Annex A · Student handout

Are these four datasets the same?

For each set, work out: mean of x, mean of y, variance of x, variance of y, correlation between x and y, and the regression line y = a + b·x. **Afterwards** (not before!) draw a scatter plot of each one.

Set I		Set II		Set III		Set IV	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

	I	**II**	**III**	**IV**
Mean of x				
Mean of y				
Variance of x				
Variance of y				
Correlation r				
Line y = a + b·x				

What do you conclude from the numbers? .....

.....

And after drawing them? .....

.....