Categorical VS Continuous variables:

Categorical variables : It is a fixed collection of values that is divided into groups.

Examples:
- Annual income: $30.000-$40.000
- Level of education: phD, MSc., BSc.
- Gender: Male, Female
- Payment status (cash or card)

Categorical variable can take on numerical values (such as “1” indicating Yes and “2” indicating No) but those numbers don’t have mathematical meaning.

Continuous variables: It is a not fixed values, that is often likely to change with time.

Examples:
- weight
- Age
- Tempreture

Calculating the correlation between two variables:

- calculating the correlation between two continuous variables, we can use (Pearson, Spearman, Kendall)

Spearman correlation : measures the strength and direction of between two ranked variables.

The Spearman Rank Correlation can take a value from +1 to -1

Example of Spearman:

The table below shows the score of 5 students in 2 different courses

Students	Math	science	d	d^2
student A	35 (3)	24 (5)	-2	4
student B	20 (5)	35 (4)	1	1
student C	49 (1)	39 (3)	-2	4
student D	44 (2)	48 (1)	1	1
student E	30 (4)	45 (2)	2	4

① Ranking the data in descending order

② $r_s = 1 - \frac{6 \sum d^2}{n(n^2-1)}$ $\Rightarrow n = \#$ of students $= \boxed{5}$

d^2 = the difference between the ranks (squared)

③ $= 4, 1, 4, 1, 4$

$\sum d^2 = 4 + 1 + 4 + 1 + 4$

$= \boxed{14}$

$r_s = 1 - \frac{6 \times 14}{5(5^2-1)}$

$= \boxed{0,3}$

معامل الارتباط الخطي: