

Correlation Coefficient

Here is the percent of people who have used marijuana and other drugs in 10 western countries.

Country	Marijuana	Other Drugs
Czech Republic	22	4
Denmark	17	3
England	40	21
Finland	5	1
Ireland	37	6
Italy	19	8
Northern Ireland	23	14
Norway	6	3
Portugal	7	3
Scotland	53	31
USA	34	24

Since the data here are points (x, y) , we can make a graph. When you are graphing data, it is called a scatterplot.

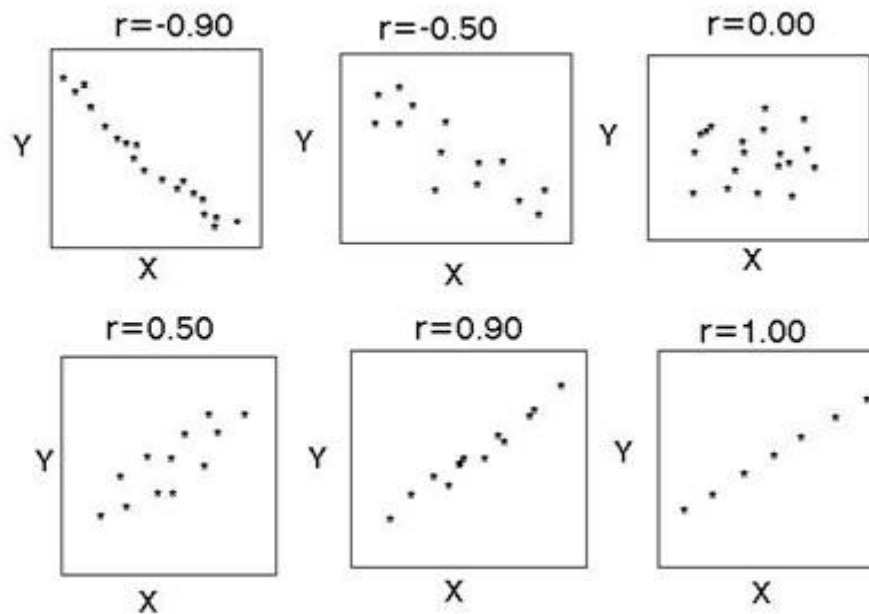


1. Name the axes, fill in the scales, and then graph the data.

We are very interested in how “linear” the data is. The correlation coefficient r is a number that measures how linear your data is.

Facts about r

1. $-1 \leq r \leq 1$
2. $r = 1$ is perfect positive linearity, $r = -1$ is perfect negative linearity
3. $r = 0$ is not linear at all



Here are some pictures showing a few different r 's.

In general, an r greater than 0.5 or less than -0.5 indicates a significant linear relationship.

2. Enter the data from #1 into a graphing calculator and then find the line of regression to find r . Is there a significant linear relationship? Is it positive or negative?

For each data set, make a scatterplot, and then find and interpret the correlation coefficient r .

3. Frying time of a tortilla chip (x) and the % water content in the chip (y).

(5, 16.3) (10, 9.7) (15, 8.1) (20, 4.2) (25, 3.4) (30, 2.9) (45, 1.9) (60, 1.3)

4. The age of injured athletes (x) and the number of days after shoulder surgery before being able to return to their sport (y)

(33, 6) (27, 2) (31, 4) (32, 4) (28, 1) (33, 3) (26, 3) (34, 4) (32, 2) (28, 3)