Quantitative methods

Lesson 9

Daróczi Gergely

Corvinus University of Budapest, Hungary

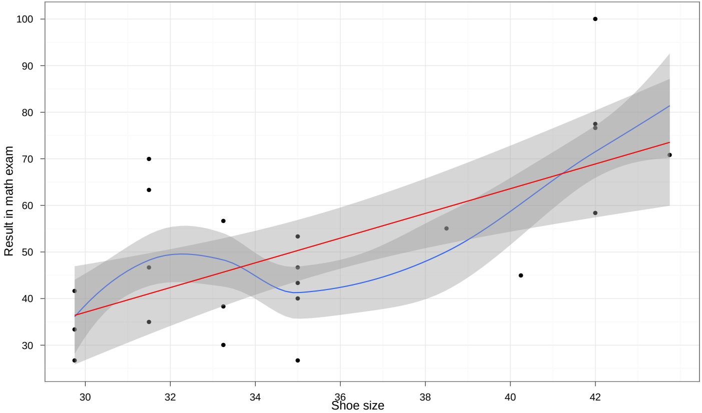
2011 April 12

Outline

Research in an elementary school

Big shoes and smart kids (example)

A small field work to examine the relationship between math skills and shoe size among juniors:



How would you comment on this very spectacular relationship?

Research in an elementary school

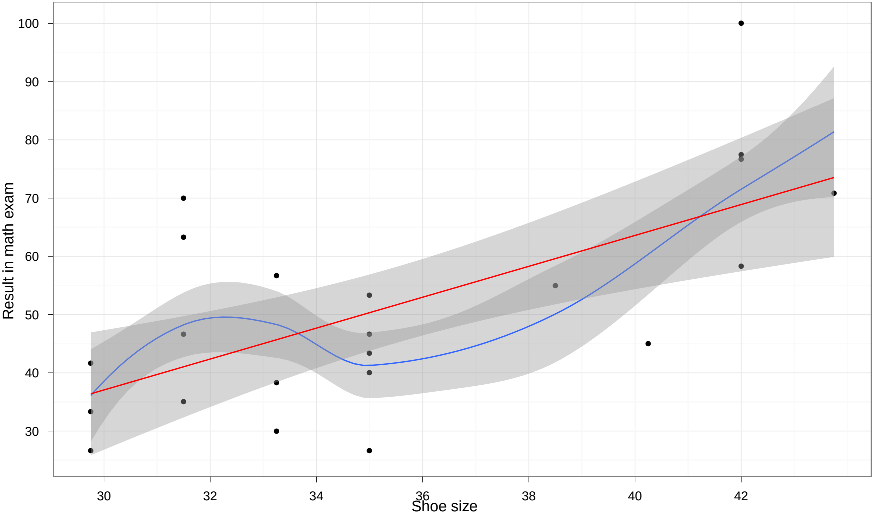
Big shoes and smart kids (example)

We made a small research on the age and shoe size of some students in an elementary shool, where we also conducted a math exam. See detailed results below:

	Shoe size	Math result	Age
1	29.75	26.67	3
2	29.75	33.33	7
3	29.75	41.67	5
4	31.50	35.00	8
5	31.50	46.67	10
6	31.50	63.33	11
7	31.50	70.00	12
8	33.25	30.00	7.
9	33.25	38.33	7
10	33.25	56.67	12
11	35.00	26.67	6
12	35.00	40.00	8
13	35.00	43.33	6
14	35.00	46.67	10
15	35.00	53.33	11
16	38.50	55.00	9
17	40.25	45.00	9
18	42.00	58.33	9
19	42.00	76.67	16
20	42.00	77.50	18
21	42.00	100.00	19
22	43.75	70.83	14

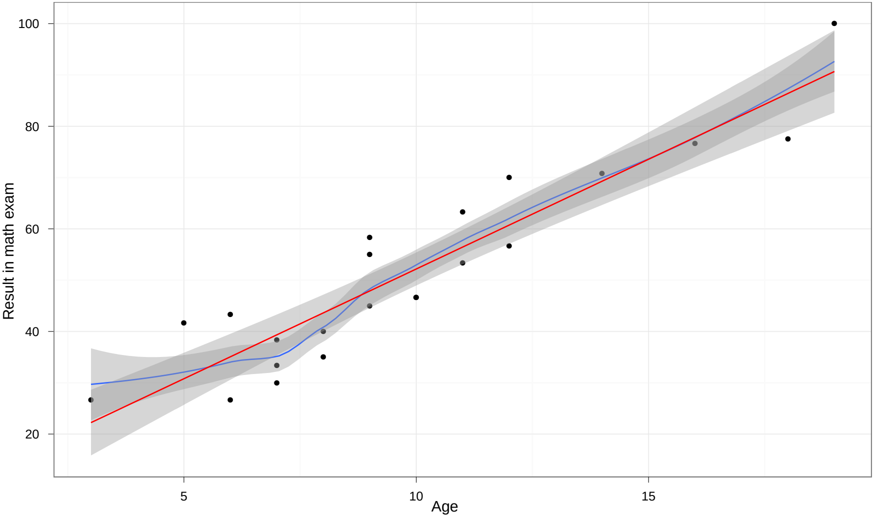
Research in an elementary school

Big shoes and smart kids (example)



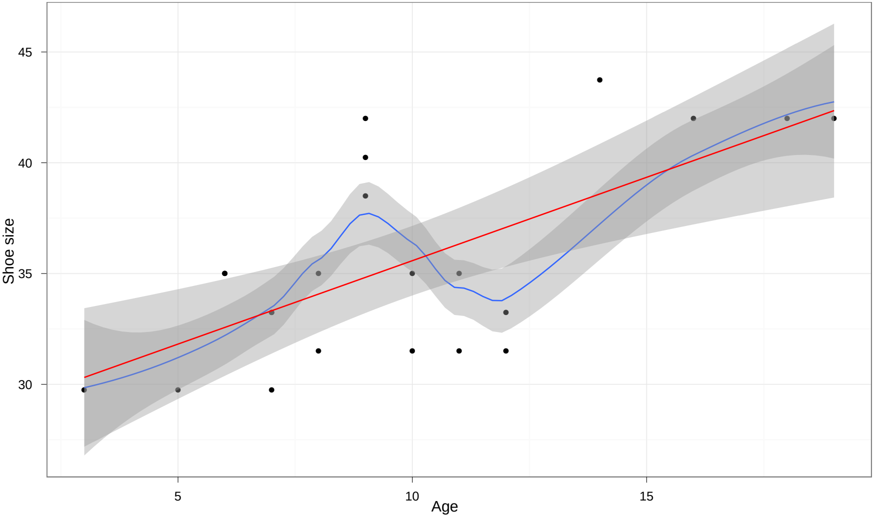
Research in an elementary school

Big shoes and smart kids (example)



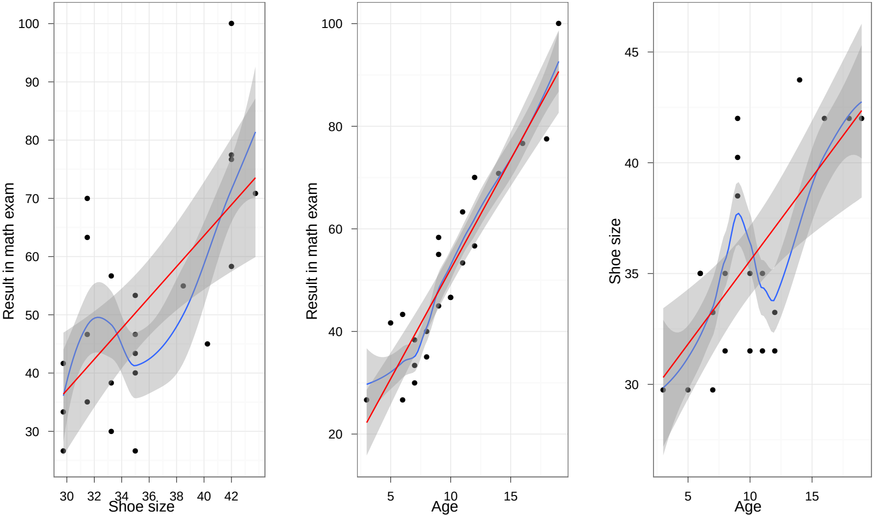
Research in an elementary school

Big shoes and smart kids (example)



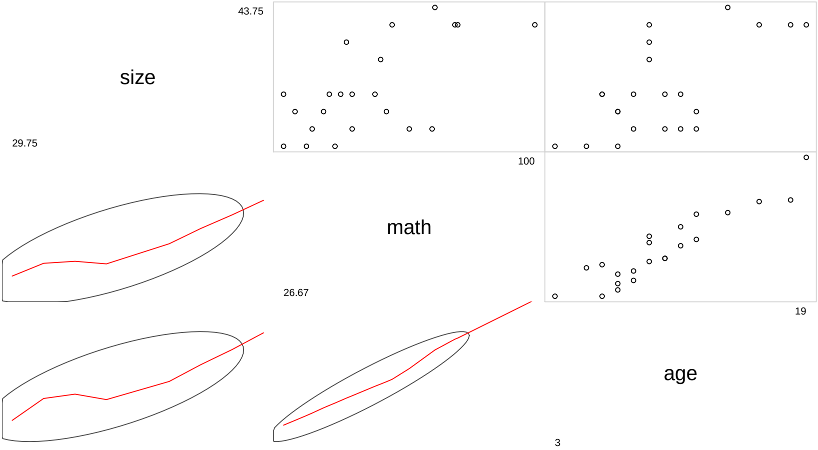
Research in an elementary school

Big shoes and smart kids (example)



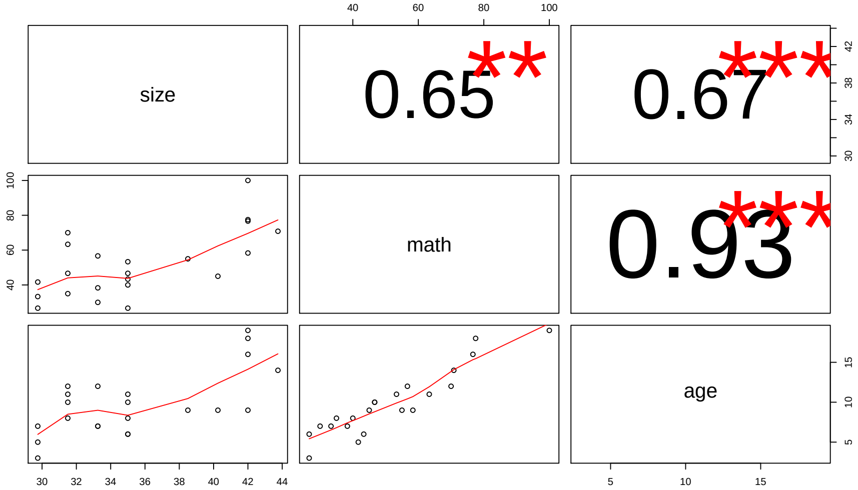
Research in an elementary school

Big shoes and smart kids (example)



Research in an elementary school

Big shoes and smart kids (example)



Partial correlation:

$$r_{math,size \cdot age} = 0.11$$

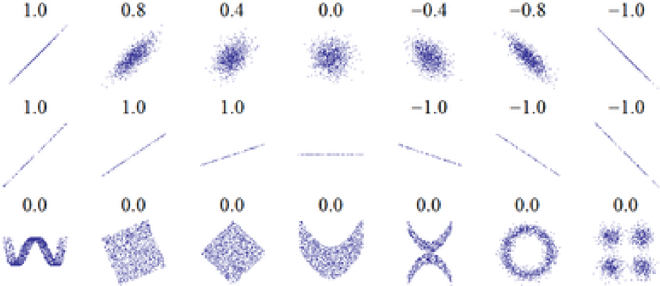
$$r_{math,age \cdot size} = 0.87$$

$$r_{size,age \cdot math} = 0.22$$

Theoretical background

Correlation

$$r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}}$$



$$\hat{r}_{XY\cdot\mathbf{Z}} = \frac{N\sum_{i=1}^{N} r_{X,i}r_{Y,i} - \sum_{i=1}^{N} r_{X,i}\sum_{i=1}^{N} r_{Y,i}}{\sqrt{N\sum_{i=1}^{N} r_{X,i}^2 - \left(\sum_{i=1}^{N} r_{X,i}\right)^2}\sqrt{N\sum_{i=1}^{N} r_{Y,i}^2 - \left(\sum_{i=1}^{N} r_{Y,i}\right)^2}}$$

so for three variables:

$$\hat{r}_{XY\cdot\mathbf{Z}} = \frac{r_{XY} - r_{X\mathbf{Z}}r_{Y\mathbf{Z}}}{\sqrt{(1-r_{X\mathbf{Z}}^2)(1-r_{Y\mathbf{Z}}^2)}}$$

Please check Chapter 8 in

Darrel Huff (1993): *How to lie with statistics*, W. W. Norton Company!

Exercises

1. Check on correlation and partial correlation in a statistics textbook or on the Internet!
2. Building upon your findings, compute the possible pairs of correlation coefficients on the below dataset!
3. Also look for partial correlation and comment on your results!

Grade (mean)	Scholarship (in HUF)	Money spent on books (in HUF)
3.05	22000	35000
3.2	25000	30000
3.35	27000	28000
3.35	24000	37000
3.45	25000	22000
3.55	28000	32000
3.7	28000	37000
3.75	30000	41000
3.8	27000	40000
3.8	29000	38000

It was a pleasure!

Daróczi Gergely
daroczi.gergely@btk.ppke.hu