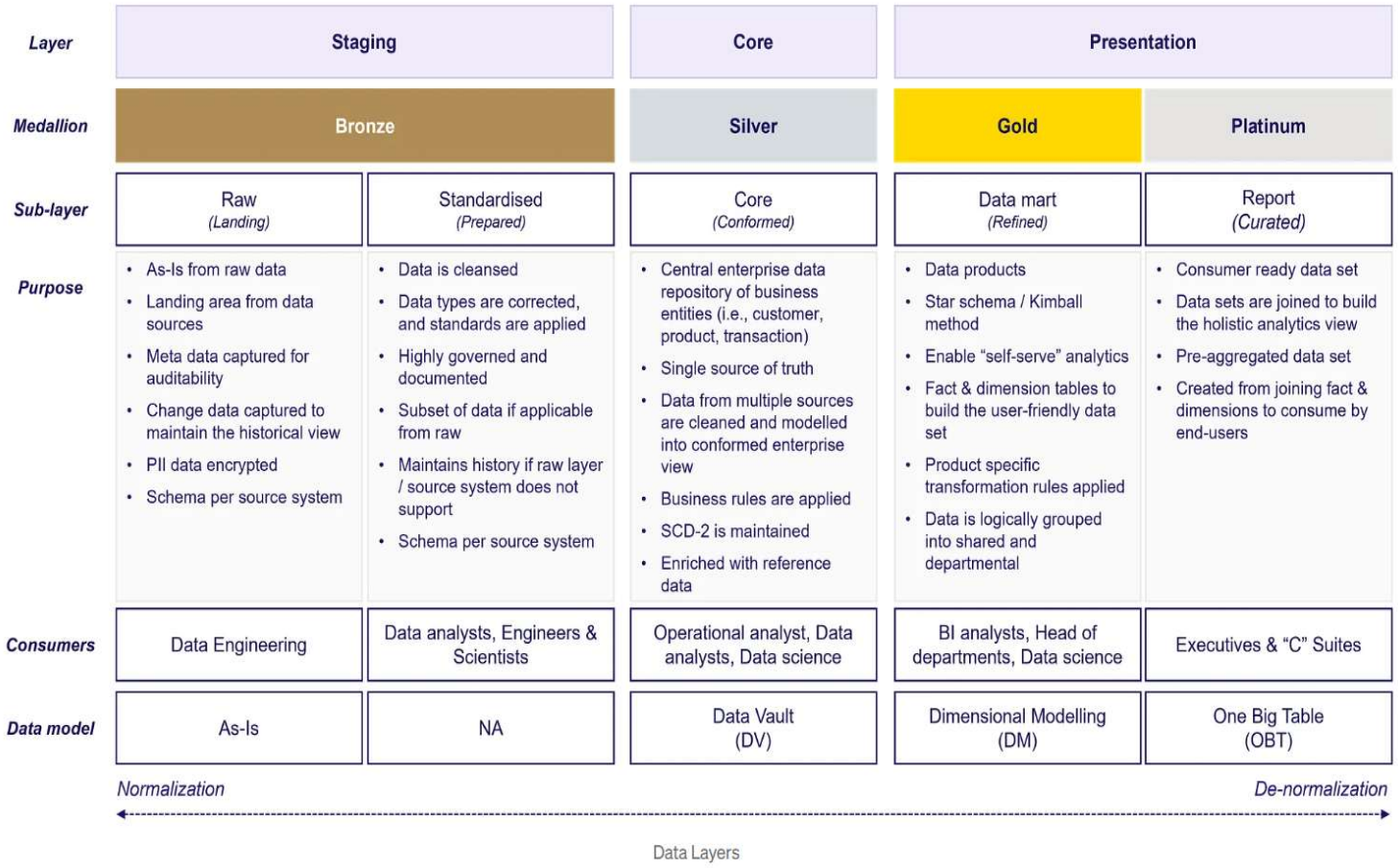
5_Data Layers

Tuesday, February 10, 2026 3:30 PM



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.