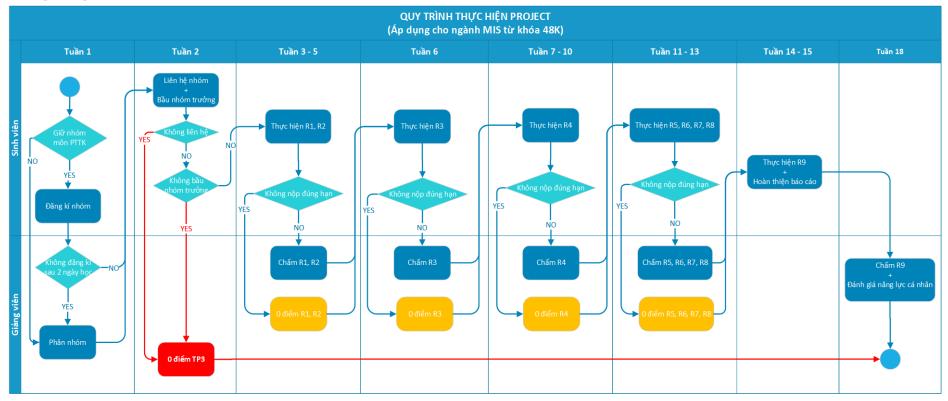
NGÀNH HỆ THỐNG THÔNG TIN QUẢN LÝ (K14 + K21)

1. Yêu cầu

STT	Y	'êu cầu	Ðiểm (11.0)	
	Nhóm mới lập	Nhóm thừa kế từ bài tập môn PTTK		
1.	Sưu tầm tất cả loại hóa đơn, chứng từ tại một cơ sở kinh doanh/sản xuất nhỏ	 Hiệu chỉnh lại các bảng sao cho ít nhất phải ở dạng chuẩn 2NF Thiết kế chi tiết các bảng 		
2.	Thiết kế cơ sở dữ liệu ở mức: • khái niệm (ERD) • logic (sơ đồ quan hệ ít nhất ở dạng chuẩn 2NF) • thiết kế chi tiết các bảng		1.0	
3.	Xây dựng cơ sở dữ liệu cho bản thiết kế ở bước trên.			
4.	 Xây dựng: Các module tạo dữ liệu dump cho các bảng trong cơ sở dữ liệu. Mỗi bảng ít nhất 1000 dòng dữ liệu; Ít nhất 10 module trong cơ sở dữ liệu để phục vụ các thao tác xử lý dữ liệu (kiểm tra sự hợp lệ của dữ liệu, xử lý các thao tác nghiệp vụ phức hợp,) 			
5.	Xây dựng cơ chế bảo mật phù hợp c	cho cơ sở dữ liệu	1.0	
6.	Xây dựng cơ chế backup dữ liệu tự	động cho cơ sở dữ liệu	1.0	
7.	Giả sử lượng dữ liệu rất lớn, không đủ lưu trữ trong một ổ đĩa. Hãy đưa ra phương án giải quyết TỐT NHẤT cho vấn đề đã đặt ra và thực hiện trên CSDL đã xây dựng. (Nêu lý do lựa chọn giải pháp và thực hiện triển khai giải pháp cho hệ thống)			
8.	Đề ra và thực hiện phương án để hạ	n chế tấn công SQL Injection	1.5	
9.	Sử dụng một ngôn ngữ lập trình bất các thao tác CRUD (Create, Read, U	t kỳ (Java, Python, C#,) để lập trình cho Jpdate, Delete).	2.0	

2. Workflow



- Mọi gian lận bị phát hiện trong quá trình thực hiện project đều bị xử phạt: 0 điểm TP3 cả nhóm và dừng project ngay lập tức
- Trong quá trình thực hiện, nếu thành viên nhóm không thực hiện nhiệm vụ được giao, nhóm trưởng thông báo cho GV để loại khỏi nhóm ngay lập tức, thành viên đó bị xử phạt: 0 điểm TP3

3. Rubric

	Trọng số	A	В	C	D	F
R1 + R2	10%	Tất cả các bảng đạt >= 2NF Tất cả quan hệ đúng Tất cả thuộc tính đúng (nằm đúng bảng) Tồn tại < 5% thuộc tính không cần thiết (tính toán được từ các cột khác)	Tất cả các bảng đạt >= 2NF Tồn tại < 10% quan hệ sai (bản số, mối quan hệ) Tất cả thuộc tính đúng (nằm đúng bảng) Tồn tại 5-10% thuộc tính không cần thiết	Tồn tại < 10% bảng chưa thỏa mãn 2NF Tồn tại 10 - 20% quan hệ sai (bản số, mối quan hệ) Tồn tại < 10% thuộc tính không hợp lý Tồn tại nhiều 10-20% thuộc tính không cần thiết	Tồn tại 10-20% bảng chưa thỏa mãn 2NF Tồn tại 20-30% quan hệ sai (bản số, quan hệ) Tồn tại 10-20% thuộc tính không hợp lý Tồn tại 20-30% thuộc tính không cần thiết	Tồn tại >= 20 bảng chưa thoả mãn 2NF Tồn tại >= 30% quan hệ sai Tồn tại >= 20% thuộc tính không hợp lý Tồn tại >= 30% thuộc tính không cần thiết
R3	5%	Script tạo CSDL đúng theo thiết kế Có các constraint ràng buộc dữ liệu cần thiết	Script tạo CSDL đúng theo thiết kế	Hoàn thiện 100% script tạo CSDL, tuy nhiên tồn tại < 10% sai khác so với thiết kế (thuộc tính, ràng buộc)	Hoàn thiện 100% script tạo CSDL, tuy nhiên tồn tại: 1) 10-20% sai khác so với thiết kế về thuộc tính hoặc ràng buộc;	Script tạo CSDL chưa hoàn thiện hoặc, Tồn tại >= 20% sai khác so với thiết kế về thuộc tính hoặc ràng buộc Tồn tại sự khác biệt về các bảng
R4	25%	Làm đúng 100% thủ tục tạo dữ liệu dump cho các bảng một cách tối ưu	Làm đúng 100% thủ tục tạo dữ liệu dump cho các bảng	Làm đúng >= 90% thủ tục tạo dữ liệu dump cho tất cả các bảng	Làm đúng >= 70% thủ tục tạo dữ liệu dump cho tất cả các bảng	Sai > 30% thủ tục tạo dữ liệu dump Tồn tại >= 40% thủ

		Có 10 thủ tục/hàm/trigger hỗ trợ các thao tác nghiệp vụ phức hợp, ví dụ: 1) kiểm tra dữ liệu mà constraint không làm được 2) các thao tác nghiệp vụ gồm nhiều hoạt động xử lý dữ liệu	Có 10 thủ tục/hàm/trigger hỗ trợ thao tác nghiệp vụ, tuy nhiên, tồn tại <10% module ở mức đơn giản (chỉ bao gồm 1 câu lệnh làm việc với dữ liệu hoặc constraint có thể giải quyết được)	Tồn tại 10-20% thủ tục/hàm/trigger không hợp lý hoặc đơn giản (có thể giải quyết được bằng constraint hoặc chỉ cần 1 câu lệnh)	Tồn tại 20-40% thủ tục/hàm/trigger không hợp lý hoặc đơn giản	tục/hàm/trigger đơn giản hoặc không hợp lý
R5	10%	Thiết lập cơ chế authentication phù hợp Tạo và phân quyền phù hợp cho người dùng (ở cấp độ server và cấp độ database)	Thiết lập cơ chế authentication phù hợp Đã tạo tài khoản người dùng và phân quyền, tuy nhiên: 1) Phân quyền của một số tài khoản	Thiết lập cơ chế authentication phù hợp Đã tạo một số tài khoản nhưng chưa hợp lý và không phân quyền cụ thể	Chưa thiết lập cơ chế authentication cho server Đã tạo tài khoản nhưng chưa phân quyền hoặc phân quyền không hợp lý	Chưa thiết lập cơ chế authentication cho server Có tạo tài khoản nhưng bị sai hoặc chưa tạo tài khoản Chưa phân quyền
		Có cơ chế bảo mật các dữ liệu nhạy cảm (quan trọng)	chưa hợp lý hoặc, 2) Một số tài khoản không hợp lý Có cơ chế bảo mật dữ liệu nhạy cảm nhưng chưa tối ưu	Chưa bảo mật dữ liệu nhạy cảm/quan trọng	Chưa bảo mật dữ liệu nhạy cảm/quan trọng	Chưa bảo mật dữ liệu nhạy cảm/quan trọng

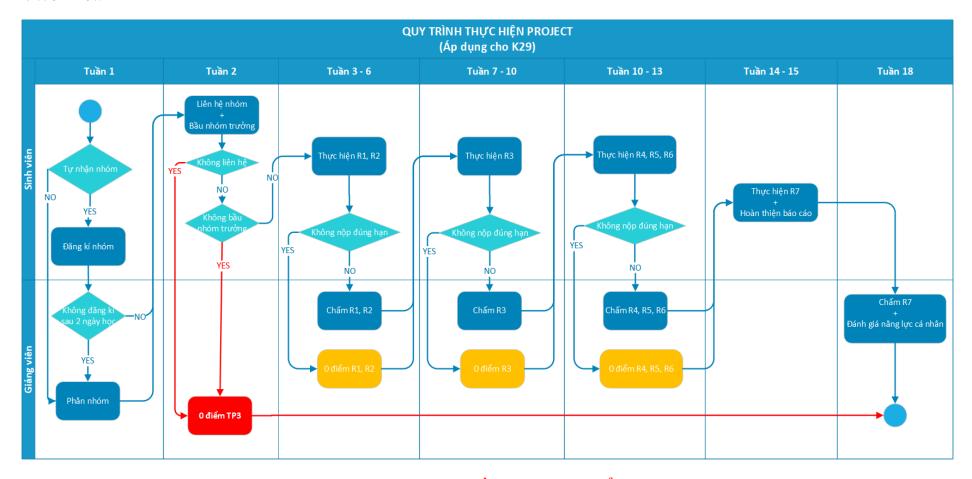
D6	100/	Thiất lân chấ đâ	Da thiất lân chấ đâ	Dã thiết lận chế độ	Da thiất lân chấ đâ	Dã thiất lân
R6	10%	Thiết lập chế độ backup tự động phù hợp: 1) Loại backup 2) Thời gian backup 3) Thiết bị lưu trữ backup 4) Quản lý các bản backup	Đã thiết lập chế độ backup tự động nhưng một trong số các yếu tố sau chưa phù hợp: 1) Loại backup 2) Thời gian backup 3) Thiết bị lưu trữ backup 4) Quản lý các bản backup	Đã thiết lập chế độ backup tự động nhưng hai trong số các yếu tố sau chưa phù hợp: 1) Loại backup 2) Thời gian backup 3) Thiết bị lưu trữ backup 4) Quản lý các bản backup	ba trong số các yếu tố sau chưa phù hợp: 1) Loại backup 2) Thời gian backup 3) Thiết bị lưu trữ	Đã thiết lập backup nhưng chưa tự động hoặc Chưa thiết lập chế độ backup
R7	15%	Trình bày giải pháp hợp lý và thực hiện một cách tối ưu	Trình bày giải pháp hợp lý và thực hiện thành công	Trình bày giải pháp chấp nhận được và thực hiện thành công	Trình bày giải pháp chấp nhận được và thực hiện mức demo	Giải pháp không hợp lý hoặc Trình bày nhưng không thực hiện giải pháp
R8	15%	Đề ra và thực hiện được giải pháp một cách tối ưu	Đề ra giải pháp hợp lý và thực hiện được	Đề ra giải pháp chấp nhận được và thực hiện được	Đề ra giải pháp chấp nhận được và demo một phần	Đề ra giải pháp nhưng không demo
R9	20%	Thực hiện được các thao tác CRUD cho tất cả các bảng Úng dụng được R4, R5, R7, R8	Thực hiện được các thao tác CRUD cho các bảng Úng dụng được một trong các yêu cầu R4, R5, R7, R8	Thực hiện được các thao tác CRUD cho >= 90% các bảng	Thực hiện được các thao tác CRUD cho >= 70% các bảng	Thực hiện được các thao tác CRUD cho dưới 70% các bảng

NGÀNH HỆ KHOA HỌC DỮ LIỆU TRONG KINH DOANH (K29)

1. Yêu cầu

STT	Yêu cầu	Ðiểm (11.0)			
1.	Cào dữ liệu từ một website (các nhóm không trùng nhau). Dữ liệu lấy về ít nhất 10.000 dòng.	2.0			
2.	. Import dữ liệu vào cơ sở dữ liệu				
3.	Viết các module bằng ngôn ngữ T-SQL để thực hiện tiền xử lý dữ liệu	3.0			
4.	Thực hiện backup dữ liệu vừa xử lý	1.5			
5.	 Thực hiện phân quyền cho cơ sở dữ liệu: Admin: có toàn quyền trên cơ sở dữ liệu DE: được phép thêm, sửa, xóa, truy vấn dữ liệu trên bảng, thêm, sửa, xóa các thủ tục hoặc hàm DA: chỉ được truy vấn dữ liệu trên bảng 	1.5			
6.	Thực hiện trực quan dữ liệu đang có trong cơ sở dữ liệu	1.0			
7.	Sử dụng ngôn ngữ Python để xây dựng một mô hình ML trên bộ dữ liệu đã có	1.0			

2. Workflow



- Mọi gian lận bị phát hiện trong quá trình thực hiện project đều bị xử phạt: 0 điểm TP3 cả nhóm và dừng project ngay lập tức
- Trong quá trình thực hiện, nếu thành viên nhóm không thực hiện nhiệm vụ được giao, nhóm trưởng thông báo cho GV để loại khỏi nhóm ngay lập tức, thành viên đó bị xử phạt: 0 điểm TP3

3. Rubric

	Trọng số	A	В	С	D	F
R1	20%	Dữ liệu lấy về >= 10.000 dòng	Dữ liệu lấy về >= 10.000 dòng	Dữ liệu lấy về >= 9.000 dòng	Dữ liệu lấy về từ 5.000- 9.000 dòng	Dữ liệu lấy về ít hơn 5.000 dòng
		Cung cấp đầy đủ source code dùng để crawl dữ liệu Thuyết minh đầy đủ, logic cho quá trình chọn lọc và crawl dữ liệu	Cung cấp đầy đủ source code để crawl dữ liệu Có thuyết minh cho các bước chọn lọc và crawl dữ liệu	Cung cấp đầy đủ source code để crawl dữ liệu Có thuyết minh nhưng không chi tiết về quá trình chọn lọc và crawl dữ liệu	Cung cấp đầy đủ source code để crawl dữ liệu Thiếu thuyết minh hoặc thuyết minh sơ sai về quá trình chọn lọc và crawl dữ liệu	Không có source code đầy đủ để crawl dữ liệu Thiếu hoặc không có thuyết minh về quá trình crawl dữ liệu
R2	10%	Khi crawl dữ liệu, import trực tiếp vào CSDL để trên cloud	Khi crawl dữ liệu, import trực tiếp vào CSDL ở trên local	Crawl dữ liệu ra file, sau đó dùng code để import dữ liệu vào CSDL	Import thủ công dữ liệu từ file trung gian sau khi crawl dữ liệu (Excel, csv,)	
R3	30%	Tất cả các thao tác xử lý dữ liệu đều được thực hiện bằng thủ tục/hàm/trigger (module) Các module được thiết kế một cách hợp lý, logic và tối ưu	>= 90% các thao tác xử lý dữ liệu được hiện bằng thủ tục/hàm/trigger (module) Các module được thiết kế một cách hợp lý và logic	xử lý dữ liệu được thực hiện bằng hàm/thủ tục/trigger (module) Tồn tại <10%		lý dữ liệu được thực hiện bằng hàm/thủ tục/trigger (module) >= 20% module

R4	15%	Thiết lập chế độ backup phù hợp: 1) Loại backup 2) Thời gian backup 3) Thiết bị lưu trữ backup 4) Quản lý backup	Thiết lập chế độ backup nhưng một trong số các yếu tố sau chưa phù hợp: 1) Loại backup 2) Thời gian backup 3) Thiết bị lưu trữ backup 4) Quản lý backup	Thiết lập chế độ backup nhưng hai trong số các yếu tố sau chưa phù hợp: 1) Loại backup 2) Thời gian backup 3) Thiết bị lưu trữ backup 4) Quản lý backup	nhưng ba trong số các yếu tố sau chưa phù hợp: 1) Loại backup 2) Thời gian backup 3) Thiết bị lưu trữ backup 4) Quản lý backup	Chế độ backup chưa phù hợp
R5	15%	Thực hiện phân quyền đầy đủ cho các yêu cầu bằng lệnh SQL Mô tả đầy đủ các thao tác khi thực hiện phân quyền Sử dụng một trong	Thực hiện phân quyền đầy đủ cho các yêu cầu bằng giao diện Mô tả đầy đủ các thao tác khi thực hiện phân quyền Sử dụng một trong	Thực hiện phân quyền đầy đủ cho các yêu cầu bằng giao diện Mô tả cơ bản các thao tác khi thực hiện phân quyền Sử dụng một trong	Thực hiện phân quyền đầy đủ cho các yêu cầu bằng giao diện Mô tả chưa rõ ràng các thao tác khi thực hiện phân quyền Sử dụng một trong	Thực hiện phân quyền nhưng chưa đúng yêu cầu Nhiều hơn 40%
		những công cụ mà sinh viên biết để trực quan dữ liệu theo nhiều góc nhìn một cách hợp lý	những công cụ mà sinh viên biết để trực quan dữ liệu theo nhiều góc nhìn, tuy nhiên < 10% biểu đồ chưa hợp lý	những công cụ mà sinh viên biết để trực quan dữ liệu theo nhiều góc nhìn, tuy nhiên < 20% biểu đồ chưa hợp lý	những công cụ mà sinh viên biết để trực quan dữ liệu theo nhiều góc	biểu đồ chưa hợp lý Thiếu phân tích, diễn giải

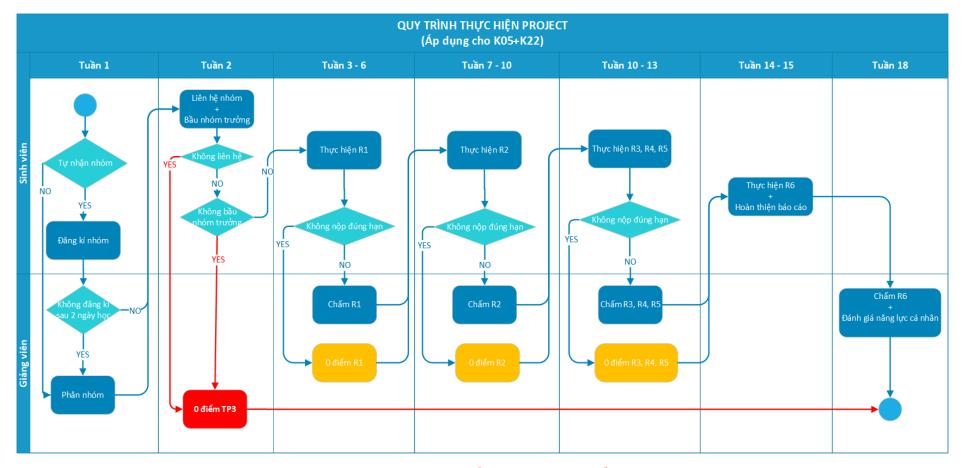
		Có phân tích, diễn giải một cách logic các trực quan dữ liệu	, ,		một cách logic các trực	
R7	10%	với performance tốt, có ý nghĩa với bài	Xây dựng ML model với performance tốt, giải quyết được bài toán liên quan tới dữ liệu nhưng chưa thật sự có ý nghĩa	với performance khá tốt Model có ý nghĩa với	với performance trung bình	• •

NGÀNH THỐNG KÊ & TMĐT (K05 + K22)

1. Yêu cầu

STT	Yêu cầu	Ðiểm (11.0)
1.	Tạo cơ sở dữ liệu cho một bài toán bất kỳ, có ít nhất 6 bảng (các nhóm không trùng nhau)	2.0
2.	Viết module tạo dữ liệu dump cho cho các bảng. Mỗi bảng ít nhất 1000 dòng dữ liệu	2.0
3.	Tạo ít nhất 10 module trong cơ sở dữ liệu để phục vụ các thao tác xử lý dữ liệu	2.5
4.	Tạo các nhóm người dùng và phân quyền phù hợp.	1.5
5.	Thực hiện backup dữ liệu tự động	1.0
6.	Sử dụng công cụ phù hợp để thực hiện các thống kê/mô tả/trực quan dữ liệu đang có trong CSDL	2.0

2. Workflow



- Mọi gian lận bị phát hiện trong quá trình thực hiện project đều bị xử phạt: 0 điểm TP3 cả nhóm và dừng project ngay lập tức
- Trong quá trình thực hiện, nếu thành viên nhóm không thực hiện nhiệm vụ được giao, nhóm trưởng thông báo cho GV để loại khỏi nhóm ngay lập tức, thành viên đó bị xử phạt: 0 điểm TP3

3. Rubric

	Trọng số	A	В	C	D	F
R1	20%	Tạo bảng theo đúng yêu cầu Các bảng có cấu trúc và mối quan hệ tối ưu	Tạo bảng theo đúng yêu cầu < 10% bảng có các cột chưa hợp lý < 10% quan hệ chưa hợp lý	yêu cầu < 20% bảng có các cột chưa hợp lý	Tạo bảng theo đúng yêu cầu < 40% bảng có các cột chưa hợp lý < 40% quan hệ chưa hợp lý	Số lượng bảng không đúng yêu cầu hoặc >= 40% bảng có các cột chưa hợp lý >= 40% quan hệ chưa hợp lý
R2	20%	Làm đúng 100% thủ tục tạo dữ liệu dump cho các bảng một cách tối ưu	Làm đúng 100% thủ tục tạo dữ liệu dump cho các bảng	Làm đúng >= 90% thủ tục tạo dữ liệu dump cho tất cả các bảng	Làm đúng >= 70% thủ tục tạo dữ liệu dump cho tất cả các bảng	Sai > 30% thủ tục tạo dữ liệu dump
R3	25%	Có 10 thủ tục/hàm/trigger hỗ trợ các thao tác nghiệp vụ phức hợp, ví dụ: 1) kiểm tra dữ liệu mà constraint không làm được 2) các thao tác nghiệp vụ gồm nhiều hoạt động xử lý dữ liệu	Có 10 thủ tục/hàm/trigger hỗ trợ thao tác nghiệp vụ, tuy nhiên, tồn tại <10% module ở mức đơn giản (chỉ bao gồm 1 câu lệnh làm việc với dữ liệu hoặc constraint có thể giải quyết được)	Tồn tại 10-20% thủ tục/hàm/trigger không hợp lý hoặc đơn giản (có thể giải quyết được bằng constraint hoặc chỉ cần 1 câu lệnh)	Tồn tại 20-40% thủ tục/hàm/trigger không hợp lý hoặc đơn giản	Tồn tại >= 40% thủ tục/hàm/trigger đơn giản hoặc không hợp lý

R4	15%	Tạo và phân quyền phù hợp cho người dùng (ở cấp độ server và cấp độ database)	Đã tạo tài khoản người dùng và phân quyền, tuy nhiên: 1) Phân quyền của một số tài khoản chưa hợp lý hoặc, 2) Một số tài khoản không hợp lý	Đã tạo một số tài khoản nhưng chưa hợp lý và không phân quyền cụ thể	Đã tạo tài khoản nhưng chưa phân quyền không hợp lý	Có tạo tài khoản nhưng bị sai hoặc chưa tạo tài khoản
R5	10%	Thiết lập chế độ backup tự động phù hợp: 1) Loại backup 2) Thời gian backup 3) Thiết bị lưu trữ backup 4) Quản lý các bản backup	Đã thiết lập chế độ backup tự động nhưng một trong số các yếu tố sau chưa phù hợp: 1) Loại backup 2) Thời gian backup 3) Thiết bị lưu trữ backup 4) Quản lý các bản backup	Đã thiết lập chế độ backup tự động nhưng hai trong số các yếu tố sau chưa phù hợp: 5) Loại backup 6) Thời gian backup 7) Thiết bị lưu trữ backup 8) Quản lý các bản backup	Đã thiết lập chế độ backup tự động nhưng ba trong số các yếu tố sau chưa phù hợp: 1) Loại backup 2) Thời gian backup 3) Thiết bị lưu trữ backup 4) Quản lý các bản backup	Đã thiết lập backup nhưng chưa tự động hoặc Chưa thiết lập chế độ backup
R6	20%	Sử dụng một trong những công cụ mà sinh viên biết để thống kê/mô tả/trực quan dữ liệu theo	Sử dụng một trong những công cụ mà sinh viên biết để thống kê/mô tả/trực quan dữ liệu theo nhiều góc nhìn, tuy	Sử dụng một trong những công cụ mà sinh viên biết để thống kê/mô tả/trực quan dữ liệu theo nhiều góc nhìn, tuy	Sử dụng một trong những công cụ mà sinh viên biết để thống kê/mô tả/trực quan dữ liệu theo nhiều góc	Nhiều hơn 40% biểu đồ/bảng biểu chưa hợp lý Thiếu phân tích, diễn giải

nhiều góc nhìn cách hợp lý	một nhiên < 10% biểu đồ chưa hợp lý		nhìn, tuy nhiên < 40% biểu đồ chưa hợp lý	
giải một cách	diễn Có phân tích, diễn giải logic một cách logic các trực quan dữ liệu	-	một cách logic các trực	