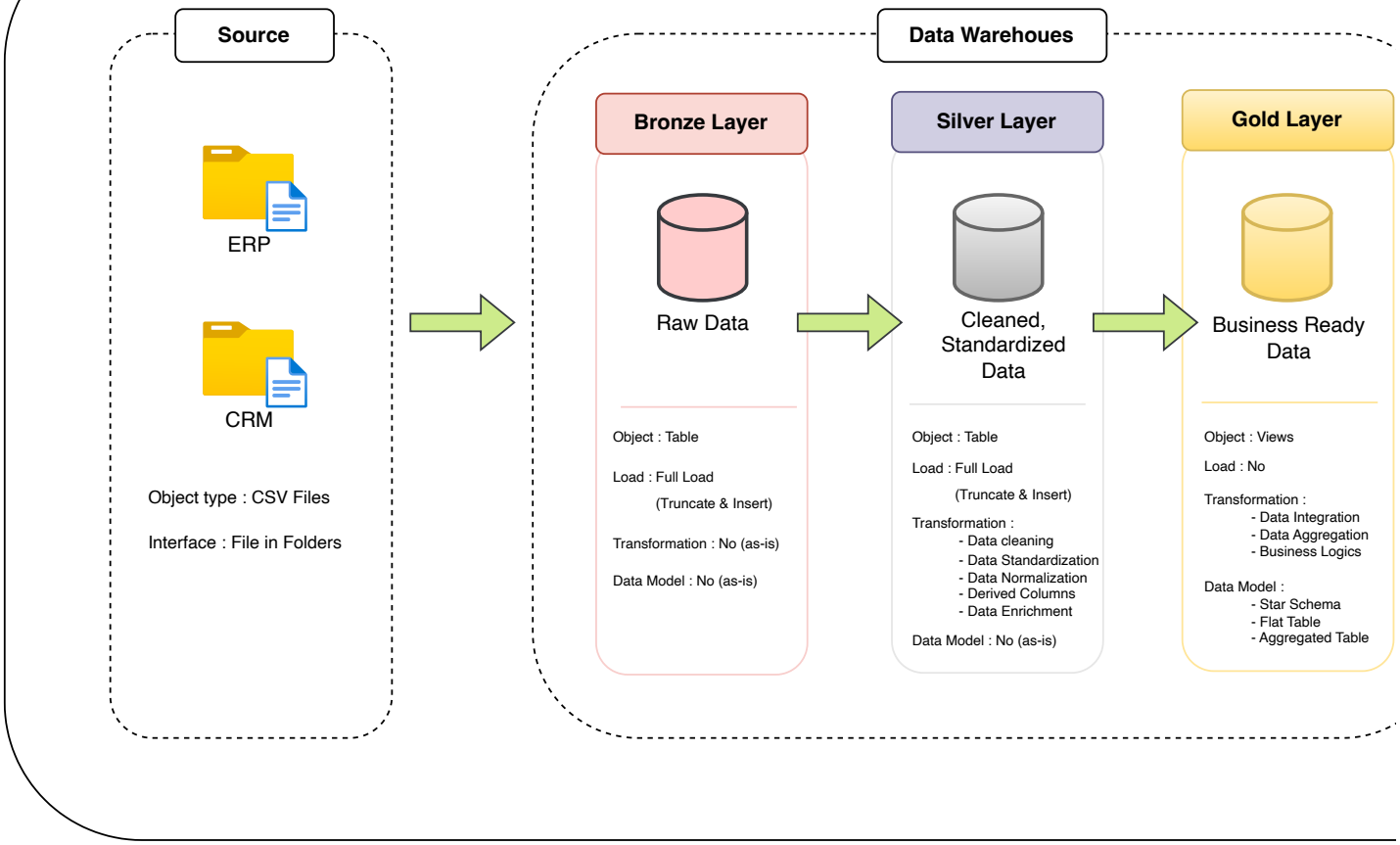
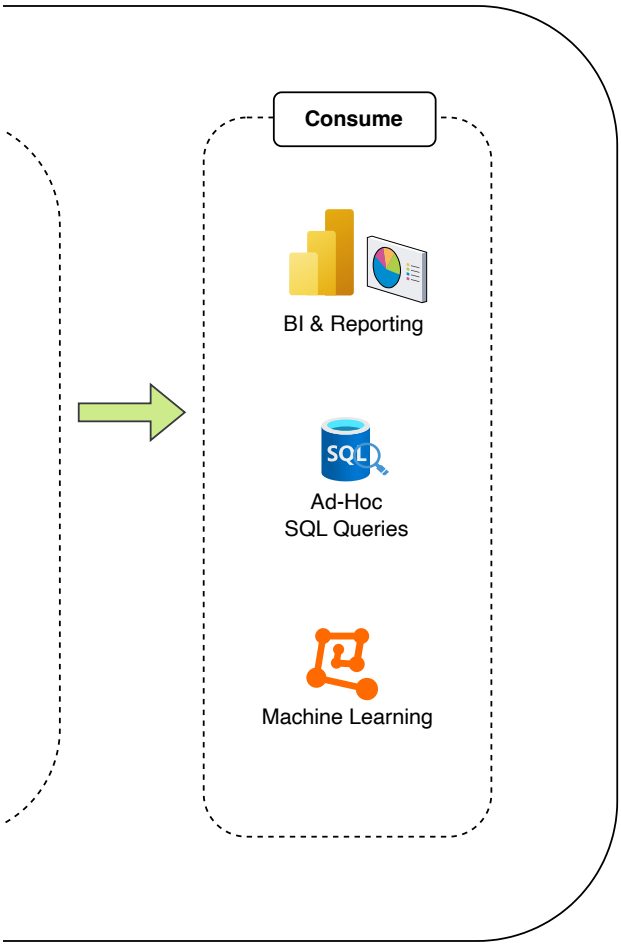
High Level Architecture




Consume
BI & Reporting
SQL
Ad-Hoc
SQL Queries
Machine Learning

Data Warehouse Naming Rules

General Principles

- Naming Convention : Use snake_case, with lower letters and Underscore(_) to separate words.
- Language : Use English for all name
- Avoid Reserved Word : Do not use SQL reserved words as object names.

Table Naming Conventions

1. Bronze Rules

- All names must start with source system name, and table name must match their original names without renaming.
- <sourcesystem>_<entity>
  - <sourcesystem> : Name of source system (ex : crm , erp)
  - <entity> : Exact table name from the source system.
  - Ex: crm_customer_info -> Customer information from the CRM system.

2. Silver Rules

- All names must start with source system name, and table name must match their original names without renaming.
- <sourcesystem>_<entity>
  - <sourcesystem> : Name of source system (ex : crm , erp)
  - <entity> : Exact table name from the source system.
  - Ex: crm_customer_info -> Customer information from the CRM system.

3. Gold Rules

- All names must use meaningful, business-aligned names for table, starting with the catogory prefix.
- <sourcesystem>_<entity>
  - <sourcesystem> : Describes the role of the table, such as "dim"(dimension) or fact(fact table)
  - <entity> : Descriptive name of the table, aligned with the business domain(ex: customer, product, sales)
  - Ex: dim_customers --> Dimension table for customer data.
    fact_sales --> Fact table of sales transactions.

Glossary of Category Pattern

Pattern	Meaning	Example(s)
dim_	Dimesion table	dim_customer, dim_product
fact_	Fact table	fact_sales
agg_	Aggregated table	agg_customes, agg_sales_monthly

Column Naming Conventions

1. Surrogate Keys

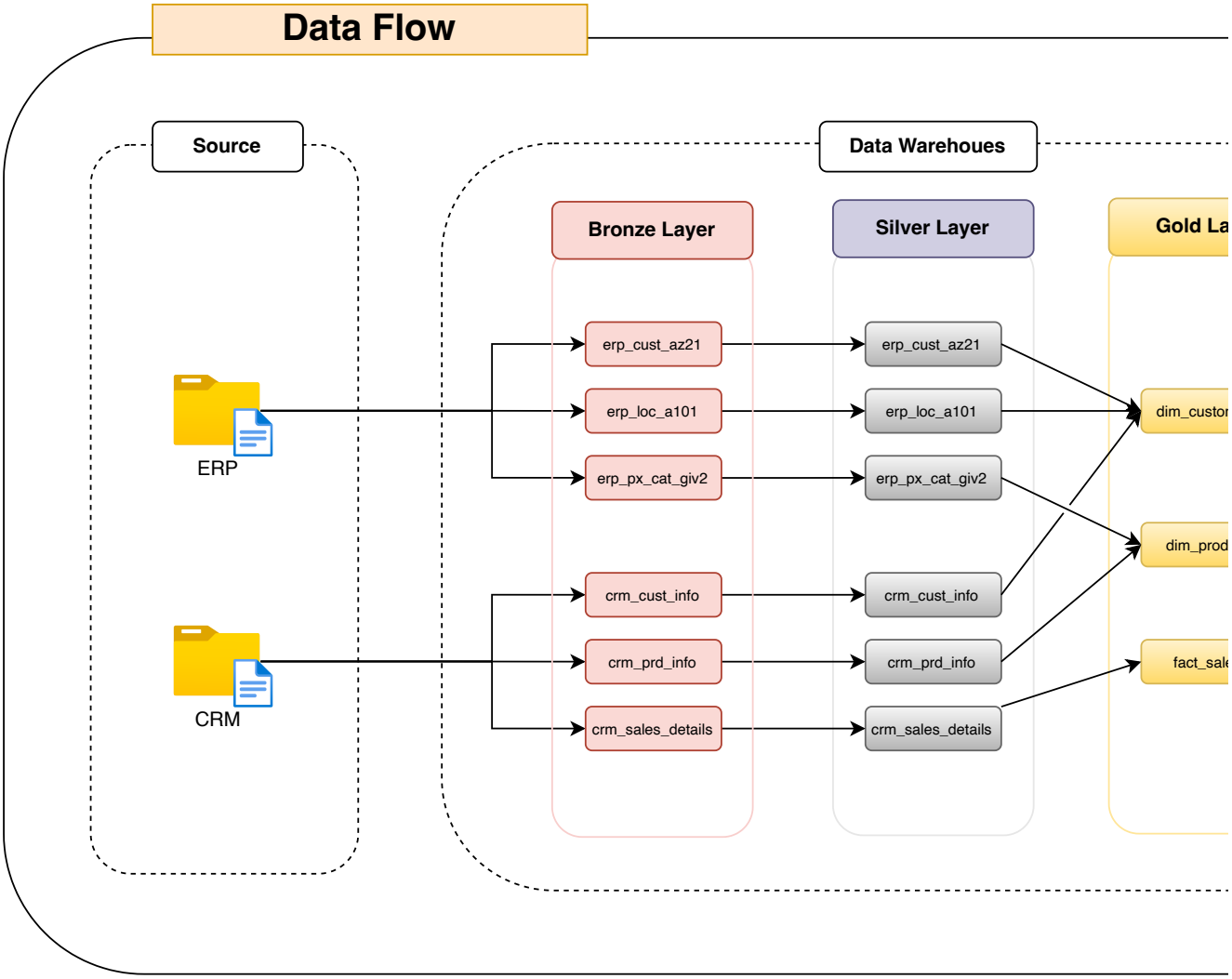
- All primary keys in dimension tables must use the suffix "_key"
- <table_name>_key
  - <table_name> : Refers to the name of the table or entity the key belongs to.
  - _key: A suffix indicating that this column is a surrogate key.
  - Ex: "customer_key" -> Surrogate key in the "dim_customers" table

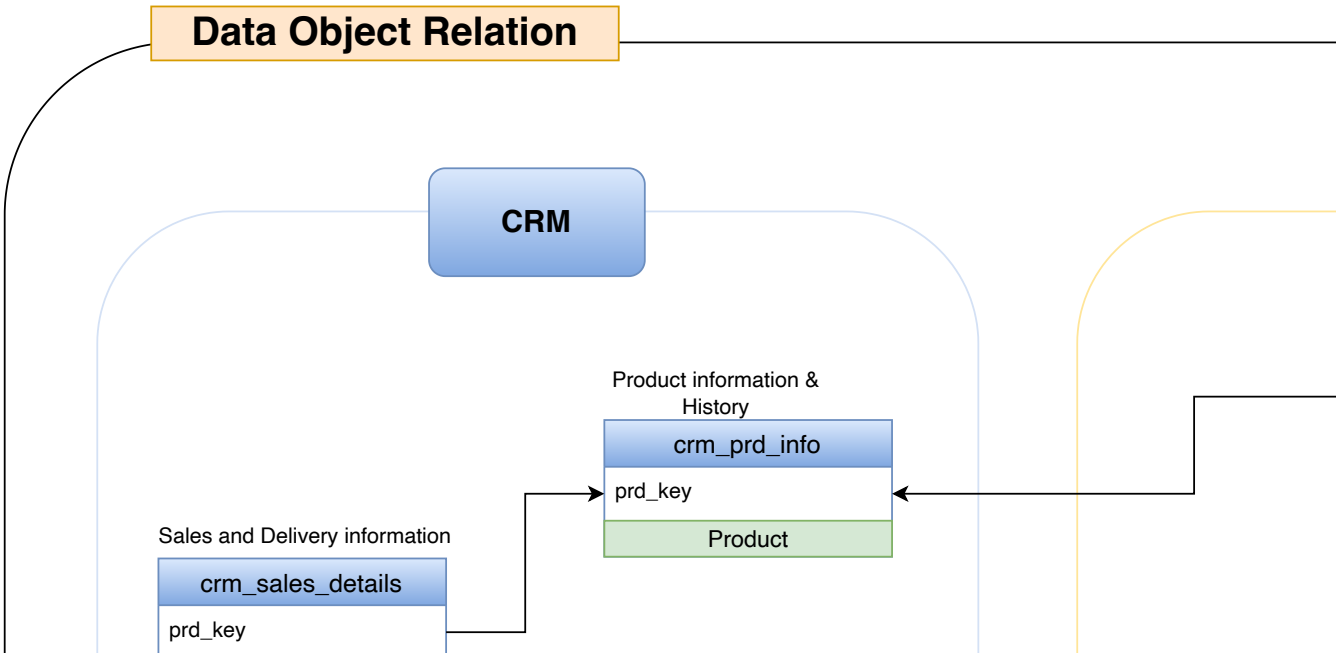
2. Technical Column

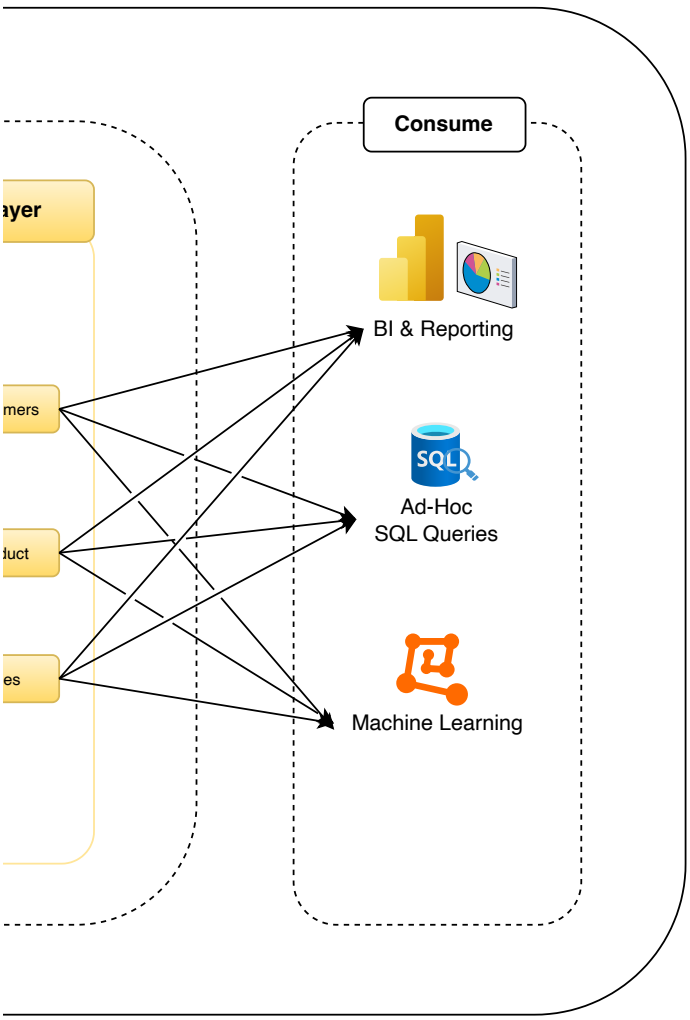
- All technical column must start with the prefix "dwh_", followed by a descriptive name indicating the column's purpose.
- dwh_<column_name>
  - dwh_: Prefix exclusive for system-generated metadata
  - <column_name> : Descriptive name indicating the column's purpose.
  - Ex: "dwh_load_date" -> System-generated column used to store the date when the record was load.

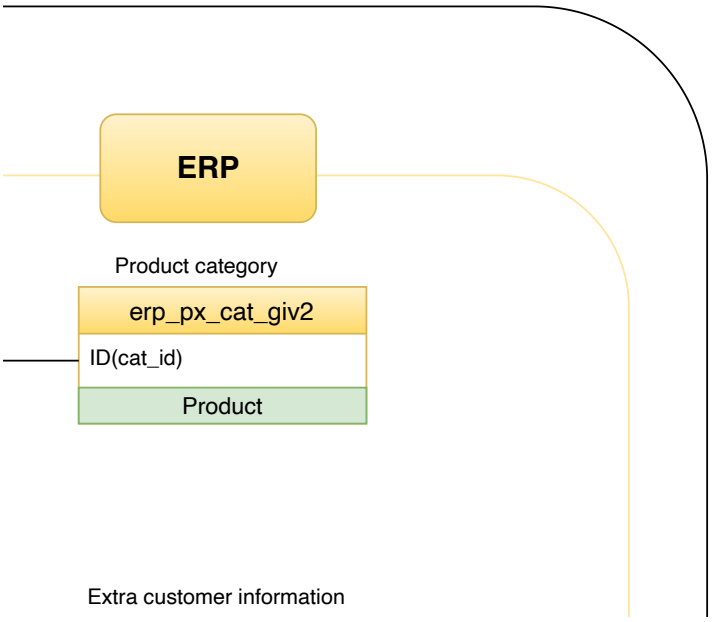
3. Store Procedure

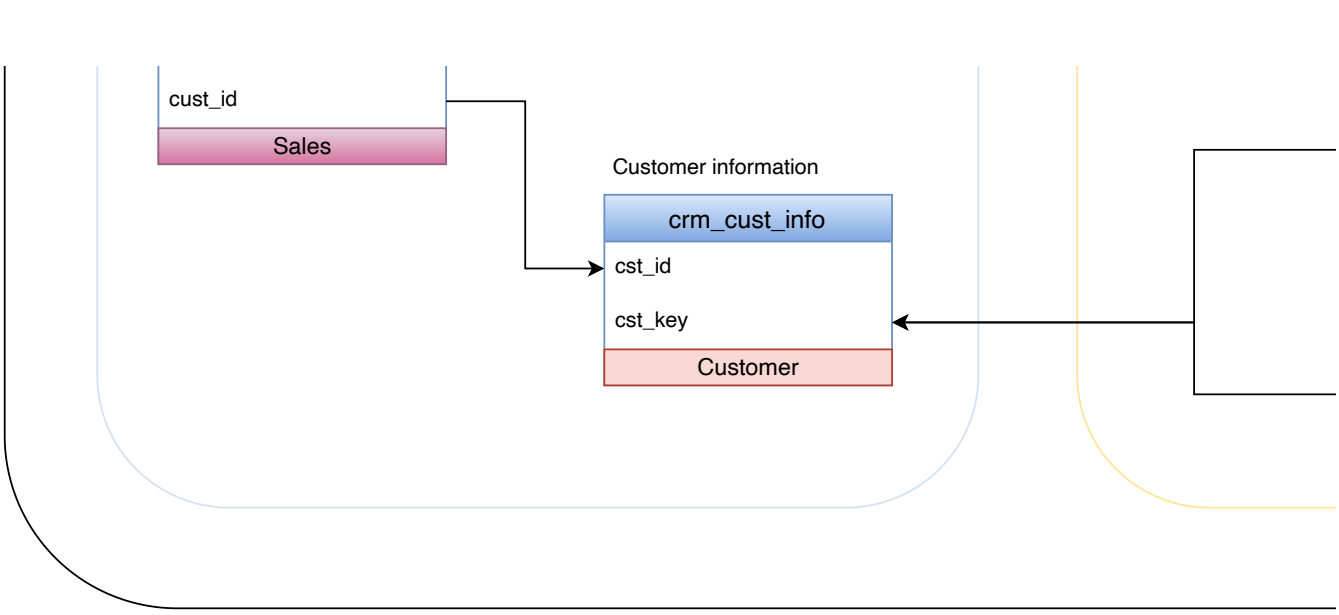
- All store procedure use for loading data must following the naming pattern.
- load_<layer>
  - load_: Prefix exclusive for procedure
  - <layer> : Represents the layer being loaded, such as "brozne", "silver", "gold"
  - Ex: load_bronze -> Store procedure for loading data in to the Bronze layer
    load_silver -> Store procedure for loading data in to the Silver layer

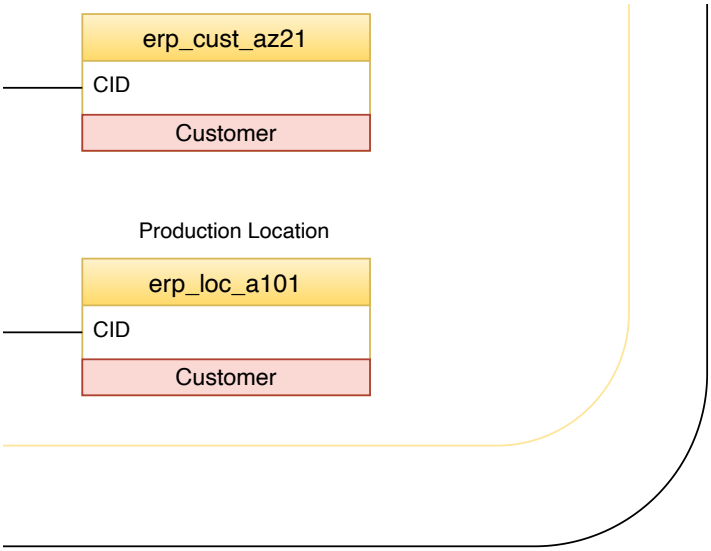
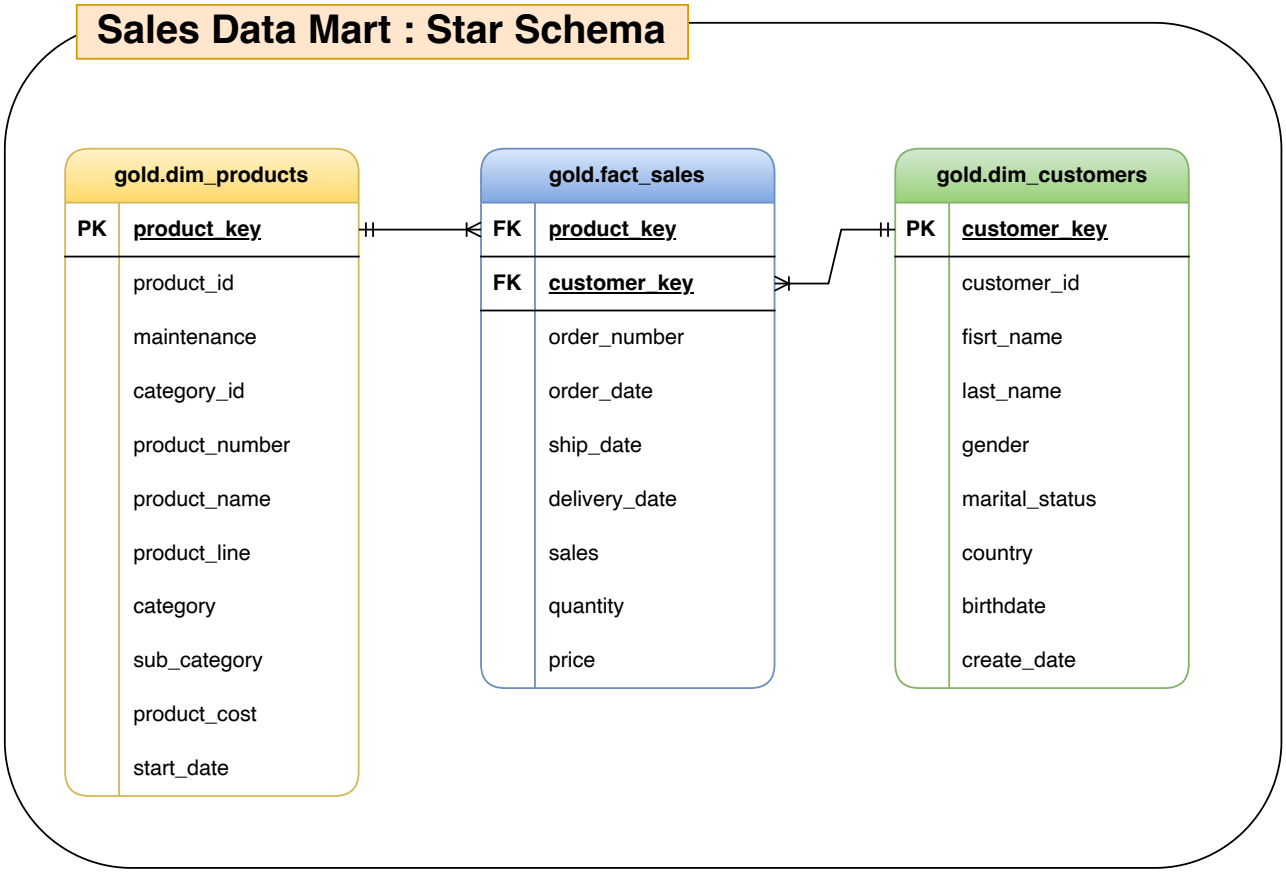

Data Flow
Source
Data Warehoues
Bronze Layer
Silver Layer
Gold La
ERP
CRM
erp_cust_az21
erp_loc_a101
erp_px_cat_giv2
crm_cust_info
crm_prd_info
crm_sales_details
dim_custor
dim_prod
fact_sale


Data Object Relation
CRM
Product information &
History
crm_prd_info
prd_key
Product
Sales and Delivery information
crm_sales_details
prd_key


Consume
ayer
mers
duct
es
BI & Reporting
Ad-Hoc
SQL Queries
Machine Learning


ERP
Product category
erp_px_cat_giv2
ID(cat_id)
Product
Extra customer information


cust_id
Sales
Customer information
crm_cust_info
cst_id
cst_key
Customer


erp_cust_az21
CID
Customer
Production Location
erp_loc_a101
CID
Customer

Sales Data Mart : Star Schema