**PHÂN TÍCH THỐNG KÊ NHIỀU CHIỀU**

**Bài tập 1: ANOVA**

Tệp số liệu được sử dụng: “HybridTest.xls”

Mục tiêu của nghiên cứu là để so sánh các loại xe ô tô điện và xe xăng truyền thống. Consumer Số liệu sau đây là số đo về lượng tiêu hao năng lượng đo bằng dặm/gallon (MPG) cho hai mẫu xe điện cỡ nhỏ, hai mẫu xe điện cỡ trung, hai mẫu xe điện SUV cỡ nhỏ và hai mẫu xe điện SUV cỡ trung. Tương tự là MPG cho tám mẫu xe xăng thông thường.

1. Thực hiện một số thống kê mô tả và nhận xét.

2. Với mức ý nghĩa alpha = 0,05, hãy kiểm định các tác động do loại xe, cỡ xe và tương tác giữa loại xe và cỡ xe.

3. Thực hiện các kiểm định so sánh cặp (multiple comparison).

**Bài tập 2: PCA và phân tích nhân tố**

Dữ liệu được sử dụng: “car\_sale”. Tệp dữ liệu này chứa các biến về doanh số, giá niêm yết và thông số kỹ thuật cho các kiểu dáng và mẫu xe khác nhau. (nhãn của các biến trong bảng “nhãn”).

Một nhà phân tích muốn dự đoán doanh số bán ô tô từ một tập hợp các yếu tố đã cho. Tuy nhiên, nhiều yếu tố có mối tương quan với nhau và nhà phân tích lo ngại rằng điều này có thể ảnh hưởng xấu đến kết quả phân tích.

Yêu cầu: Sử dụng Phân tích nhân tố với PCA để phân tích các yếu tố ảnh hưởng đến giá bán.

**Bài tập 3: Phân tích cụm**

Các nhà sản xuất ô tô cần có khả năng đánh giá thị trường hiện tại để xác định khả năng cạnh tranh cho xe của họ. Nếu ô tô có thể được nhóm theo dữ liệu có sẵn thì việc này sẽ có thể được thực hiện tự động bằng cách sử dụng phân tích cụm.

Thông tin về các kiểu dáng và nhãn hiệu xe khác nhau có trong tập dự liệu “car\_sale” (giống như dữ liệu cho Bài tập 2). Sử dụng quy trình Phân tích cụm phân cấp (hierarchical) để nhóm các loại ô tô bán chạy nhất theo giá cả và đặc tính kĩ thuật của chúng.

**Bài tập 4: Phân tích khác biệt**

Một hãng hàng không thu thập dữ liệu về nhân viên của họ đang làm 3 loại công việc: 1) nhân viên dịch vụ khách hàng, 2) công nhân kĩ thuật và 3) nhân viên điều phối. Giám đốc Nhân sự của hãng muốn biết liệu ba cách phân loại công việc này có phù hợp với các loại tính cách khác nhau của nhân hay không. Mỗi nhân viên được thực hiện một loạt bài kiểm tra tâm lý để đánh giá, đo lường về 3 loại tính cách: (i) tính cách hướng ngoại; (ii) tính cách hòa đồng và (iii) tính bảo thủ (khó thay đổi hoặc bị thuyết phục bởi người khác.

Dự liệu trong tệp “discriminant\_example” gồm biến: JOB (loại công việc được phân loại); OUTDOOR: Tính hướng ngoại; SOCIAL: Tính hòa đồng và CONSERVATIVE: Tính bảo thủ.

Yêu cầu: Sử dụng pp phân tích khác biệt để phân loại nhân viên của hang thành 3 nhóm.

YÊU CẦU CHUNG:

Viết báo cáo chi tiết cá nhân. Nộp báo cáo và tệp code R tương ứng`

Tham khảo bài giảng và các tài liệu tham khảo khác để tự tìm hiểu cách viết báo cáo và code phù hợp.

Hạn nộp: 0h đêm ngày 4/12/2023. Cả lớp gửi file của mình (báo cáo và code R) cho lớp trưởng (Thương). Sau đó thương gửi vào email của thầy thành một thư mục (file rar), (thenm@neu.edu.vn hoặc manhthe@gmail.com). Các tệp phải bắt đầu bằng tên đầy đủ của SV (ví dụ “Nguyen\_Manh\_The\_Report” và “Nguyen\_Manh\_The\_script”).