Statistical correlation: Is a statistical technique used to determine if two variables are related. There is said to be a relationship between the two variables X and Y if the assignment of the modalities of X and Y does not occur randomly. That is to say:

- Y depending on X means that knowing the values of X allows predicting the values of Y.
- If Y depends on X: Y = f(X): Y is the dependent variable (to be explained), X is the independent variable (explanatory).

The correlation coefficient: Measures the strength of a relationship between two variables.

• **Properties of *r*(XY)**: This coefficient varies between -1 and +1. Its interpretation is as follows:

- If ***r*** is close to 0, there is no linear relationship between X and Y.
- If ***r*** is close to -1, there is a strong negative linear relationship between X and Y.
- If ***r*** is close to 1, there is a strong positive linear relationship between X and Y.

• The sign of ***r*** indicates the direction of the relationship.

• The absolute value of ***r*** indicates the strength of the relationship.

1. **The Pearson linear correlation coefficient** : allows for the detection of the presence or absence of a linear relationship between two quantitative variables.

$$r=\frac{n\sum xy-(\sum x)(\sum y)}{\sqrt{(n\sum x^2-(\sum x)^2)}\sqrt{(n\sum y^2-(\sum y)^2}}$$

Example 1 :
Let's consider the following table of grades for a set of students in mathematics and physics. Are mathematics and physics lineary related ?

Mathematics	11	11	12	13	13	14	15	16	17	18
Physics	10	12	11	13	14	15	16	17	18	19

X	Y	X*X	Y*Y	XY
11,00	10,00	121,00	100,00	110,00
11,00	12,00	121,00	144,00	132,00
12,00	11,00	144,00	121,00	132,00
13,00	13,00	169,00	169,00	169,00
13,00	14,00	169,00	196,00	182,00
14,00	15,00	196,00	225,00	210,00
15,00	16,00	225,00	256,00	240,00
16,00	17,00	256,00	289,00	272,00
17,00	18,00	289,00	324,00	306,00
18,00	19,00	324,00	361,00	342,00
∑X	∑Y	∑X*X	∑Y*Y	∑X*Y
140,00	145,00	2014,00	2185,00	2095,00

r*=(10*2095,00 -140,00 *145,00) /(Root (10*2014,00 -140,00 *140,00) *Root (10*2185,00 -145,00 *145,00))=*0,97384425

2. The Spearman's rank correlation coefficient

The Spearman's coefficient is based on the study of the rank difference between the attributes of individuals for both variables X and Y.

$$rs(X,Y) = 1 - \frac{6 \cdot \sum_{i=1}^{N}[r(X_i) - r(Y_i)]^2}{N^3 - N}$$

Example 2 : Are Result1 and Result2 lineary related

Result1	Result2
Doubled	Acceptable
Excellent	Good
Good	Very good
Acceptable	Doubled
Good	Acceptable
Very good	Excellent
Good	Acceptable
Acceptable	Doubled

Result1	Result2	Rang1		Rang2		Rang1-Rang2	
Doubled	Acceptable	1,00	1,00	5,00	4,00	-3,00	9,00
Excellent	Good	8,00	8,00	6,00	6,00	2,00	4,00
Good	Very good	4,00	5,00	7,00	7,00	-2,00	4,00
Acceptable	Doubled	2,00	2,50	2,00	1,50	1,00	1,00
Good	Acceptable	5,00	5,00	4,00	4,00	1,00	1,00
Very good	Excellent	7,00	7,00	8,00	8,00	-1,00	1,00
Good	Acceptable	6,00	5,00	3,00	4,00	1,00	1,00
Acceptable	Doubled	3,00	2,50	1,00	1,50	1,00	1,00

Rs(result1 ,result2)=1-(6*22)/(8*(8*8-1))

=0.74