

Frequently Asked Questions

1. What is a scatter diagram?

Answer:

It is a graphical technique used to analyze the relationship between two variables. Two sets of data are plotted on a graph with the Y-axis being used for the variable to be predicted and X axis being used for the variable to make prediction.

2. What are the different reasons that a scatter diagram shows correlation between two variables?

Answer:

A Scatter Diagram shows correlation between two items for three reasons:

- There is a cause and effect relationship between the two measured items, where one is causing the other (at least in part).
- The two measured items are caused by a third item. For example, a Scatter Diagram that shows a correlation between cracks and transparency of glass utensils, where changes in both are caused by changes in furnace temperature.
- Complete coincidence. It is possible to find high correlation of unrelated items, such as the number of ants crossing a path and newspaper sales.

Scatter Diagrams may thus be used to give evidence for a cause and effect relationship.

3. What are the factors that affect the cause-effect relationship when the data is represented in a scatter diagram?

Answer:

- **The proximity of the cause and effect:** There is better chance of a high correlation if the cause is directly connected to the effect than if it is at the end of a chain of causes. Thus, a root cause may not have a clear relationship with the end effect.
- **Multiple causes of the effect:** When measuring one cause, other causes are making the effect vary in an unrelated way. Other causes may also be having a greater effect, swamping the actual effect of the cause in question.
- **Natural variation in the system:** The effect may not react in the same way each time, even to a close major cause.

4. What do you mean by correlation?

Answer:

The term correlation or co-variation indicates the relationship between two such variables in which with changes in the values of one variable, the values of the other variable also changes.

5. Give any two definitions of correlation.

Answer:

“If two or more quantities vary in sympathy so that movements in the one tend to be accompanied by corresponding movements in the other then they are said to be correlated.”-L.R. Connor

“When the relationship of a quantitative nature, the appropriate statistical tool for discovering and measuring the relationship and expressing it in a brief formula is known as correlation.” –Croxtton & Cowden

6. What are the uses of correlation?

Answer:

- The study of correlation reduces the range of uncertainty associated with decision-making.
- Correlation analysis is very helpful in understanding economic behavior.
- Correlation study helps us to estimate the likely change in a variable with a particular amount of change in related variable.
- Inter-relationship studies between different variables are very helpful tools in promoting research and opening new frontiers of knowledge.

7. What are different types of correlation?

Answer:

Different types of correlation are:

- Positive or Negative Correlation
- Simple, Multiple or Partial Correlation
- Linear or Non-linear Correlation

8. Explain Positive or Negative Correlation.

Answer:

Correlation can be, positive or negative. When values of the variables move in the same direction, that is, when an increase in the value of one variable is associated with an increase in the value of the other variable or a decrease in the value of one variable is associated with the decrease in the value of the other variable, correlation is to be positive.

On the other hand if the values of two variables move in opposite directions, so that with an increase in the values of one variable the value of the other variable decrease, or with a decrease in the values of one variable the values of the other variable increases, correlation is said to be negative.

9. Explain Simple, multiple or partial correlation.

Answer:

Correlation can be, simple, multiple or partial.

In simple correlation, we study only two variables – say price and demand.

In multiple correlation, we study together the relationship between three or more factors like production rainfall and use of fertilizers.

In partial correlation even though more than two factors involved but correlation is studied only between two factors and the other factors are assumed to be constant.

The correlation between two variables is said to non-linear or curvilinear if corresponding to a unit change in the value of one variable the other variable does not change at a constant rate but at a fluctuating rate.

10. Explain Linear and non-linear correlation.

Answer:

The correlation between two variables is said to be linear if corresponding to a unit change in the value of one variable there is a constant change in the value of the other variable. That is in case of linear correlation the relation between the variable x and y is of the type $y=a+bx$

11. Explain the four parameters of Linearity.

Answer:

Linearity has four parameters:

1. **Correlation** - Measures how well the data line up. The more the data resembles a straight line, the higher the correlation to each other.
2. **Slope** - Measures the steepness of the data. The steeper the data slope, assuming the correlation is good, the greater the importance of the relationship. A change in the "X" variable will have a larger impact on the "Y" variable, and you will begin to see a pattern that represents the Moderate Correlation diagram above.
3. **Direction** - The "X" variable can have a positive or a negative impact on the "Y" variable. As one factor goes up, the other goes down. In a positive correlation, both factors will move in the same direction. In the graph examples below, you can see that the positive correlation moves from the lower left, toward the upward right. The negative correlation moves from the lower right, toward the upward left.
4. **Y Intercept** – It is a line drawn through the data crosses the "Y" axis. For a positive correlation, it represents the minimum "Y" value; for a negative correlation, it presents the maximum "Y" value.

12. Give some examples for positive correlation.

Answer:

- Age of husband and age of wife
- Demand of the commodity and its supply
- Increase in rainfall (up to a point) and production of rice
- Increase in heat and temperature
- Increase in advertisement cost and volume of sales
- Increase in height and weight

13. Write merits and demerits of scatter diagram.

Answer:

Following are the merits of the scatter diagram:

- It is very to draw a scatter diagram
- It can be easily understood and interpreted as we can easily identify the whether the scattering of the points are linear or non-linear.
- Values of extreme items do not affect this method. Such points are always isolated in the diagram

Following gives the demerits of the scatter diagram:

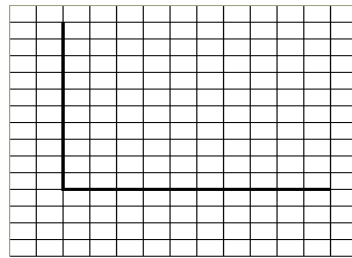
- The major limitation of the method is that it only gives a visual picture of the relationship of two variables. It only tells us whether there is correlation between the variables, and if so, then in which direction, positive or negative.
- It does not give an idea about the precise degree of relationship, as it is not amenable to mathematical treatment.

14. Explain the method of constructing a scatter diagram.

Answer:

The Basic Scatter Diagram Layout:

Step 1 - Draw an "L" form just like you did for the pareto diagram (see the below figure). Make your scale units at even multiples, such as 10, 20, etc. so as to have an even scale system.



Step 2 - On the Horizontal axis (Known as the "X" axis, from Left to Right) you place the Independent or "cause" variable.

Step 3 - On the Vertical axis (Known as the "Y" axis, from Bottom to Top) you place the Dependent or "effect" variable.

Step 4 - Plot your data points at the intersection of your data plots of the X and Y values. For Example = X = 5, Y = 2. Go right 5 spaces, and then go up 2 spaces to plot the point.

15. Estimate the coefficient of correlation for the following data regarding ages of husbands and wives with the help of scatter diagram.

| | | | | | | | | | | |
|-----------------------|----|----|----|----|----|----|----|----|----|----|
| Age of Husband | 23 | 27 | 28 | 28 | 29 | 30 | 31 | 33 | 35 | 36 |
| Age of Wife | 18 | 20 | 20 | 27 | 21 | 29 | 27 | 29 | 28 | 29 |

Answer:

Estimating the coefficient of correlation with the help of a scatter diagram is very easy because no mathematical calculations are required. The major shortcoming of this method is that degree of correlation cannot be determined.

Observe that there is high and positive correlation between the ages of husbands and the ages of wives.