Missing Values: Decision Guide

Notes for SHS • Topic 83 • Fictional data only • 25 November 2026

Identify and define

Blank, N/A and other sentinels are context-dependent; zero can be valid.

Mini example: two missing scores in ten imported rows = 20%.

Investigate why

MCAR, MAR and MNAR are modelling assumptions, not patterns proven by looking at blanks.

Choose an honest action

Retain missing values; recover from source when authorized; filter for specified analyses; impute with justified method; disclose limitations.

Compare rather than conceal

Raw 18 missing scores in 252 imported rows; analysis-ready 12 missing in 240 unique rows.

Different denominators do not imply like-for-like percentages.

Avoid leakage

Split training and test first; fit learned imputation statistics from training data only.