử dụng lại điểm tổng của sinh viên về tiêu chí kỹ thuật ở trên, dùng vòng lặp để xác định những sinh viên có điểm lớn nhất.

```
> max14 =0
> for (m in 1:196) if (b13[m]> max14){max14<- b13[m]}
> for (n in 1:196) if (b13[n]== max14){print (kythuat13[n,1])}
> "P1.15"
```

Kết quả in ra màn hình là sinh viên thứ 15 trong sheet E1.

15. Liệt kê danh sách sinh viên mà có tổng điểm các tiêu chí về kỹ năng mềm là nhỏ nhất

```
> cau15 <-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = 9, colNames = TRUE)
> knm15 <- data.frame(cau15[hang,1:11])
> c15 = array(dim=196)
> for (i in 1:196) c15[i]<- sum(knm15[i,2:11])
> min15=1000
> for (k in 1:196) if(c15[k]<min15){min15<- c15[k]}
> for (l in 1 :196) if (c15[l]== min15) {print(knm15[l,1])}
> "P6.2"
```

Kết quả in ra màn hình là sinh viên thứ 2 của sheet E6.

16. Liệt kê danh sách sinh viên mà có tổng điểm các tiêu chí về kỹ năng mềm là lớn nhất

```
>max16=0
> for (o in 1:196) if(c15[o]>max16){max16<- c15[o]}
> for (u in 1 :196) if (c15[u] == max16) {print(knm15[u,1])}
> "P7.1"
```

Kết quả in ra màn hình là sinh viên thứ nhất của sheet E7.

17. Liệt kê danh sách sinh viên mà có tổng điểm các tiêu chí về tính kỷ luật là nhỏ nhất

```
>kyluat_c17 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet="KL")
>tong_kyluat_c17 <- rowSums(kyluat_c17[,2:4])
>sv_c17 <- kyluat_c17[,1]
>thongke_kyluat_c17 <- data.frame(sv_c17,tong_kyluat_c17)
>tk_nho_c17 <- thongke_kyluat_c17 [order(>thongke_kyluat_c17$tong_kyluat_c17),]
```

Kết quả:

|     | sv    | tong_ky luat |
|-----|-------|--------------|
| 83  | P5.27 | 0            |
| 17  | P1.17 | 1            |
| 98  | P6.7  | 1            |
| 167 | P8.9  | 1            |
| 12  | P1.12 | 2            |
| 77  | P5.21 | 2            |
| 90  | P5.34 | 2            |
| 93  | P6.2  | 2            |
| 129 | P7.11 | 2            |
| 157 | P7.39 | 2            |
| 28  | P1.28 | 3            |
| 47  | P2.17 | 3            |
| 84  | P5.28 | 3            |
| 88  | P5.32 | 3            |
| 104 | P6.13 | 3            |
| 111 | P6.20 | 3            |
| 162 | P8.4  | 3            |
| 165 | P8.7  | 3            |
| 3   | P1.3  | 4            |
| 8   | P1.8  | 4            |
| 10  | P1.10 | 4            |
| 16  | P3.16 | 4            |

Showing 1 to 23 of 196 entries, 2 total columns

18. Liệt kê danh sách sinh viên mà có tổng điểm các tiêu chí về tính kỷ luật là lớn nhất

```
>tk_lon_c18 <- thongke_ky luat_c17[order(-thongke_ky luat_c17$tong_ky luat_c17),]
```

Kết quả:

|     | sv    | tong_ky luat |
|-----|-------|--------------|
| 49  | P2.19 | 15           |
| 14  | P1.14 | 14           |
| 23  | P1.23 | 14           |
| 87  | P5.31 | 14           |
| 143 | P7.25 | 14           |
| 188 | P8.30 | 14           |
| 62  | P5.6  | 13           |
| 69  | P5.13 | 13           |
| 15  | P1.15 | 12           |
| 29  | P1.29 | 12           |
| 58  | P5.2  | 12           |
| 75  | P5.19 | 12           |
| 102 | P6.11 | 12           |
| 119 | P7.1  | 12           |
| 122 | P7.4  | 12           |
| 132 | P7.14 | 12           |
| 154 | P7.36 | 12           |
| 61  | P5.5  | 11           |
| 63  | P5.7  | 11           |
| 64  | P5.8  | 11           |
| 76  | P5.20 | 11           |
| 100 | P6.17 | 11           |

Showing 1 to 23 of 196 entries, 2 total columns

19. Kiểm tra đánh giá tương thích giữa các cặp tiêu chí tương ứng:

- Khả năng ứng dụng kiến thức tin học trong công việc so với sử dụng Internet để phục vụ công việc
- Nhận thức về trách nhiệm của sinh viên thực tập so với tính kỷ luật trong công việc
- Khả năng trao đổi ý tưởng và chia sẻ công việc so với Giúp đỡ các bạn cùng nhóm trong công việc
- Khả năng trao đổi ý tưởng và chia sẻ công việc so với đóng góp kỹ thuật vào dự án chung của nhóm
- Khả năng sử dụng công cụ và công nghệ thích hợp tại đơn vị so với khả năng tự học
- Khả năng thiết kế ứng dụng dựa trên yêu cầu của khách hàng so với làm chủ ít nhất một ngôn ngữ lập trình
- Khả năng suy nghĩ logic và/hoặc vận dụng kiến thức toán học trong công việc so với sinh viên có thể đọc hiểu một phần thiết kế của một dự án thực tế

Nhóm thống nhất với những cặp tiêu chí tương ứng, sai lệch từ 2 điểm trở lên sẽ được tính là sai lệch lớn. Từ đó ta có kết quả như đã tính.

```
>e1_c19 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet = "E1")
>e2_c19 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet = "E2")
>e5_c19 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet = "E5")
>e6_c19 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet = "E6")
>e7_c19 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet = "E7")
>e8_c19 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet = "E8")
>hieu_e1_c19 <- data.frame();
>len <- length(e1_c19[,1]); tc <- data.frame(c(2, 8, 7, 7, 4, 5, 3), >c(15, 26, 22, 23, 13, 17, 20));
>for (i in 1: len) {
+for (j in 1:7) {
+ hieu_e1_c19[i,j] <- abs(e1_c19[i,tc[j,1]] - e1_c19[i,tc[j,2]])
+}
+}
>s1sv_e1_c19 <- c();
>for (i in 1:7) {
+ s1sv_e1_c19[i] <- length(hieu_e1_c19[hieu_e1_c19[i]>=2,i])
+}
>hieu_e2_c19 <- data.frame();
>len <- length(e2_c19[,1]); tc <- data.frame(c(2, 8, 7, 7, 4, 5, 3), >c(15, 26, 22, 23, 13, 17, 20));
>for (i in 1: len) {
+for (j in 1:7) {
+ hieu_e2_c19[i,j] <- abs(e2_c19[i,tc[j,1]] - e2_c19[i,tc[j,2]])
+}
+}
>s1sv_e2_c19 <- c();
>for (i in 1:7) {
+ s1sv_e2_c19[i] <- length(hieu_e2_c19[hieu_e2_c19[i]>=2,i])
+}
>hieu_e5_c19 <- data.frame();
>len <- length(e5_c19[,1]); tc <- data.frame(c(2, 8, 7, 7, 4, 5, 3), >c(15, 26, 22, 23, 13, 17, 20));
>for (i in 1: len) {
+for (j in 1:7) {
+ hieu_e5_c19[i,j] <- abs(e5_c19[i,tc[j,1]] - e5_c19[i,tc[j,2]])
+}
+}
>s1sv_e5_c19 <- c();
>for (i in 1:7) {
+ s1sv_e5_c19[i] <- length(hieu_e5_c19[hieu_e5_c19[i]>=2,i])
+}
>hieu_e6_c19 <- data.frame();
>len <- length(e6_c19[,1]); tc <- data.frame(c(2, 8, 7, 7, 4, 5, 3), >c(15, 26, 22, 23, 13, 17, 20));
>for (i in 1: len) {
+for (j in 1:7) {
+ hieu_e6_c19[i,j] <- abs(e6_c19[i,tc[j,1]] - e6_c19[i,tc[j,2]])
+}
+}
>s1sv_e6_c19 <- c();
>for (i in 1:7) {
+ s1sv_e6_c19[i] <- length(hieu_e6_c19[hieu_e6_c19[i]>=2,i])
+}
>hieu_e7_c19 <- data.frame();
>len <- length(e7_c19[,1]); tc <- data.frame(c(2, 8, 7, 7, 4, 5, 3), >c(15, 26, 22, 23, 13, 17, 20));
>for (i in 1: len) {
+for (j in 1:7) {
+ hieu_e7_c19[i,j] <- abs(e7_c19[i,tc[j,1]] - e7_c19[i,tc[j,2]])
+}
+}
>s1sv_e7_c19 <- c();
>for (i in 1:7) {
```

```
+slsv_e7_c19[i] <- length(hieu_e7_c19[hieu_e7_c19[,i]>=2,i])
}
>hieu_e8_c19 <- data.frame();
>len <- length(e8_c19[1]); tc <- data.frame(c(2, 8, 7, 7, 4, 5, 3), >c(15, 26, 22, 23, 13, 17, 20));
>for (i in 1: len) {
 +for (j in 1:7) {
 +hieu_e8_c19[i,j] <- abs(e8_c19[i,tc[j,1]] - e8_c19[i,tc[j,2]])
 }
}
>slsv_e8_c19 <- c();
>for (i in 1:7) {
 +slsv_e8_c19[i] <- length(hieu_e8_c19[hieu_e8_c19[,i]>=2,i])
}
```

Kết quả Sheet 1 ( vì số lượng dữ liệu lớn nên các sheet còn lại được trình bày rõ hơn dưới dạng biểu đồ câu 32)

|    | V1 | V2 | V3 | V4 | V5 | V6 | V7 |
|----|----|----|----|----|----|----|----|
| 1  | 1  | 2  | 3  | 1  | 0  | 0  | 1  |
| 2  | 2  | 1  | 1  | 1  | 0  | 0  | 1  |
| 3  | 0  | 1  | 1  | 1  | 2  | 3  | 2  |
| 4  | 4  | 5  | 1  | 1  | 2  | 1  | 1  |
| 5  | 5  | 0  | 3  | 2  | 1  | 5  | 0  |
| 6  | 3  | 0  | 3  | 4  | 0  | 3  | 1  |
| 7  | 4  | 1  | 2  | 1  | 0  | 2  | 3  |
| 8  | 1  | 1  | 1  | 2  | 1  | 2  | 2  |
| 9  | 1  | 2  | 3  | 1  | 1  | 4  | 2  |
| 10 | 2  | 2  | 0  | 2  | 0  | 3  | 2  |
| 11 | 1  | 4  | 2  | 2  | 0  | 3  | 2  |
| 12 | 1  | 2  | 3  | 0  | 1  | 5  | 2  |
| 13 | 5  | 0  | 3  | 1  | 2  | 0  | 5  |
| 14 | 2  | 4  | 0  | 2  | 0  | 1  | 4  |
| 15 | 2  | 4  | 5  | 0  | 0  | 1  | 2  |
| 16 | 1  | 3  | 1  | 4  | 0  | 2  | 0  |
| 17 | 1  | 2  | 3  | 2  | 0  | 1  | 2  |
| 18 | 5  | 4  | 2  | 2  | 3  | 3  | 0  |
| 19 | 3  | 0  | 1  | 1  | 1  | 4  | 1  |
| 20 | 2  | 0  | 0  | 2  | 3  | 3  | 1  |
| 21 | 2  | 4  | 1  | 0  | 1  | 0  | 1  |

Showing 1 to 23 of 30 entries, 7 total columns

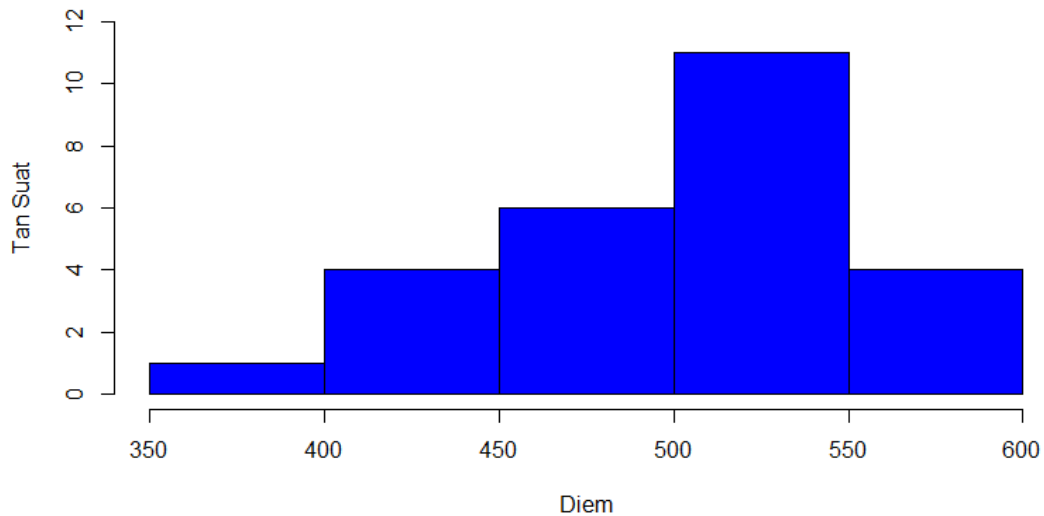
## Biểu đồ

20. Hãy vẽ biểu đồ hiển thị phân phối tổng điểm các tiêu chí của sinh viên

```
>cau20<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "TONG HOP")
>dulieu <- data.frame(cau20[1:196,2:27])
>diemTongTC = c()
>for (i in 1:26){
 >values= sum(dulieu[1:196,i])
 >diemTongTC[i] <- values
}
>hist(diemTongTC,xlab="Diem",ylab="Tan Suat",
 +main="Biểu đồ hiển thị phân phối tổng điểm các tiêu chí của SV",
```

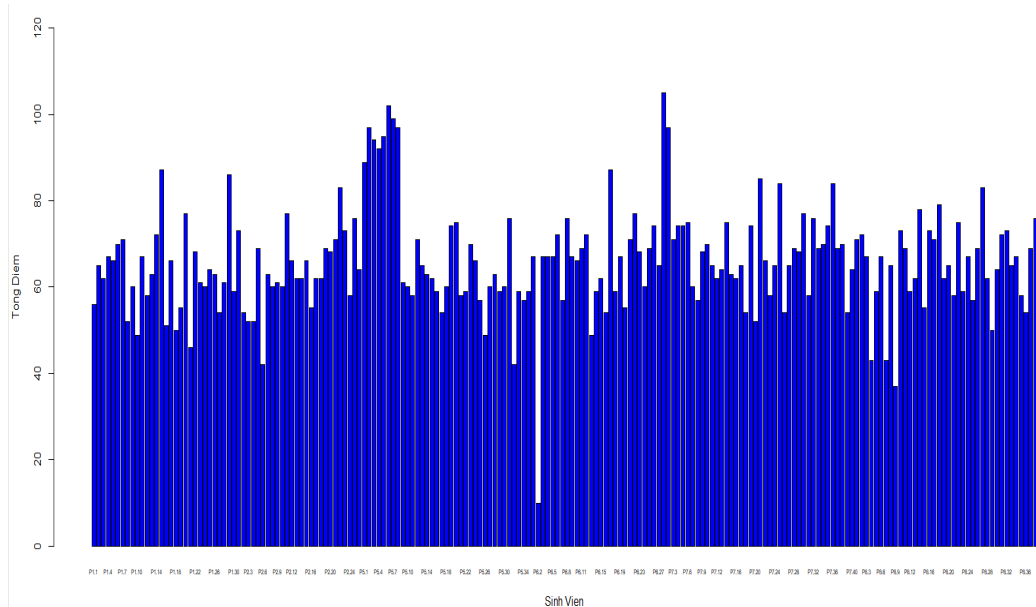
```
+ylim=c(0,12),col="blue")
```

**Biểu đồ hiển thị phân phối tổng điểm các tiêu chí của SV**



21. Vẽ phổ tổng điểm các tiêu chí theo sinh viên

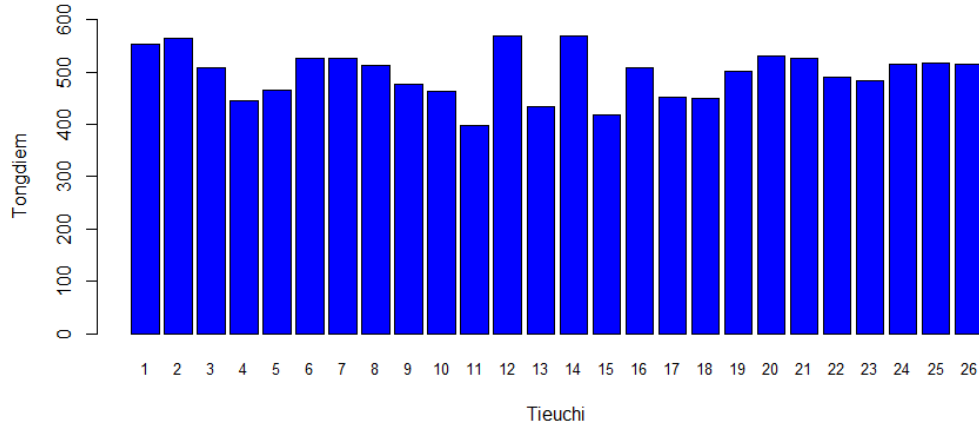
```
>cau21<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "TONG HOP")
>dulieu <- data.frame(cau21[1:196,2:27])
>sum= c(rowSums(dulieu))
>tong_diem_SV = data.frame(SV = cau21[1:196,1], tongDiemSV = sum)
>barplot(tong_diem_SV$tongDiemSV,names.arg=tong_diem_SV$SV,xlab="Sinh
Vien",ylab="Tong Diem",cex.names=0.5,ylim=c(0,120),col = "blue")
```



22. Vẽ phổ tổng điểm các sinh viên theo tiêu chí

```
>data22<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "TONG HOP")
>i<-{1:196}
>k<-{2:27}
>cau22<-data.frame(data22[i,k])
```

```
>c=array(dim=26)
>for(m in 1:26) c[m]<-sum(cau22[i,m])
>k<-c(1:26)
> barplot(c,names.arg =k,xlab="Tieuchi",ylab="Tongdiem",ylim=c(0,600),col="blue",
cex.names = 0.8)
```



23. Hãy vẽ đồ thị dotplot cho tiêu chí ‘01’. Khả năng ứng dụng kiến thức tin học trong công việc” và cho giải thích về hình dạng (shape), tâm (center), độ biến thiên (variability), outlier. của đồ thị cho từng sheets riêng lẻ, abcd

- (a) shape: symmetric, skewed
- (b) center: median
- (c) variability: dùng giá trị nhỏ nhất và lớn nhất khi mô tả sự thay đổi
- (d) outlier: quan sát nằm phía ngoài tổng thể mẫu

Để tìm outlier cho đồ thị dotplot ta dùng công thức:

$$Lowfence = Q1 - IQR$$

$$Highfence = Q3 - IQR$$

$$IQR = Q3 - Q1$$

Lowfence là biên dưới của đồ thị, Highfence là biên trên. Những giá trị ngoài lowfence và highfence là outlier của đồ thị.

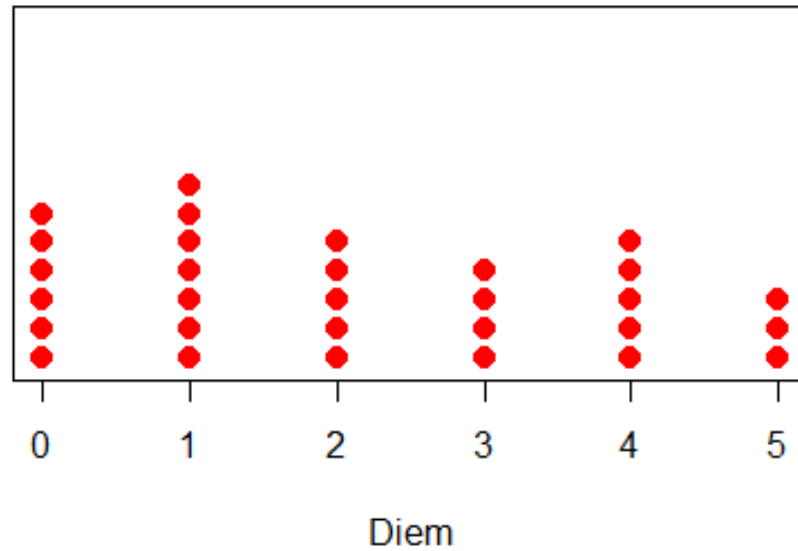
```
>e1_c23 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet = "E1")
>e2_c23 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet = "E2")
>e5_c23 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet = "E5")
>e6_c23 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet = "E6")
>e7_c23 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet = "E7")
>e8_c23 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", sheet = "E8")
>qe1_c23 <- quantile(as.vector(table(e1_c23[,2])))
>q1_e1 <- qe1_c23[2]
>q3_e1 <- qe1_c23[4]
>out_low_e1 <- q1_e1 - 1.5*(q3_e1-q1_e1)
>out_high_e1 <- q3_e1 + 1.5*(q3_e1-q1_e1)
>qe2_c23 <- quantile(as.vector(table(e2_c23[,2])))
>q1_e2 <- qe2_c23[2]
>q3_e2 <- qe2_c23[4]
>out_low_e2 <- q1_e2 - 1.5*(q3_e2-q1_e2)
```

```
>out_high_e2 <- q3_e2 + 1.5*(q3_e2-q1_e2)
>qe5_c23 <- quantile(as.vector(table(e5_c23[,2])))
>q1_e5 <- qe5_c23[2]
>q3_e5 <- qe5_c23[4]
>out_low_e5 <- q1_e5 - 1.5*(q3_e5-q1_e5)
>out_high_e5 <- q3_e5 + 1.5*(q3_e5-q1_e5)
>qe6_c23 <- quantile(as.vector(table(e6_c23[,2])))
>q1_e6 <- qe6_c23[2]
>q3_e6 <- qe6_c23[4]
>out_low_e6 <- q1_e6 - 1.5*(q3_e6-q1_e6)
>out_high_e6 <- q3_e6 + 1.5*(q3_e6-q1_e6)
>qe7_c23 <- quantile(as.vector(table(e7_c23[,2])))
>q1_e7 <- qe7_c23[2]
>q3_e7 <- qe7_c23[4]
>out_low_e7 <- q1_e7 - 1.5*(q3_e7-q1_e7)
>out_high_e7 <- q3_e7 + 1.5*(q3_e7-q1_e7)
>qe8_c23 <- quantile(as.vector(table(e8_c23[,2])))
>q1_e8 <- qe8_c23[2]
>q3_e8 <- qe8_c23[4]
>out_low_e8 <- q1_e8 - 1.5*(q3_e8-q1_e8)
>out_high_e8 <- q3_e8 + 1.5*(q3_e8-q1_e8)
>stripchart(e1_c23[,2],method="stack",
+pch=20,cex=2,las=1,col="red",at=c(0.05),
+main = "Sheet E1 - TC01", xlab = "Diem")
>stripchart(e2_c23[,2],method="stack",
+pch=20,cex=2,las=1,col="blue",at=c(0.05),
+main = "Sheet E2 - TC01", xlab = "Diem")
>stripchart(e5_c23[,2],method="stack",
+pch=20,cex=2,las=1,col="green",at=c(0.05),
+main = "Sheet E5 - TC01", xlab = "Diem",offset = 1/4)
>stripchart(e6_c23[,2],method="stack",
+pch=20,cex=2,las=1,col="purple",at=c(0.05),
+main = "Sheet E6 - TC01", xlab = "Diem")
>stripchart(e7_c23[,2],method="stack",
+pch=20,cex=2,las=1,col="orange",at=c(0.05),
+main = "Sheet E7 - TC01", xlab = "Diem")
>stripchart(e8_c23[,2],method="stack",
+pch=20,cex=2,las=1,col="magenta",at=c(0.05),
+main = "Sheet E8 - TC01", xlab = "Diem",offset = 1/4)
```

Kết quả



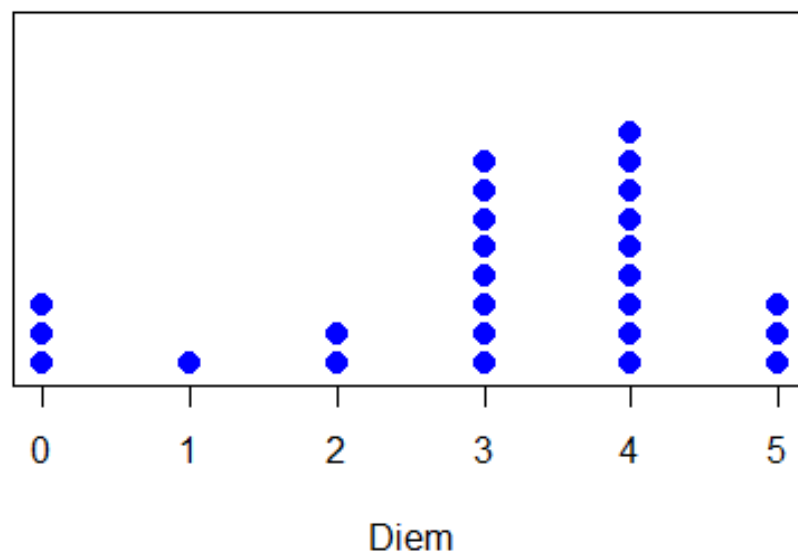
## Sheet E1 - TC01



Sheet E1

- (a) shape: không đối xứng
- (b) center: median = 5
- (c) variability: min = 3, max = 7
- (d) outlier: không có giá trị nào ( $> 8$  hoặc  $< 2$ )

## Sheet E2 - TC01

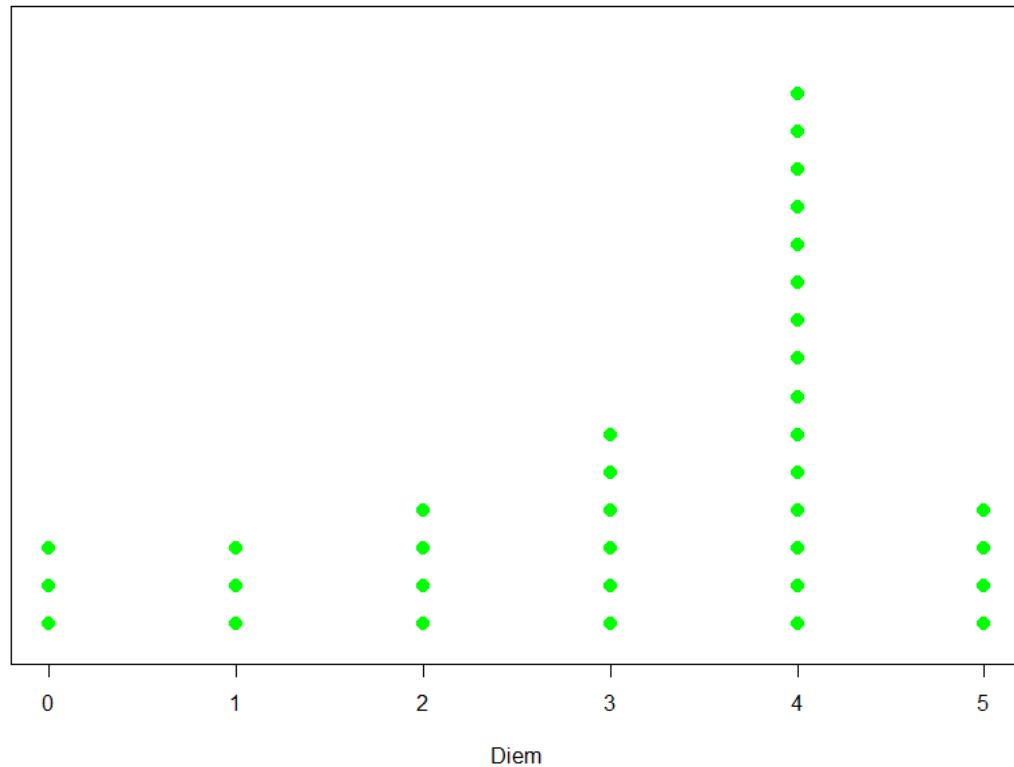


Sheet E2

- (a) shape: không đối xứng
- (b) center: median = 3

- (c) variability: min = 1, max: = 9  
(d) outlier: không có giá trị nào ( $> 13.5$  hoặc  $< -4.5$ )

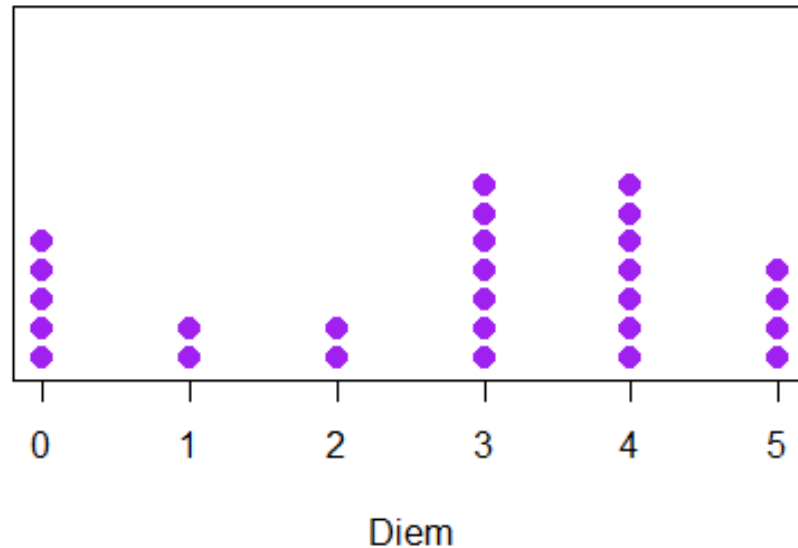
**Sheet E5 - TC01**



Sheet E5

- (a) shape: không đối xứng  
(b) center: median = 4  
(c) variability: min = 3, max = 15  
(d) outlier: có 1 giá trị outlier ( $15 > 8.88$ ), không có giá trị nào  $< -0.125$

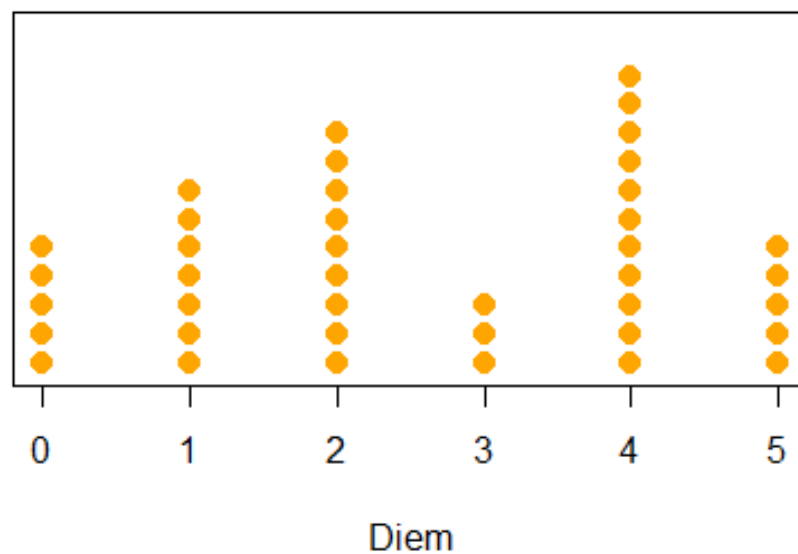
## Sheet E6 - TC01



Sheet E6

- (a) shape: không đối xứng
- (b) center: median = 4.5
- (c) variability: min = 0, max = 5
- (d) outlier: không có giá trị nào ( $> 12.5$  hoặc  $< -3.5$ )

## Sheet E7 - TC01

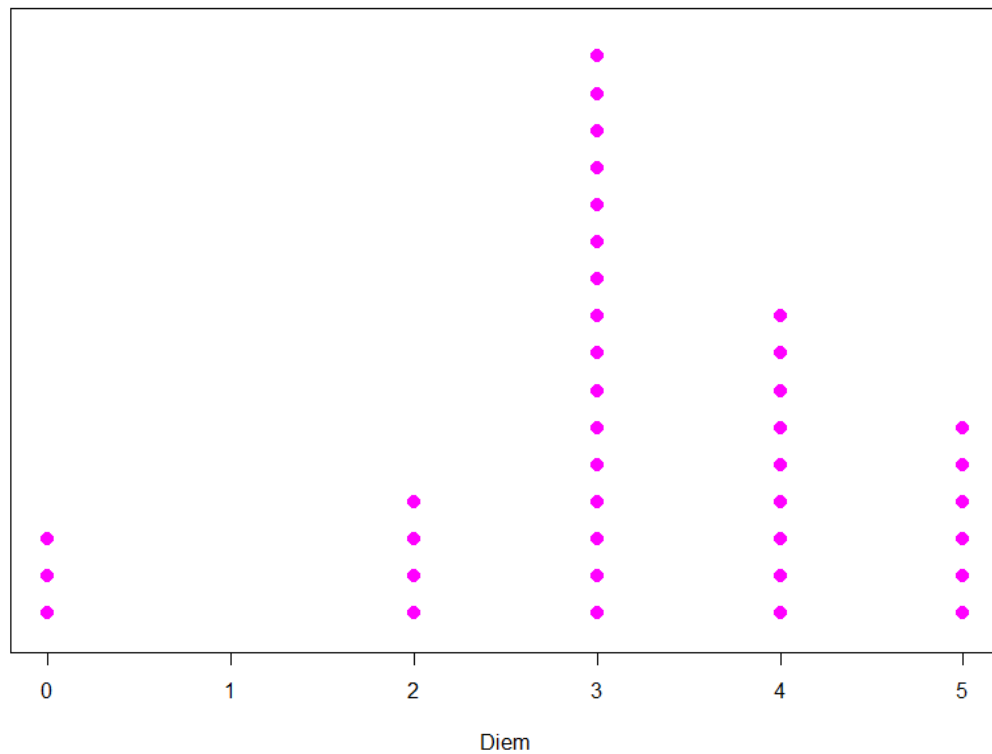


Sheet E7

- (a) shape: không đối xứng
- (b) center: median = 4

- (c) variability: min = 3, max = 11  
(d) outlier: không có giá trị nào ( $> 13.5$  hoặc  $< -0.25$ )

**Sheet E8 - TC01**



Sheet E8

- (a) shape: không đối xứng  
(b) center: median = 5  
(c) variability: min = 0, max = 16  
(d) outlier: có 1 giá trị outlier ( $16 > 15.75$ ), không có giá trị nào  $< -4.25$

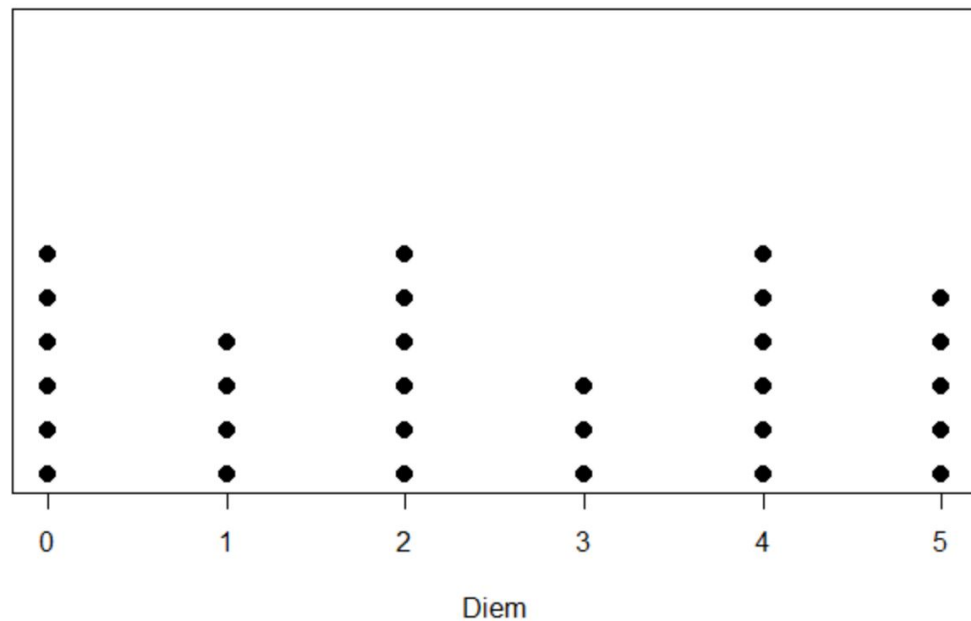
24. Thử so sánh phân phối của 2 biến tiêu chí bằng cách vẽ đồ thị dotplot và thảo luận shape, outliers, center, và variability. Từ 2 tiêu chí được tính ra cho nhóm cho từng sheets riêng lẻ. Vẽ đồ thì với Sheet E1.

#Tiêu chí 6

```
> cau24e1<- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = 1, colNames = TRUE)
> stripchart(cau24e1[,7], method = "stack",
+ pch = 20, cex = 2 , las=1,at= c(0), offset = 1/5,
+ main= "SHEET E1 TC6", xlab = "Diem")
> lf6e1 = quantile(cau24e1[,7],probs = 0.25)-1.5*IQR(cau24e1[,7])
> hf6e1 = quantile(cau24e1[,7],probs = 0.75)+1.5*IQR(cau24e1[,7])
```

Tính được giá trị lf6e1 = -3.5 và hf6e1 = 8.5

## SHEET E1 TC6



Đồ thị cho tiêu 6 sheet E1:

+ Không có outliner.

+ Center = 2

+ Shape: đồ thị tương đối cân bằng

+ Validity : min = 0 , max =5

#Tiêu chí 16

```
> stripchart(cau24e1[,17], method = "stack",
+ pch = 20, cex=2, las=1,at=c(0), offset = 1/5,
+ main="SHEET E1 TC16", xlab = "DIEM")
```

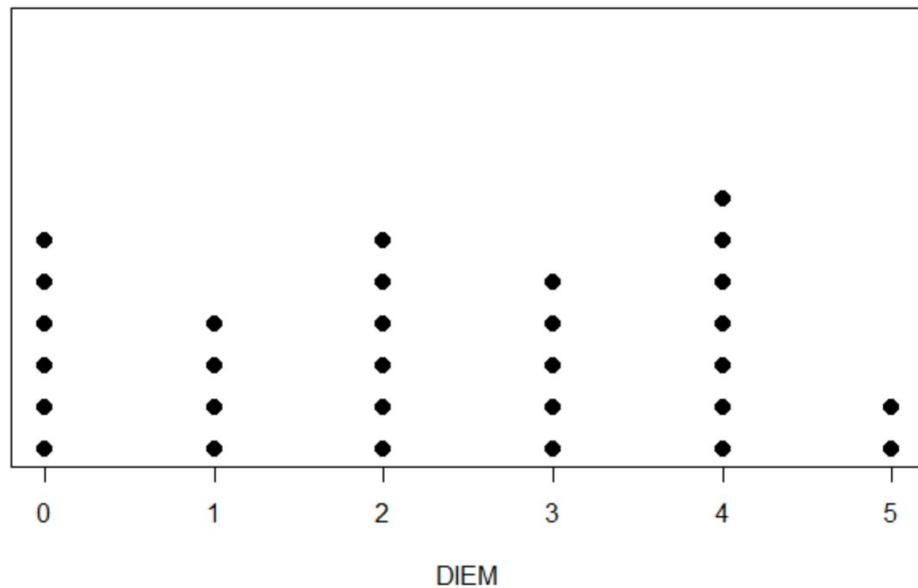
# Tính biên cho đồ thị để xác định outliners

```
> lf16e1= quantile(cau24e1[,17],probs = 0.25)-1.5*IQR(cau24e1[,17])
```

```
> hf16e1= quantile(cau24e1[,17],probs = 0.75)+1.5*IQR(cau24e1[,17])
```

Tính được giá trị lf16e1 = -3.5 và hf16e1 = 8.5

## SHEET E1 TC16



Đồ thị cho tiêu 16 sheet E1:

- + Không có outlier.
- + Center = 2
- + Shape: đồ thị tương đối cân bằng
- + Validity: min = 0, max = 5

Vẽ đồ thị với sheet E2

#Tiêu chí 6

```
> cau24e2<- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = 2, colNames = TRUE)
```

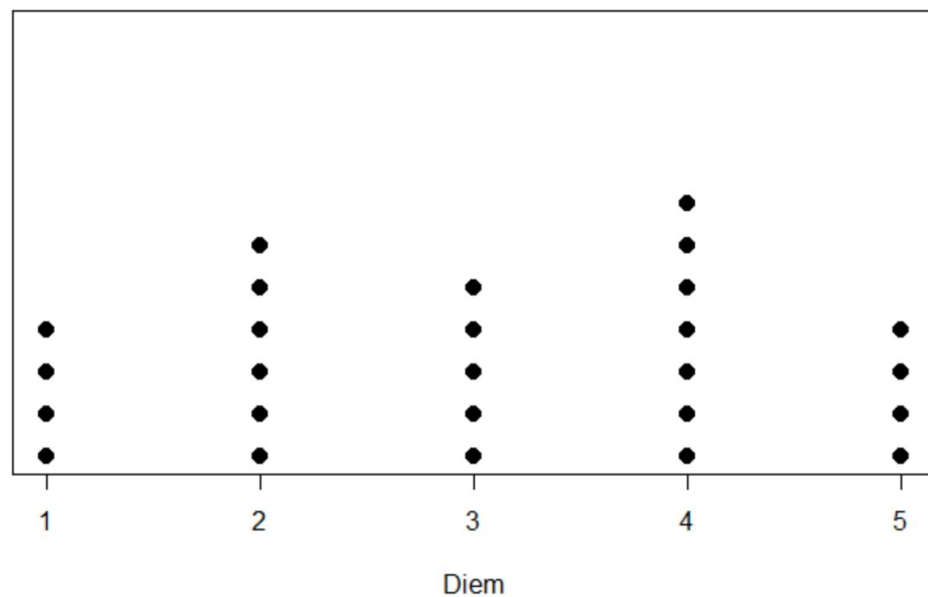
```
> stripchart(cau24e2[,7], method = "stack",
+ pch = 20, cex=2 , las=1,at= c(0), offset = 1/5,
+ main= "SHEET E2 TC6", xlab = "Diem")
```

```
> lf6e2= quantile(cau24e2[,7],probs = 0.25)-1.5*IQR(cau24e2[,7])
```

```
> hf6e2= quantile(cau24e2[,7],probs = 0.75)+1.5*IQR(cau24e2[,7])
```

Tính được giá trị lf6e2 = -1 và hf6e2 = 7

## SHEET E2 TC6



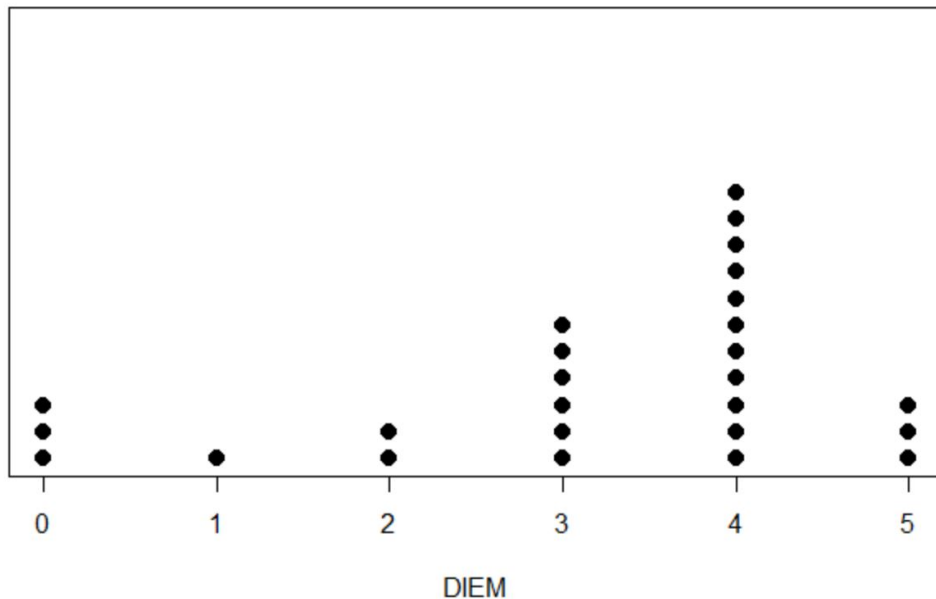
Đồ thị cho tiêu 6 sheet E2:

- + Không có outliner.
- + Center = 3
- + Shape: Đồ thị tương đối cân bằng
- + Vadility: min =0, max =5

#tiêu chí 16

```
> stripchart(cau24e2[,17], method = "stack",
+ pch = 20, cex=2, las=1,at=c(0), offset = 1/8,
+ main="SHEET E2 TC16", xlab = "DIEM")
> lf16e2= quantile(cau24e2[,17],probs = 0.25)-1.5*IQR(cau24e2[,17])
> hf16e2= quantile(cau24e2[,17],probs = 0.75)+1.5*IQR(cau24e2[,17])
Tính được giá trị lf16e2 = 1.5 và hf16e2 = 5.5
```

## SHEET E2 TC16



Đồ thị cho tiêu 16 sheet E2:

- + Outliner : Các giá trị đánh 0 và 1 là outliner của đồ thị.
- + Center = 4 . Cho thấy được mức đánh giá chủ yếu ở xung quanh giá trị 4.
- + Shape: đồ thị có dạng nghiêng trái
- + Validity: min = 0 , max = 5

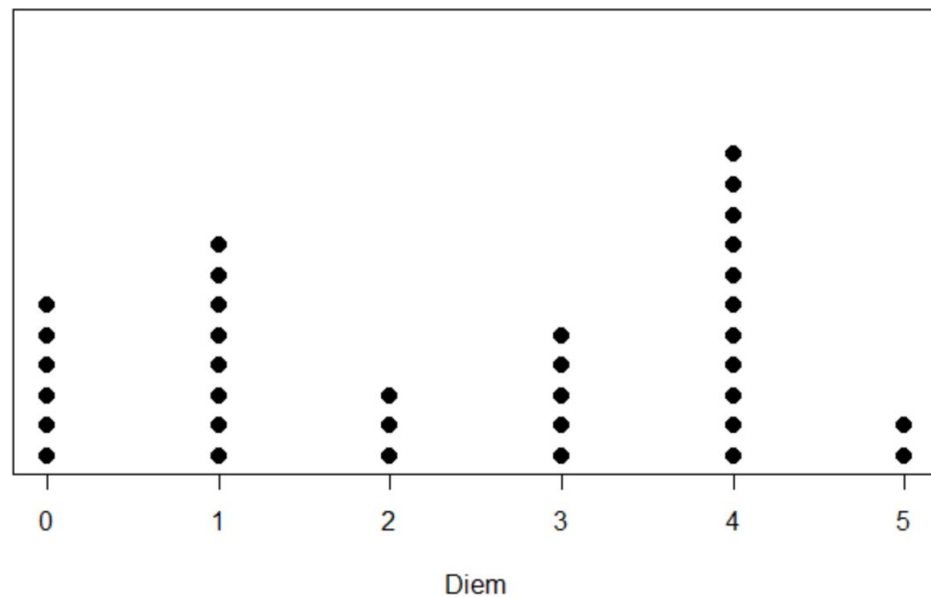
Vẽ đồ thị cho sheet E5

#Tiêu chí 6

```
> cau24e5<- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = 3, colNames = TRUE)
> stripchart(cau24e5[,7], method = "stack",
+ pch = 20, cex = 2 , las=1,at= c(0), offset = 1/7,
+ main= "SHEET E5 TC6", xlab = "Diem")
> lf6e5= quantile(cau24e5[,7],probs = 0.25)-1.5*IQR(cau24e5[,7])
> hf6e5= quantile(cau24e5[,7],probs = 0.75)+1.5*IQR(cau24e5[,7])
Ta tính được giá trị lf6e5 = -3.5 và hf6e5 = 8.5
```



### SHEET E5 TC6



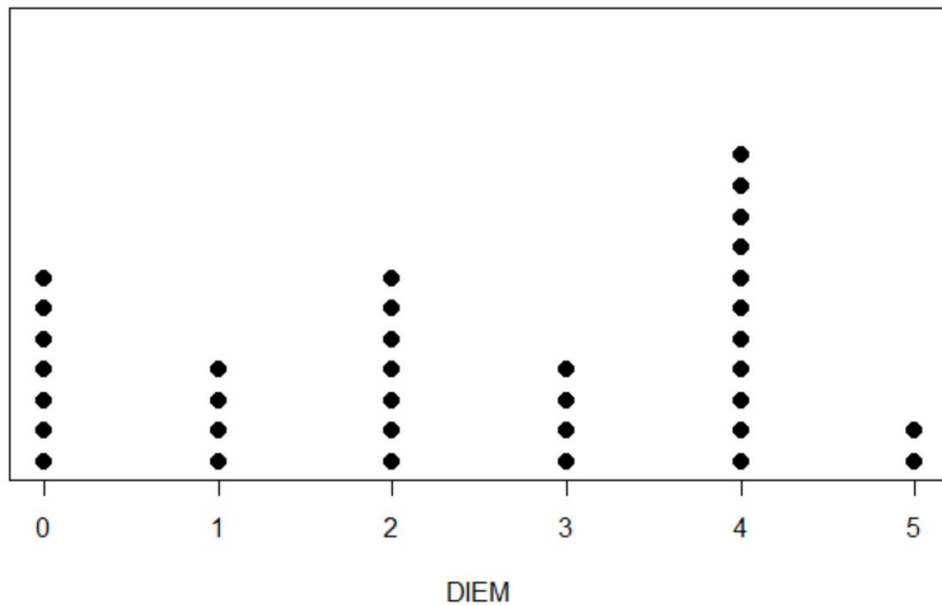
Đồ thị cho tiêu 6 sheet E5:

- + Không có outliner.
- + Center = 3
- + Shape: Đồ thị không cân đối, phân bố chủ yếu ở 4 hoặc 0 và 1 .
- + Validity : min = 0 , max = 5

#Tiêu chí 16

```
> stripchart(cau24e5[,17], method = "stack",
+ pch = 20, cex=2, las=1, at=c(0), offset = 1/7,
+ main="SHEET E5 TC16", xlab = "DIEM")
> lf16e5= quantile(cau24e5[,17],probs = 0.25)-1.5*IQR(cau24e5[,17])
> hf16e5= quantile(cau24e5[,17],probs = 0.75)+1.5*IQR(cau24e5[,17])
Tính được giá trị lf16e5 = -3.5 và hf16e5 = 8.5
```

### SHEET E5 TC16



Đồ thị cho tiêu 16 sheet E5:

- + Không có outlier.
- + Center = 2
- + Shape: Đồ thị không cân đối
- + Validity: min = 0 , max = 5;

Vẽ đồ thị cho sheet E6

#Tiêu chí 6

```
> cau24e6<- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = 4, colNames = TRUE)
```

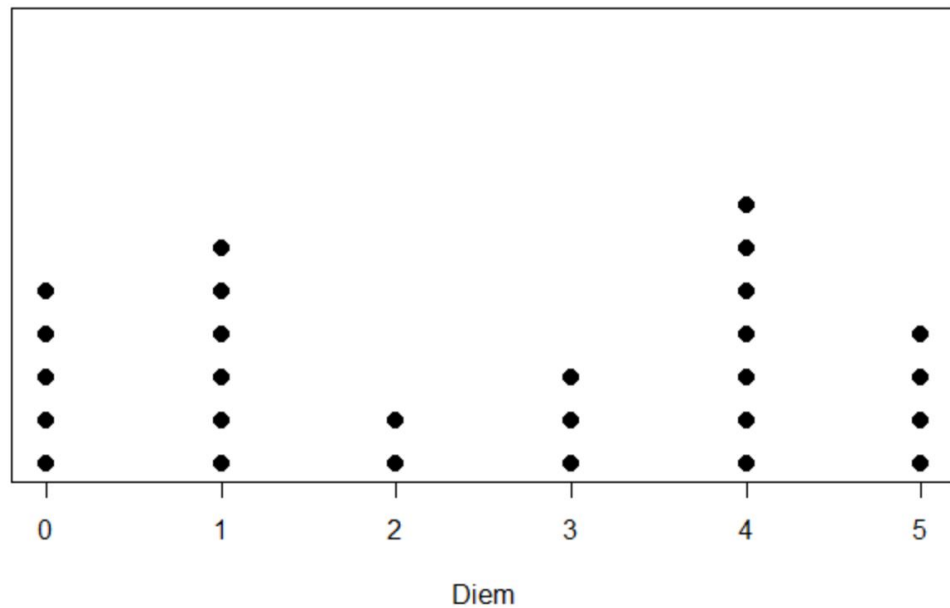
```
> stripchart(cau24e6[,7], method = "stack",
+pch = 20, cex = 2 , las=1,at= c(0), offset = 1/5,
+main= "SHEET E6 TC6", xlab = "Diem")
```

```
> lf6e6= quantile(cau24e6[,7],probs = 0.25)-1.5*IQR(cau24e6[,7])
```

```
> hf6e6= quantile(cau24e6[,7],probs = 0.75)+1.5*IQR(cau24e6[,7])
```

Tính được lf6e6 = -3.5 và hf6e6 = 8.5

### SHEET E6 TC6



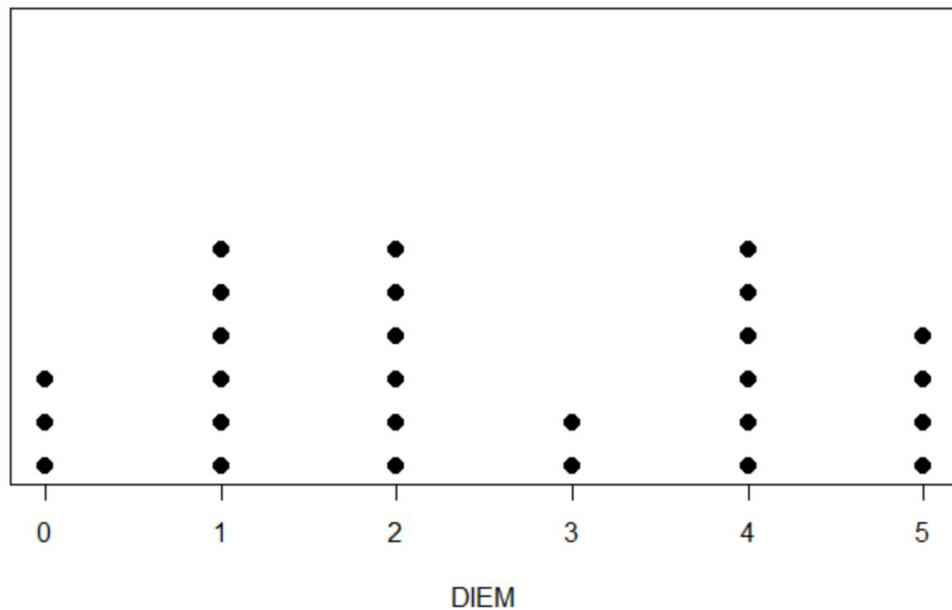
Đồ thị cho tiêu 6 sheet E6:

- + Không có outliner.
- + Center = 3
- + Shape : Đồ thị không đối xứng
- + Vadility: min = 0 , max = 5

#Tiêu chí 16

```
> stripchart(cau24e6[,17], method = "stack",
 +pch = 20, cex=2, las=1,at=c(0), offset = 1/5,
 +main="SHEET E6 TC16", xlab = "DIEM")
> lf16e6= quantile(cau24e6[,17],probs = 0.25)-1.5*IQR(cau24e6[,17])
> hf16e6= quantile(cau24e6[,17],probs = 0.75)+1.5*IQR(cau24e6[,17])
Tính được lf16e6 = -3.5 và hf16e6 = 8.5
```

## SHEET E6 TC16



Đồ thị cho tiêu 16 sheet E6:

- + Không có outliner.
- + Center = 2
- + Shape: Đồ thị dạng không cân đối
- + Vadility: min =0, max =5

>

Vẽ đồ thị cho sheet E7

```
#Tiêu chí 6 > cau24e7<- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = 5, colNames = TRUE)
```

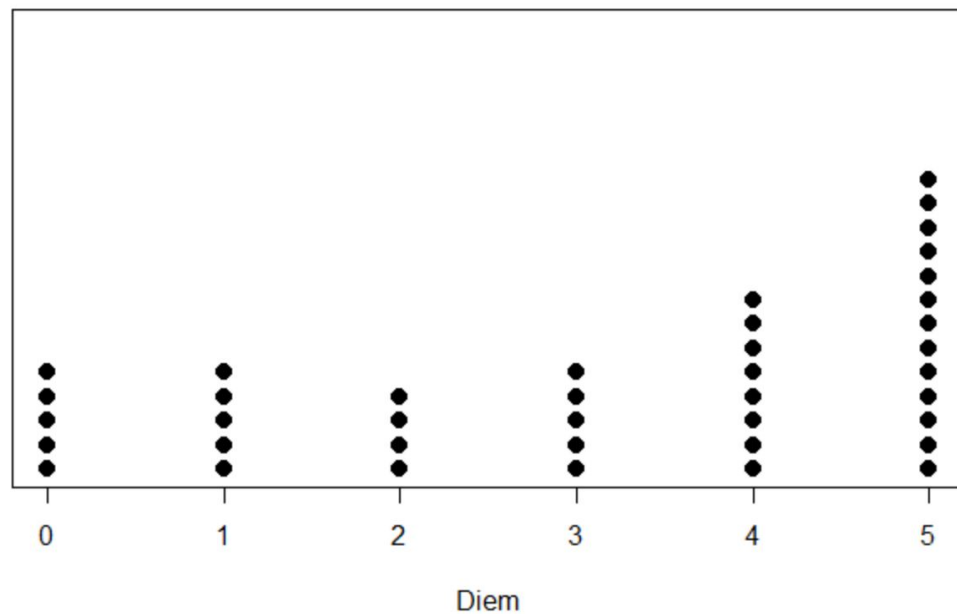
```
> stripchart(cau24e7[,7], method = "stack",
+pch = 20, cex =2 , las=1,at= c(0), offset = 1/9,
+main= "SHEET E7 TC6", xlab = "Diem")
```

```
> lf6e7= quantile(cau24e7[,7],probs = 0.25)-1.5*IQR(cau24e7[,7])
```

```
> hf6e7= quantile(cau24e7[,7],probs = 0.75)+1.5*IQR(cau24e7[,7])
```

Tính được lf6e7 = -3.12 và hf6e7 = 9.88

### SHEET E7 TC6

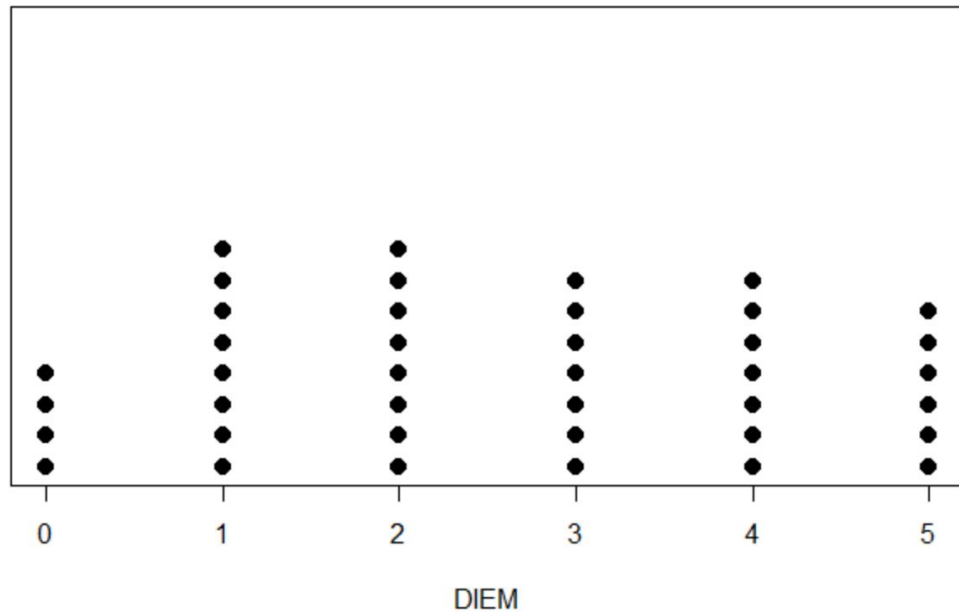


Đồ thị cho tiêu 6 sheet E7:

- + Không có outlier.
- + Center = 4
- + Shape: Đồ thị có dạng nghiêng trái
- + Vadility : min = 0 và max = 5

```
#Tiêu chí 16 > stripchart(cau24e7[,17], method = "stack",
+pch = 20, cex=2, las=1,at=c(0), offset = 1/7,
+ main="SHEET E7 TC16", xlab = "DIEM")
> lf16e7= quantile(cau24e7[,17],probs = 0.25)-1.5*IQR(cau24e7[,17])
> hf16e7= quantile(cau24e7[,17],probs = 0.75)+1.5*IQR(cau24e7[,17])
Tính được lf16e7 = 3.5 và hf16e7 = 8.5
```

### SHEET E7 TC16



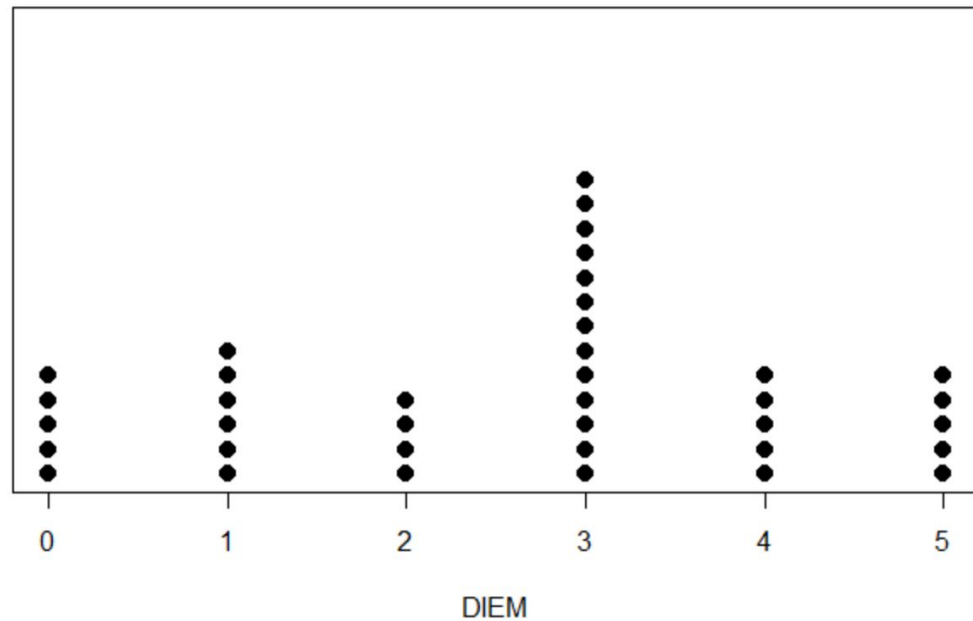
Đồ thị cho tiêu 16 sheet E7:

- + Không có outliner.
- + Center = 2.5
- + Shape: Đồ thị không đối xứng
- + Vadility: min = 0 và max = 5

Vẽ đồ thị sheet E8

```
#Tiêu chí 6 > cau24e8<- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = 6, colNames = TRUE)
> stripchart(cau24e8[,7],method = "stack",
+ pch = 20,offset = 1/9,cex = 2, las=1,at=c(0),
+ main="SHEET E8 TC6", xlab = "DIEM")
> lf6e8= quantile(cau24e8[,7],probs = 0.25)-1.5*IQR(cau24e8[,7])
> hf6e8= quantile(cau24e8[,7],probs = 0.75)+1.5*IQR(cau24e8[,7])
Tính được lf6e8 = -3.12 và hf6e8 = 7.88
```

## SHEET E8 TC6



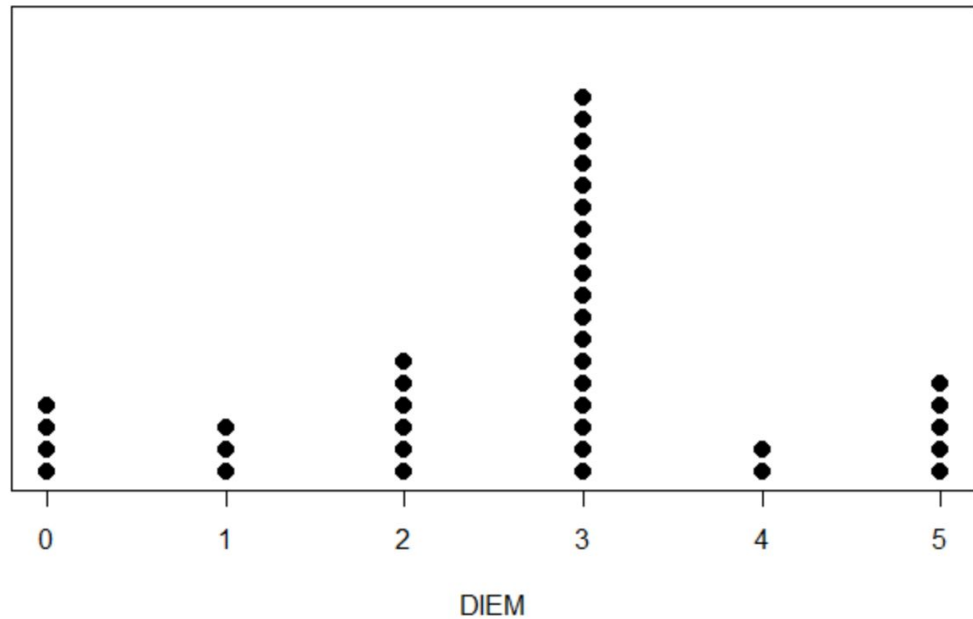
Đồ thị cho tiêu 6 sheet E8:

- + Không có outlier.
- + Center = 3
- + Shape: Đồ thị có dạng đối xứng
- + Validity: min = 0 và max = 5

#Tiêu chí 16

```
> stripchart(cau24e8[,17],
+ method = "stack",pch = 20,cex=2, offset = 1/15,
+ las=1,at=c(0),main="SHEET E8 TC16", xlab = "DIEM")
> lf16e8= quantile(cau24e8[,17],probs = 0.25)-1.5*IQR(cau24e8[,17])
> hf16e8= quantile(cau24e8[,17],probs = 0.75)+1.5*IQR(cau24e8[,17])
Tính được lf16e8 = 0.5 và hf16e8 = 4.5
```

### SHEET E8 TC16

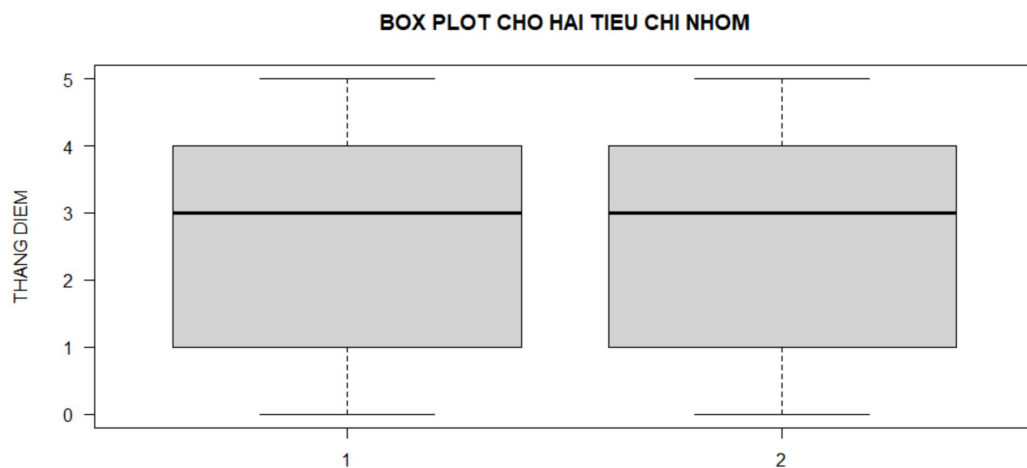


Đồ thị cho tiêu 16 sheet E8:

- + Outliner: Giá trị 0 và 5 là outliner của đồ thị
- + Center = 3
- + Shape: Đồ thị có dạng nghiêng trái. Phần lớn giá trị đánh giá ở mức 3.
- + Validity: min = 0 và max = 5

25. Vẽ box-and-whisker cho 2 tiêu chí ở câu trên của nhóm

```
> cau25tcn <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",
> boxplot(cau25tcn$TC6, cau25tcn$TC16,
+ main="BOX PLOT CHO HAI TIEU CHI NHOM",
+ ylab="THANG DIEM", ylim=c(0,5), las=1, names = TRUE)
```



Nhận xét:

- Với boxplot cho hai tiêu chí của nhóm (6 và 16), phân phối của hai tiêu chí gần như tương đồng nhau.



- Tứ phân vị 1, 3 và giá trị median là giống nhau.

26. Vẽ scatterplot cho tổng điểm trung bình cho mỗi sheets của tập các tiêu chí về kỹ thuật, làm tròn 2 số thập phân

```
>sheet1<-texttt read.xlsx("D:/R/BTLCTRR.xlsx",sheet = "E1")
>sheet2<-texttt read.xlsx("D:/R/BTLCTRR.xlsx",sheet = "E2")
>sheet5<-texttt read.xlsx("D:/R/BTLCTRR.xlsx",sheet = "E5")
>sheet6<-texttt read.xlsx("D:/R/BTLCTRR.xlsx",sheet = "E6")
>sheet7<-texttt read.xlsx("D:/R/BTLCTRR.xlsx",sheet = "E7")
>sheet8<-texttt read.xlsx("D:/R/BTLCTRR.xlsx",sheet = "E8")

>s1<-data.frame(sheet1[,8],sheet1[,2],sheet1[,3],
+sheet1[,4],sheet1[,5],sheet1[,11],sheet1[,13],sheet1[,14],
+sheet1[,15],sheet1[,16],sheet1[,17],sheet1[,19],sheet1[,22])

>s2<-data.frame(sheet1[,8],sheet1[,2],sheet1[,3],
+sheet1[,4],sheet1[,5],sheet1[,11],sheet1[,13],sheet1[,14],
+sheet1[,15],sheet1[,16],sheet1[,17],sheet1[,19],sheet1[,22])

>s5<-data.frame(sheet1[,8],sheet1[,2],sheet1[,3],
+sheet1[,4],sheet1[,5],sheet1[,11],sheet1[,13],sheet1[,14],
+sheet1[,15],sheet1[,16],sheet1[,17],sheet1[,19],sheet1[,22])

>s6<-data.frame(sheet1[,8],sheet1[,2],sheet1[,3],
+sheet1[,4],sheet1[,5],sheet1[,11],sheet1[,13],sheet1[,14],
+sheet1[,15],sheet1[,16],sheet1[,17],sheet1[,19],sheet1[,22])

>s7<-data.frame(sheet1[,8],sheet1[,2],sheet1[,3],
+sheet1[,4],sheet1[,5],sheet1[,11],sheet1[,13],sheet1[,14],
+sheet1[,15],sheet1[,16],sheet1[,17],sheet1[,19],sheet1[,22])

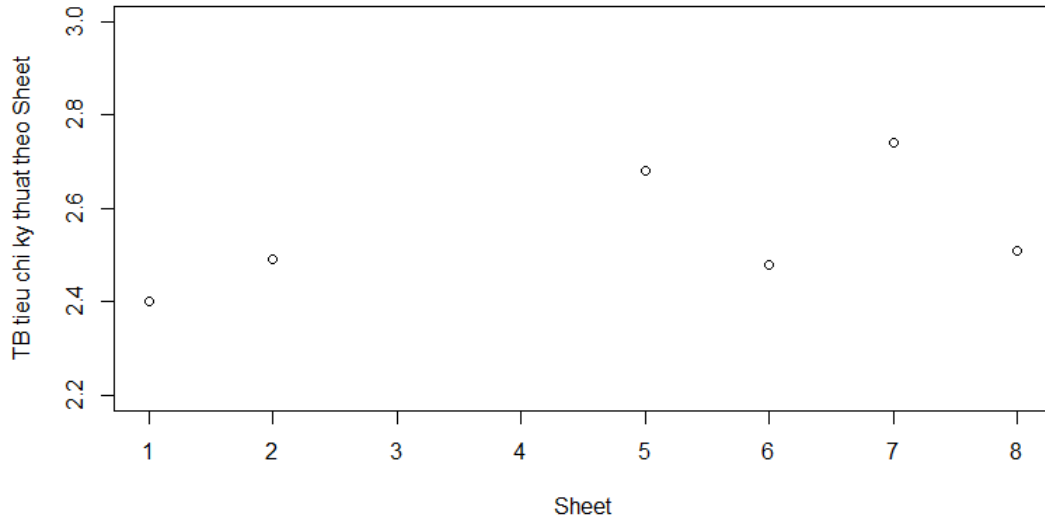
>s8<-data.frame(sheet1[,8],sheet1[,2],sheet1[,3],
+sheet1[,4],sheet1[,5],sheet1[,11],sheet1[,13],sheet1[,14],
+sheet1[,15],sheet1[,16],sheet1[,17],sheet1[,19],sheet1[,22])

sum1=sum2=sum5=sum6=sum7=sum8=0;
for(i in 1:30) sum1=sum1+sum(s1[i,])
tb1=signif(sum1/(30*13),digits = 3)
for(i in 1:26) sum2=sum2+sum(s2[i,])
tb2=signif(sum2/(26*13),digits = 3)
for(i in 1:35) sum5=sum5+sum(s5[i,])
tb5=signif(sum5/(35*13),digits = 3)
for(i in 1:27) sum6=sum6+sum(s6[i,])
tb6=signif(sum6/(27*13),digits = 3)
for(i in 1:40) sum7=sum7+sum(s7[i,])
tb7=signif(sum7/(40*13),digits = 3)
for(i in 1:38) sum8=sum8+sum(s8[i,])
tb8=signif(sum8/(38*13),digits = 3)
tb<-c(tb1,tb2,tb5,tb6,tb7,tb8)
```

```
tt<-c(1,2,5,6,7,8)
```

```
plot(x=tt, y=tb,ylab = "TB tiêu chí kỹ thuật theo Sheet",ylim = c(2.2,3), main = "Ve Scatterplot
tung sheet theo tiêu chí kỹ thuật",xlab = "Sheet")
```

**Ve Scatterplot tung sheet theo tiêu chí kỹ thuật**



Nhận xét:

- Điểm trung bình các Sheet 1, 2, 5, 6, 8 đồng đều nhau, trong khoảng 2.4-2.8.
- Kết quả trung bình thấp, ở mức cận trung bình.
- Qua đồ thị thấy được về mặt kỹ thuật của sinh viên thực tập chưa được đánh giá cao. Ở mức cận trung bình.

27. Vẽ scatterplot cho tổng điểm trung bình cho mỗi sheets của tập các tiêu chí về kỹ năng mềm, làm tròn 2 số thập phân

```
>sheet1<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "E1")
>sheet2<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "E2")
>sheet5<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "E5")
>sheet6<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "E6")
>sheet7<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "E7")
>sheet8<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "E8")
>s1<-data.frame(sheet1[,6],sheet1[,8],sheet1[,9],sheet1[,10],
+sheet1[,12],sheet1[,18],sheet1[,21],sheet1[,23],
+sheet1[,24],sheet1[,26])

>s2<-data.frame(sheet1[,6],sheet1[,8],sheet1[,9],sheet1[,10],
+sheet1[,12],sheet1[,18],sheet1[,21],sheet1[,23],
+sheet1[,24],sheet1[,26])

>s5<-data.frame(sheet1[,6],sheet1[,8],sheet1[,9],sheet1[,10],
+sheet1[,12],sheet1[,18],sheet1[,21],sheet1[,23],
+sheet1[,24],sheet1[,26])

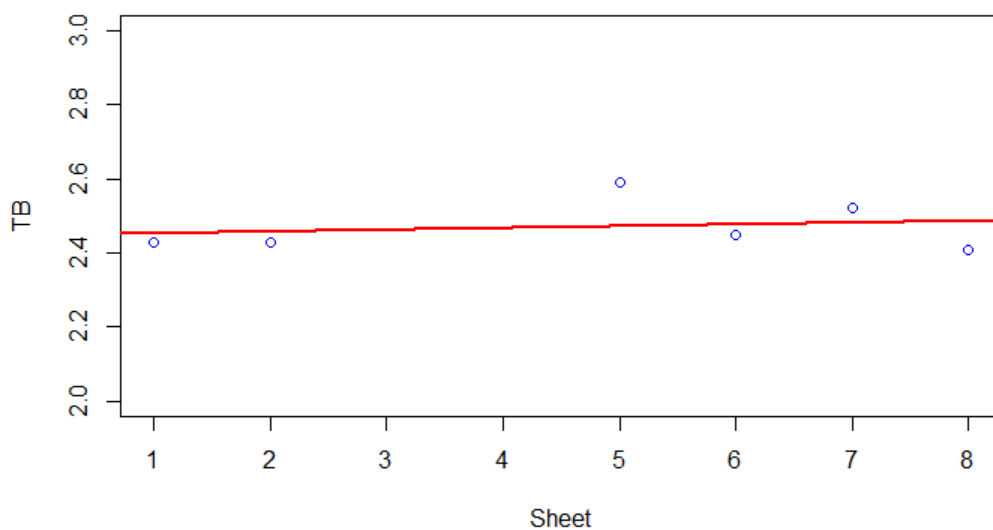
>s6<-data.frame(sheet1[,6],sheet1[,8],sheet1[,9],sheet1[,10],
+sheet1[,12],sheet1[,18],sheet1[,21],sheet1[,23],
+sheet1[,24],sheet1[,26])
```

```
>s7<-data.frame(sheet1[,6],sheet1[,8],sheet1[,9],sheet1[,10],
+sheet1[,12],sheet1[,18],sheet1[,21],sheet1[,23],
+sheet1[,24],sheet1[,26])

>s8<-data.frame(sheet1[,6],sheet1[,8],sheet1[,9],sheet1[,10],
+sheet1[,12],sheet1[,18],sheet1[,21],sheet1[,23],
+sheet1[,24],sheet1[,26])

>sum1=sum2=sum5=sum6=sum7=sum8=0;
>for(i in 1:30) sum1=sum1+sum(s1[i,]) tb1=signif(sum1/90,digits = 3)
>for(i in 1:26) sum2=sum2+sum(s2[i,]) tb2=signif(sum2/(26*3),digits = 3)
>for(i in 1:35) sum5=sum5+sum(s5[i,]) tb5=signif(sum5/(35*3),digits = 3)
>for(i in 1:27) sum6=sum6+sum(s6[i,]) tb6=signif(sum6/(27*3),digits = 3)
>for(i in 1:40) sum7=sum7+sum(s7[i,]) tb7=signif(sum7/(40*3),digits = 3)
>for(i in 1:38) sum8=sum8+sum(s8[i,]) tb8=signif(sum8/(38*3),digits = 3)
>tb<-c(tb1,tb2,tb5,tb6,tb7,tb8)
>tt<-c(1,2,5,6,7,8)
>plot(x=tt, y=tb,ylab = "TB",ylim = c(2,3), main = "Ve Scatterplot tung sheet theo tieu chi ve
ky nang mem",xlab = "Sheet",col="blue")
>abline(lm(tb ~ tt),col="red",lwd=2)
```

**Ve Scatterplot tung sheet theo tieu chi ve ky nang mem**



Nhận xét:

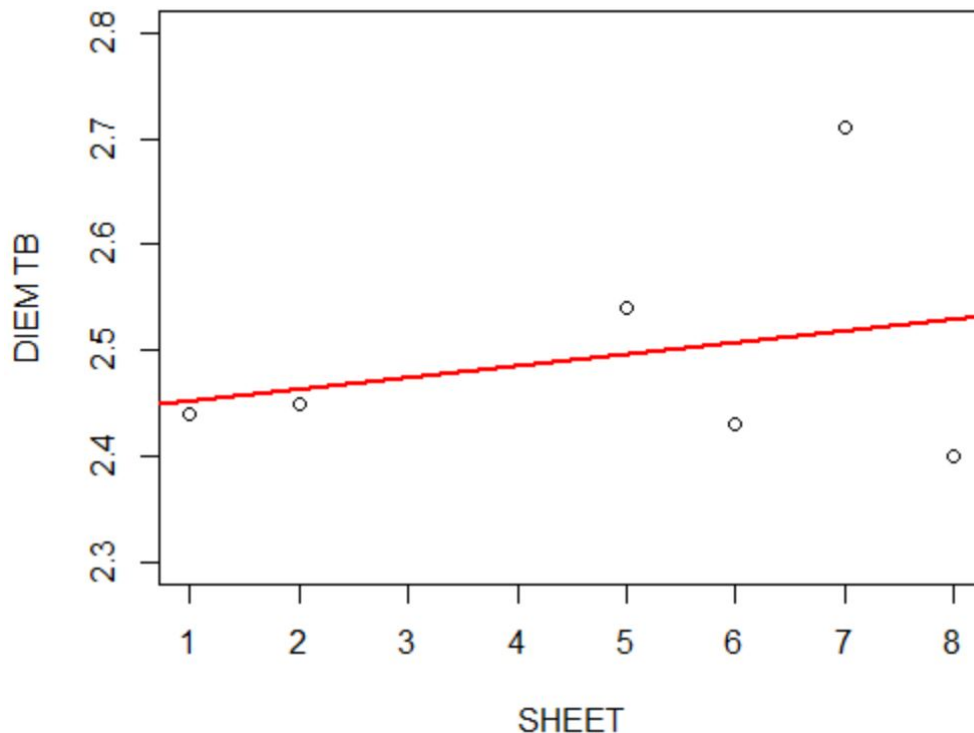
- Điểm số trung bình tiêu chí kỹ năng mềm cho 6 sheet đồng đều nhau, ở khoảng 2.4-2.6 .
- Điểm số trung bình không cao, theo thang đánh giá ở mức rất yếu chưa đến trung bình.
- Qua đồ thị thấy được phần nào về kỹ năng mềm của sinh viên chưa được các nhà tuyển dụng đánh giá cao. Là một yếu tố cần phải cải thiện.

28. Vẽ scatterplot cho tổng điểm trung bình cho mỗi sheets của tập các tiêu chí về tính kỷ luật, làm tròn 2 số thập phân.

```
> cau28e1 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",
+sheet = "E1",cols = c(8,16,21), colNames = TRUE)
> cau28e2 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",
```

```
+sheet = "E2",cols = c(8,16,21), colNames = TRUE)
> cau28e5 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",
+ sheet = "E5",cols = c(8,16,21), colNames = TRUE)
> cau28e6 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",
+ sheet = "E6",cols = c(8,16,21), colNames = TRUE)
> cau28e7 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",
+ sheet = "E7",cols = c(8,16,21), colNames = TRUE)
> cau28e8 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx",
+ sheet = "E8",cols = c(8,16,21), colNames = TRUE)
> s28e1 = sum(cau28e1)
> s28e2 = sum(cau28e2)
> s28e5 = sum(cau28e5)
> s28e6 = sum(cau28e6)
> s28e7 = sum(cau28e7)
> s28e8 = sum(cau28e8)
> tb28e1 = signif(s28e1/(30*3),digits = 3)
> tb28e2 = signif(s28e2/(26*3),digits = 3)
> tb28e5 = signif(s28e5/(35*3), digits = 3)
> tb28e6 = signif(s28e6/(27*3), digits = 3)
> tb28e7 = signif(s28e7/(40*3), digits = 3)
> tb28e8 = signif(s28e8/(38*3), digits = 3)
> tb28=c(tb28e1, tb28e2, tb28e5, tb28e6, tb28e7, tb28e8)
> tt28=c(1,2,5,6,7,8)
> plot(x=tt28, y=tb28,ylab = "DIEM TB",
+ ylim = c(2.3, 2.8),xlab = "SHEET",
+ main = "SCATTERPLOT TIEU CHI KI LUAT CHO TUNG SHEET")
```

## SCATTERPLOT TIEU CHI KI LUAT CHO TUNG SHEET



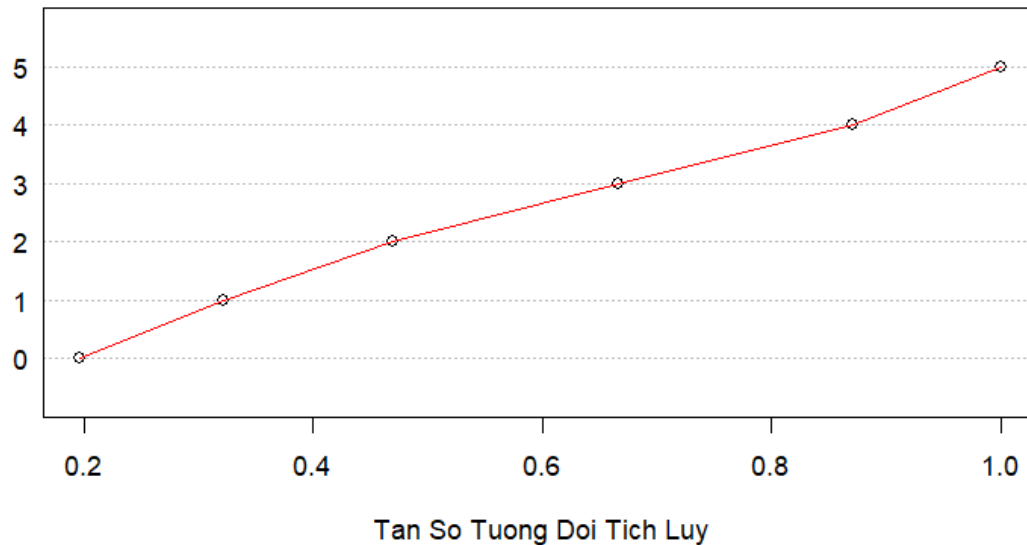
- Điểm trung bình các Sheet 1, 2, 5, 6, 8 đồng đều nhau, trong khoảng 2.4-2.55.

- Cao nhất là sheet E7 2.7 điểm
- Kết quả trung bình thấp, ở mức cận trung bình.

29. Vẽ biểu đồ tần số tương đối tích lũy của điểm của sinh viên trên tập các tiêu chí về kỹ thuật

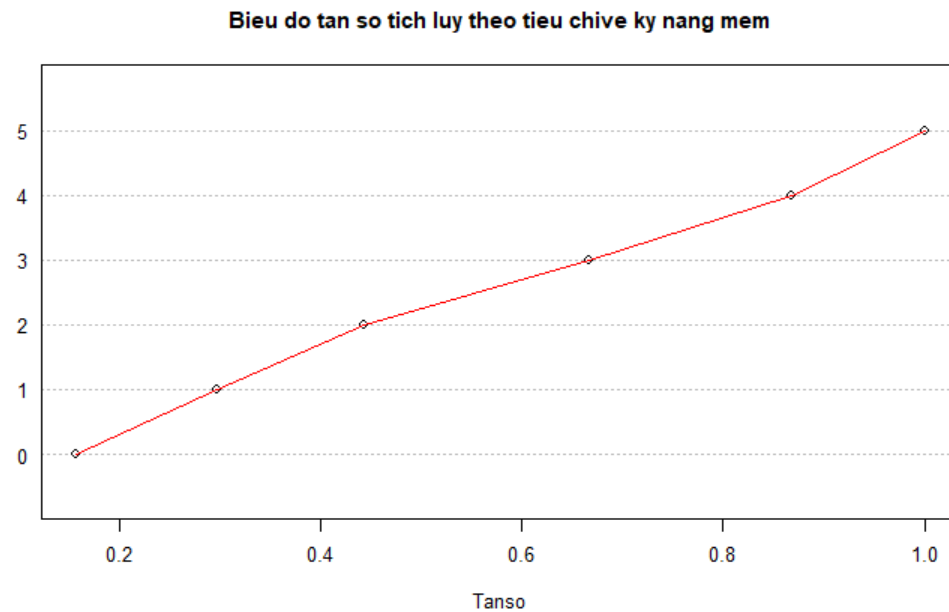
```
>kt6sheet <- read_xlsx(C:/BTL/file/BTLCTRR.xlsx",sheet = "KT6SHEET")
>kt6sheet_final <- data.frame(kt6sheet[1:196,2:14])
>f_ktsheet1 = data.frame(table(kt6sheet_final[,1]))
>f_ktsheet2 = data.frame(table(kt6sheet_final[,2]))
>f_ktsheet3 = data.frame(table(kt6sheet_final[,3]))
>f_ktsheet4 = data.frame(table(kt6sheet_final[,4]))
>f_ktsheet5 = data.frame(table(kt6sheet_final[,5]))
>f_ktsheet6 = data.frame(table(kt6sheet_final[,6]))
>f_ktsheet7 = data.frame(table(kt6sheet_final[,7]))
>f_ktsheet8 = data.frame(table(kt6sheet_final[,8]))
>f_ktsheet9 = data.frame(table(kt6sheet_final[,9]))
>f_ktsheet10 = data.frame(table(kt6sheet_final[,10]))
>f_ktsheet11 = data.frame(table(kt6sheet_final[,11]))
>f_ktsheet12 = data.frame(table(kt6sheet_final[,12]))
>f_ktsheet13 = data.frame(table(kt6sheet_final[,13]))
>freq <-c() >for (i in 1:nrow(f_ktsheet1)) freq[i] <- f_ktsheet1[i,2] + f_ktsheet2[i,2]
+ f_ktsheet3[i,2]+ f_ktsheet4[i,2]+ f_ktsheet5[i,2]+ f_ktsheet6[i,2]+ f_ktsheet7[i,2]+
f_ktsheet8[i,2]+ f_ktsheet9[i,2]+ f_ktsheet10[i,2]+ f_ktsheet11[i,2]+ f_ktsheet12[i,2]+
f_ktsheet13[i,2]
f_tuongdoi <- c() sum_f<- sum(freq)
>for (i in 1:length(freq)) f_tuongdoi[i] <- freq[i]/sum_f
f_tuongDoiTichLuy <- c() f_tuongDoiTichLuy[1]<-f_tuongdoi[1] >for (i in
2:length(f_tuongdoi)) f_tuongDoiTichLuy[i] = f_tuongDoiTichLuy[i-1]+f_tuongdoi[i]
>diem <- f_ktsheet1[,1]
>dotchart(f_tuongDoiTichLuy, labels = diem, + cex = 1.6, xlab = "Tan So Tuong Doi Tich
Luy", ylab="Diem")
>lines(f_tuongDoiTichLuy,y=1:6,col="red")
>f_table <- data.frame(diem,freq,f_tuongdoi, f_tuongDoiTichLuy)
```

**biểu đồ tần số tích lũy về kỹ thuật**



30. Vẽ biểu đồ tần số tương đối tích lũy của điểm của sinh viên trên tập các tiêu chí về kỹ năng mềm

```
>data30<-read.xlsx("C:/BTL/file/BTLCTRR.xlsx",sheet = "TONG HOP")
>i<-1:196
>k<-2:27
>cau30<-data.frame(data30[i,k])
>dem0=dem1=dem2=dem3=dem4=dem5=0;
>tc<-c(cau30[,6],cau30[,8],cau30[,9],cau30[,10],
+cau30[,12],cau30[,18],cau30[,21],cau30[,23],
+cau30[,24],cau30[,26])
>for (i in 1:1960) if(tc[i]==0){dem0=dem0+1}
+else if(tc[i]==1){dem1=dem1+1}
+else if(tc[i]==2){dem2=dem2+1}
+else if(tc[i]==3){dem3=dem3+1}
+else if(tc[i]==4){dem4=dem4+1}
+else if(tc[i]==5){dem5=dem5+1}
>k<-c(dem0/1960,(dem0+dem1)/1960,(dem0+dem1+dem2)/1960,
+(dem0+dem1+dem2+dem3)/1960,(dem0+dem1+dem2+dem3+dem4)/1960,
+(dem0+dem1+dem2+dem3+dem4+dem5)/1960)
>data.frame("Diem"=0:5,"TSTL"=k)
Diem TSTL
1 0 0.1571429
2 1 0.2964286
3 2 0.4433673
4 3 0.6668367
5 4 0.8673469
6 5 1.0000000
>dotchart(x=k,labels = 0:5,xlab = "Tanso",cex=0.8,main = "Biểu đồ tần số tích lũy theo tiêu
chỉ về kỹ năng mềm")
>lines(x=k,y=1:6,col="red")
```



31. Vẽ biểu đồ tần số tương đối tích lũy của điểm của sinh viên trên tập các tiêu chí về tính kỷ luật

```
>kyluat_c31 <- read.xlsx("C:/BTL/file/BTLCTRR.xlsx", >sheet = "KL")
>tanso_tc1_c31<- data.frame(table(kyluat_c31[,2]))
>tanso_tc2_c31 <- data.frame(table(kyluat_c31[,3]))
>tanso_tc3_c31 <- data.frame(table(kyluat_c31[,4]))
>diem_kyluat_c31 <- tanso_tc1_c31$Var1
>tanso_c31 <- c();
>for (i in 1:length(diem_kyluat_c31))
 {tanso_c31 [i] <- tanso_tc1_c31 [,2][i] +
 tanso_tc2_c31 [,2][i] + tanso_tc3_c31 [,2][i]}

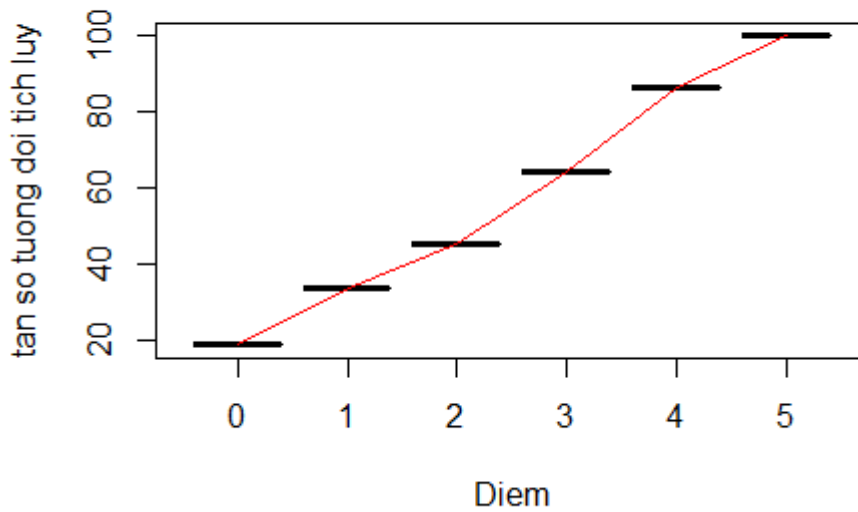
>tanso_tuongdoi_c31 <- c();
>tong_tanso_c31 <- sum(tanso_c31);
>for (i in 1:length(tanso_c31))
 {tanso_tuongdoi_c31 [i] <- tanso_c31 [i]/tong_tanso_c31}
>ts_tuongdoi_tichluy_c31 <- c();
>ts_tuongdoi_tichluy_c31 [1] <- tanso_tuongdoi_c31 [1];
>ts_phantram_c31 <- c();
>ts_phantram_c31 [1] <- ts_tuongdoi_tichluy_c31 [1]*100;
>for (i in 2:length(tanso_c31))
 {ts_tuongdoi_tichluy_c31 [i] <- ts_tuongdoi_tichluy_c31 [i-1] + tanso_tuongdoi_c31 [i]
 ts_phantram_c31 [i] <- ts_tuongdoi_tichluy_c31 [i]*100}

>cau31 <- data.frame(diem_kyluat_c31,tanso_c31,
 +tanso_tuongdoi_c31,ts_tuongdoi_tichluy_c31,
 +ts_phantram_c31)
>plot(diem_kyluat_c31,ts_phantram_c31,
 +main = "Tan so tuong doi tich luy - Ky luat", xlab="Điểm", ylab="Tần số")
>lines(diem_kyluat_c31,ts_phantram_c31, col="red")
```

|   | diem_kyluat | tanso | tanso_tuongdoi | ts_tuongdoi_tichluy | ts_phantram |
|---|-------------|-------|----------------|---------------------|-------------|
| 1 | 0           | 110   | 0.1870748      | 0.1870748           | 18.70748    |
| 2 | 1           | 89    | 0.1513605      | 0.3384354           | 33.84354    |
| 3 | 2           | 69    | 0.1173469      | 0.4557823           | 45.57823    |
| 4 | 3           | 111   | 0.1887755      | 0.6445578           | 64.45578    |
| 5 | 4           | 131   | 0.2227891      | 0.8673469           | 86.73469    |
| 6 | 5           | 78    | 0.1326531      | 1.0000000           | 100.00000   |

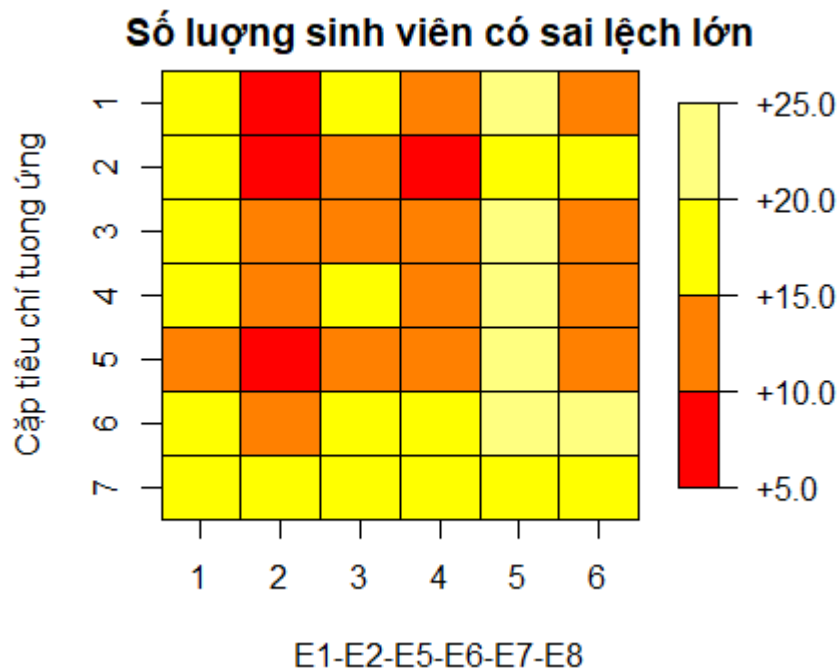


### Tan so tương doi tích luy - Tiêu chí Ky Luat



32. Vẽ biểu đồ cho thấy số lượng sinh viên bị sai lệch lớn trong đánh giá các cặp tiêu chí tương thích cho từng trường hợp ở câu hỏi trên cho 6 sheets

```
>library(plot.matrix)
>m_c32 <- matrix(c(slsv_e1_c19,slsv_e2_c19,slsv_e5_c19,
+slsv_e6_c19,slsv_e7_c19,slsv_e8_c19), ncol = 6)
>plot(m_c32, xlab="E1-E2-E5-E6-E7-E8", +ylab="Cặp tiêu chí tương ứng",
+main ="Số lượng sinh viên có sai lệch lớn")
```



## 6 Kết luận

Thông qua bài tập lớn, các thành viên trong nhóm học tập được cách sử dụng R trong phân tích số liệu. Học tập được cách làm việc nhóm, phân chia và hoàn thành nhiệm vụ cá nhân. Trình bày báo cáo, bài thuyết trình bằng cách sử dụng latex thông qua trang online Overleaf. Trang bị kiến thức cần thiết cho những môn học tiếp theo. Trong quá trình hoàn thành bài tập lớn, cảm ơn sự giúp đỡ và giải đáp thắc mắc của thầy Nguyễn Ngọc Lễ và thầy Huỳnh Tường Nguyên.

## Tài liệu

- [Dal] Dalgaard, P. *Introductory Statistics with R*. Springer 2008.
- [K-Z] Kenett, R. S. and Zacks, S. *Modern Industrial Statistics: with applications in R, MINITAB and JMP*, 2nd ed., John Wiley and Sons, 2014.
- [Ker] Kerns, G. J. *Introduction to Probability and Statistics Using R*, 2nd ed., CRC 2015.