

The logo of Galgotias University is a stylized circular emblem with three curved, overlapping bands in shades of yellow, blue, and red, creating a sense of motion or a globe.

UNIT I

DATA INTEGRATION

GALGOTIAS
UNIVERSITY

Correlation Analysis (Numeric Data)

Correlation coefficient (also called **Pearson's product moment coefficient**)

$$r_{A,B} = \frac{\sum_{i=1}^n (a_i - \bar{A})(b_i - \bar{B})}{(n-1)\sigma_A\sigma_B} = \frac{\sum_{i=1}^n (a_i b_i) - n\bar{A}\bar{B}}{(n-1)\sigma_A\sigma_B}$$

where n is the number of tuples, \bar{A} and \bar{B} are the respective means of A and B, σ_A and σ_B are the respective standard deviation of A and B, and $\sum(a_i b_i)$ is the sum of the AB cross-product.

If $r_{A,B} > 0$, A and B are positively correlated (A's values increase as B's). The higher, the stronger correlation.

$r_{A,B} = 0$: independent; $r_{AB} < 0$: negatively correlated



Correlation (viewed as linear relationship)

- Correlation measures the linear relationship between objects
- To compute correlation, we standardize data objects, A and B, and then take their dot product

$$a'_k = (a_k - \text{mean}(A)) / \text{std}(A)$$

$$b'_k = (b_k - \text{mean}(B)) / \text{std}(B)$$

$$\text{correlation}(A, B) = A' \bullet B'$$

- Covariance is similar to correlation

where n is the number of tuples, \bar{A} and \bar{B} are the respective mean or **expected values** of A and B, σ_A and σ_B are the respective standard deviation of A and B.

- **Positive covariance:** If $\text{Cov}_{A,B} > 0$, then A and B both tend to be larger than their expected values.
- **Negative covariance:** If $\text{Cov}_{A,B} < 0$ then if A is larger than its expected value, B is likely to be smaller than its expected value.
- **Independence:** $\text{Cov}_{A,B} = 0$ but the converse is not true:
 - Some pairs of random variables may have a covariance of 0 but are not independent. Only under some additional assumptions (e.g., the data follow multivariate normal distributions) does a covariance of 0 imply independence

Software Process and Process Models (Heading)



GALGOTIAS
UNIVERSITY



GALGOTIAS
UNIVERSITY

DSDM framework

References:



GALGOTIAS
UNIVERSITY



Thank You