

A3 DATABASES · A3.4 · HL ONLY

Alternative databases and data warehouses

Beyond relational: **NoSQL** database types, **data warehouses** and ETL, OLAP and data mining, and **distributed** databases. HL only.

01 Relational vs NoSQL

● Relational

Linked tables, fixed schema, SQL, strong integrity (ACID). Best for structured, related data.

● NoSQL

Flexible schema, scales out across machines. Best for big, changing, or unstructured data.

02 NoSQL types

Key-value Simple key to value pairs (Redis).

Document JSON-like documents (MongoDB).

Column Data in column families (Cassandra).

Graph Nodes and edges (Neo4j).

03 Data warehouse & ETL

E

Extract

Pull data out of each source system (sales, web, CRM).

T

Transform

Clean and convert it into one consistent format.

L

Load

Load it into the warehouse: a central, historical, append-only store for analysis.

04 OLAP & data mining

OLAP	Complex analysis of historical data.
OLTP	Everyday transactions (the contrast).
Data mining	Finds patterns and trends in the data.
Warehouse	Central, historical, append-only.

05 Distributed databases

● Replication	● Sharding
Full copy on every node. Availability and fast reads; writes must sync.	Data split across nodes. Scales storage and writes; a lost node loses its shard.

06 Know the difference

Relational vs NoSQL	Fixed table schema and SQL versus a flexible, non-table model that scales out.	MODEL
OLAP vs OLTP	Analysing large historical data versus handling everyday transactions.	USE
Replication vs sharding	Full copies on each node versus splitting the data across nodes.	DISTRIBUTE
Warehouse vs database	Historical store built for analysis versus a live store for daily operations.	PURPOSE

FINAL PASS BEFORE THE EXAM

Rapid exam tips

Eight things that lose marks in Paper 1 if you slip on them. A3.4 is HL only. Skim before you walk in.

01

NoSQL means "not only SQL" — a different model, not "better than relational".

02

The four NoSQL types: **key-value**, **document**, **column-family**, **graph**.

03

A **data warehouse** is historical, integrated, append-only, and built for analysis.

04

ETL = Extract, Transform, Load data into the warehouse.

05

OLAP analyses historical data; **OLTP** handles everyday transactions.

06

Data mining searches the data for patterns and trends.

07

Replication = full copies (availability); **sharding** = split data (scale).

08

This whole subtopic is **HL only**.