

Project

Data Warehouse -

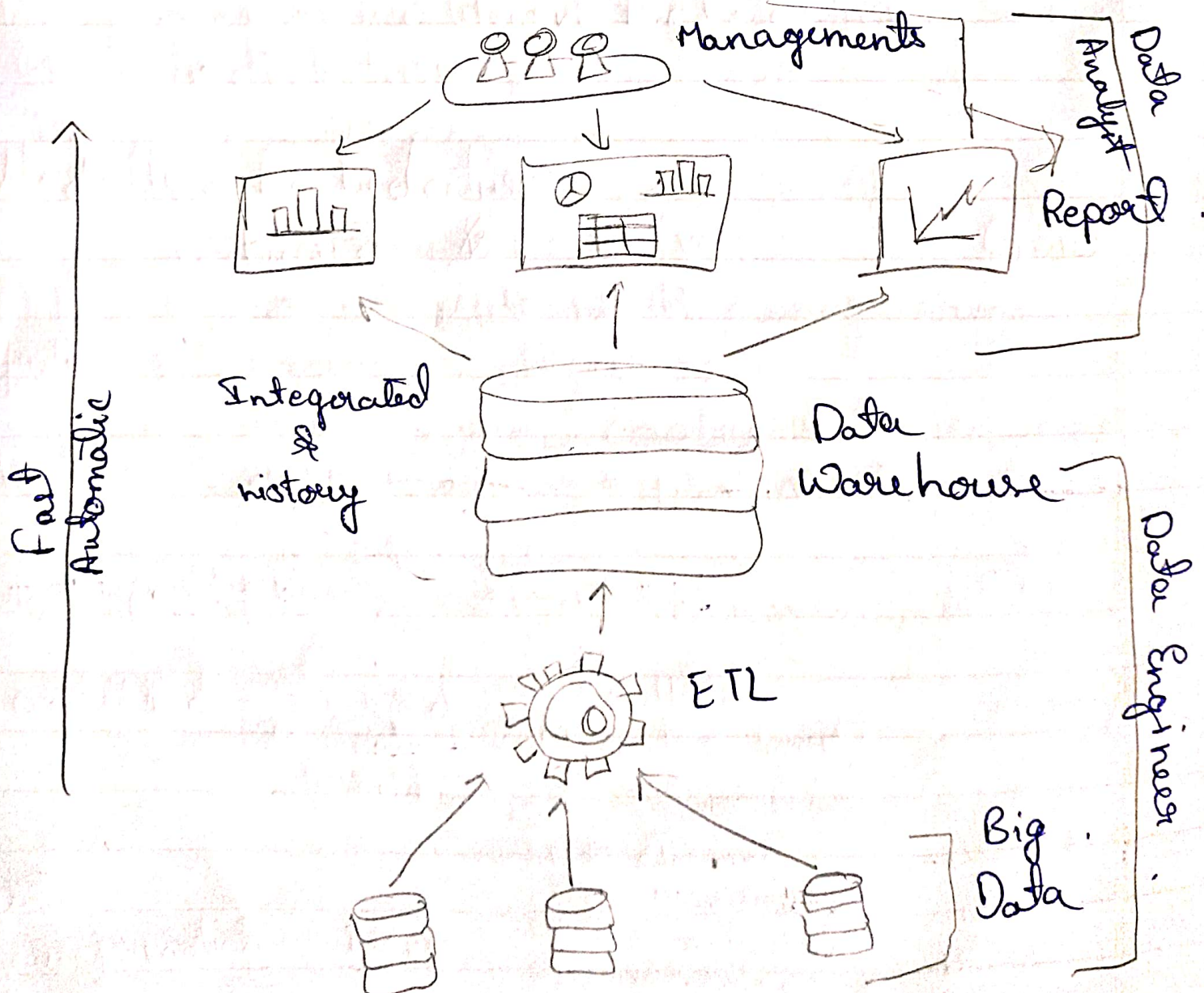
it focus on business area like sales, production

it integral multiple sources

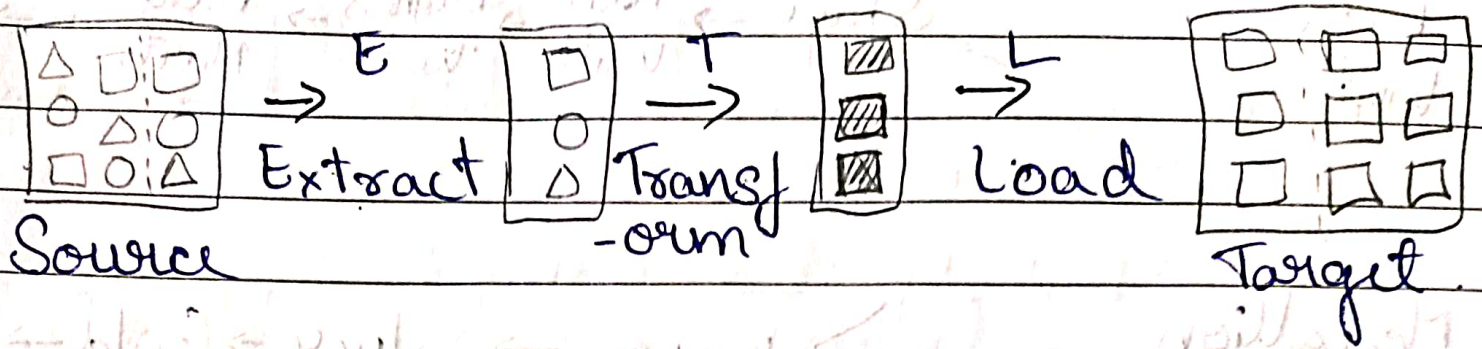
to keep historical data inside a warehouse

A Subject-oriented, Integrated, Time-variant and non-volatile collection of data in a support of management decision making process.

It means once the data enter the warehouse it cannot be deleted or modified.



ETL Process



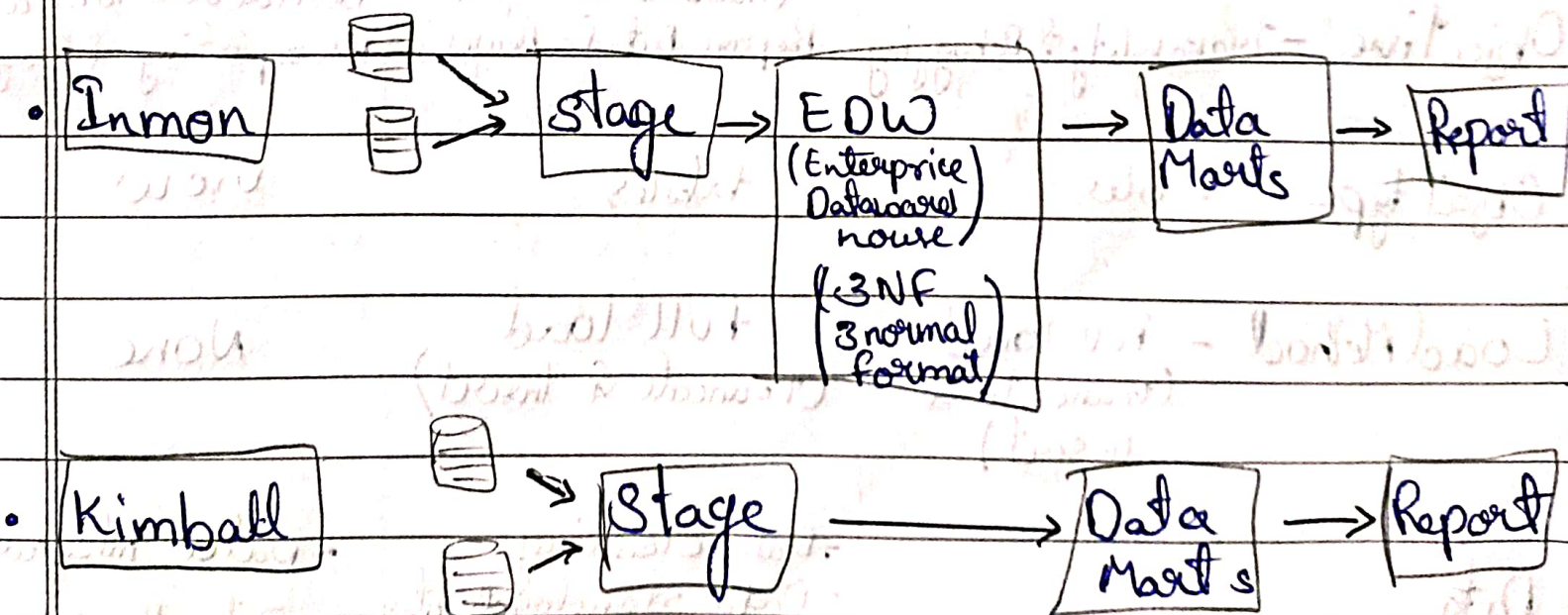
OLTP (Online Transactional Processing)

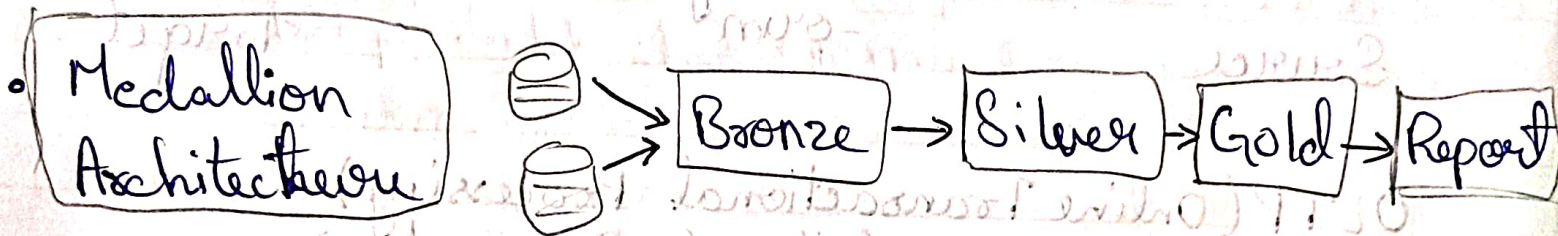
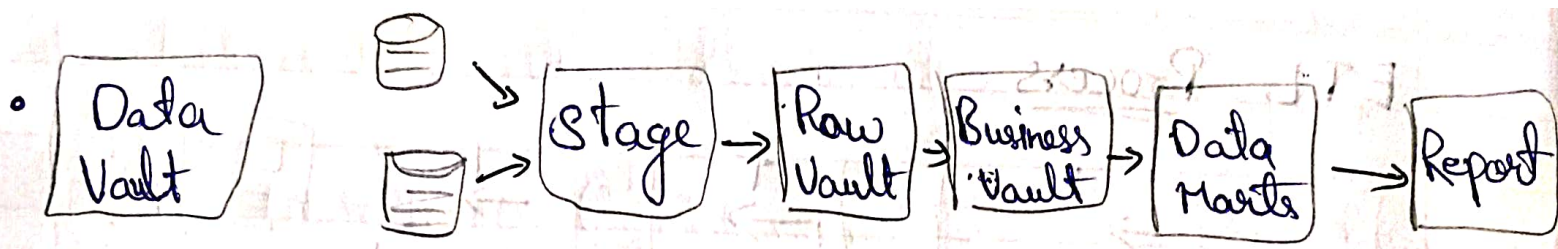
OLAP (Online Analytical Processing)

Data Architecture:-

- Data warehouse
- Data lake
- Data lakehouse
- Data mesh

Data warehouse:- to build data Architecture this are the methods





We go use Medallion Architecture to build data architecture.

Medallion Architecture

Bronze

Silver

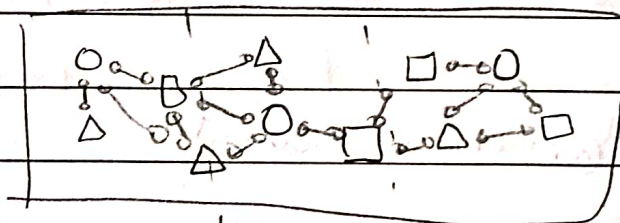
Gold

Definition - Raw, Unprocessed data as-is from sources	Clean & Standardized data (Intermediate layer)	Business Ready data
Objective - Traceability & Debugging	Prepare Data for Analysis	Provide data to be consumed for reporting & Analytics
Object type - tables	tables	View
Load Method - Full load (truncate & Insert)	Full load (truncate & Insert)	None
Data Transformation - None	- Data Cleaning - Data Standardization - Data Normalization - Derived Columns - Data Enrichment	- Data Integration - Data Aggregation - Business logic & Rules

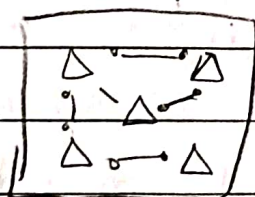
	Bronze	Silver	Gold
Data Modeling -	None	None	<ul style="list-style-type: none"> - Start Schema - Aggregated Object - Flat tables
target Audience -	Data Engineers	<ul style="list-style-type: none"> - Data Engineers - Data Analysts 	<ul style="list-style-type: none"> - Data Analysts - Business Users

SOC (Separation of concerns)

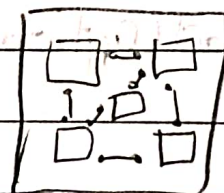
wrong
without SOC



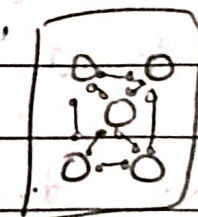
correct
with SOC



Module A



Module B



Module C