

## Correlation with SPSS

### Correlation

Kiểm tra mối quan hệ giữa hai biến:

- Có liên quan với nhau không (associated/ related or not?)
- Liên quan như thế nào (how related?)
- Có biến thiên cùng với nhau không (as X increases, Y also increases and vice versa?)

Ví dụ: Có mối quan hệ giữa số giờ tự học ở nhà và điểm số bài thi hay không?

### Thiết kế tương quan

TA KHÔNG:

- Không tác động lên biến độc lập (IV) và đo lường biến phụ thuộc (DV)
- Không chia người tham gia vào các điều kiện thí nghiệm hoặc nhóm khác nhau

MÀ TA SẼ:

- Đo lường hai biến và đi tìm mối quan hệ giữa chúng

Ví dụ: Có mối quan hệ giữa trình độ học vấn và thu nhập không? – Ta thu thập dữ kiện về trình độ học vấn và thu nhập của người tham gia (chứ không can thiệp lên biến nào cả). Đây là một thiết kế cắt ngang tại một thời điểm.

### Hệ số tương quan

Cung cấp cho ta các thông tin quan trọng và hữu ích:

- Hướng (direction) của mỗi quan hệ:
  - Hệ số tương quan dương (+) biểu hiện mỗi quan hệ thuận
  - Hệ số tương quan âm (-) biểu hiện mỗi quan hệ nghịch
- Độ mạnh (strength) của mỗi quan hệ:
  - Dao động trong khoảng -1 đến +1
  - -1 = tương quan nghịch hoàn hảo
  - 0 = không có tương quan
  - +1 = tương quan thuận hoàn hảo

VD:

$r = .92$  là một tương quan thuận mạnh

$r = -.13$  là một tương quan nghịch trung bình – yếu

$r = -.88$  là một tương quan nghịch mạnh

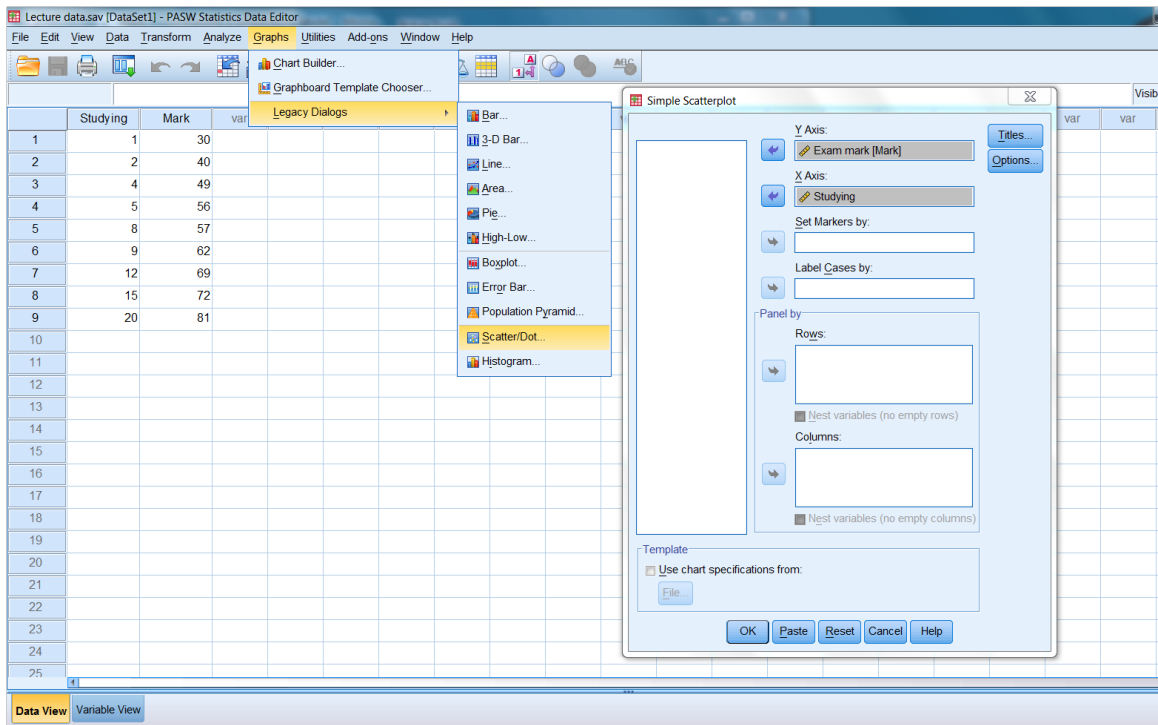
### Raw data

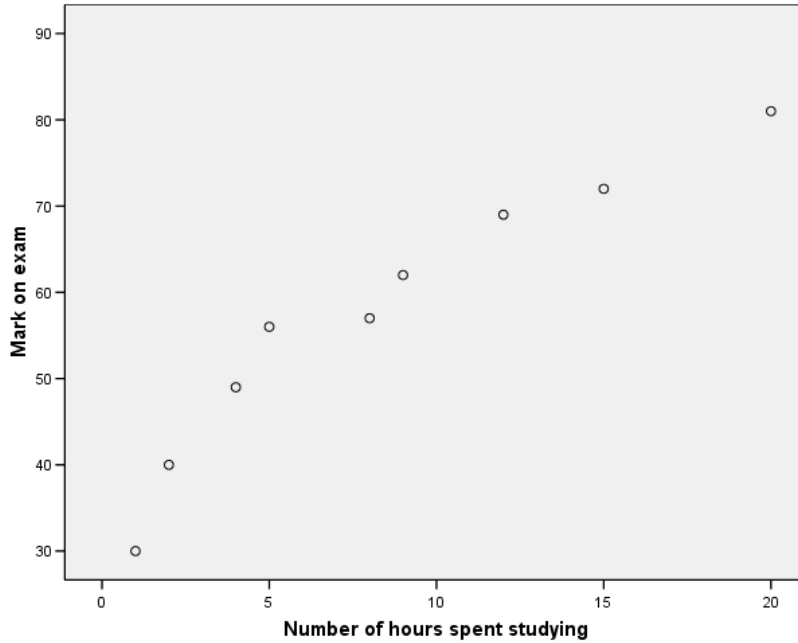
Giả thuyết: Số giờ tự học có liên quan với điểm số trong kỳ thi. Để kiểm chứng, ta thu thập dữ kiện về hai biến số: số giờ tự học và điểm.

Participant	Studying	Exam mark
1	1	30
2	2	40
3	4	49

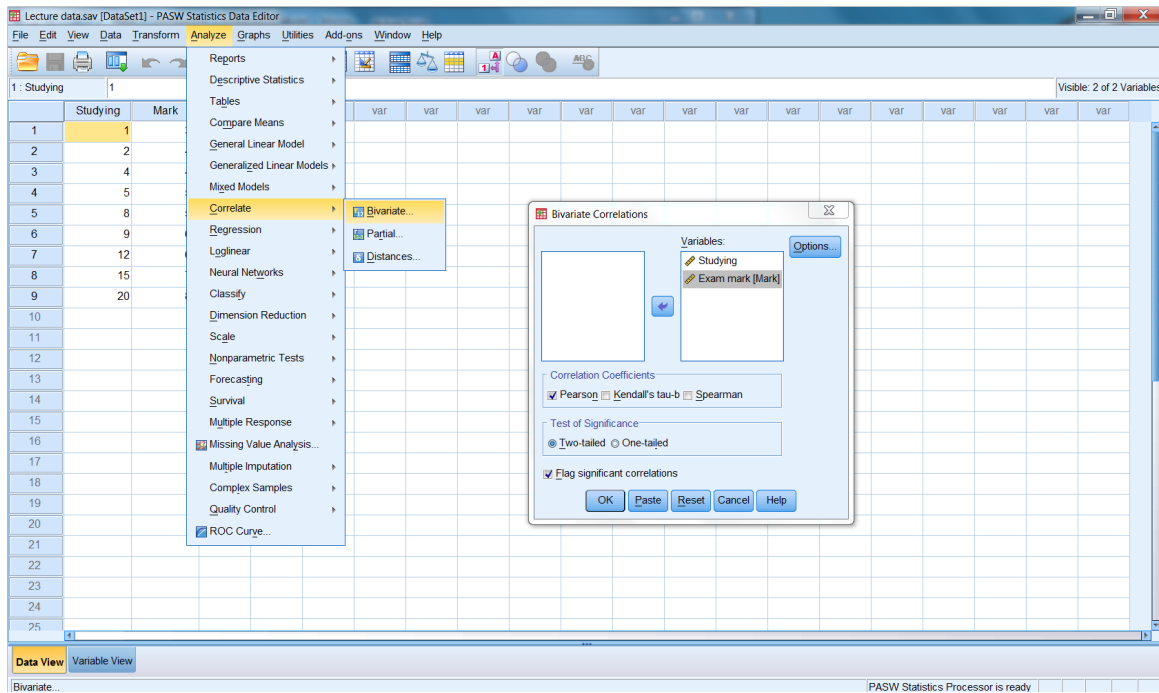
4	5	56
5	8	57
6	9	62
7	12	69
8	15	72
9	20	81

### Plot with SPSS





### Pearson's r correlation



### Correlation coefficient: Pearson's r in SPSS

Descriptive Statistics

	Mean	Std. Deviation	N
Number of hours spent studying	8.44	6.307	9
Mark on exam	57.33	16.016	9

**Correlations**

		Number of hours spent studying	Mark on exam
Number of hours spent studying	Pearson Correlation	1	.954**
	Sig. (2-tailed)		.000
	N	9	9
Mark on exam	Pearson Correlation	.954**	1
	Sig. (2-tailed)	.000	
	N	9	9

\*\* . Correlation is significant at the 0.01 level (2-tailed).

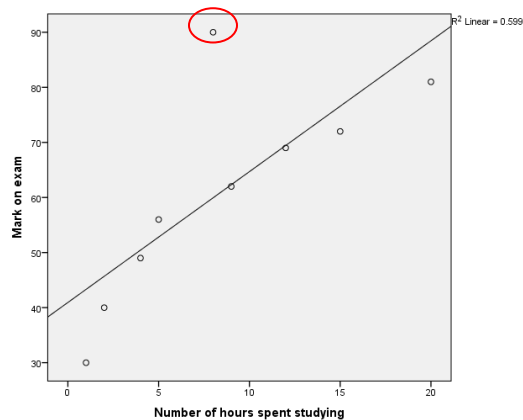
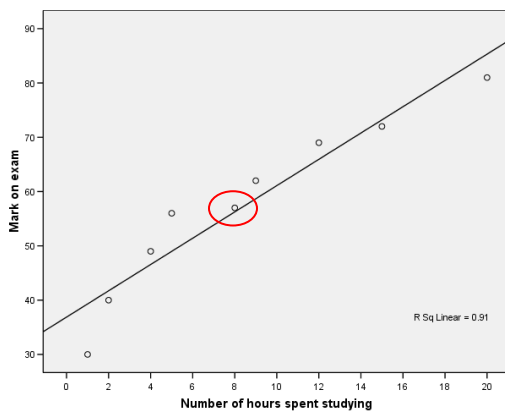
Báo cáo: Có mối quan hệ thuận có ý nghĩa thống kê giữa số giờ tự học và điểm số trong kỳ thi,  $r(7) = .95, p < .001$ .

**Các giả định của Pearson's r**

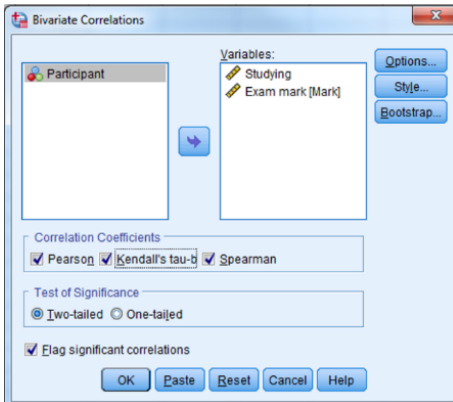
- 1) Dữ kiện khoảng hoặc tỉ lệ
- 2) Phân bố chuẩn
- 3) Mối quan hệ tuyến tính
- 4) Coi chừng: giá trị ngoại biên – Extreme Outliers/ Influential cases

Participant	Studying	Exam mark
1	1	30
2	2	40
3	4	49
4	5	56
5	8	57
6	9	62
7	12	69
8	15	72
9	20	81

Participant	Studying	Exam mark
1	1	30
2	2	40
3	4	49
4	5	56
5	8	90
6	9	62
7	12	69
8	15	72
9	20	81



\*Khi giả định của Pearson's r không thỏa mãn, ta dùng **Spearman's rho** (hoặc Kendall's  $\tau$ )



**Correlations**

		Studying	Exam mark
Spearman's rho	Studying	1.000	.833**
	Correlation Coefficient		
	Sig. (2-tailed)	.	.005
	N	9	9
Exam mark	Exam mark	.833**	1.000
	Correlation Coefficient		
	Sig. (2-tailed)	.005	.
	N	9	9

\*\* . Correlation is significant at the 0.01 level (2-tailed).

### Thực hành

Giả thuyết: Sau một trận động đất, các nạn nhân được đánh giá trên một thang đo mức độ mạo hiểm. Nhà nghiên cứu sau đó thu thập lại dữ kiện về số lần họ tìm kiếm sự giúp đỡ. Ông ta đặt câu hỏi nghiên cứu rằng liệu mức độ mạo hiểm có liên quan đến số lần tìm kiếm sự giúp đỡ của các nạn nhân động đất.

1) Hãy nhập dữ liệu sau vào SPSS

Người tham gia	Mức độ mạo hiểm	Tìm giúp đỡ
1	8	7
2	3	2
3	9	7
4	7	6
5	2	3
6	3	2
7	9	7
8	8	6
9	6	5
10	7	9

2) Vẽ biểu đồ tán xạ

3) Vẽ đường line of best fit

Click đúp vào scatterplot để mở hộp Chart Editor

Click Elements/ Fit Line at Total.

4) Hãy dùng kiểm định Pearson để kiểm định giả thuyết. Hãy báo cáo kết quả theo APA format dưới dạng văn bản.