

# CORRELATION ANALYSIS

---

# Major Points - Correlation

- Questions answered by correlation
- Types of correlation
- Scatterplots
- The correlation coefficient

# The Question

- Are two variables related?
  - Does one increase as the other increases?
    - e. g. skills and income
  - Does one decrease as the other increases?
    - e. g. health problems and nutrition
- How can we get a numerical measure of the degree of relationship?

# Correlation

- ❑ Correlation is a statistical tool that helps to measure and analyse the degree of relationship between two variables.
- ❑ Correlation analysis deals with the association between two or more variables.

# Definition

- The variables are said to be correlated if the changes in one variable results in a corresponding change in the other variable. That is, when two variables move together we say they are correlated.
- Boddington states that “ whenever some definite connection exists between the two or more groups, classes or series or data there is said to be correlation”.
- Bowely defines correlation as, “ when two quantities are so related that the fluctuations in one are in sympathy with the fluctuations of the other, that an increase or decrease of the one is found in connection with the increase or decrease of the other and greater the magnitude of change in the other, the quantities are said to be correlated”

# Types of Correlation

The different types of correlation are

- Positive and Negative correlation
- Linear and Non-linear correlation
- Simple, Multiple and Partial correlation.

# Positive Correlation

When the values of two variables move same direction, correlation is said to be positive

ie; an increase in the value of one variable results into an increase in the other variable also or if decrease in the value of one variable results into a decrease in the other variable also correlation is said to be positive.

Examples:

- The heights and weights of a group of persons
- The income and expenditure of a certain class of people
- The fertilizer used and the production of certain crop

# Negative Correlation

When the values of two variables move opposite direction, correlation is said to be negative.

ie; an increase in the value of one variable results into an decrease in the other variable also or if decrease in the value of one variable results into a increase in the other variable also correlation is said to be positive.

Example:

- The price and demand of a commodity, as the price of a product increases the demand for that product decreases.
- The price and supply of a commodity.



# Simple Correlation Coefficient

Simple correlation coefficient is a quantitative measure of the strength and direction of linear relationship between two numerically measured variables.

# Properties of Correlation Coefficient

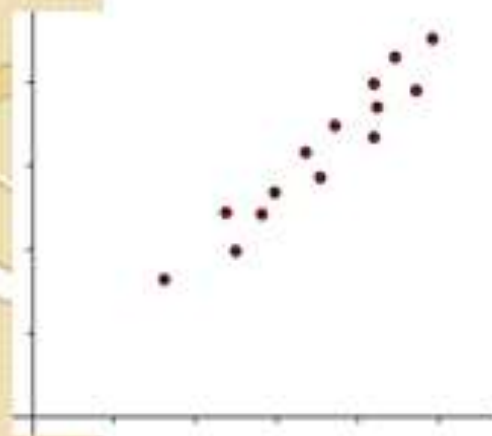
- Coefficient of correlation is a pure number lying between -1 and +1.
- When the correlation is negative, it lies between -1 and 0.
- When the correlation is positive, it lies between 0 and 1.
- When the correlation of coefficient is zero, it indicates that there is no correlation between the variables.
- When the correlation coefficient is 1, there is perfect correlation.
- The geometric mean of the two regression coefficients is equal to the correlation coefficient
- the correlation coefficient is independent of the origin and change of scale

# METHODS FOR STUDYING CORRELATION

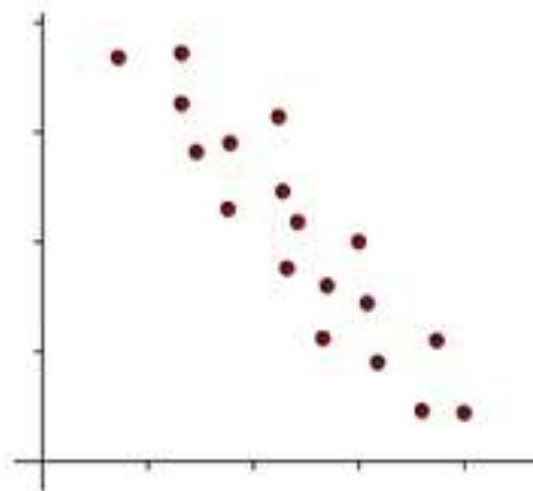
Correlation between two variables can be measured by both graphic and algebraic method. **Scatter diagram** and **correlation graph** are the two important graphic methods while **coefficient of correlation** is an algebraic method used for measuring correlation.

## a) Scatter diagram

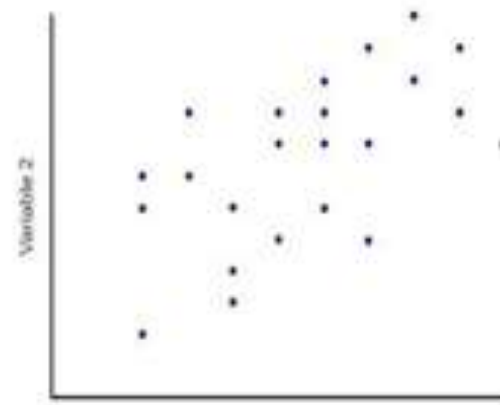
This is a graphical method of studying the correlation between two variables. One of the variable is shown on the X- axis and the other on the Y-axis. Each pair of values is plotted on the graph by means of a dot mark. After all the items are plotted we get as many dots on the graph paper as the number of points. If these points show some trend either upward or downward, the two variables are said to be correlated. If the point do not show any trend, the two variables are not correlated



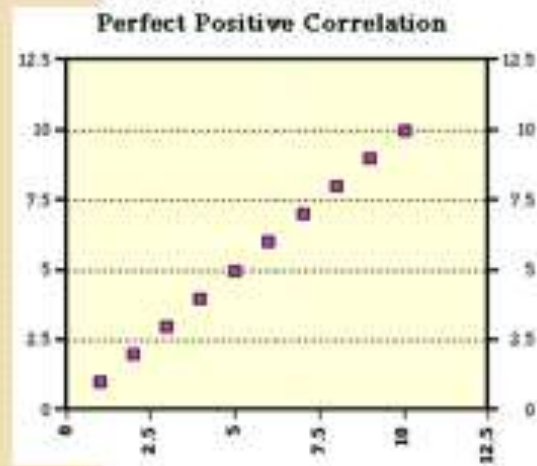
positive correlation



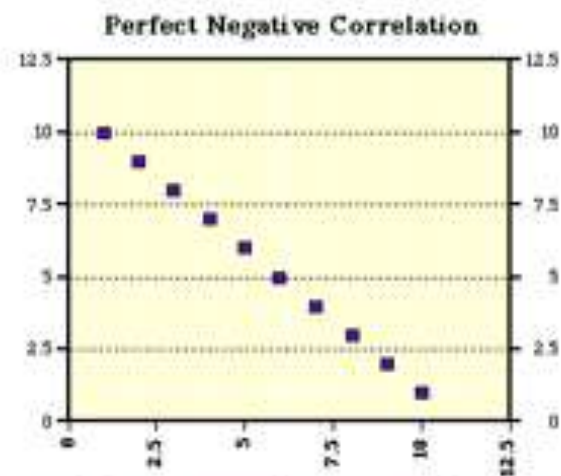
Negative correlation



zero correlation



perfect positive correlation



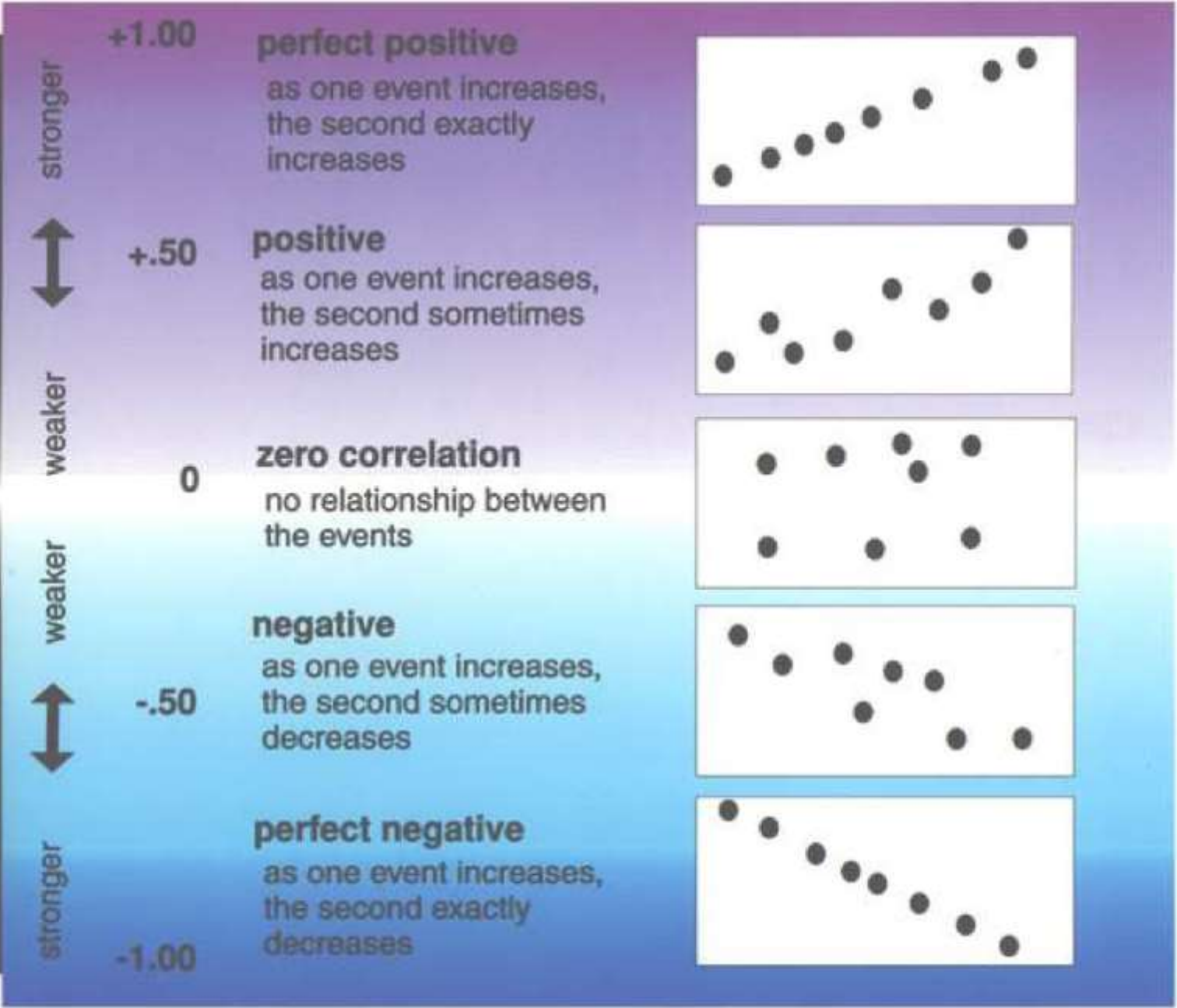
perfect negative correlation

Correlation

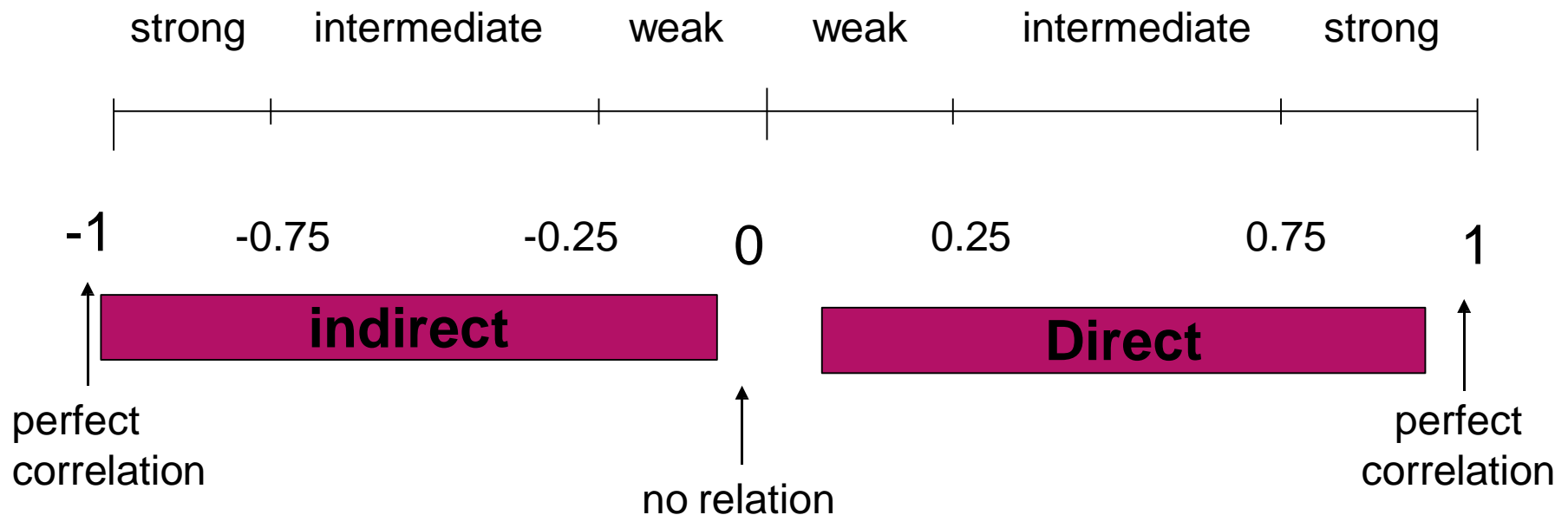
High positive correlation

Zero correlation

High negative correlation



- The value of  $r$  ranges between ( -1) and ( +1)
- The value of  $r$  denotes the strength of the association as illustrated by the following diagram.



- ➡ If  $r = \text{Zero}$  this means no association or correlation between the two variables.
- ➡ If  $0 < r < 0.25$  = weak correlation.
- ➡ If  $0.25 \leq r < 0.75$  = intermediate correlation.
- ➡ If  $0.75 \leq r < 1$  = strong correlation.
- ➡ If  $r = 1$  = perfect correlation.