



Bài giảng
**HƯỚNG DẪN
THỰC HÀNH
SPSS**

Phan Thanh Hồng - Nguyễn Thị Nhung

Tháng 5 năm 2010

Mục lục

	Trang
Bài 1. Xử lý dữ liệu trên SPSS	1
1.1 Giới thiệu và cài đặt phần mềm SPSS	1
1.1.1 Giới thiệu phần mềm SPSS	1
1.1.2 Cài đặt SPSS	1
1.1.3 Khởi động SPSS	1
1.2 Nhập dữ liệu trong SPSS	2
1.3 Đọc dữ liệu từ những file có sẵn trong SPSS	9
1.3.1 Đọc dữ liệu từ file .sav	9
1.3.2 Đọc dữ liệu từ file .xls	10
1.3.3 Đọc dữ liệu từ file.txt	11
1.4 Lọc dữ liệu	14
1.5 Mã hóa dữ liệu	20
1.5.1 Mã hóa dữ liệu trong nội bộ biến	20
1.5.2 Mã hóa dữ liệu vào biến mới	23
1.6 Làm sạch dữ liệu	26
1.6.1 Dùng bảng tần số	27
1.6.2 Dùng lệnh Sort case để tìm dữ liệu sai trên cửa sổ Data View	28
1.6.3 Dùng bảng phối hợp hai biến hay ba biến	28
Bài 2. Lập bảng tần số và tính các đại lượng thống kê mô tả	30
2.1 Tóm tắt dữ liệu bằng bảng tần số	30
2.1.1 Bảng tần số cho dữ liệu định tính và dữ liệu định lượng có ít biểu hiện	30
2.1.2 Phân tổ dữ liệu	31
2.2 Lập bảng tổng hợp nhiều biến	32
2.2.1 Bảng kết hợp hai biến định tính	32
2.2.2 Bảng kết hợp ba biến định tính	34
2.2.3 Bảng kết hợp một biến định tính và một biến định lượng	35
2.3 Tính các đại lượng thống kê mô tả	36
2.3.1 Bài tập ôn luyện	40
Bài 3. Biểu đồ	41

3.1	Một số loại biểu đồ	41
3.2	Biểu đồ thanh	41
3.2.1	Cách vẽ biểu đồ thanh	41
3.2.2	Ví dụ: vẽ biểu đồ thanh biểu diễn một biến	43
3.2.3	Ví dụ: vẽ biểu đồ thanh biểu diễn một biến được phân tách bởi một biến khác	43
3.3	Biểu đồ tròn	45
3.3.1	Cách vẽ biểu đồ tròn	45
3.4	Biểu đồ gấp khúc, diện tích (Line, Area Chart)	46
3.4.1	Ví dụ	46
3.4.2	Ví dụ	47
3.5	Biểu đồ tần số (Histogram)	48
3.5.1	Cách vẽ biểu đồ tần số (Histogram)	49
3.6	Biểu đồ hộp và râu (Boxplot)	51
3.6.1	Cách vẽ biểu đồ hộp và râu (Boxplot)	51
3.7	Biểu đồ tán xạ (Scatter plot)	54
3.7.1	Cách vẽ biểu đồ tán xạ (Scatter plot)	54
Bài 4.	Kiểm định mối quan hệ giữa hai biến	56
4.0.2	Ví dụ: kiểm định mối liên hệ hai biến định danh, hay định danh-thứ bậc	56
4.0.3	Kiểm định mối liên hệ hai biến thứ bậc	59
Bài 5.	Kiểm định về trung bình	61
5.1	Kiểm định về trung bình một tổng thể	61
5.2	Kiểm định về sự khác nhau của hai trung bình của hai tổng thể khi mẫu độc lập	62
5.3	Kiểm định về sự khác nhau của hai trung bình của hai tổng thể khi mẫu cặp	64
Bài 6.	Phân tích phương sai (ANOVA)	67
6.0.1	Ví dụ	67
6.0.2	Phân tích sâu ANOVA	69
Tài liệu tham khảo		71

Bài 1

XỬ LÝ DỮ LIỆU TRÊN SPSS

1.1 Giới thiệu và cài đặt phần mềm SPSS

Giới thiệu phần mềm SPSS

SPSS (viết tắt của Statistical Package for the Social Sciences) là một chương trình máy tính phục vụ công tác thống kê. SPSS được sử dụng rộng rãi trong công tác thống kê xã hội.

Thế hệ đầu tiên của SPSS được đưa ra từ năm 1968. Thế hệ mới nhất là thế hệ 18 được giới thiệu từ tháng 8 năm 2008, có cả phiên bản cho các hệ điều hành Microsoft Windows, Mac, và Linux / UNIX.

SPSS là một hệ thống phần mềm thống kê toàn diện được thiết kế để thực hiện tất cả các bước trong các phân tích thống kê từ những tính toán trong thống kê mô tả (liệt kê dữ liệu, lập bảng tần số, lập biểu đồ, tính các đại lượng thống kê mô tả,...) đến thực hiện những bài toán trong thống kê suy diễn (kiểm định, tương quan, hồi qui, ...)

Cài đặt SPSS

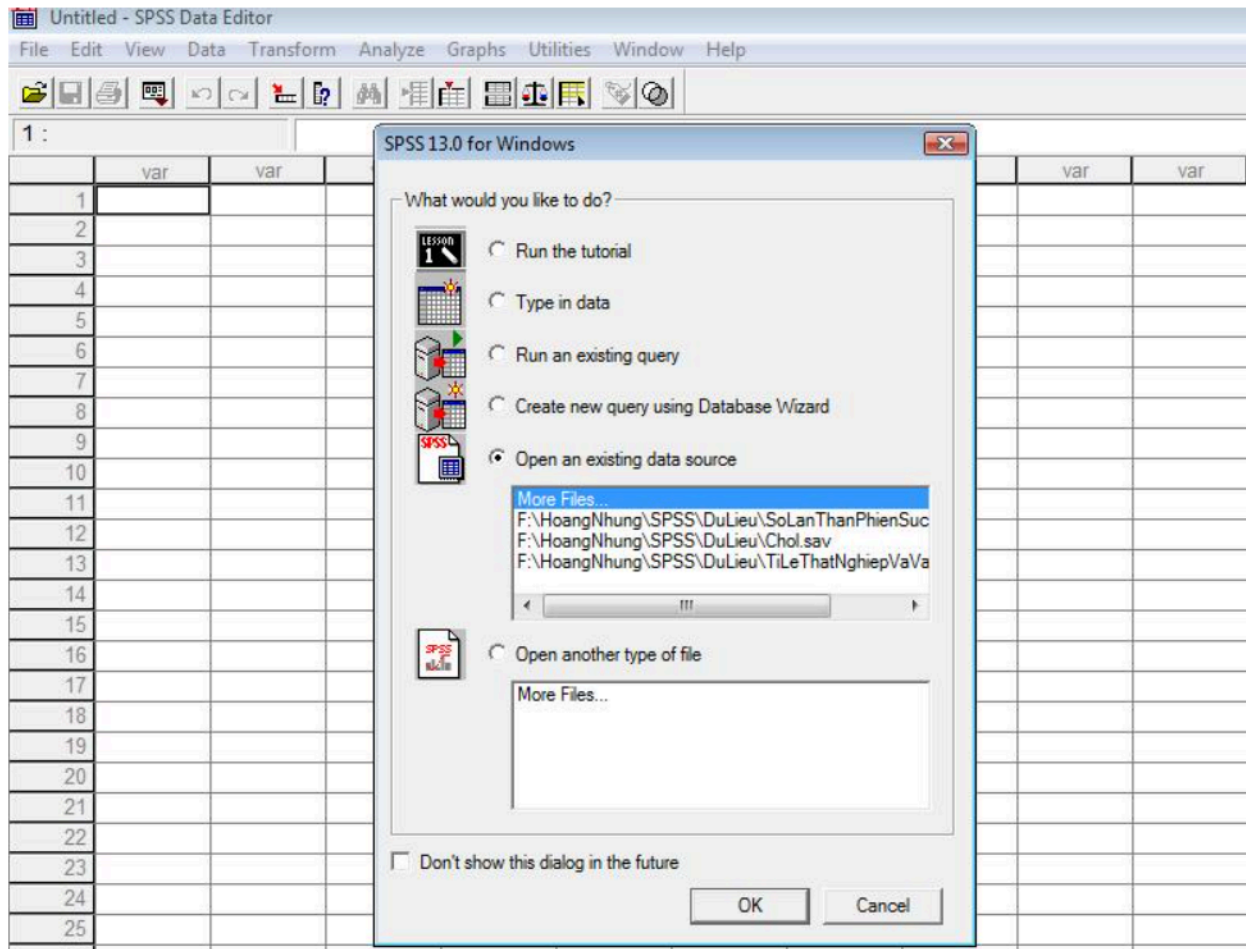
Cho đến nay, SPSS có rất nhiều phiên bản khác nhau, phiên bản mới nhất (5/2010) là phiên bản 18.0. Với những tính toán thống kê phổ biến ta có thể cài SPSS với một trong các phiên bản 13.0, 14.0 hay 16.0, ... Tất cả các phân tích trong tài liệu này được thực hiện trên phiên bản SPSS 13.0.

Khởi động SPSS

Để khởi động SPSS ta có thể thực hiện theo những cách sau:

- Kích đúp vào biểu tượng SPSS trên màn hình;
- Vào **Start** → **Program** → **SPSS for Windows** → **SPSS 13.0 for Windows**

Sau khi khởi động, ta được giao diện sau:



1.2 Nhập dữ liệu trong SPSS

Để nhập một tập dữ liệu vào SPSS, ta thực hiện như sau:

Bước 1: Tại cửa sổ **SPSS Data Editor**, ta ấn vào nút **Variable View** để khai báo thông tin về các thuộc tính của từng biến trong file dữ liệu.

Trong một bảng Variable View

- Các hàng là các biến;
- Các cột là các thuộc tính của biến.

Các thuộc tính của một biến bao gồm:

- Tên biến (**Name**)
- Loại dữ liệu (**Type**)
- Số lượng con số hoặc chữ (**Width**)
- Số lượng chữ số thập phân (**Decimals**)
- Mô tả biến/nhãn biến (**Label**) và nhãn trị số biến (**Values**)