13. Presentation of bivariate data through scatter-plot diagrams and calculations of covariance.

Bivariate Data/ Bivariate Analysis

Bivariate analysis is one of the statistical analysis where two variables are observed. One variable here is dependent while the other is independent. These variables are usually denoted by X and Y. So, here we analyse the changes occurred between the two variables and to what extent.

The term bivariate analysis refers to as the analysis of two variables . the objective of bivariate analysis to understand the relationship between two variables. There are three common way to analysis the bivariate analysis –

1. Scatter plots

2. Correlation Coefficient

3. Simple linear Regression(SLR)

Bivariate frequency distribution

A series of statistical data showing the frequency of two variables simultaneously is called Bivariate frequency distribution. In other words, the frequency distribution of two variable is called Bivariate frequency distribution. For example: sales and advertisement expenditure , weight and height of an individual.

Why bivariate frequency distribution is significant in business research ?

1. Decision Making

2. Market-segmentation

3. Risk-assessment

4. Resource allocation

how to implement in excel

= COVARIANCE.P(array1,array2)

The COVARIANCE.P function used the following arguments array1, this is range or array of integer value. array2 is also the second range or values.

Few things to remember about argument

1. If the given array contain text or logical value then are ignore by the Covariance function in excel.

2. The data should contain numbers, names, array or references that are numeric .IF the some cell do not contain numeric data they are ignored.

3. The data set should be same size with the same number of data points.

4. The data set should not be empty nor should the standard Deviation of the value equal .

$$Cov(X,Y) = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{n}$$

X and Y are the sample mean of the two set of values and n is the sample size.

5. Covariance is measure to indicate the extent to which two random variable in tandem.

6. Correlation is the measure used to represent how strongly two random variable are strongly related to each other.

7. Covariance is nothing but a measure of correlation.

8. Correlation referred to the scaled form of covariance.

9. Covariance can ary between $-\infty$ to $+\infty$ and correlation range between -1 to +1 .

10. Covariance indicate the direction of the linear relationship between variables .

11. Correlation on the other hand measure both the strength and direction of the linear relationship between two variables.

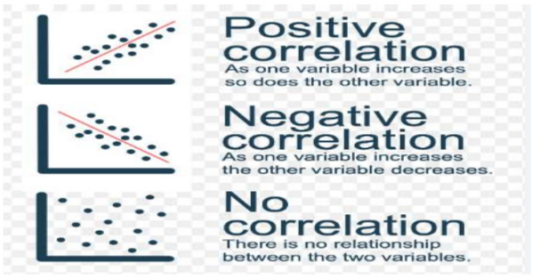
12. Covariance is affected by change in scale.

13. Correlation is not affected by the change in scale.

Pearson Correlation Coefficient formula

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

r = correlation coefficient
x_i = values of the x-variable in a sample
$\bar{x}$ = mean of the values of the x-variable
y_i = values of the y-variable in a sample
$\bar{y}$ = mean of the values of the y-variable



= PEARSON(arry:arry2)

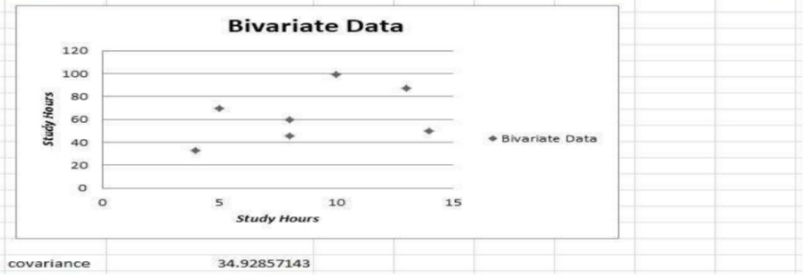
Scatter plots

Scatter plots are the graphs that present the relationship between two variables in a data-set. It represents data points on a two-dimensional plane or on a Cartesian system. The independent variable or attribute is plotted on the X-axis, while the dependent variable is plotted on the Y-axis. These plots are often called scatter graphs or scatter diagrams.

Scatter plots instantly report a large volume of data. It is beneficial in the following situations –

- For a large set of data points given
- Each set comprises a pair of values
- The given data is in numeric form

Study Hours	10	13	14	8	5	4	8
Marks	99	87	50	60	70	33	46



covariance	34.92857143

Click insert tab along the top ribbon then click scatter chart within chart group.

CORRELATION in excel and COVARIANCE in excel –

= CORREL(hours,score)

= COVARIANCE.P(hours,score)