Correlation

Department of Computer Applications

Course: 23CAT705- RESEARCH METHODOLOGY

UNIT I : Analytical Statistics

Class: II MCA / III SEMESTER

Understanding the impact of business initiatives on desired outcomes can help companies identify trends, uncover hidden opportunities, refine strategies, allocate resources more efficiently, and make better decisions overall.

Metrics

- Customer satisfaction survey
- Employee survey
- Customer experience (CX) programs
- Market survey

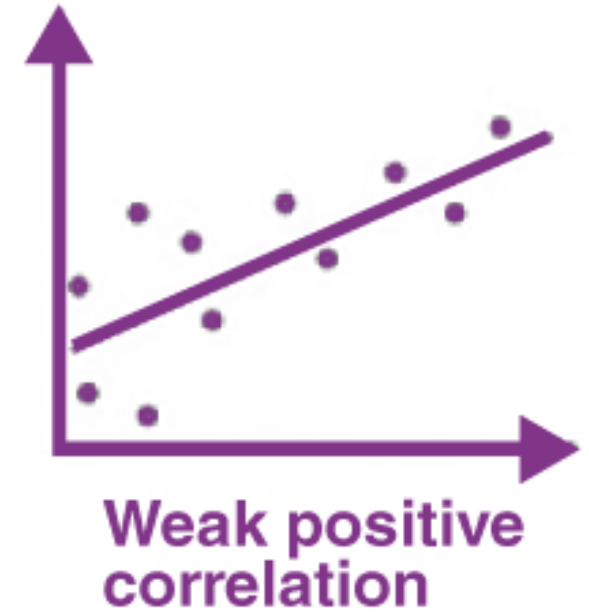
Correlation

- Examines between two or more variables for relationship
- *Determine the effect of independent variable(s)*
- Represented using the symbol r, and it ranges from -1 to +1.
- Correlation is an analysis of co-variation between two or more variables

If the coefficient is close to 0 then the relation between the relationship between the two numbers is less and when the relationship is far away from 0 then the relationship is strong between the two variables.

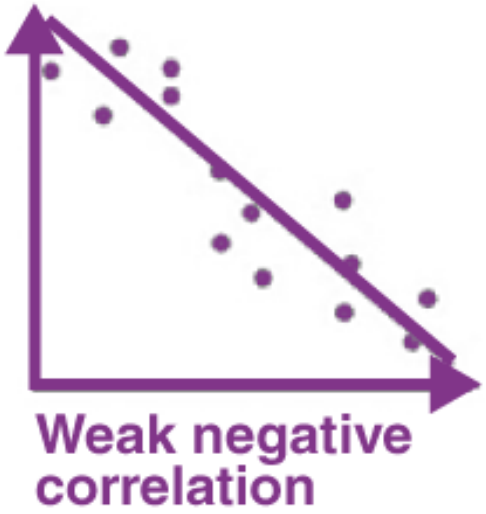
$$r_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}$$

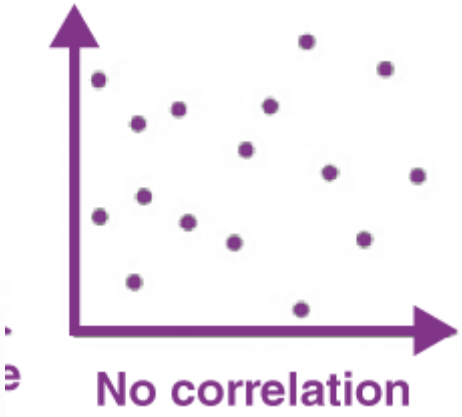
Positive Correlation



- If an increase (or a decrease) in one variable corresponds to an increase (or a decrease) in the other

Negative Correlation





- if an increase (or a decrease) corresponds to a decrease (or an increase) in the other.

Illustration - Correlation

Example: Poor suburbs are more likely to have high pollution

- Do poor people make pollution?
- Are polluted suburbs the only place poor people can afford?
- Is it a common link, such as factories with low paying jobs and lots of pollution?

Example: strong positive correlation between "Studying an external course" and Sick Days

- Studying makes them sick?
- Sick people study a lot?
- Did they lie about being sick so they can study more?

Correlation

Simple Correlation

- It involves the study of only two variables.
- Example, correlation between the price and demand of a product

Multiple Correlation

- Study of three or more variables simultaneously
- Example, Relationship between the yield of wheat per acre and both amount of rainfall and the amount of fertilizers used

Partial Correlation

- It involves the study of three or more variables, but considers only two variables to be influencing each other
- Example: It involves the study of three or more variables, but considers only two variables to be influencing each other

Example

2 Subtract Mean

3 Calculate ab, a^2 and b^2

Temp °C	Sales	"a"	"b"	a×b	a^2	b^2
14.2	$215	-4.5	-$187	842	20.3	34,969
16.4	$325	-2.3	-$77	177	5.3	5,929
11.9	$185	-6.8	-$217	1,476	46.2	47,089
15.2	$332	-3.5	-$70	245	12.3	4,900
18.5	$406	-0.2	$4	-1	0.0	16
22.1	$522	3.4	$120	408	11.6	14,400
19.4	$412	0.7	$10	7	0.5	100
25.1	$614	6.4	$212	1,357	41.0	44,944
23.4	$544	4.7	$142	667	22.1	20,164
18.1	$421	-0.6	$19	-11	0.4	361
22.6	$445	3.9	$43	168	15.2	1,849
17.2	$408	-1.5	$6	-9	2.3	36
18.7	**$402**			**5,325**	**177.0**	**174,757**

1 Calculate Means

4 Sum Up

5 $$\frac{5{,}325}{\sqrt{177.0 \times 174{,}757}} = 0.9575$$

Linear Correlation

- If the amount of change in one variable bears a constant ratio to the amount of change in the other variable, then correlation is said to be linear.

Example

Raw material	: X	10	20	30	40	50	60
Finished Product	: Y	2	4	6	8	10	12

Non-Linear Correlation

- If the amount of change in one variable does not bear a constant ratio to the amount of change in the other variable, then correlation is said to be linear.

Example

Raw material	: X	10	20	30	40	50	60
Finished Product	: Y	2	6	7	4	8	5

Measuring Correlation

Scatter Diagram

Given data are plotted on a graph paper in the form of dot representing X and Y variables

Karl Pearson's Coefficient of Correlation

$$r = \frac{\sum XY}{\sqrt{\sum X^2 \sum Y^2}}$$

where

$X = x - \bar{x}$

$Y = y - \bar{y}$

References

1. Kothari, C.R. &Garg, G. (2019). *Research Methodology: Methods and Techniques*. New Age International Publishers, New Delhi

2. Goode, W.J. &Hatt, P.K. (2022). Methods in Social Research. McGraw Hill, London

3. Bhandarkar, P.L. & Wilkinson, T.S. (2016). Methodology and Techniques of Social Research. Himalaya Publishing House, Mumbai.

Thank You