



# Data Warehouse Project

## Todo List

Aa	Task name	👤 Assignee	📅 Due date	⚙️ Status
	<u>Untitled</u>			Not started

## SQL Data Warehouse

In this project we are doing

Pull Extraction

Extraction Type: Full Extraction

Extraction Technique: File parsing

Transformation

Batch Processing

Full Load: Truncate & Insert

SCD1

## Building a Datawarehouse

### Objective

### Specifications

Data Sources : ERP and CRM csv files

Data Quality; cleanse

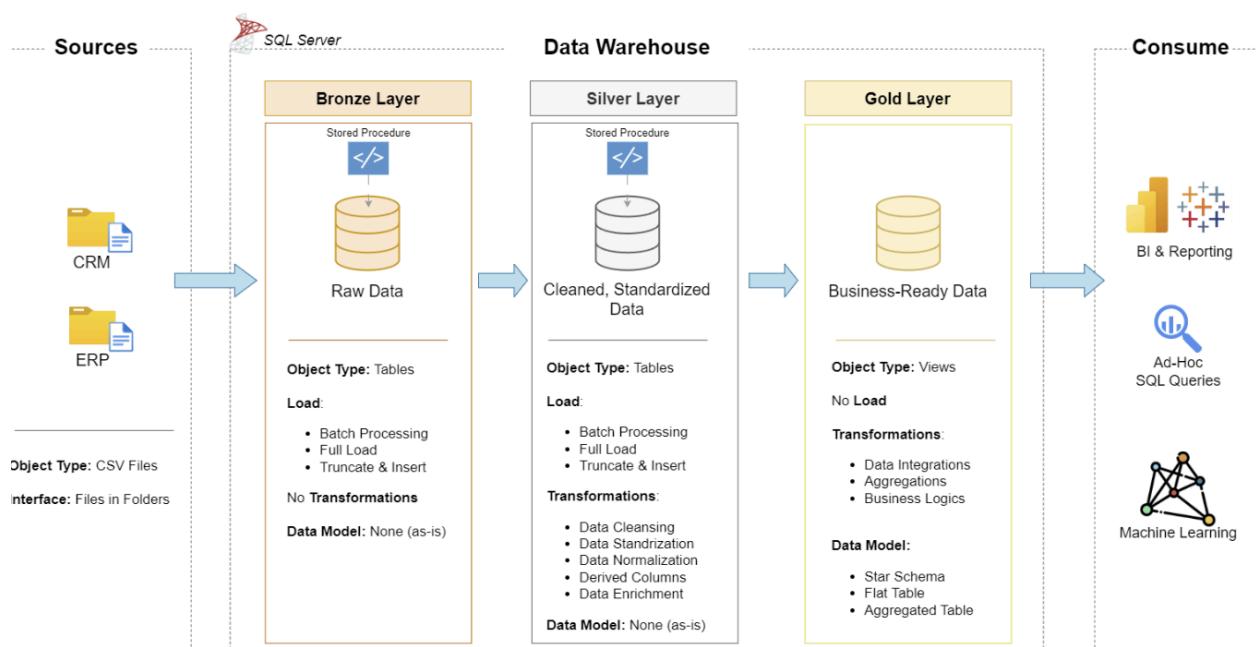
Integration: combine sources

Scope

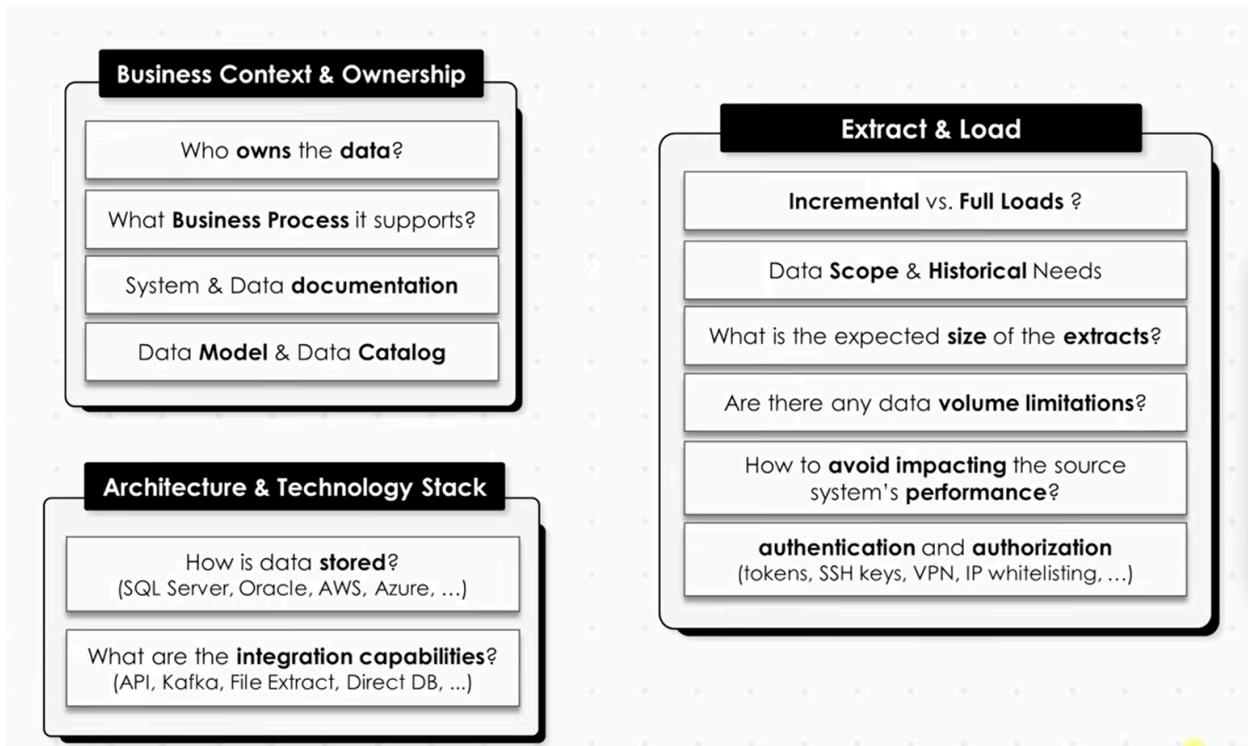
Documentation

# Separation of Concerns:

Dividing the complex into small modules and not duplicates



# Data Architecture



## Data Warehouse

Inmon: Stage → EDW → Data Marts

Kimnaff: Stage → Data Marts

Data Vault: Stage → Raw Vault → Business Vault → Data Marts

Medallion Archi : Bronze → Silver → Gold

	Bronze Layer	Silver Layer	Gold Layer
Definition	Raw, unprocessed data as-is from sources	Clean & standardized data	Business-Ready data
Objective	Traceability & Debugging	(Intermediate Layer) Prepare Data for Analysis	Provide data to be consumed for reporting & Analytics
Object Type	Tables	Tables	Views
Load Method	Full Load (Truncate & Insert)	Full Load (Truncate & Insert)	None
Data Transformation	None (as-is)	<ul style="list-style-type: none"> <li>- Data Cleaning</li> <li>- Data Standardization</li> <li>- Data Normalization</li> <li>- Derived Columns</li> <li>- Data Enrichment</li> </ul>	<ul style="list-style-type: none"> <li>- Data Integration</li> <li>- Data Aggregation</li> <li>- Business Logic &amp; Rules</li> </ul>
Data Modeling	None (as-is)	None (as-is)	<ul style="list-style-type: none"> <li>- Start Schema</li> <li>- Aggregated Objects</li> <li>- Flat Tables</li> </ul>
Target Audience	- Data Engineers	<ul style="list-style-type: none"> <li>- Data Analysts</li> <li>- Data Engineers</li> </ul>	<ul style="list-style-type: none"> <li>- Data Analysts</li> <li>- Business Users</li> </ul>

Bronze Layer :

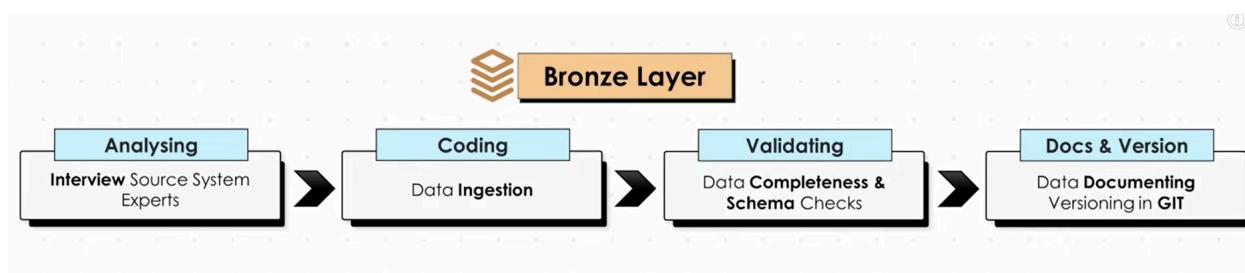
Def: Raw, Unprocessed data as is from sources

Obj: Traceability & Debugging

obj Type: Tables

Load Method : Full Load( Truncate & insert)

Target Audiences: Data Engineers



## DDL: Data Definition Language

Silver layer:

Def: Clean & standardized data

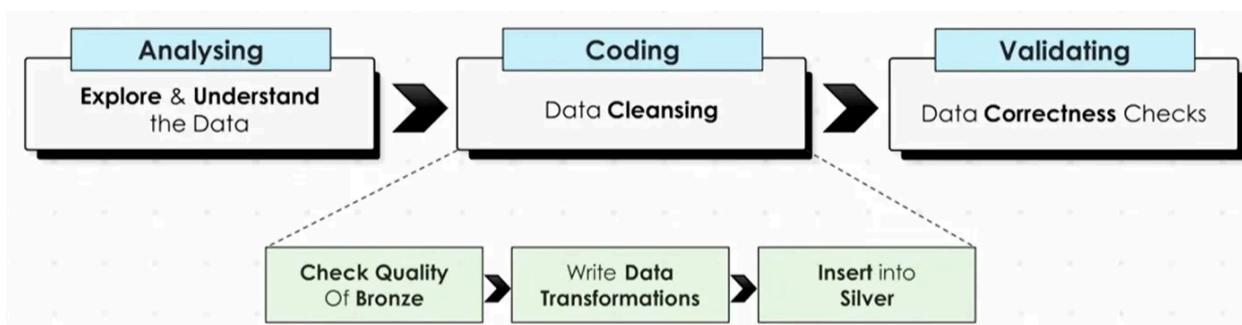
Obj: prepare Data for Analysis

Type: Tables

Load Method: Full Load( Truncate & insert)

Data Transformation: Data Cleaning, Data Standardization, Data Normalization, Derived Columns, Data Enrichment

TA: Data Engineers, Data Analysts



Gold Layer:

Def: Business - Ready data

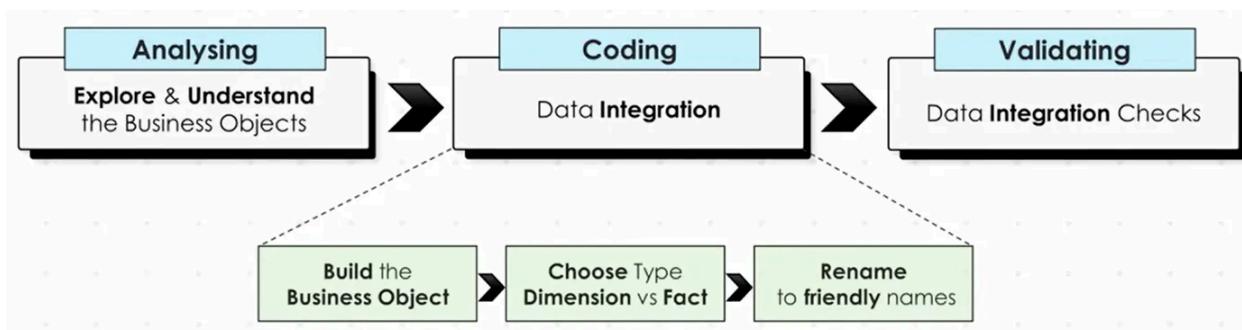
obj: Provide data to be consumed for reporting & analytics

Type: Views

Data Transformation: Data Integration, Data Aggregation, Business Logic & Rules

Data Modeling: Start Schema, Aggregated Objects, Flat Tables

TA: Data Analysts, Business Users

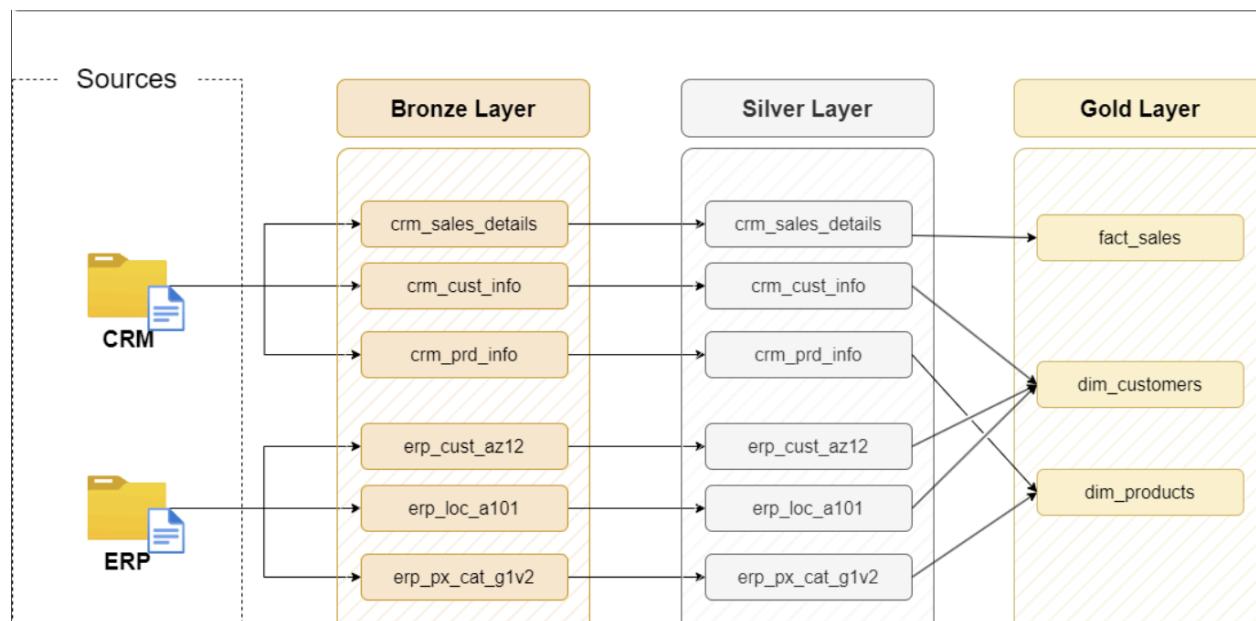


Star Schema

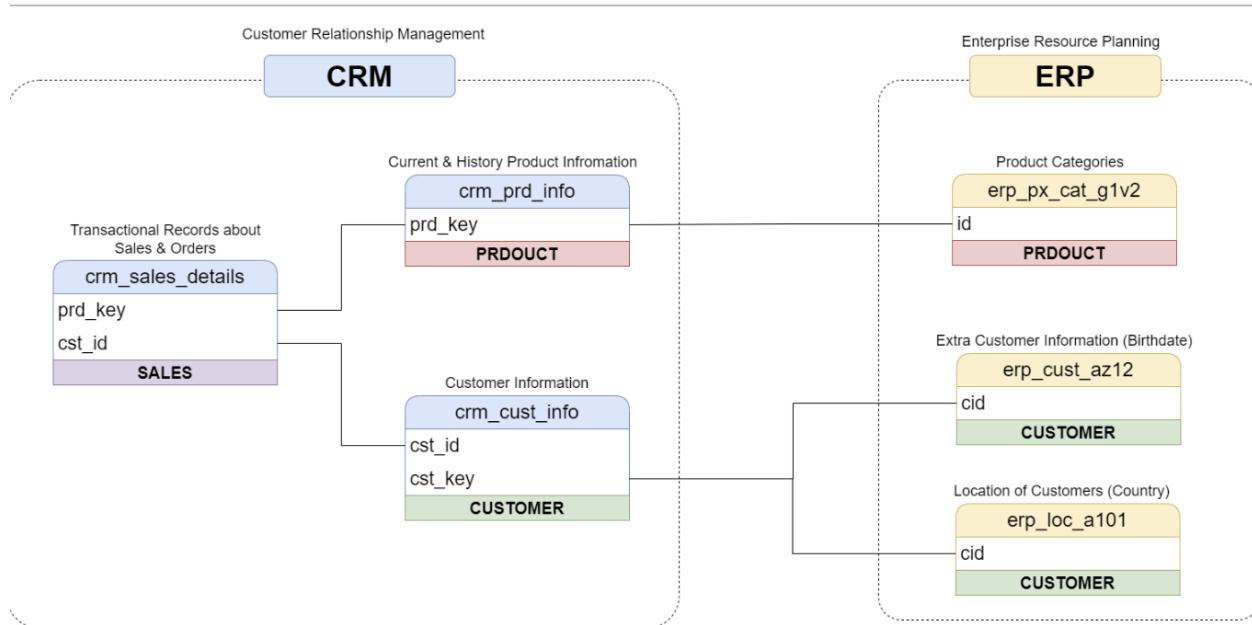
SnowFlake Schema

## Surrogate Key

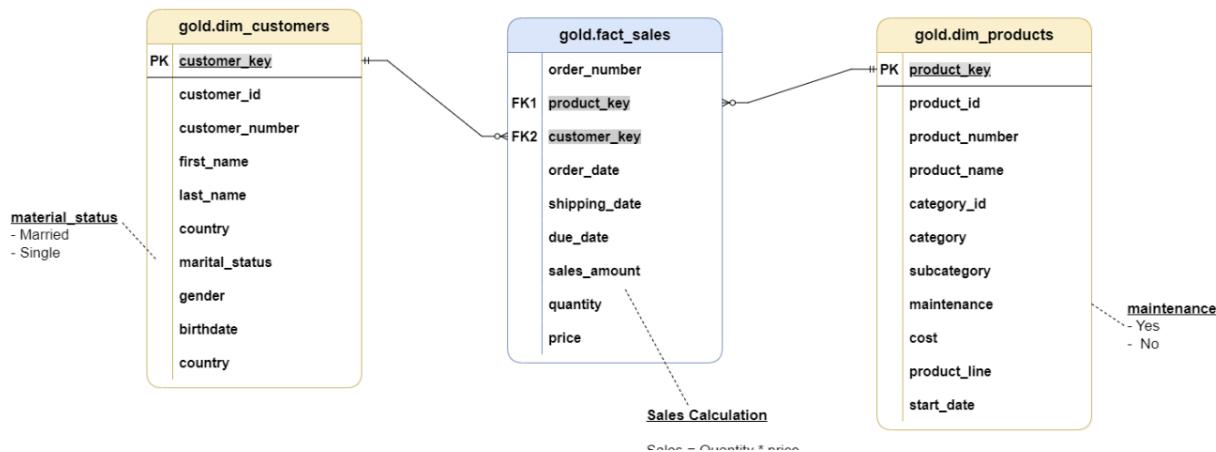
# Data Flow



# Data Integration



# Data Model



# Data Lake

## Date LakeHouse

## Data Mesh

