

<b>TÊN ĐỒ ÁN: SỬ DỤNG NGÔN NGỮ R ĐỂ THĂM DÒ VÀ TRỰC QUAN DỮ LIỆU VỀ GIÁ CHỨNG KHOÁN CÁC MÃ LIÊN QUAN TỚI NGÂN HÀNG QUỐC DOANH</b>				
<b>Nhóm đăng ký đồ án: nhóm số:..... DHKL16A2HN</b>				
Họ và tên các thành viên nhóm	Mã Sinh viên	Điện thoại		
Nguyễn Văn A				
<p><b>Tóm tắt nội dung đồ án: Sử dụng ngôn ngữ R để thăm dò và trực quan dữ liệu về giá chứng khoán các mã liên quan tới ngân hàng quốc doanh</b></p> <p>1. Giới thiệu:</p> <ul style="list-style-type: none"> <li>- Đề xuất mục tiêu và tầm quan trọng của việc phân tích dữ liệu giá chứng khoán.</li> <li>- Giới thiệu tóm tắt về phương pháp phân tích thăm dò và trực quan dữ liệu.</li> </ul> <p>2. Thu thập dữ liệu:</p> <p>Sử dụng nguồn dữ liệu phổ biến nào (API, website, v.v.).</p> <p>Kỹ thuật thu thập dữ liệu (điều tra, phỏng vấn, khảo sát, cào,...)</p> <p>Làm sạch dữ liệu và xử lý các giá trị thiếu: Kỹ thuật nào? thu gọn còn bao nhiêu quan sát ?, bao nhiêu biến?, tên các biến, ý</p> <p>Chuyển đổi dữ liệu phù hợp cho phân tích.</p> <p>3. Thống kê mô tả:</p> <ul style="list-style-type: none"> <li>- Các đại lượng thống kê cơ bản : trung bình, trung vị, độ lệch chuẩn, v.v về chỉ số giá (giá mở, giá cao, giá thấp, giá đóng cửa, v.v.).</li> <li>- Biểu đồ tóm tắt dữ liệu (boxplot, histogram, v.v.).</li> <li>- Phân tích phân phối dữ liệu.</li> </ul> <p>Sử dụng ước lượng và kiểm định nào? để phân tích mối liên hệ giữa giá chứng khoán và các yếu tố khác (khối lượng giao dịch, tin tức thị trường, v.v.).</p> <p>4. Phân tích mối liên hệ giữa các biến:</p> <p>Biểu đồ (scatter plot) và ma trận tương quan (correlation matrix) phân tích cho biến nào?</p> <p>Hệ số tương quan (correlation coefficient) và độ đồng biến (covariance) sẽ cho kết quả gì?</p> <p>5. Trực quan hóa dữ liệu:</p> <p>Các loại biểu đồ và đồ thị phổ biến (bar chart, pie chart, line chart, v.v.) sử dụng cho dữ liệu nào? Kết quả dự kiến là gì</p> <p>6. Một số chủ đề nâng cao: lựa chọn một trong các chủ đề nâng cao sau để phân tích dữ liệu:</p> <p>Phân tích thành phần chính (PCA).</p> <p>Phân tích nhóm (cluster analysis).</p> <p>Khai phá dữ liệu (data mining).</p>				

Bảng phân công công việc				
Stt	Nội dung công việc	Kết quả đạt được	Thời gian bắt đầu, kết thúc	Người, đơn vị thực hiện

1	Thu thập dữ liệu:	Sử dụng nguồn dữ liệu phổ biến nào (API, website, v.v.). Kỹ thuật thu thập dữ liệu (điều tra, phỏng vấn, khảo sát, cào,...) Làm sạch dữ liệu và xử lý các giá trị thiếu: Kỹ thuật nào? thu gọn còn bao nhiêu quan sát ?, bao nhiêu biến?, tên các biến, ý Chuyển đổi dữ liệu phù hợp cho phân tích.	Từ .....đến	
2	Thống kê mô tả:	- Các đại lượng thống kê cơ bản : trung bình, trung vị, độ lệch chuẩn, v.v về chỉ số giá (giá mở, giá cao, giá thấp, giá đóng cửa, v.v.). - Biểu đồ tóm tắt dữ liệu (boxplot, histogram, v.v.). - Phân tích phân phối dữ liệu. Sử dụng ước lượng và kiểm định nào? để phân tích mối liên hệ giữa giá chứng khoán và các yếu tố khác (khối lượng giao dịch, tin tức thị trường, v.v.).	Từ .....đến	
3	Phân tích mối liên hệ giữa các biến:	Biểu đồ (scatter plot) và ma trận tương quan (correlation matrix) phân tích cho biến nào? Hệ số tương quan (correlation coefficient) và độ đồng biến (covariance) sẽ cho kết quả gì?	Từ .....đến	
4	Trực quan hóa dữ liệu:	Các loại biểu đồ và đồ thị phổ biến (bar chart, pie chart, line chart, v.v.) sử dụng cho dữ liệu nào? Kết quả dự kiến là gì	Từ .....đến	
5	Tạo ra một số chương trình tính chính, đánh giá, trực quan dữ liệu ban đầu để tiền xử lý dữ liệu trích xuất.	Viết chương trình trực quan hóa dữ liệu trích xuất bằng hai ngôn ngữ lập trình R, Python.	Từ .....đến	
6	Một số chủ đề nâng cao:	lựa chọn một trong các chủ đề nâng cao sau để phân tích dữ liệu: Phân tích thành phần chính (PCA). Phân tích nhóm (cluster analysis). Khai phá dữ liệu (data mining).	Từ .....đến	
7	Viết báo cáo	Chương 1: giới thiệu tổng quan về thư viện ggplot và phân tích thống kê Chương 2: Phân tích thăm dò dữ liệu Chương 3: Trực quan dữ liệu		

**Giảng viên hướng dẫn**

***Trần Chí Lê***

*Ngày ... tháng .. năm ...*

**Các thành viên tham gia đồ án**

***(Ký và ghi rõ họ tên)***

***Nguyễn Văn A***

***Nguyễn Văn B***

***Nguyễn Văn C***