

DA0-002 Starter Glossary

Twelve concise terms from all five Data+ study domains.

Use this starter list for active recall: cover the definition, explain the term in your own words, then check your answer.

12 starter terms

Term	Definition
Correlation	A measure describing the direction and strength of an association between two variables.
Dashboard	A consolidated visual display of selected metrics and information used to monitor performance or status.
Data lake	A scalable repository that can retain large amounts of raw data in multiple formats for later processing.
Data lineage	A record of where data came from and how it moved or changed throughout its lifecycle.
Data masking	The replacement or concealment of sensitive values so unauthorized viewers cannot see the original information.
Data profiling	The examination of a dataset's values, patterns, types, and quality before it is transformed or analyzed.
Data warehouse	A curated analytical store that organizes integrated, historical data for reporting and analysis.
Deduplication	The process of identifying and resolving repeated records that represent the same entity or event.
Imputation	A method of replacing missing values with reasonable estimates based on a defined rule or model.
Least privilege	An access principle that grants only the permissions required to perform an authorized task.
Median	The middle value after observations are placed in order; it is less affected by extreme values than the mean.
Structured data	Data organized according to a consistent model, such as named columns with defined data types.

The complete subscriber glossary contains 153 terms and is available inside Data Plus Study.

An affordable, independently developed study platform for CompTIA Data+ DA0-002 preparation. Exam objectives and policies can change; confirm current details with the exam provider.