Correlation

Need for correlation

- To find the association between the variables
- To find the degree of the association

What is correlation?

- The relationship between two variables is called their correlation.
- Positive Correlation: As one variable becomes large, the other also becomes large, and vice versa.
- Negative Correlation: As one variable becomes small, the other becomes large, and vice versa.

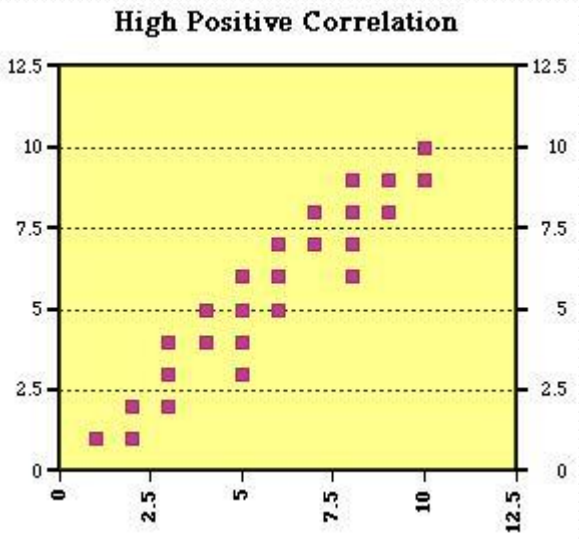
Other types of correlation

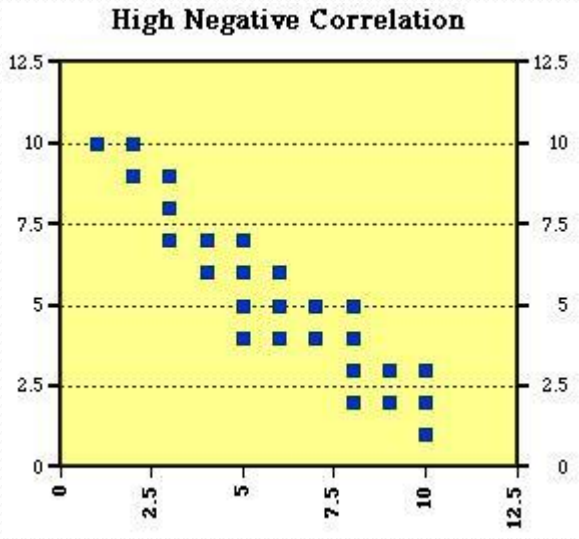
- Linear: Corresponding to a unit change in one variable, there is a constant change in the other variable.
- Non-Linear: Corresponding to a unit change in one variable, the other variable doesn't change at a constant rate but it changes at a fluctuating rate.

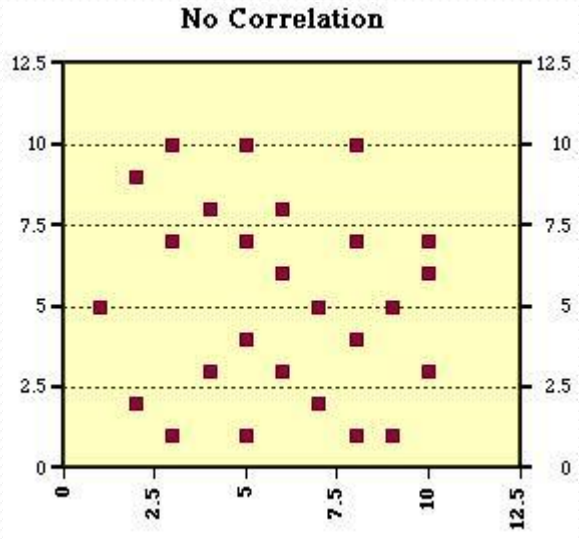
How to find the correlation?

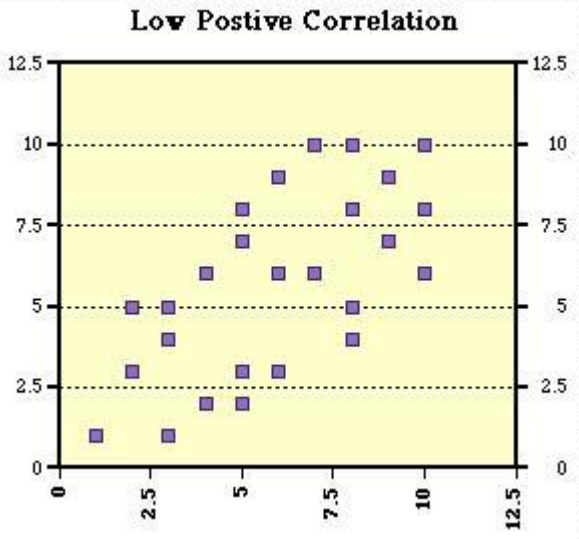
- Scatter plots show how much one variable is affected by another.
- Correlation Coefficients give the degree of correlation.

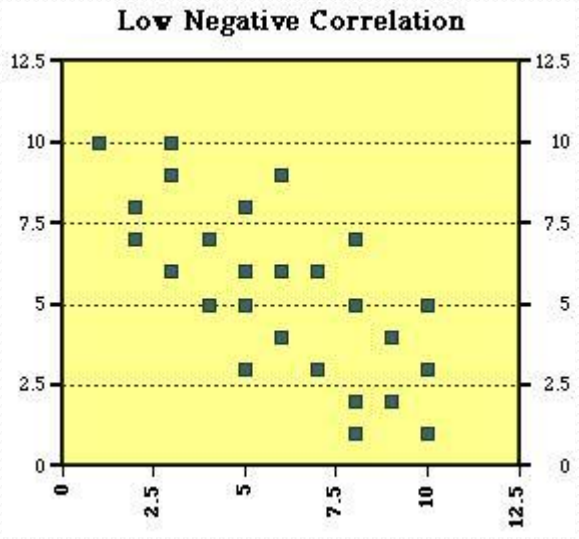
Example – Numerical Data (Two variables) Scatter Plot











Karl Pearson's Coefficient

- It is calculated a formula involving variance and covariance values.

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \qquad -1 \leq \rho \leq 1$$