

Livrable Urbaniste – Vue Transformation

SmartShop 360

23 février 2026

Objectif

Distinguer clairement :

- **SI Opérationnel** (sources de données externes)
- **SI Décisionnel** (entrepôt PostgreSQL + vues analytiques)

Mettre en évidence le **Référentiel Produit (MDM)** comme pivot entre les silos.

Description de la transformation

1. Les sources `online_retail_II.csv` (ERP) et `labeledReview.datasetFix.json` (avis) arrivent dans l'ETL.
2. `incremental.py` compare les hashes SHA-256 et évite les rechargements inutiles.
3. `cleaning.py` et `data_quality.py` nettoient/valident les données.
4. `mdm_mapping.py` produit le pivot `product_mapping` (TF-IDF + fallback rank).
5. `run_etl.py` charge les tables et publie les vues `v_product_kpi`, `v_customer_kpi`, `v_alerts`, `v_data_quality`.
6. Les couches UI, Agent IA, Analytics/ML et monitoring consomment le SI décisionnel.

Séparation SI Opérationnel / SI Décisionnel

- **Opérationnel** : fichiers sources (CSV/JSON), données hétérogènes non jointes.
- **Pivot MDM** : `product_mapping` relie les identifiants ERP et Avis.
- **Décisionnel** : PostgreSQL 16 (tables `products`, `customers`, `sales_facts`, `review_facts`) + vues KPI.

Diagramme

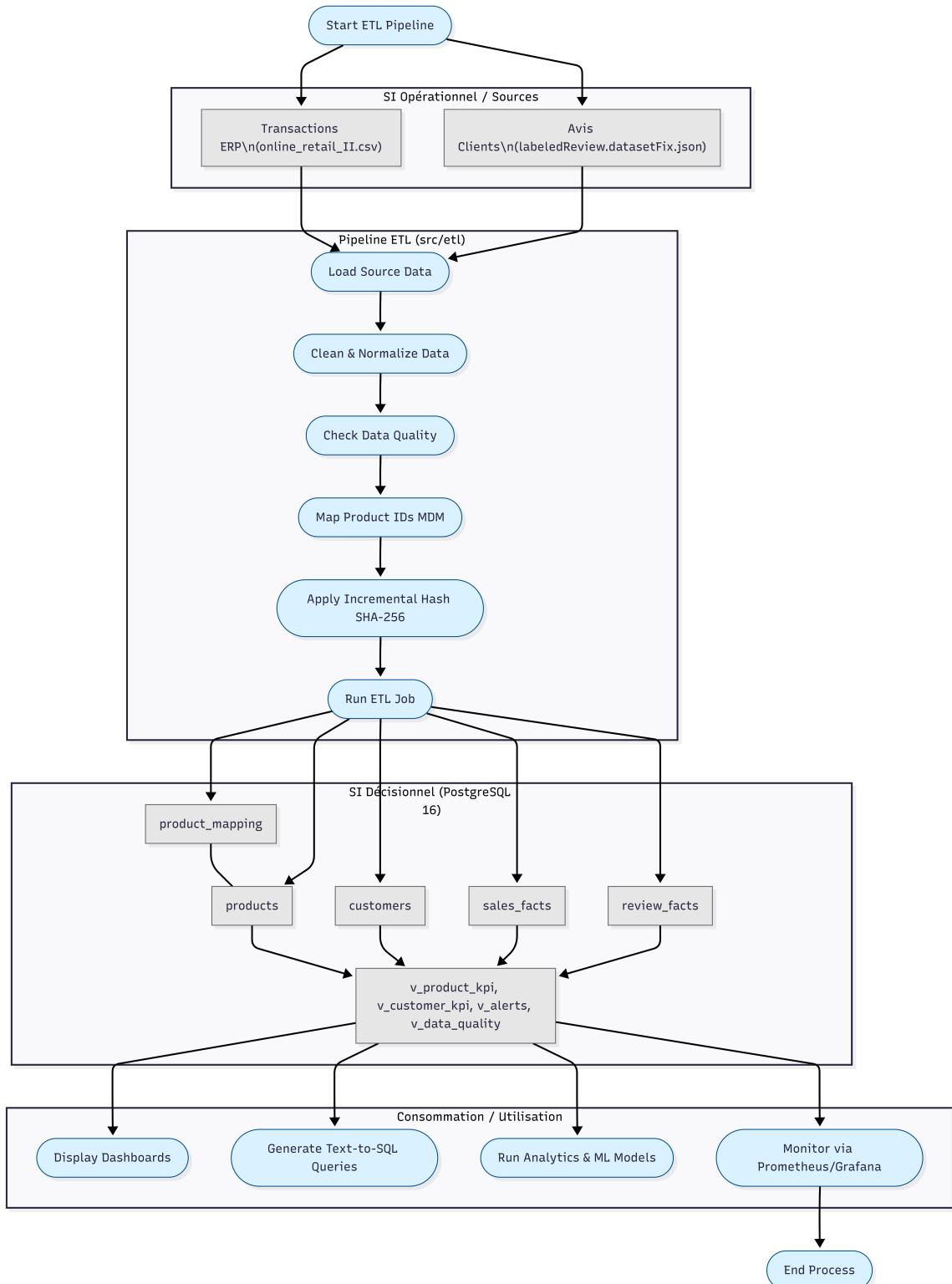


Diagramme Mermaid associé

Voir le fichier : 02_vue_transformation.mmd