Trường Đại học Bách Khoa - ĐHQG Tp.HCM
Khoa: Phòng Đào tạo Sau đại học
Khoa/Bộ môn quản lý MH: Phòng Đào tạo Sau đại học

Tp.HCM, ngày ..... tháng ..... năm .....

Đề cương môn học Sau đại học

PHÂN TÍCH VÀ XỬ LÝ DỮ LIỆU ĐA CHIỀU (MULTIVARIATE ANALYSIS)

Mã số MH: GK5926

Số tín chỉ:	**Tc (LT.BT&TH.Tự Học): 3**						**ECTS: 6**				
Số tiết -Tổng:	**75**	LT:	**30**	BT:	**0**	TH:	**0**	ĐA:		BTL/TL:	**45**
Đánh giá:	Bài tập:			20%							
	Bài tập lớn:			20%							
	Thi cuối kỳ:			60%							
- Môn tiên quyết:											
- Môn học trước:											
- Môn song hành:											
- CTĐT ngành (Mã ngành):	Bản Đồ, Viễn Thám Và Hệ Thống Thông Tin Địa Lý (8440214)										
- Ghi chú khác:											

1. <u>Mục tiêu môn học:</u>

Cung cấp những kiến thức cơ bản về xử lý số liệu thống kê đa chiều và phương pháp lựa chọn những phương pháp phân tích thống kê mô tả một cách thích hợp số liệu thực nghiệm, để từ đó học viên có thể tiến hành phân tích và đưa ra những giải thích kết quả một cách xác đáng.

Aims:

Provides the fundamental knowledge of handling multi-dimensional statistics and statistical analysis to describe appropriately experimental data, so that students can conduct analysis explain and give an accurate result.

2. <u>Nội dung tóm tắt môn học:</u>

Nội dung môn học đề cập đến các phương pháp phân tích thống kê mô tả đa chiều nhằm phát hiện cấu trúc dữ liệu, các đặc trưng của số liệu thực nghiệm và mối quan hệ giữa những đặc trưng của số liệu thực nghiệm. Phương pháp phân tích phương sai một biến và nhiều biến độc lập sẽ được giới thiệu ở đầu chương trình tiếp theo là các phương pháp phân tích thành phần chính, phân tích tương quan và phương pháp phân nhóm. Sinh viên sẽ được hướng dẫn và nâng cao kỹ năng sử dụng phần mềm R và các packages (SensoMineR, FactoMineR,…) thông qua các bài tập tình huống. Kỹ năng giải quyết vấn đề sẽ được nâng cao thông qua việc xây dựng và triển khai các dự án phân tích dữ liệu xuất phát từ tình huống thực tế. Các phương pháp biểu diễn, khai thác kết quả bằng đồ thị thông qua các packages hỗ trợ như ggplot2 cũng sẽ được giới thiệu để nâng cao kỹ năng trình bày kết quả của học viên.

Course outline:

This course on Multivariate Analysis presents the tools and concepts of multivariate data analysis with a strong focus on applications. The aim of this course is to present multivariate data analysis in a way that is understandable for non-mathematical students and practitioners who are confronted by statistical data analysis. This is achieved by focusing on the practical relevance. The practical examples presented and may be recalculated and modified by the students using a standard statistical software R.

3. Tài liệu học tập:

[1] Abdi H. (1987), Introduction au traitement statistique des données expérimentales,Presse Universitaire de Grenoble, 420pp.
[2] Brigitte Escofier, Jerome Pages (1998), Analyse factorielles simples et multipes : objectifs, methodes et interpretation, Dunod, 284pp.
[3] F. Husson, S. Lê & J. Pagès (2017) *Exploratory Multivariate Analysis by Example Using R*. 2nd edition. Chapman & Hall/CRC Computer Science & Data Analysis.

[4] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, 2017, An Introduction to statistical learning, Springer, 426pp

[5] Cornillon, P-A., Guyader, A., Husson, F., Jégou, N., Josse, J., Kloareg, M., Matzner-Lober, E. & Rouvière, L.
(2012). *R for Statistics*. Chapman & Hall/CRC Computer Science & Data Analysis.

4. Các hiểu biết, các kỹ năng cần đạt được sau khi học môn học:

STT	Chuẩn đầu ra môn học (CĐRMH)	Công cụ đánh giá CĐRMH	Đóng góp CĐR Chương trình (CĐRCT)		
			Ứng dụng	Nghiên cứu	

Learning outcomes:

No.	Course learning outcomes (CLO)	CLO assessment	Matching with PLO		
			Coursework	Research	

Bảng ánh xạ chuẩn đầu ra môn học và chuẩn đầu ra chương trình ứng dụng:

	Chuẩn đầu ra của chương trình (CĐRCT)										
Chuẩn đầu ra môn học (CĐRMH)	a	b	c	d	e	f	g	h	i	j	k

Bảng ánh xạ chuẩn đầu ra môn học và chuẩn đầu ra chương trình nghiên cứu:

	Chuẩn đầu ra của chương trình (CĐRCT)										
Chuẩn đầu ra môn học (CĐRMH)	a	b	c	d	e	f	g	h	i	j	k

5. Hướng dẫn cách học - chi tiết cách đánh giá môn học:

Sinh viên cần đọc bài giảng của giảng viên được cập nhật trên elearning, các tài liệu khoa học, các bài đọc trước và làm báo cáo tiểu luận đầy đủ.

Tham gia đầy đủ chương trình

Làm bài tập về nhà

Tham gia các bài tập lớn

Learning strategies & Assessment Scheme:

Students need to read lectures updated on elearning website, scientific documents, and complete assignments and reports.

Full participation

Do homework

Praticipation in group projet

6. Nội dung chi tiết:

Tuần/ Buổi	Chủ đề (chương)	Nội dung	Chuẩn đầu ra môn học	Tài liệu
1	Dẫn nhập phân tích đa biến	Ứng dụng của phân tích đa biến Cấu trúc dữ liệu Khái niệm khoảng cách Đại số ma trận và các vectơ ngẫu nhiên		[1], [3]
2,3	ANOVA và MANOVA	Model và công thức tính các giá trị trung bình bình phương Sự tương tác giữa các yếu tố So sánh bội Trường hợp kích thước mẫu không đồng đều ANOVA& MANOVA 3,4 yếu tố với các phép đo lặp lại		[1],[4], [5]
4	Hồi quy bội	Ước lượng hệ số Lựa chọn biến Lựa chọn mô hình tuyến tính và điều chỉnh Dự báo		[3], [4], [5]
5,6,7	Phân tích thành phần chính	Đặc điểm của dữ liệu xử lý Cấu trúc dữ liệu Các phần tử phụ Giới thiệu phương pháp Khái niệm khoảng cách giữa hai đơn vị thống kê Ví dụ. Mô tả phương pháp. Các định nghĩa Vectơ chính và tính toán các thành phần chính Giải thích kết quả Biểu diễn hình học. Các mặt phẳng chính Một số thông số giúp đỡ cho việc giải thích kết quảVí dụ với R		[2],[3], [5]

Tuần/ Buổi	Chủ đề (chương)	Nội dung	Chuẩn đầu ra môn học	Tài liệu
8,9	Phân tích tương quan đa biến	Đặc điểm dữ liệu xử lý Cấu trúc dữ liệu Các phần tử phụ Mô tả phương pháp Các bảng xác suất Khái niệm khoảng cách giữa các profil Mô tả tập hợp của các profil. Các định nghĩa Đặc tính nhị phân Giải thích kết quả Biểu diễn đồ thị. Các mặt phẳng chính Một số thông số giải thích kết quả Ví dụ với R		[2],[3],[5]
10,11, 12	Phân tích đa nhân tố	Cấu trúc dữ liệu Cân bằng và PCA tổng thể Nghiên cứu nhóm (trình bày nhóm, biểu diễn điểm riêng phần, phân tích phân biệt) Các chủ đề khác (dữ liệu định tính, dữ liệu nhóm) Giải thích kết quả Ví dụ với R		[2],[3],[5]
13,14, 15	Phân loại và kỹ thuật phân nhóm	**Phân loại** Giới thiệu Phân loại hai tập hợp, phân loại hai tập hợp đa biến Đánh giá các hàm phân loại Hàm phân loại của Fisher **Phân nhóm** Giới thiệu Các phép đo sự đồng dạng Phương pháp phân nhóm có thứ bậc Phương pháp phân nhóm không thứ bậc MDS (Multidimensional Scaling) Biểu diễn đồ thị dữ liệu và kết quả		[2],[3],[4]

7. <u>Giảng viên tham gia giảng dạy:</u>

CBGD chính:	PGS.TS Nguyễn Hoàng Dũng

CBGD tham gia:	PGS.TS Nguyễn Thống

XÁC NHẬN CỦA HỘI ĐỒNG XÂY DỰNG CHƯƠNG TRÌNH ĐÀO TẠO VÀ KHOA

Tp. Hồ Chí Minh, ngày ..... tháng ..... năm .....

GIẢNG VIÊN LẬP ĐỀ CƯƠNG