

5_Data Layers

Tuesday, February 10, 2026 3:30 PM

Layer	Staging		Core	Presentation	
Medallion	Bronze		Silver	Gold	Platinum
Sub-layer	Raw (Landing)	Standardised (Prepared)	Core (Conformed)	Data mart (Refined)	Report (Curated)
Purpose	<ul style="list-style-type: none"> As-Is from raw data Landing area from data sources Meta data captured for auditability Change data captured to maintain the historical view PII data encrypted Schema per source system 	<ul style="list-style-type: none"> Data is cleansed Data types are corrected, and standards are applied Highly governed and documented Subset of data if applicable from raw Maintains history if raw layer / source system does not support Schema per source system 	<ul style="list-style-type: none"> Central enterprise data repository of business entities (i.e., customer, product, transaction) Single source of truth Data from multiple sources are cleaned and modelled into conformed enterprise view Business rules are applied SCD-2 is maintained Enriched with reference data 	<ul style="list-style-type: none"> Data products Star schema / Kimball method Enable "self-serve" analytics Fact & dimension tables to build the user-friendly data set Product specific transformation rules applied Data is logically grouped into shared and departmental 	<ul style="list-style-type: none"> Consumer ready data set Data sets are joined to build the holistic analytics view Pre-aggregated data set Created from joining fact & dimensions to consume by end-users
Consumers	Data Engineering	Data analysts, Engineers & Scientists	Operational analyst, Data analysts, Data science	BI analysts, Head of departments, Data science	Executives & "C" Suites
Data model	As-Is	NA	Data Vault (DV)	Dimensional Modelling (DM)	One Big Table (OBT)
<div> <div>Normalization</div> <div>De-normalization</div> </div> <div> <div>←</div> <div>→</div> </div>					
Data Layers					

1. Staging Layer (Bronze)

Raw / Landing

This is the very first place where data lands after coming from source systems.

Key points

- Data is stored **exactly as received** (no cleaning or changes).
- Used for **auditing, debugging, and historical tracking**.
- Metadata and change data (CDC) are captured.
- Sensitive data (PII) is encrypted.
- Each source system gets its **own schema** to avoid mixing data.

Purpose

- Preserve original data for traceability.
- Provide a safe backup if downstream layers fail.

2. Core Layer (Silver)

This layer has two sub-layers: **Standardised** and **Conformed**.

Standardised / Prepared

Data becomes clean, consistent, and usable.

Key points

- Fixes data types, formats, and naming standards.
- Removes duplicates and applies basic quality rules.
- Highly governed and documented.
- Maintains history if the raw layer or source doesn't support it.
- Still organized by source system.

Purpose

- Convert messy raw data into clean, reliable datasets.

Core / Conformed

This is the **enterprise-wide** view of data.

Key points

- Combines data from multiple systems (SAP + Salesforce + POS).

- Creates unified business entities (Customer, Product, Transaction).
- Applies business rules and logic.
- Maintains SCD-2 for historical tracking.
- Enriched with reference/master data.
- Acts as the **single source of truth**.

Purpose

- Provide a consistent, trusted foundation for analytics and data marts.

3. Presentation Layer (Gold & Platinum)

This layer is designed for business consumption.

Data Mart / Refined (Gold)

Business-friendly datasets organized for analytics.

Key points

- Built using **Star Schema** (facts + dimensions).
- Department-specific or shared data products.
- Product-specific transformation rules applied.
- Supports self-service analytics for analysts and BI teams.

Purpose

- Make data easy to query, understand, and analyze.

Report / Curated (Platinum)

Final, ready-to-consume datasets.

Key points

- Pre-aggregated and optimized for dashboards.
- Facts and dimensions are joined to create a complete view.
- Used by executives, leadership, and BI tools.

Purpose

- Deliver fast, ready-made insights for reporting and decision-making.

Data Modeling Across Layers

- **Raw:** As-Is (no modeling)
- **Standardised:** N/A (cleaning only)
- **Core:** Data Vault (enterprise modeling)
- **Data Mart:** Dimensional Modeling (Kimball)
- **Report:** One Big Table (OBT) for fast consumption

Normalization → De-Normalization

- Early layers (Raw, Standardised, Core) are **more normalized** → structured, clean, consistent.
- Later layers (Data Mart, Report) are **more de-normalized** → easier for analytics and reporting.