

correlation
co-relation
joint

r

TABLE 1
Galton's correlation table

Height of the midparent in inches	Height of the adult child														Total no. of adult children
	<61.7	62.2	63.2	64.2	65.2	66.2	67.2	68.2	69.2	70.2	71.2	72.2	73.2	>73.7	
>73.0	—	—	—	—	—	—	—	—	—	—	—	1	3	—	4
72.5	—	—	—	—	—	—	—	1	2	1	2	7	2	4	19
71.5	—	—	—	—	1	3	4	3	5	10	4	9	2	2	43
70.5	1	—	1	—	1	1	3	12	18	14	7	4	3	3	68
69.5	—	—	1	16	4	17	27	20	22	25	20	11	4	5	183
68.5	1	—	7	11	16	28	31	34	48	21	18	4	3	—	219
67.5	—	3	5	14	15	26	33	28	38	19	11	4	—	—	211
66.5	—	3	3	5	2	17	17	14	13	4	—	—	—	—	78
65.5	1	—	9	5	7	11	14	7	7	5	2	1	—	—	66
64.5	1	1	4	4	1	5	5	—	2	—	—	—	—	—	23
<64.0	1	—	2	4	1	2	2	1	1	—	—	—	—	—	14
Totals	5	7	32	59	48	117	138	120	167	99	64	41	17	14	928

Y

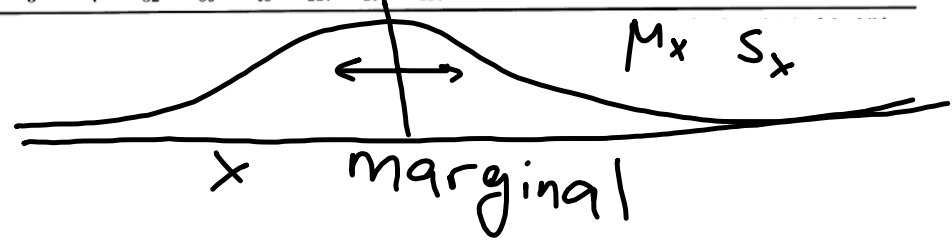
M_y

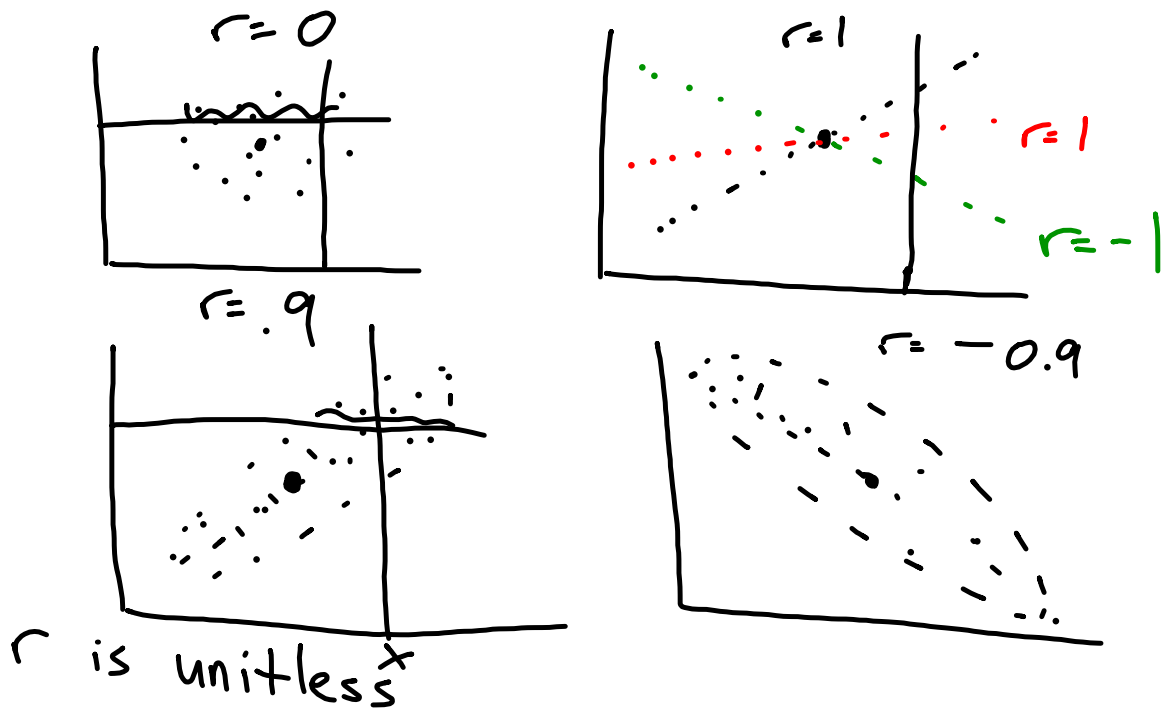
s_y

variance

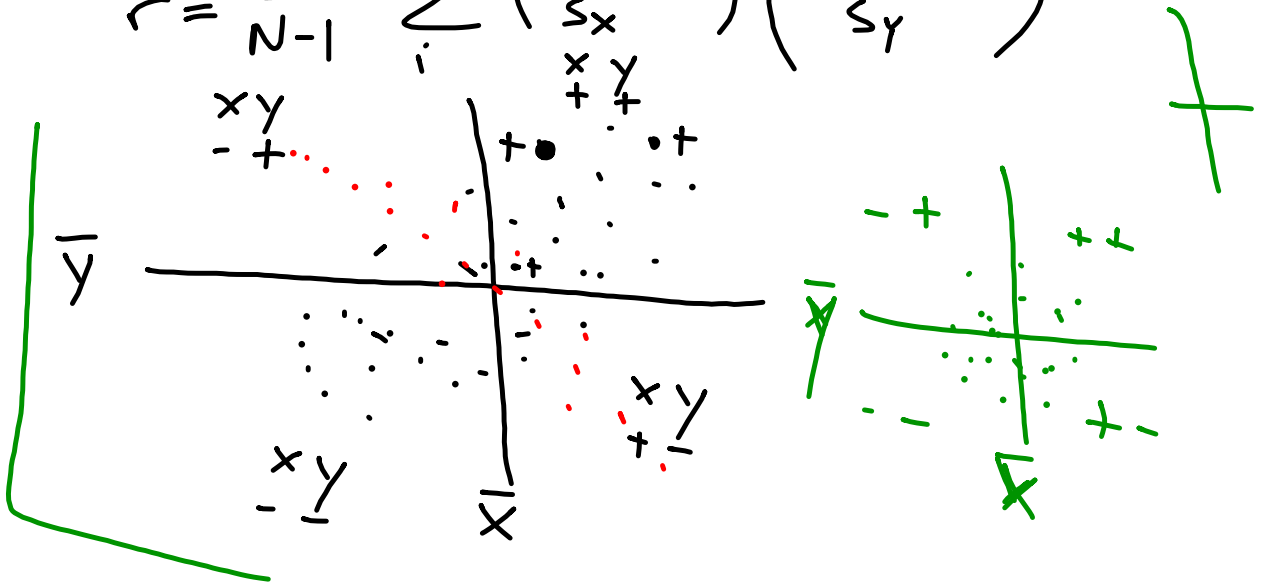
M_x S_x

X marginal

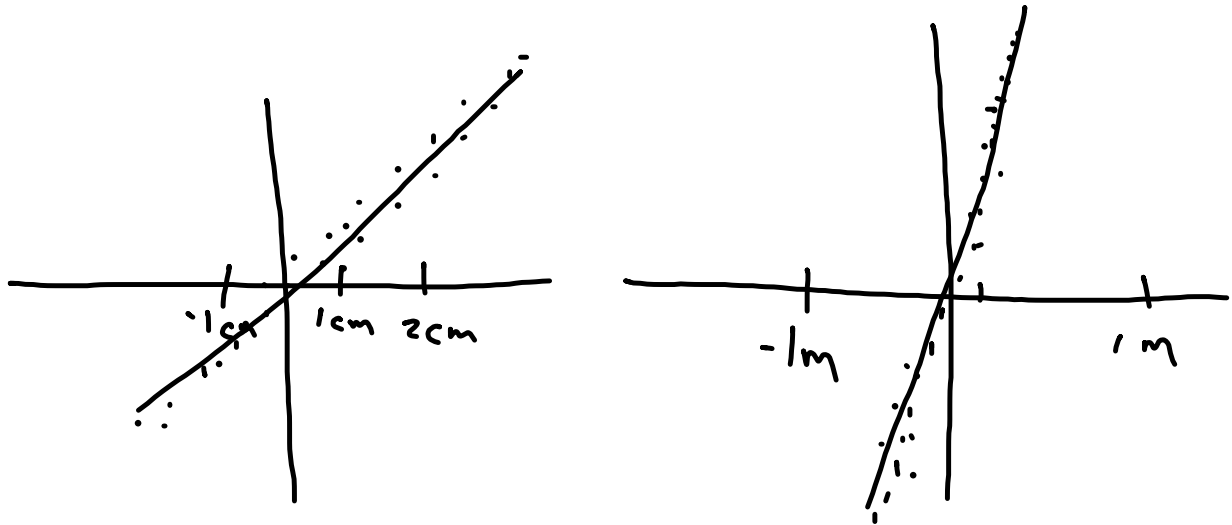


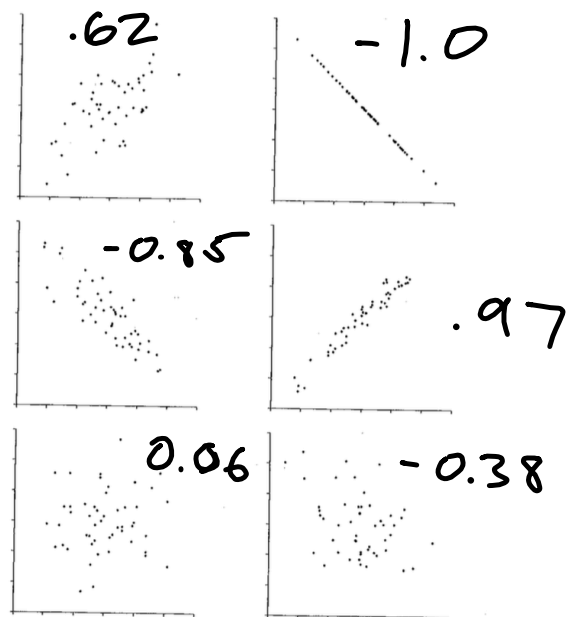


$$r = \frac{1}{N-1} \sum_i \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right)$$



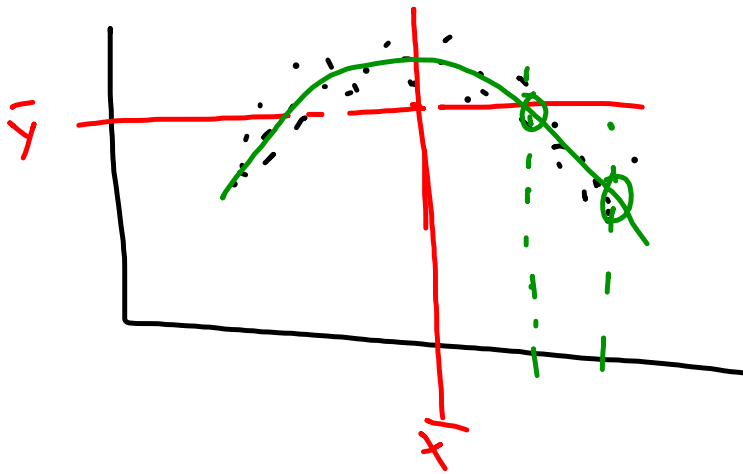
scaling doesn't matter

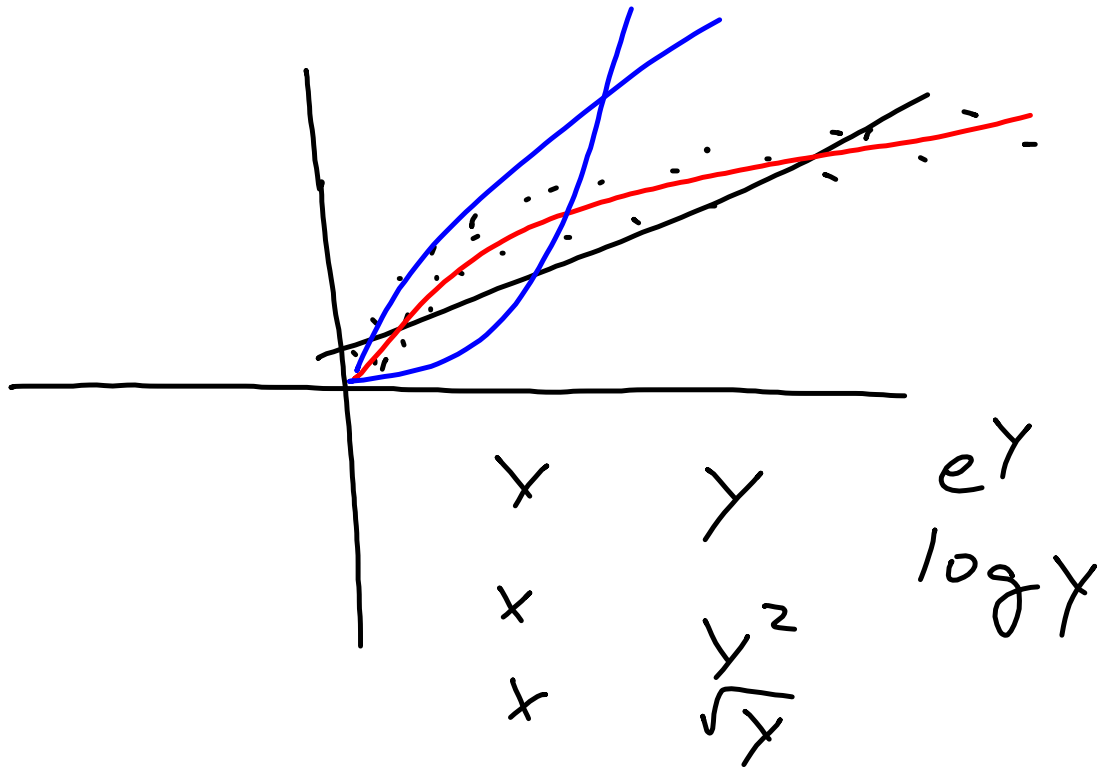




correlation is linear

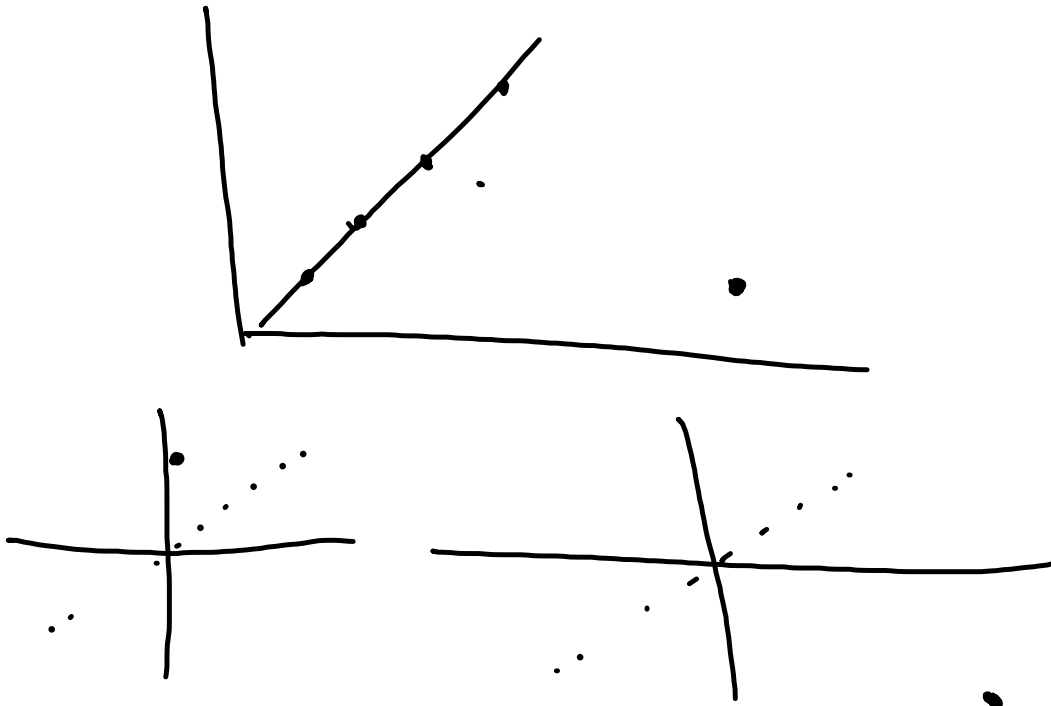
$$r=0$$

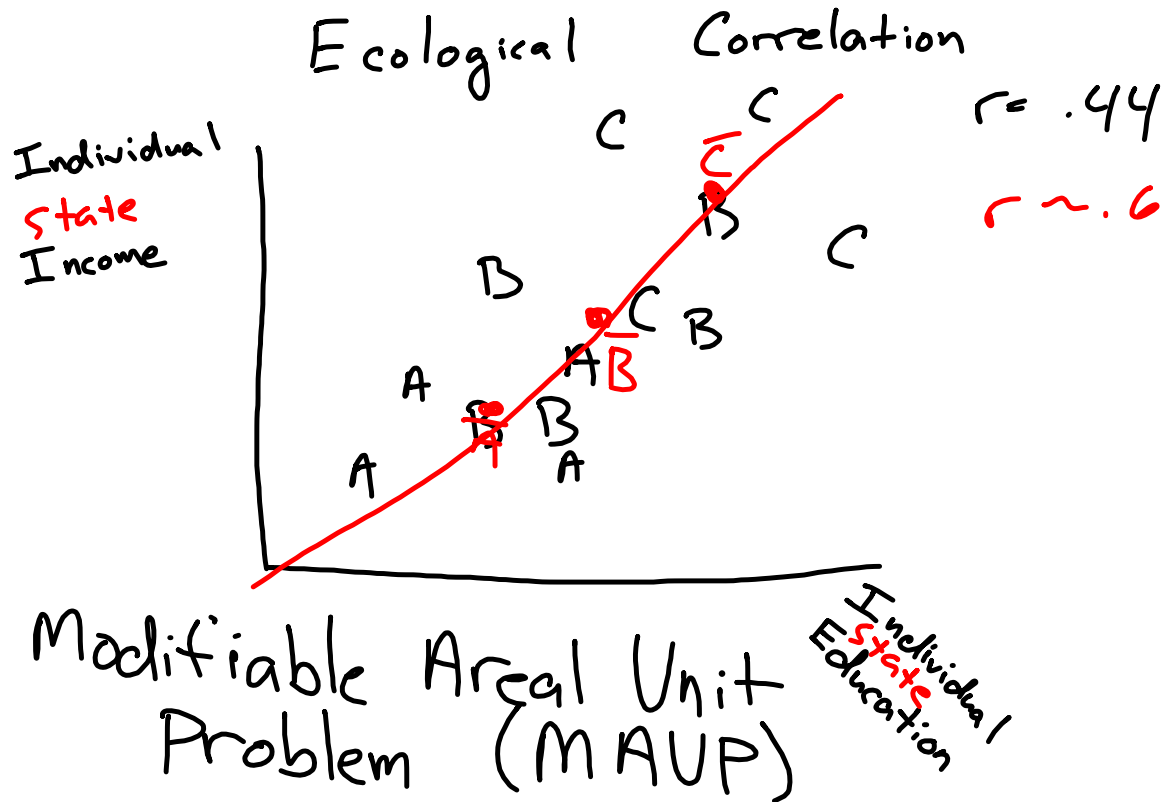






Outliers can be a problem





correlation \neq causation
 correlation = associated
 Children 5-17

