Semester : V Subject : Statistics for AIDS Academic Year 20 - 20

CORRELATION:

Correlation analysis compares two variables and is called as bivariate analysis. There are different types of Correlation:

* Positive Correlation
* Negative Correlation.
* No Correlation.

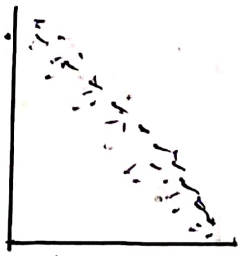
Exploratory data analysis in many modeling projects involves examining correlation among predictors, and between predictors and a target variable.

Positive Correlation → When X increases then Y also increases.

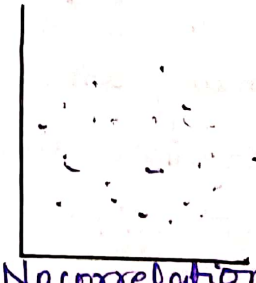
Negative Correlation → When X ~~decreases~~ increases then Y also decreases. One variable increases and other variable decreases.

No correlation → ~~When X increases and Y decreases.~~

Positive correlation



Negative correlation.



No correlation

→ There is no linear relationship between two variables.

Semester V Subject Statistics for AIDS Academic Year 2023-2024.

SCATTERPLOTS:

The standard way to visualize the relationship between two measured data variables is with a scatterplot. The x-axis represents one variable, the y-axis another, and each point on the graph is a record.

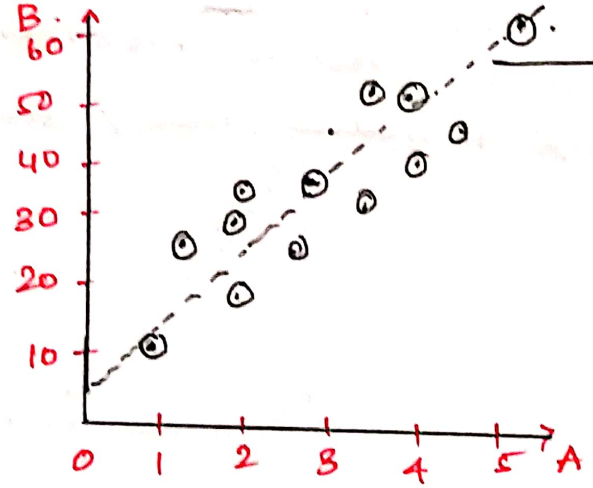
Example:-

Consider the below points and create scatterplot.

A	1	2	3	4	5	6
B	10	25	35	58	60	200

A will be plotted on x-axis and B will be plotted on Y-axis.

(6,200) → outlier.



Consider there are so many other plots. When x increases y also increases, a straight line is derived. This is known as line of best fit.