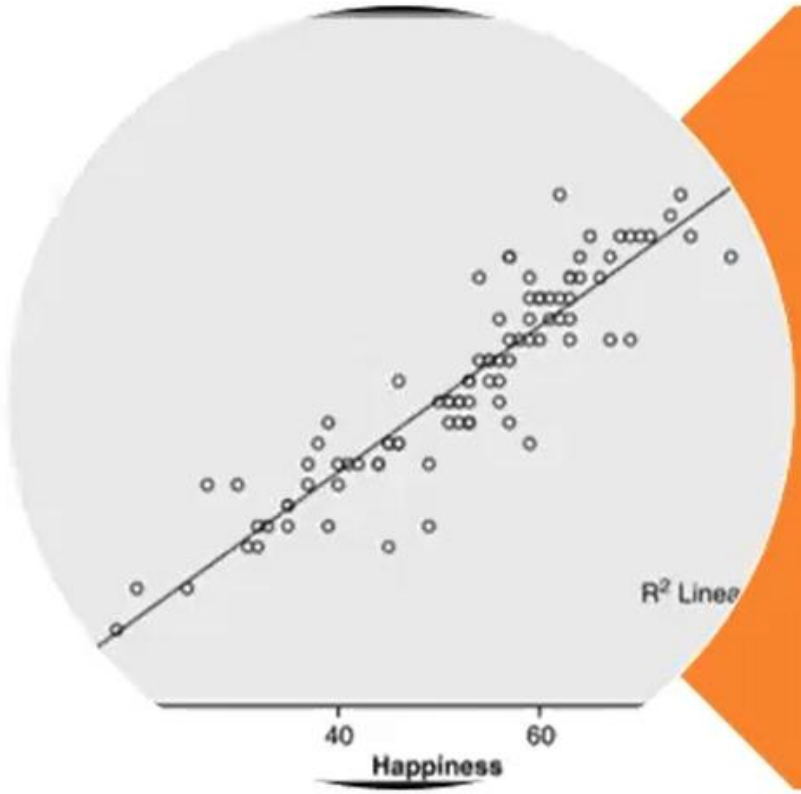
Data Science
Bivariate

Bivariate Data Analysis

X, Y Columns-2



Bivariate Data Analysis

- Covariance
- Correlation
- Collinearity
- Multicollinearity
- Variance Inflation Factor
- Homoscedasticity
- Heteroscedasticity

COVARIANCE

COVARIANCE

Large Positive Covariance

Nearly Zero Covariance

Large Negative Covariance

COVARIANCE

Population Covariance Formula

$$Cov(x,y) = \frac{\Sigma(x_i-\bar{x})(y_i-\bar{y})}{N}$$

Sample Covariance

$$Cov(x,y) = \frac{\Sigma(x_i-\bar{x})(y_i-y)}{N-1}$$

Notations in Covariance Formulas

- x_i = data value of x
- y_i = data value of y
- $\bar{x}$ = mean of x
- $\bar{y}$ = mean of y
- N = number of data values.

CORRELATION

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

Where:

- r_{xy} – the correlation coefficient of the linear relationship between the variables x and y
- x_i – the values of the x-variable in a sample
- $\bar{x}$ – the mean of the values of the x-variable
- y_i – the values of the y-variable in a sample
- $\bar{y}$ – the mean of the values of the y-variable

#learnaiwithramisha

ALL ABOUT CORRELATION

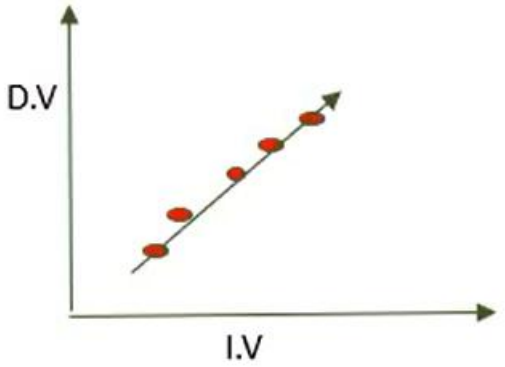
SEE THE PICTURE AND TELL THE STORY

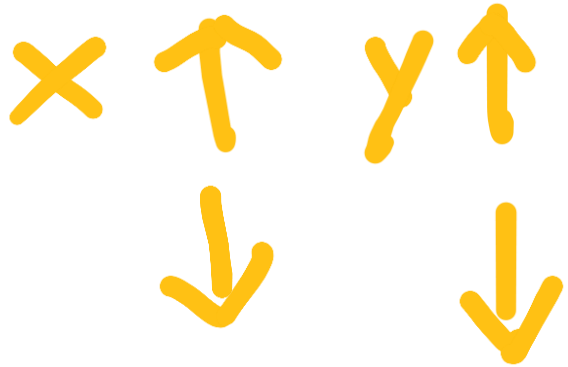
TYPES OF CORRELATION:: LINEAR TYPE WITH TWO VARIABLES

POSITIVE CORRELATION

Independent Variable is directly proportional to Dependant Variable

I.V ↑ D.V ↑ I.V α D.V

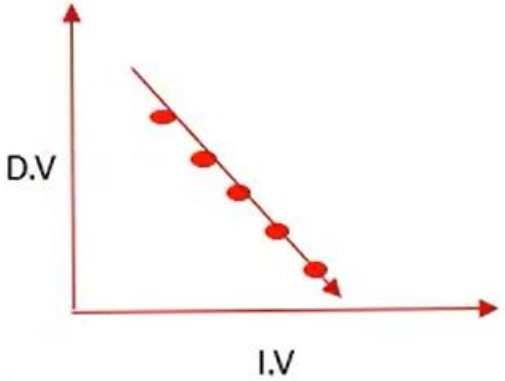


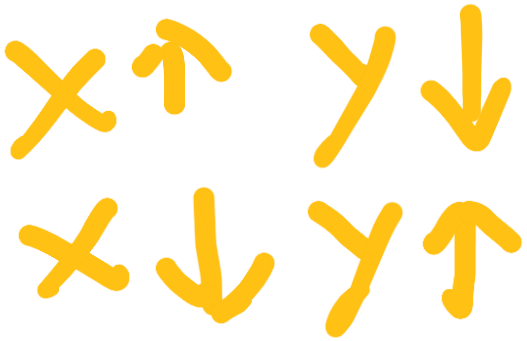


NEGATIVE CORRELATION

Independent Variable is indirectly proportional to Dependant Variable

I.V ↑ D.V ↓

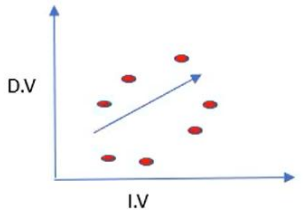




ZERO CORRELATION

ZERO CORRELATION

No pattern between Independent Variable And Dependant Variable



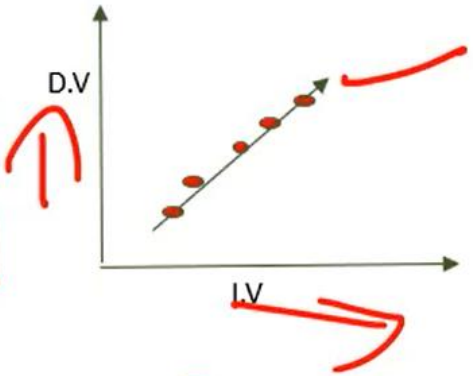
#learnaiwithramisha

TYPES OF CORRELATION:: LINEAR TYPE WITH TWO VARIABLES

POSITIVE CORRELATION

Independent Variable is directly proportional to Dependant Variable

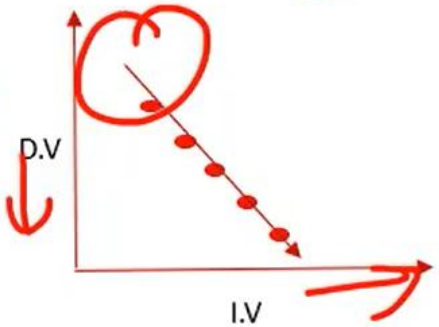
I.V ↑ D.V ↑ I.V α D.V



NEGATIVE CORRELATION

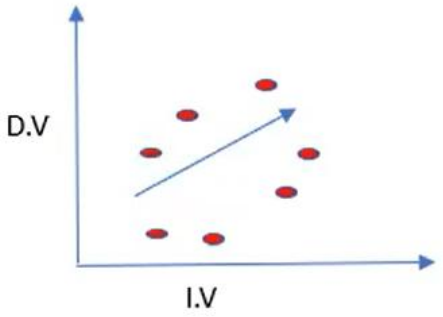
Independent Variable is indirectly proportional to Dependant Variable

I.V ↑ D.V ↓ I.V α 1/D.V



ZERO CORRELATION

No pattern between Independent Variable And Dependant Variable

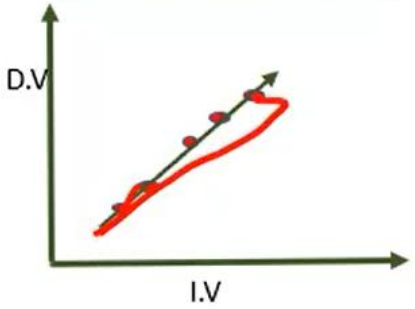


ALL ABOUT CORRELATION

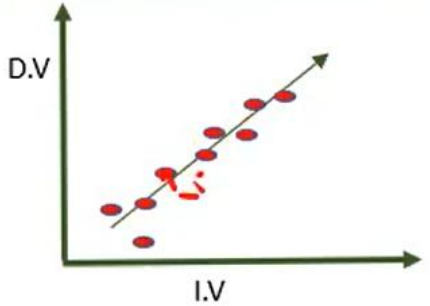
SEE THE PICTURE AND TELL THE STORY

Correlation explained in terms of degree(amount of change from one point to another point))

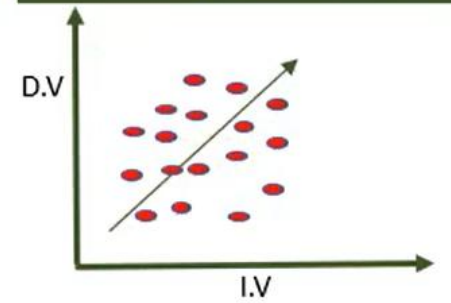
1. Perfect positive Correlation(r= 1)



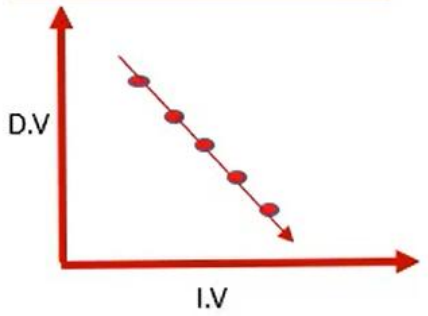
2. High Degree of +Ve Correlation (r= + High or 0.95):



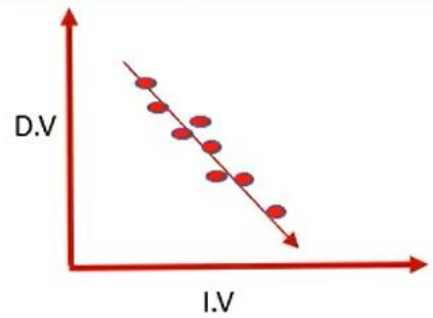
3. Low degree of +Ve Correlation (r= + Low or 0.54):



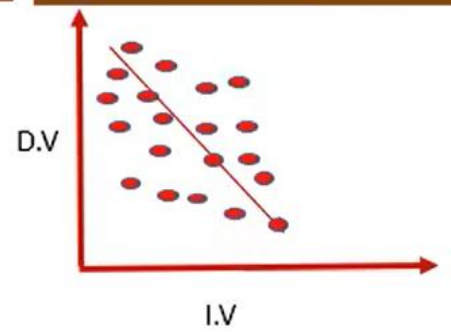
4. Perfect Negative Correlation (r=-1)



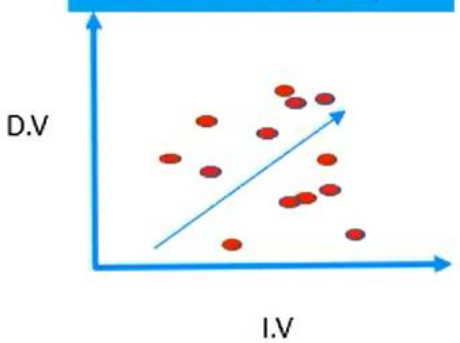
5.High Degree of –Ve Correlation (r= – low or -0.54):



6. Low Degree of –Ve Correlation (r= - high or -0.94)



7.No Correlation (r= 0):



Pause