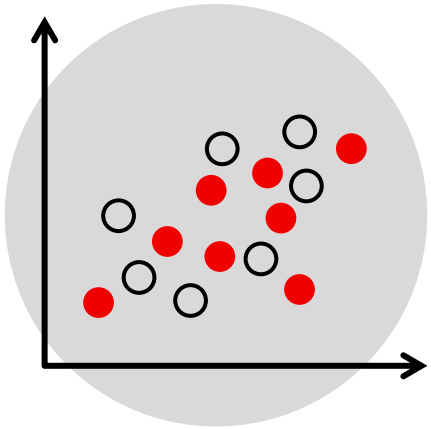
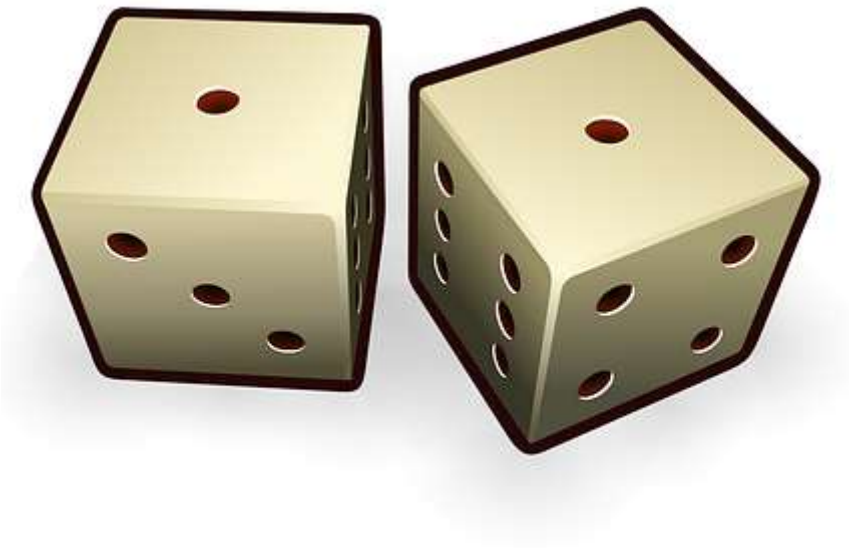
Continuous Improvement Toolkit

SCATTER DIAGRAM



SCATTER DIAGRAM

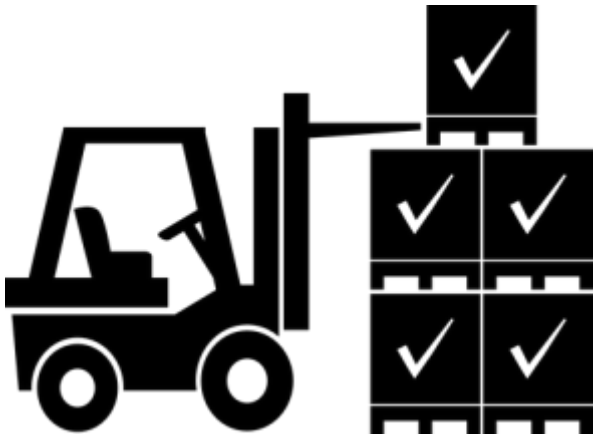
Many situations require the investigating whether a **relationship** exists between two or more variables.



SCATTER DIAGRAM

A line manager may want to check the relationship between the number of **training hours** and **productivity of employees**.

He may then want to check if the **number of defects** is a function of the **experience** of the person causing it.



SCATTER DIAGRAM

Other Examples

The relationship between **equipment downtime** and its **cost of maintenance**.

SCATTER DIAGRAM

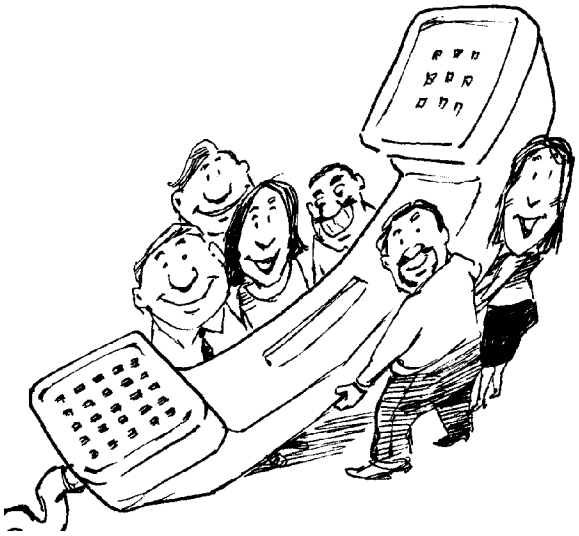
Other Examples

The relationship between **driving speed** and **fuel consumption**.

SCATTER DIAGRAM

Other Examples

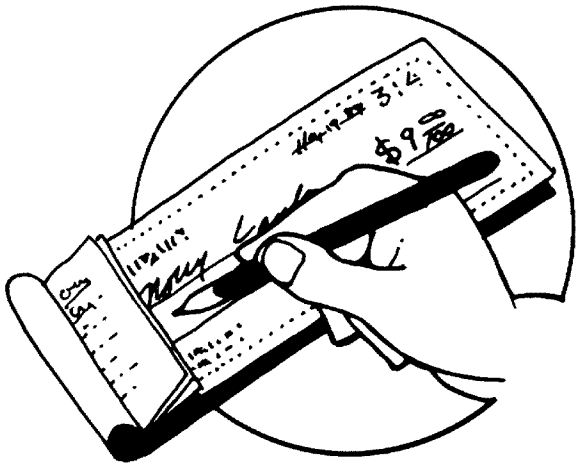
The relationship between the **number of people working on a shift** and the **average answer time** in a call center.



SCATTER DIAGRAM

Other Examples

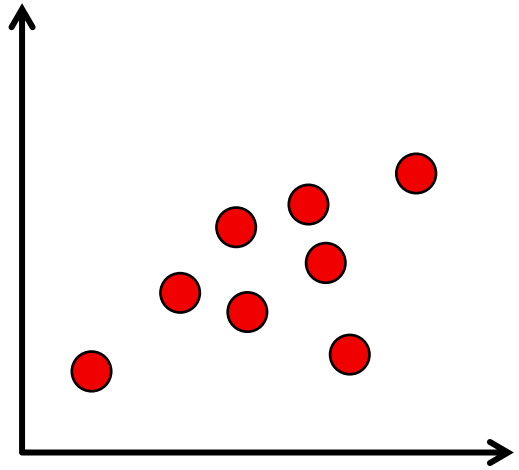
The relationship between the **number of years of education** someone has and the **annual income** of that person.



SCATTER DIAGRAM

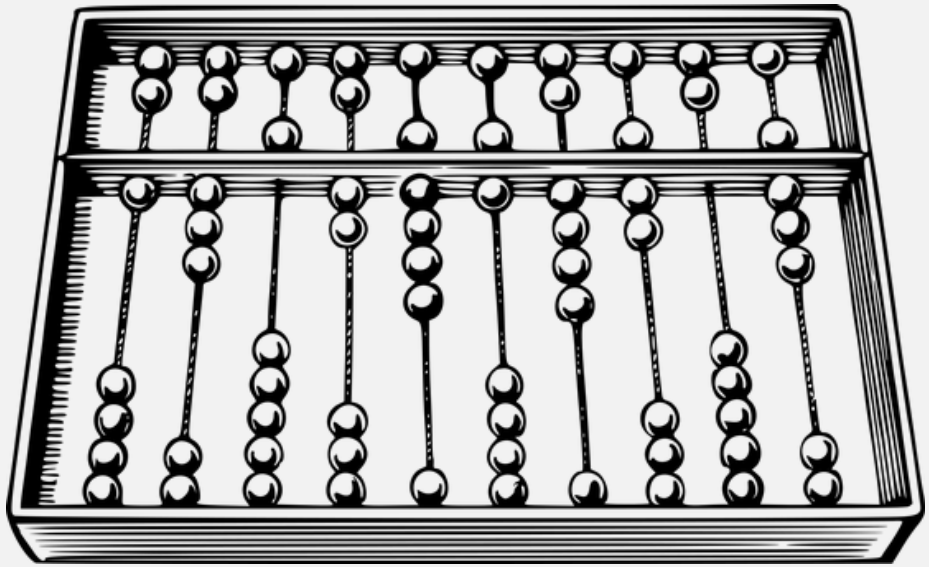
A **scatter diagram** is a diagram that shows whether two variables are correlated or related to each other.

It shows patterns in the relationship that cannot be seen by just looking at the data.



SCATTER DIAGRAM

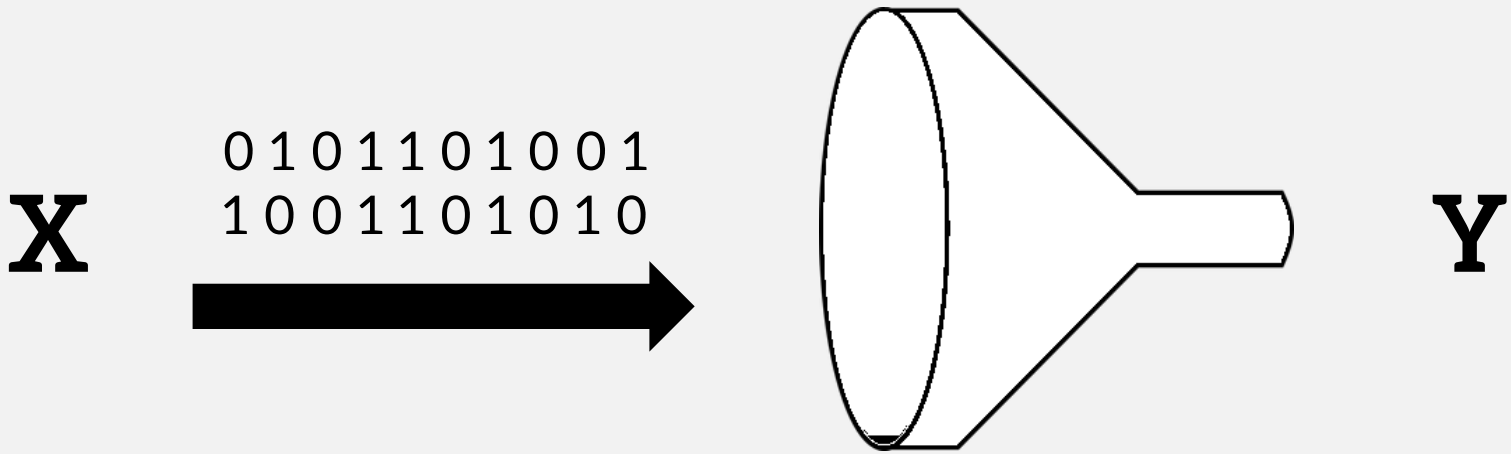
It works with both **continuous** and **count** data.



SCATTER DIAGRAM

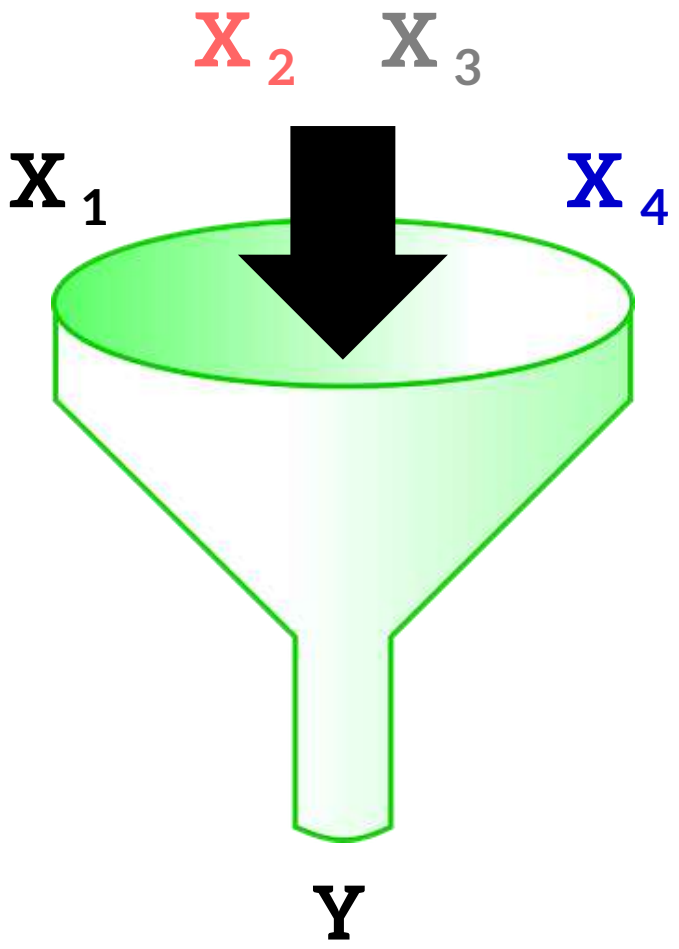
Primarily used to visually investigate the relationship between **two variables** and determine the strength of the relationship.

Often an **output** and an **input** variables



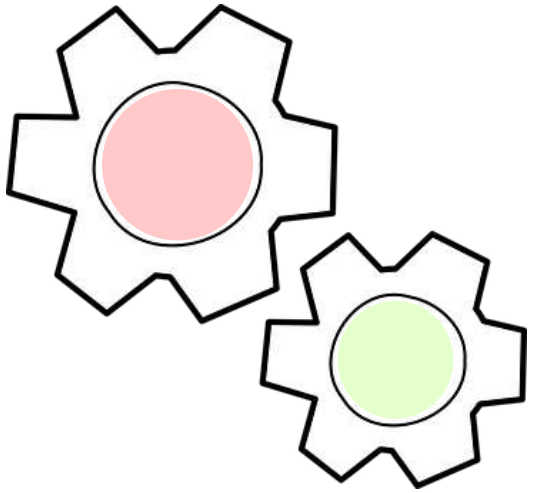
SCATTER DIAGRAM

- This is useful to verify that any change in the input variable will influence the output variable.
- It helps **detecting** the primary factors that are really causing a problem and hence eliminating non-critical factors from consideration.



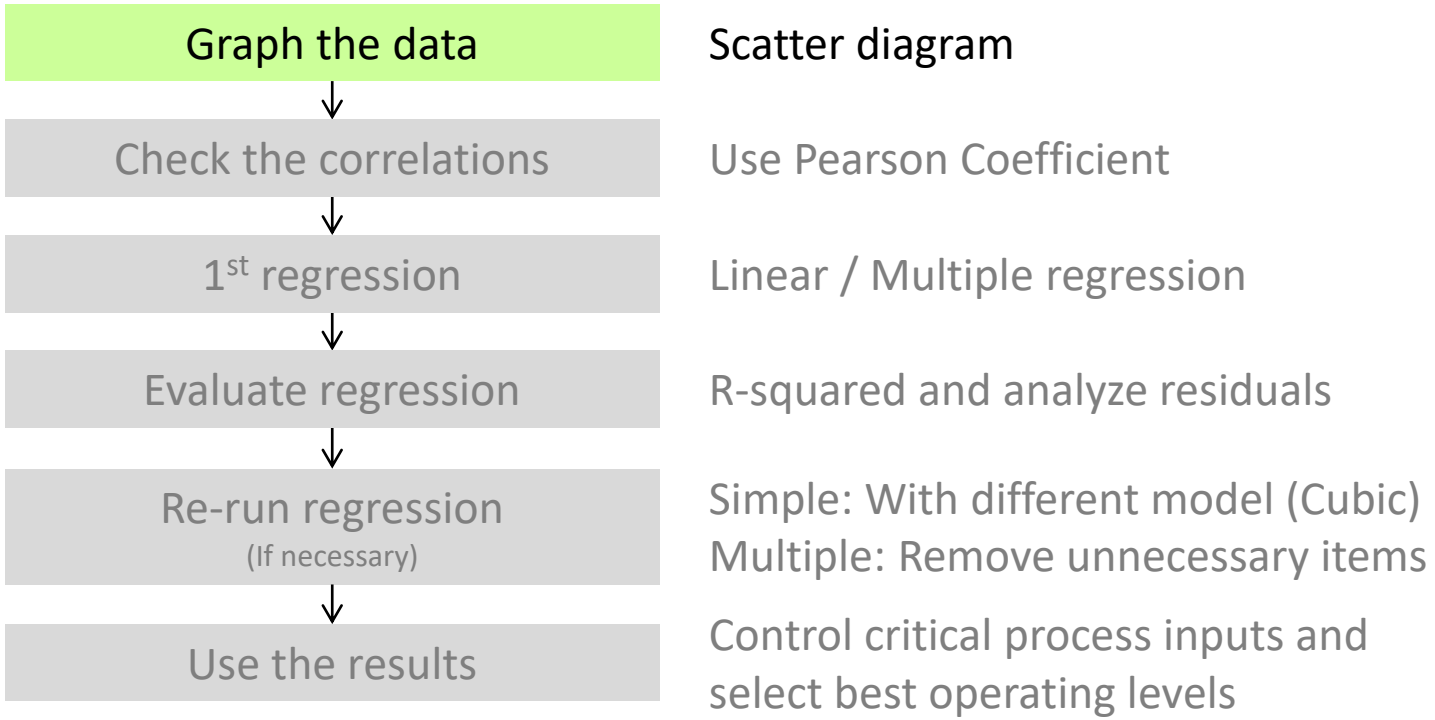
SCATTER DIAGRAM

- Often used as a **first step** when analyzing the correlation between pairs of variables and before conducting advanced statistical techniques.
- Often used with advanced **statistical tools** (e.g., regression) to support or reject hypotheses about the data.

X_3  Y

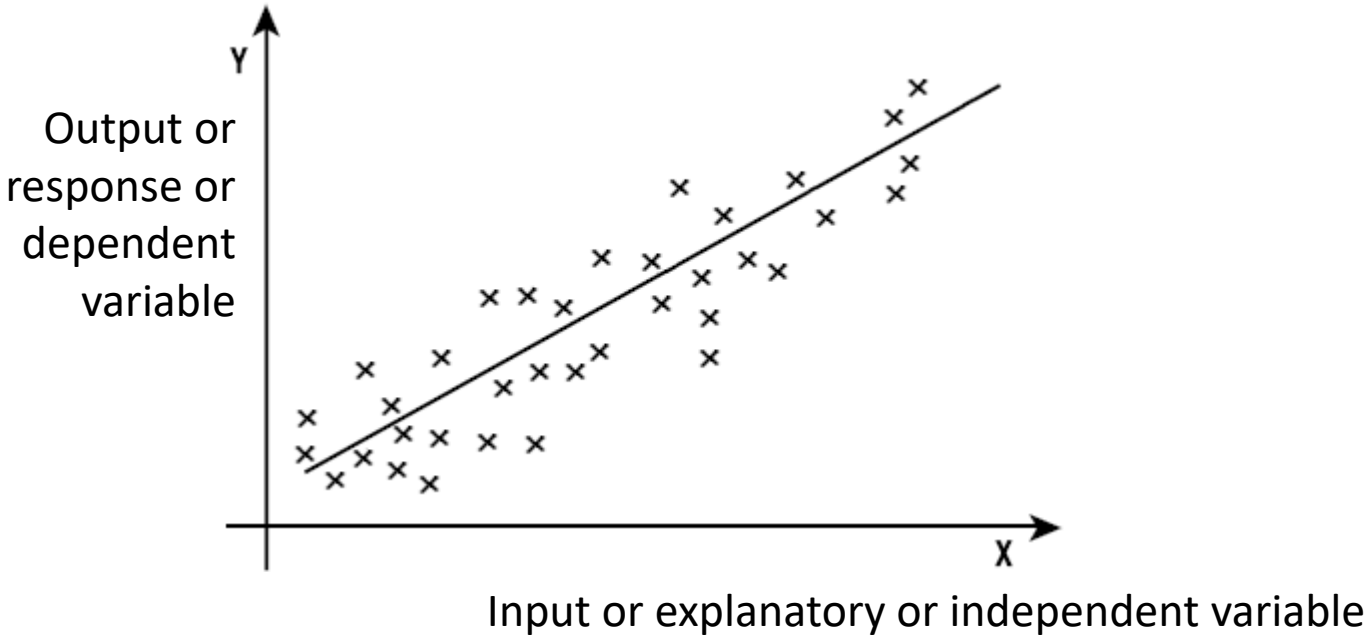
SCATTER DIAGRAM

Where do scatter diagrams fit?



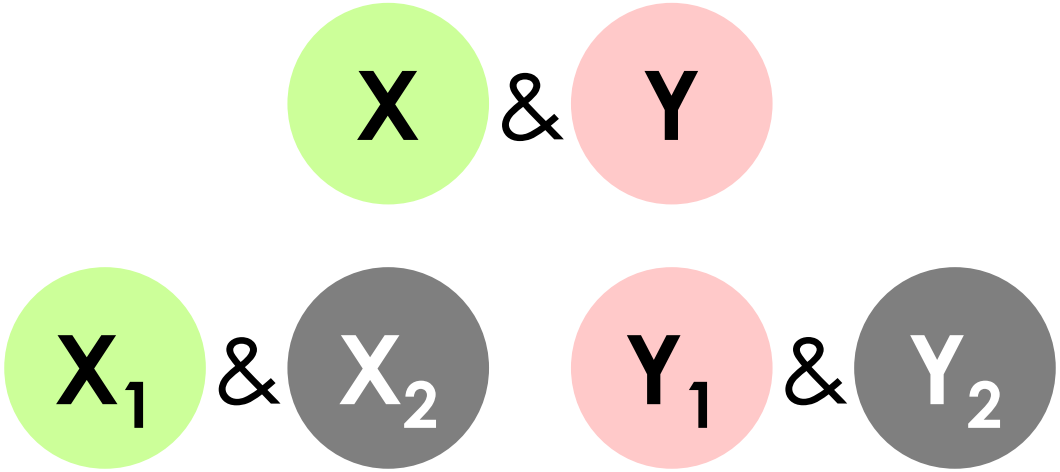
SCATTER DIAGRAM

The input variable is normally placed on the **horizontal axis** while the output variable is placed on the **vertical axis**.



SCATTER DIAGRAM

You may also study the relationship between **two input or output variables**.

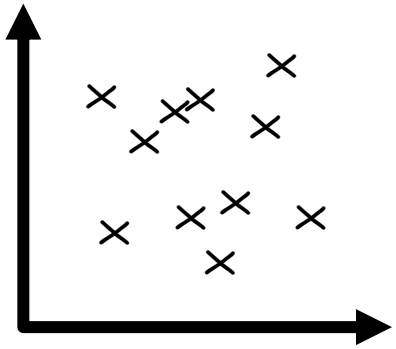


In this case, it doesn't matter which variable goes on the horizontal axis and which goes on the vertical axis

SCATTER DIAGRAM

Scatter diagrams can indicate several types of correlation.

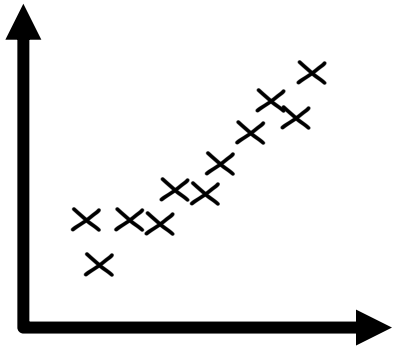
No correlation when the data points are scattered randomly without showing any particular pattern.



SCATTER DIAGRAM

Scatter diagrams can indicate several types of correlation.

A **positive correlation** occurs when the values of one variable increase as the values of the other variable increase.

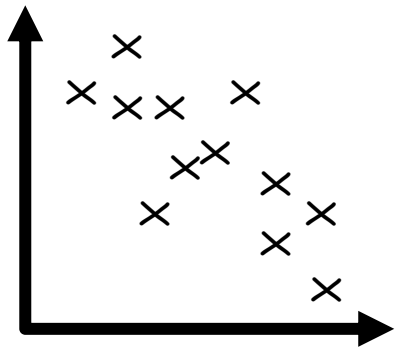


The fitted line slopes from bottom left to top right

SCATTER DIAGRAM

Scatter diagrams can indicate several types of correlation.

A **negative correlation** occurs when the values of one variable increase as the values of the other variable decrease.

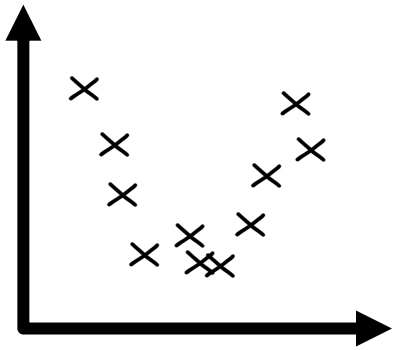


The fitted line slopes from upper left to lower right

SCATTER DIAGRAM

Scatter diagrams can indicate several types of correlation.

Scatter diagrams can also indicate **nonlinear** relationships between variables.



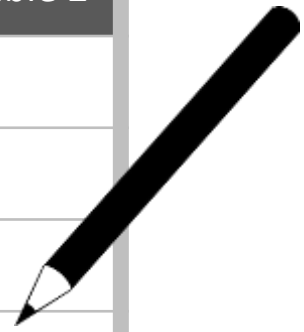
SCATTER DIAGRAM

How to Construct a Scatter Diagram?

Collect the two paired sets of data

Create a summary table of the data

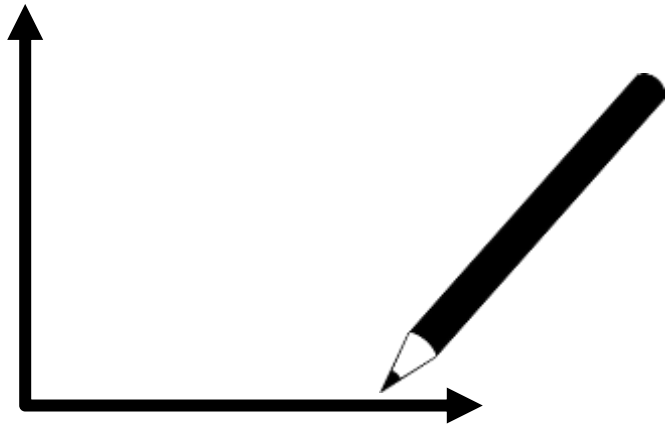
Variable 1	Variable 2



SCATTER DIAGRAM

How to Construct a Scatter Diagram?

Draw and **label** the horizontal and vertical axes with variable names and scale values

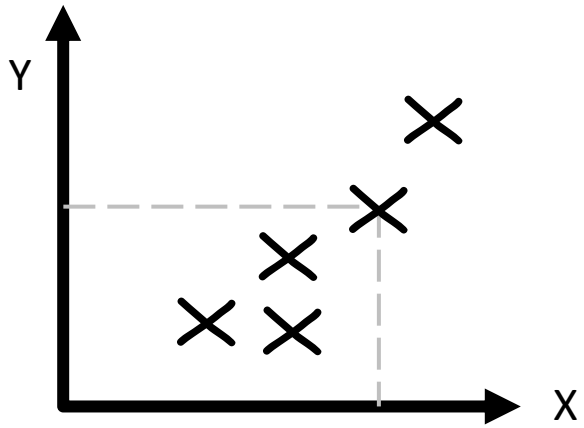


SCATTER DIAGRAM

How to Construct a Scatter Diagram?

Plot the data pairs on the diagram by placing a dot at the **intersection** of each data pair

Look at how the pattern appears and how the two variables vary together

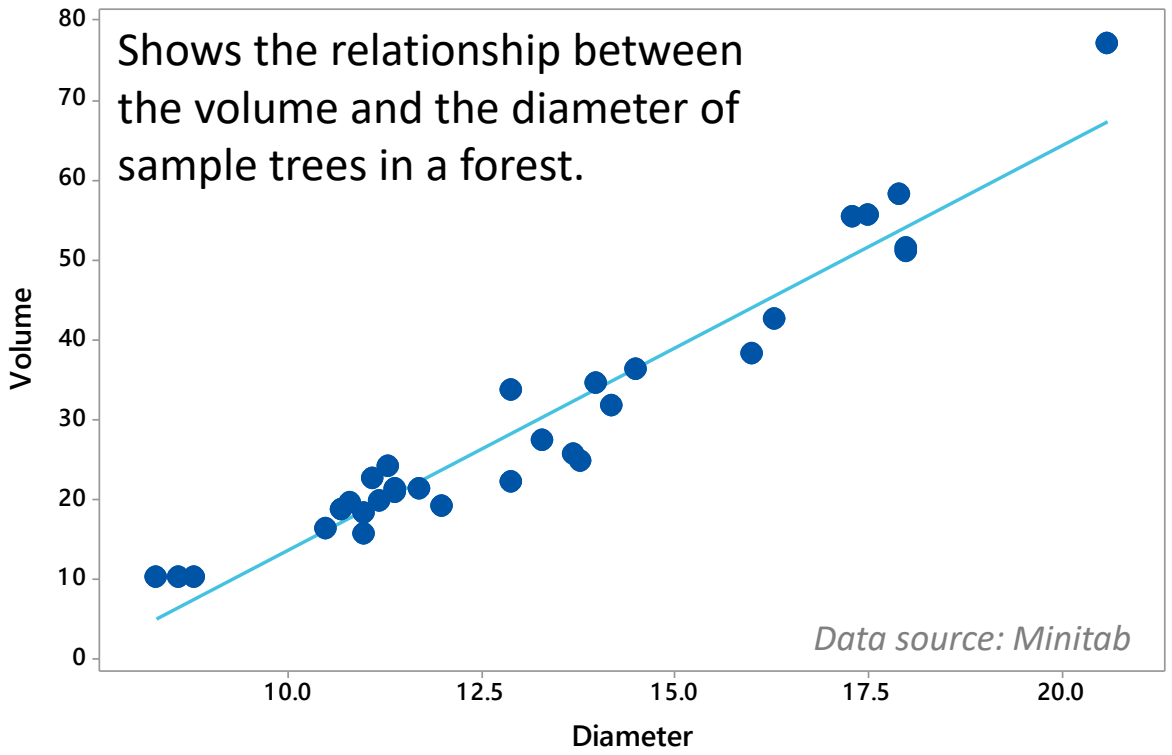


SCATTER DIAGRAM

Example – Forest Trees

Diameter	Height	Volume
8.3	70	10.3
8.6	65	10.3
8.8	63	10.2
10.5	72	16.4
10.7	81	18.8
10.8	83	19.7
11	66	15.6
11	75	18.2
11.1	80	22.6
11.2	75	19.9
11.3	79	24.2
11.4	76	21

Both are output variables

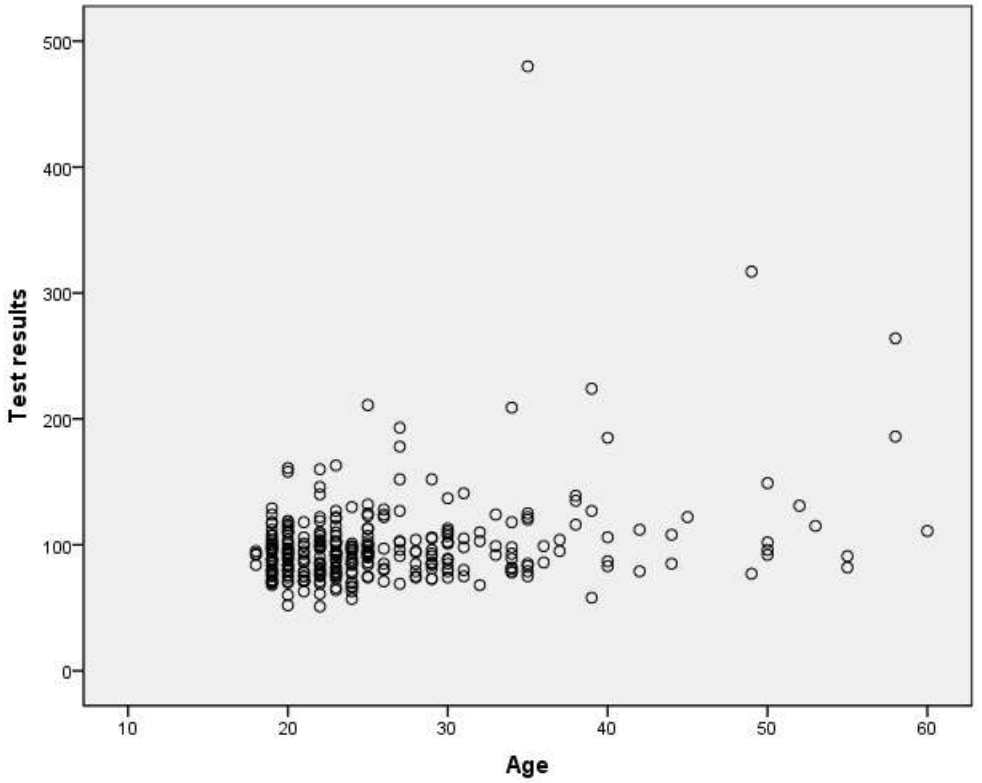


SCATTER DIAGRAM

Example – Presence of Diabetes at a Workplace

An analysis that was conducted for diagnosing the presence of diabetes at a workplace.

The population was generally **young** (75.8% were below thirty).

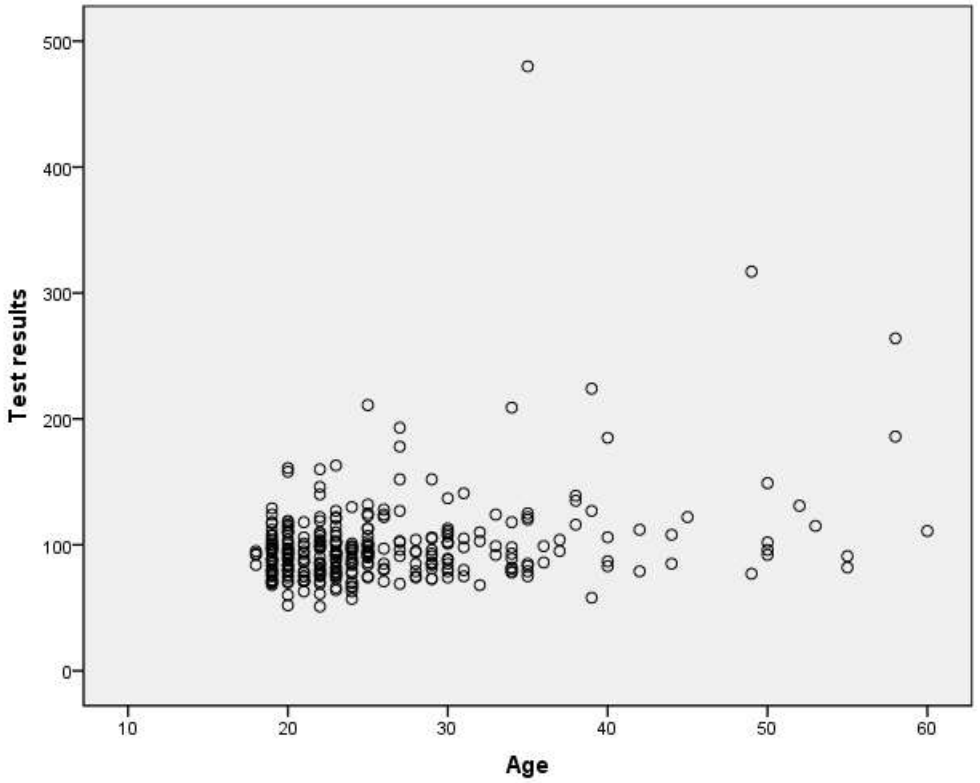


SCATTER DIAGRAM

Example – Presence of Diabetes at a Workplace

This scatter diagram illustrates that there is no obvious relationship between age and glucose levels.

High glucose levels are found in all ages, and normal glucose levels are found in higher ages.

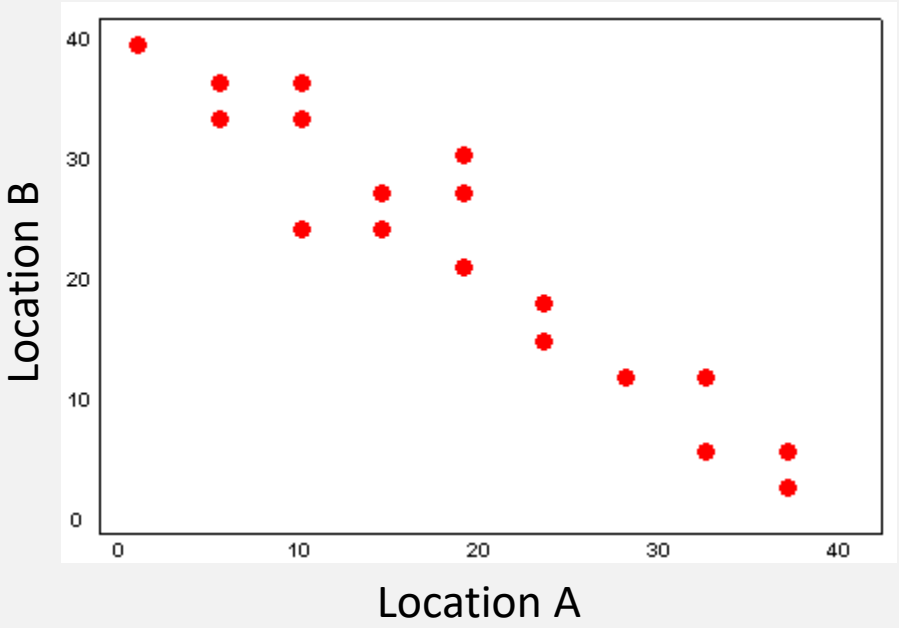


SCATTER DIAGRAM

Example – The number of sales per month generated at two locations.

The plotted points form a negative slope.

The sales at location “B” is inversely related to the sales at location “A”.

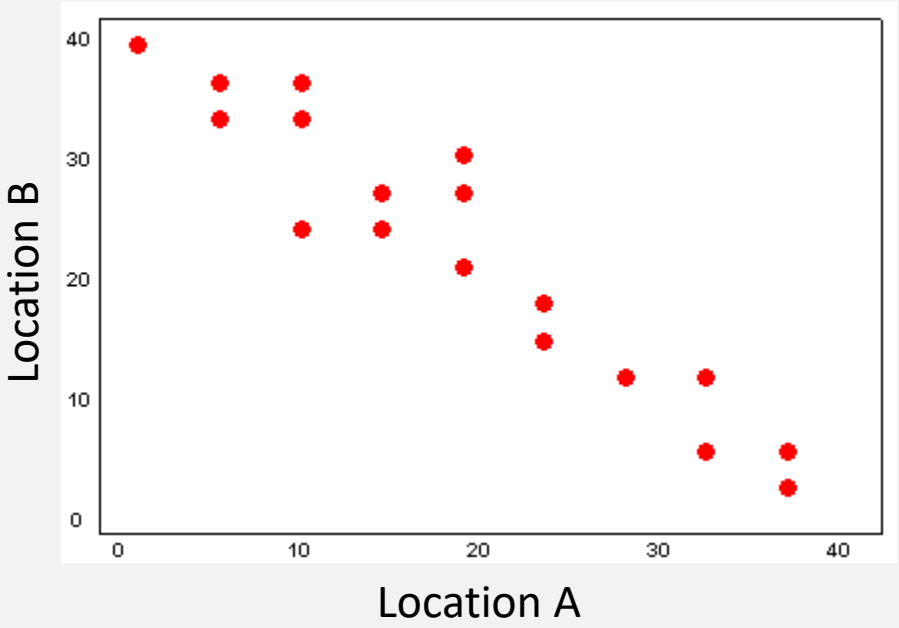


SCATTER DIAGRAM

Example – The number of sales per month generated at two locations.

Does it mean that location "A" caused the decrease in sales at location "B", or vice versa?

Answer: Not necessarily, unless the two locations are direct competitors.



SCATTER DIAGRAM

Matrix Plot

Summarizes the relationship between **pairs of multiple variables** in one graph.

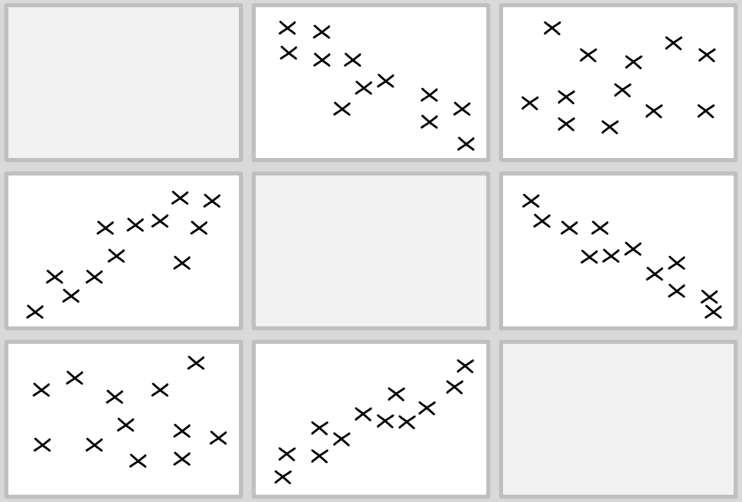
3 variables

Allows to visually assess the variables that might be related

SCATTER DIAGRAM

Matrix Plot

Produces a scatter diagram for **every combination** of variables**.**

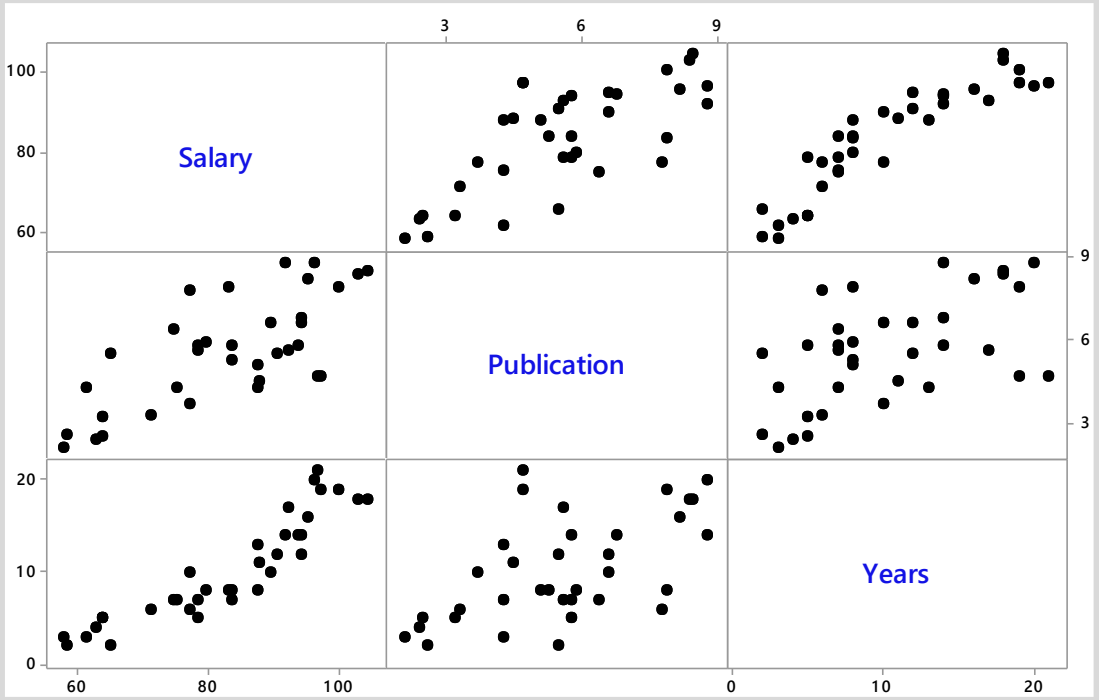


Potential correlations between pairs of variables can then be identified

SCATTER DIAGRAM

Matrix Plot - **Example**

Salary	Gender	Publication	Years
88	F	4.5	11
77.3	M	7.8	6
75.3	M	4.3	7
96.4	M	8.8	20
87.7	M	5.1	8
58.1	F	2.1	3
63.1	F	2.4	4
58.5	M	2.6	2
95.4	F	8.2	16
92	F	8.8	14
94.5	M	6.6	12
103	M	8.4	18

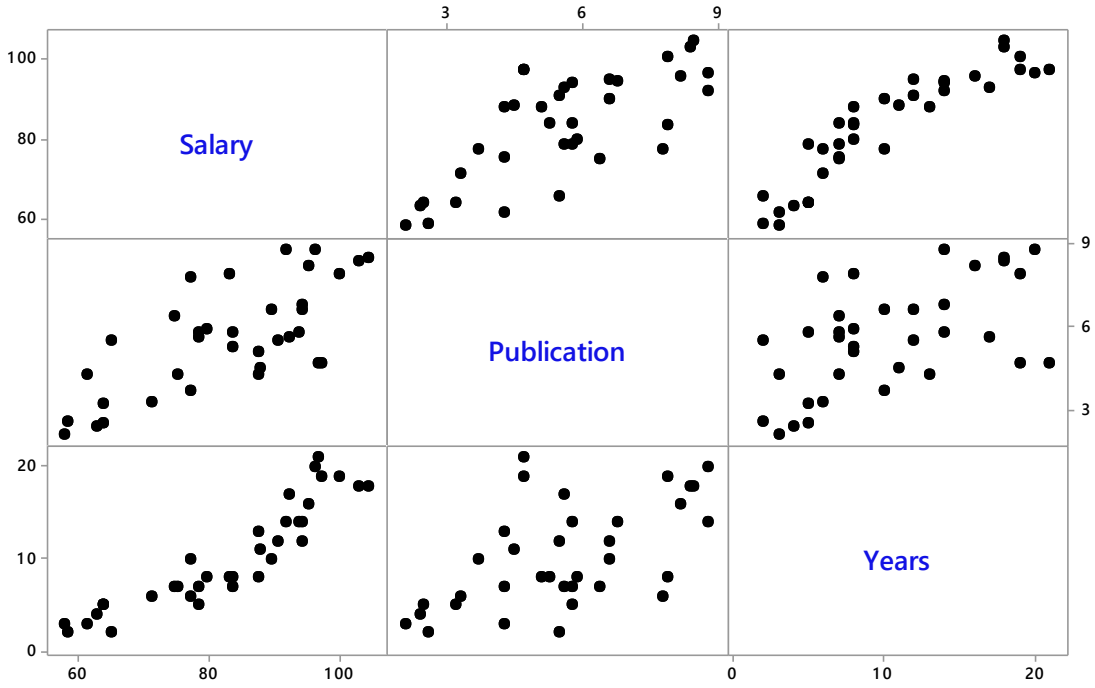


# of publications doesn't appear to be correlated with the years of experience

There may be a relationship between the years of experience and salaries

SCATTER DIAGRAM

Matrix Plot - Example



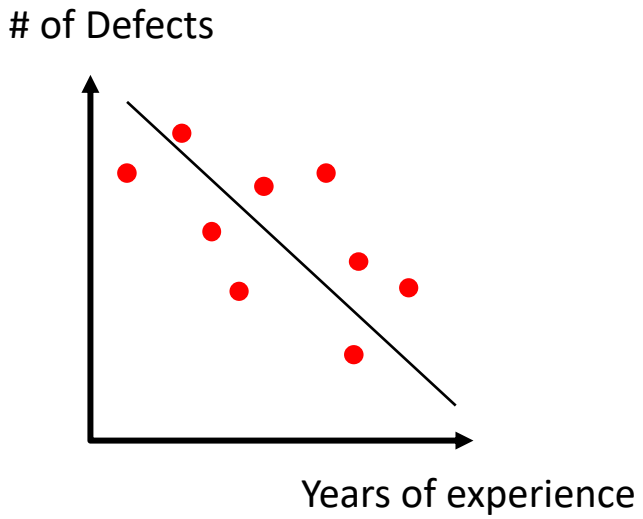
Is there a correlation between the **number of publication** & **salaries?**

SCATTER DIAGRAM

Further Information

When the relationship is not so clear, **Correlation** can be used to help validate if a relationship exists between the variables.

Regression techniques go a step further by defining the relationship in a mathematical format.



SCATTER DIAGRAM

Further Information

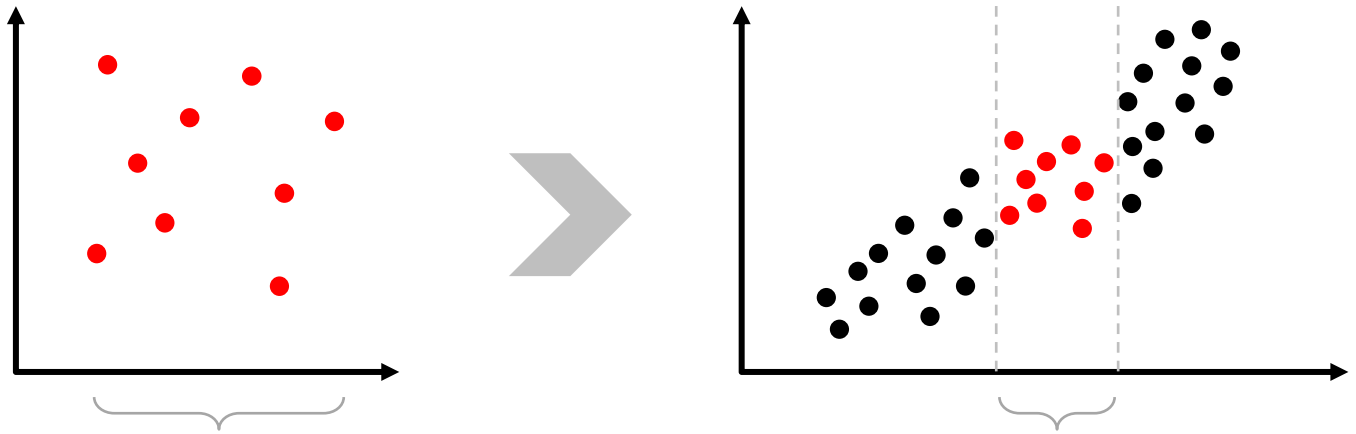
- Be careful before concluding that there is a direct cause-and-effect relationship between the variables.
- There might be a **third factor** that is causing the change in the two variables.

$$Y = f(x)$$

SCATTER DIAGRAM

Further Information

No correlation on the other hand does not mean there is **no cause-and-effect relationship**. There might be a relationship over a wider range of data or a different portion of the range.

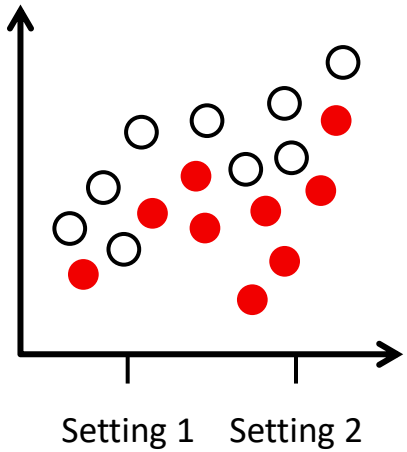


SCATTER DIAGRAM

Further Information

You can also illustrate a stratification factor in scatter diagrams.

For example, the relationship between a process output and a process input for two different settings.



SCATTER DIAGRAM

One of the 7 Basic Tools of Quality

SCATTER DIAGRAM

Made with ❤ by

CITOOLKIT

The Continuous Improvement Toolkit

www.citoolkit.com