SQL Projects

ASHESH SAMRAJYA

Data Warehousing

"Organize, Structure, Prepare"

- ETL/ELT Processing
- Data Architecture
- Data Integration
- Data Cleansing
- Data Load
- Data Modeling

Exploratory Data Analysis (EDA)

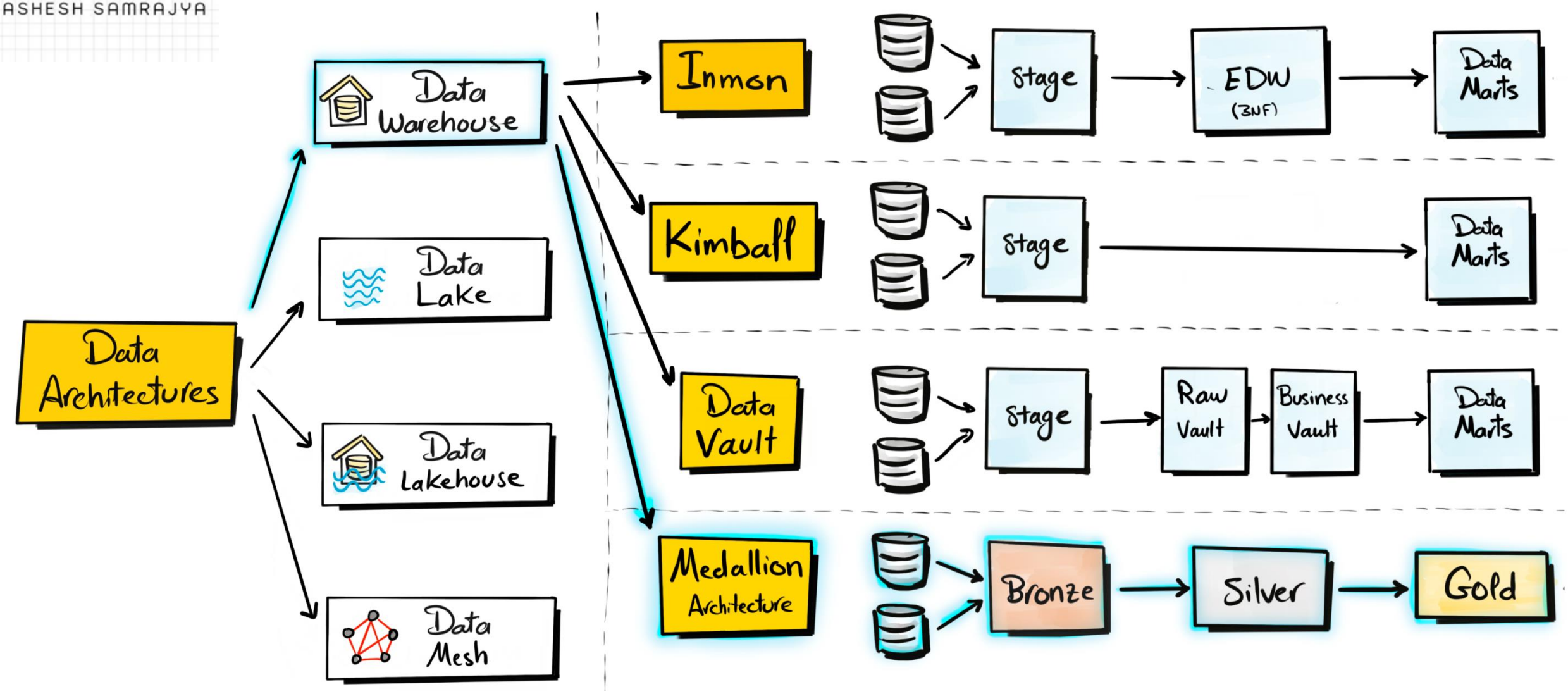
"Understand Data"

- Basic Queries
- Data Profiling
- Simple Aggregations
- Subquery

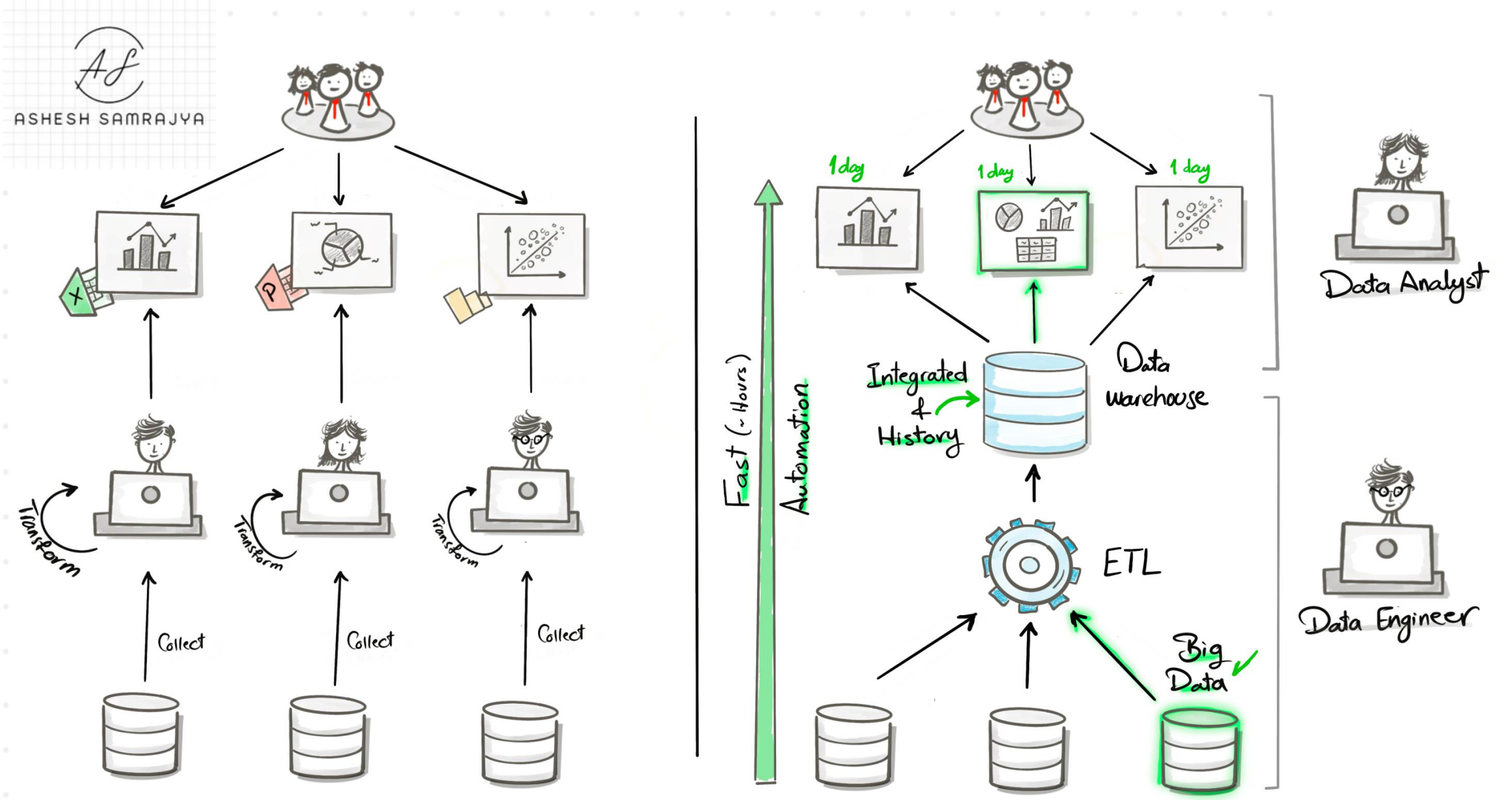
Advanced Data Analytics

"Answer Business Questions"

- Complex Queries
- Window Functions
- CTE
- Subqueries
- Reports

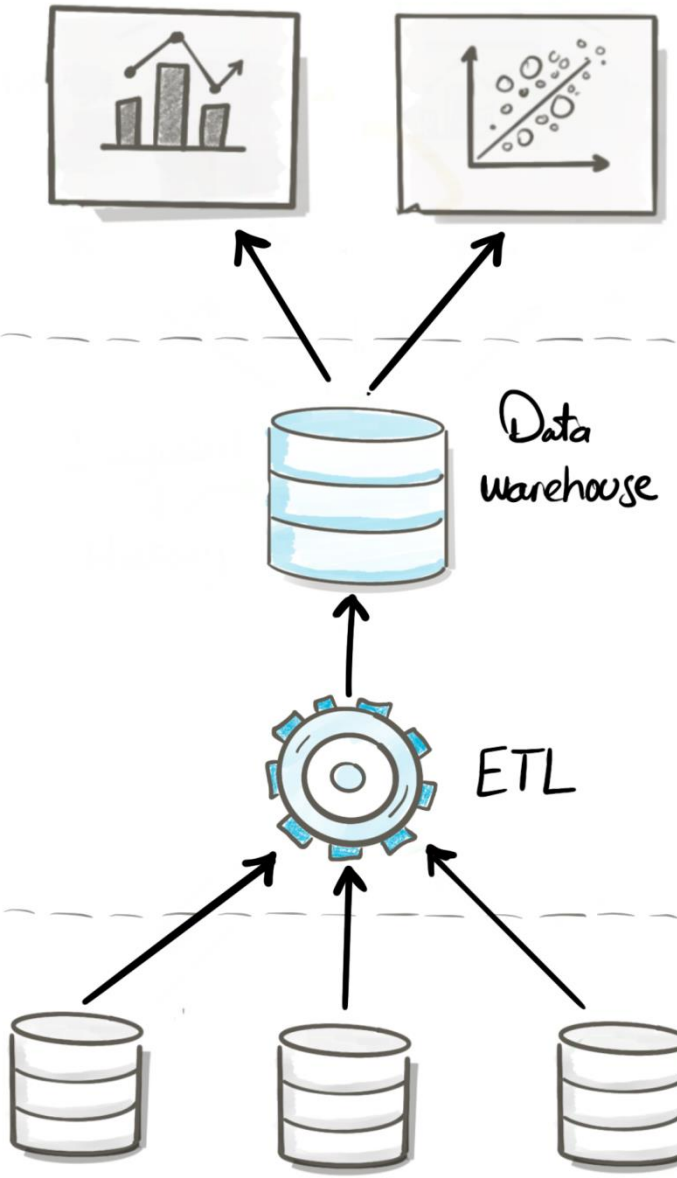

ASHESH SAMRAJYA
Data Architectures
Data Warehouse
Data Lake
Data Lakehouse
Data Mesh
Inmon
Stage
EDW
(3NF)
Data Marts
Kimball
Stage
Data Marts
Data Vault
Stage
Raw Vault
Business Vault
Data Marts
Medallion Architecture
Bronze
Silver
Gold

	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)	- Data **Cleaning** - Data **Standardization** - Data **Normalization** - **Derived Columns** - Data **Enrichment**	- Data **Integration** - Data **Aggregation** - **Business Logic & Rules**
Data Modeling	None (as-is)	None (as-is)	- Start Schema - Aggregated Objects - Flat Tables
Target Audience	- Data Engineers	- Data Analysts - Data Engineers	- Data Analysts - Business Users

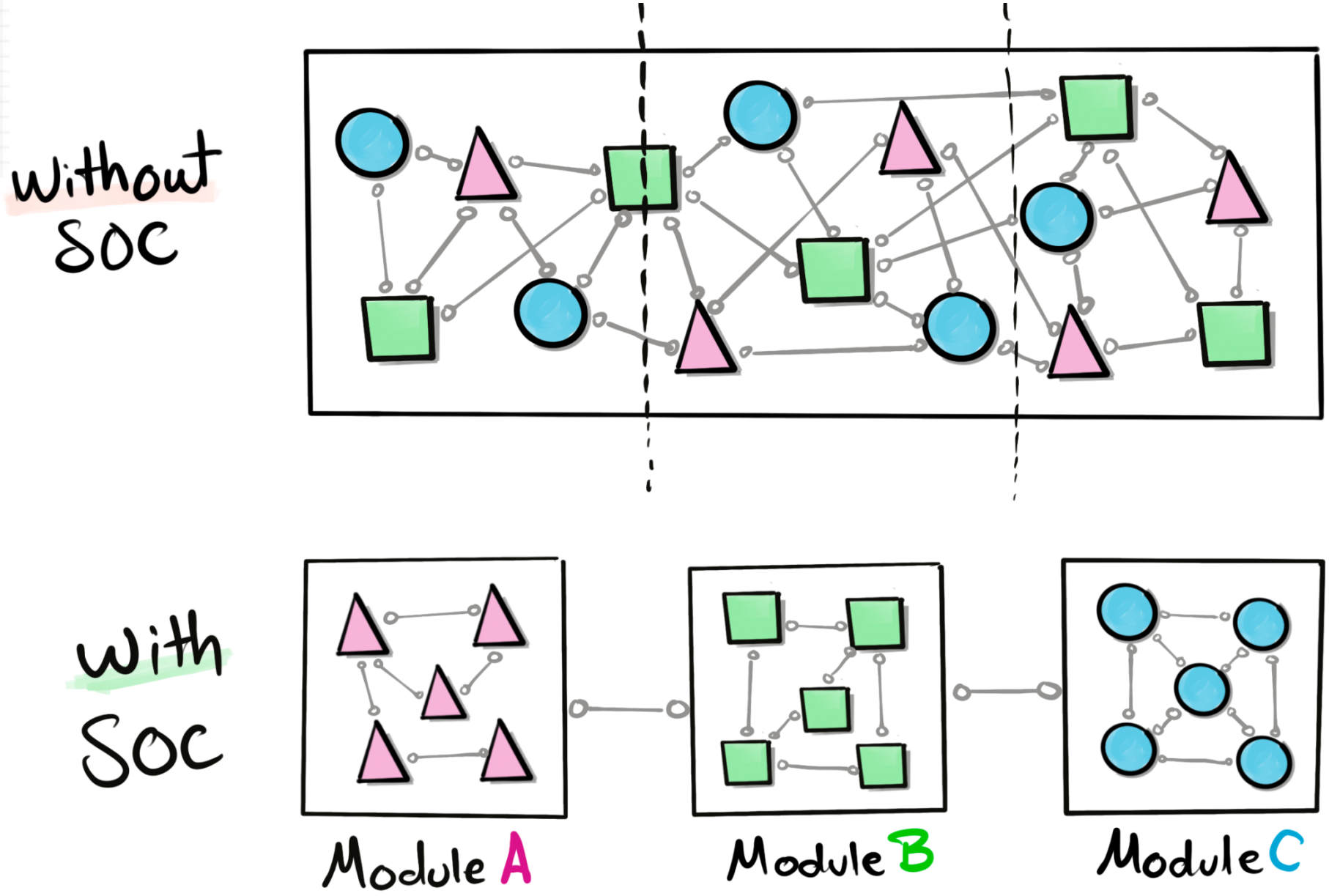


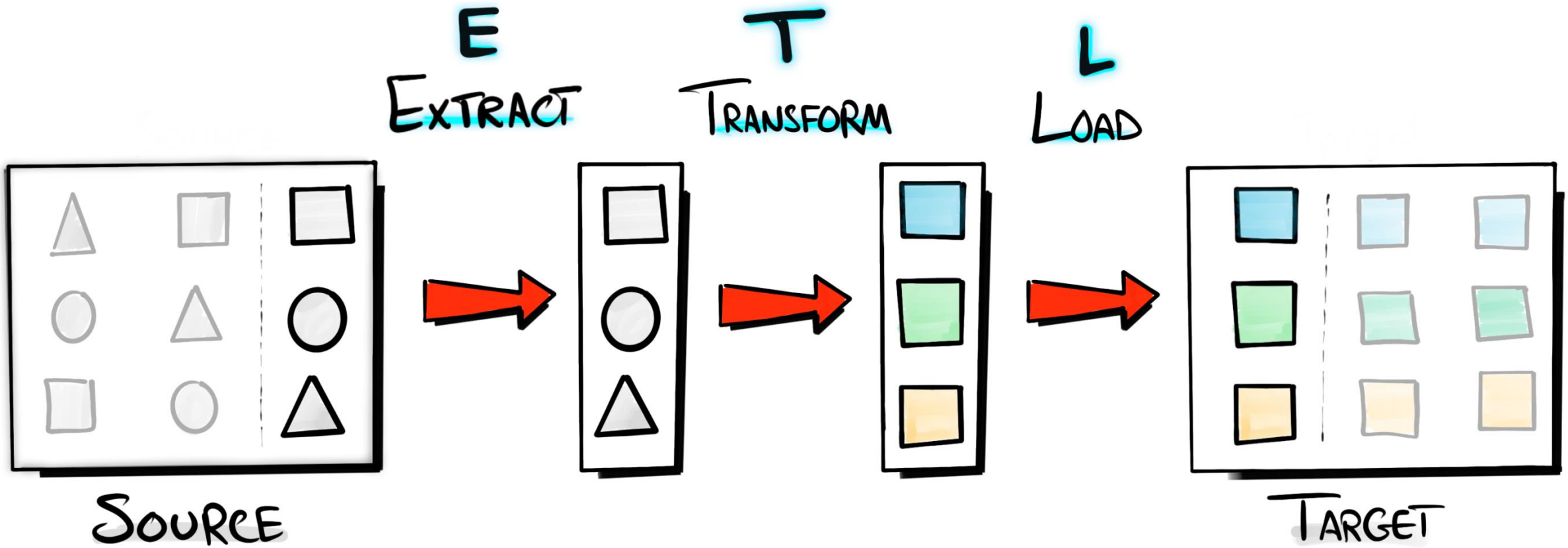
SQL Data Warehouse Project | Why we need it?

ASHESH SAMRAJYA

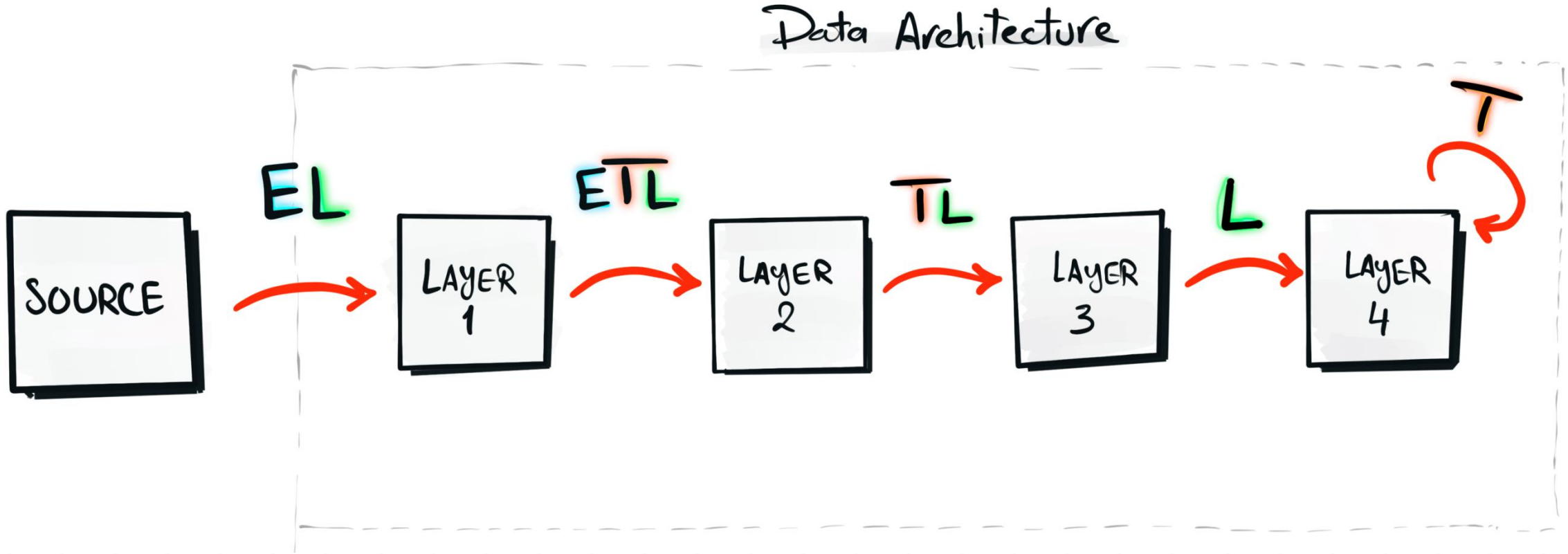


ASHESH SAMRAJYA

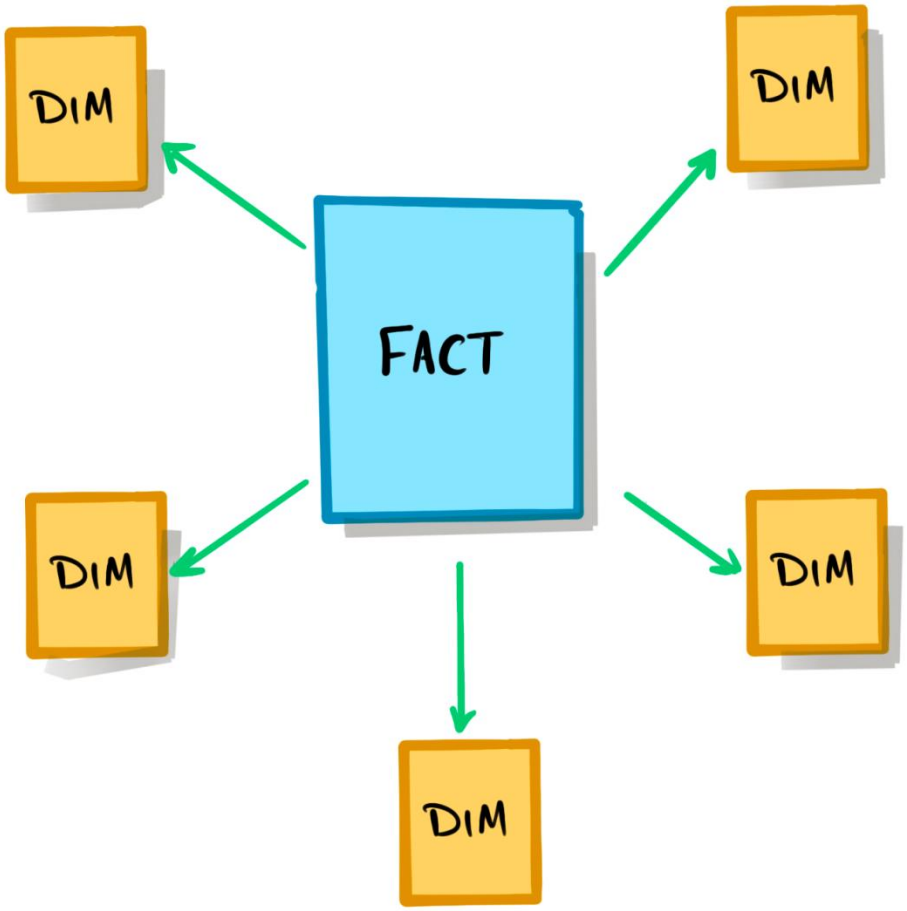



E
EXTRACT
T
TRANSFORM
L
LOAD
SOURCE
TARGET

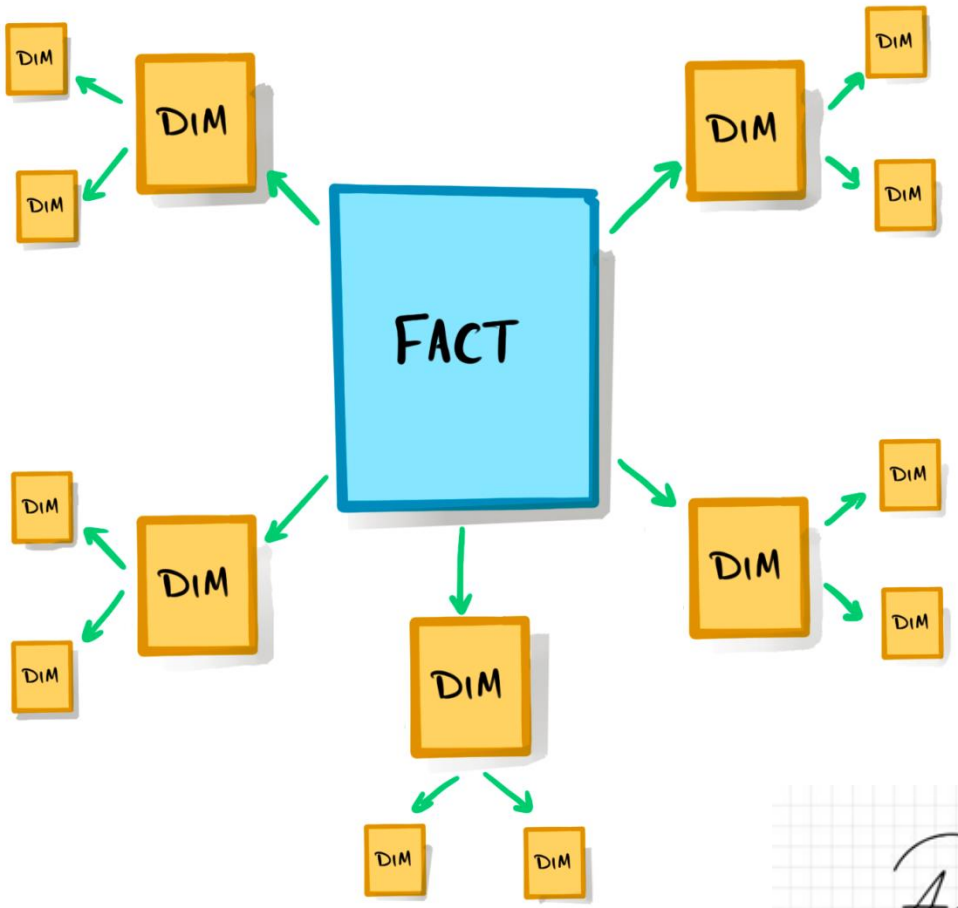
ASHESH SAMRAJYA



STAR SCHEMA



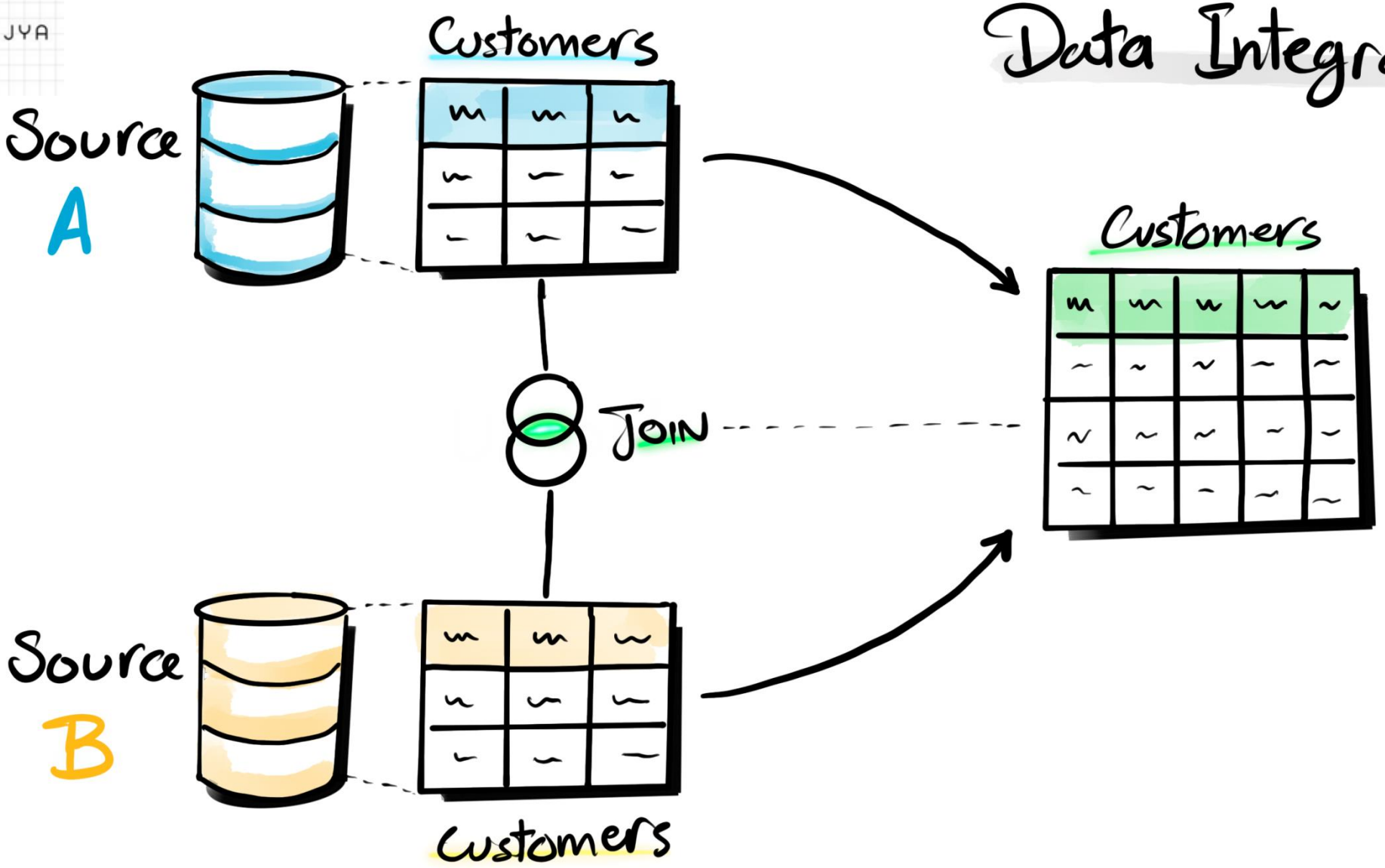
SNOWFLAKE SCHEMA



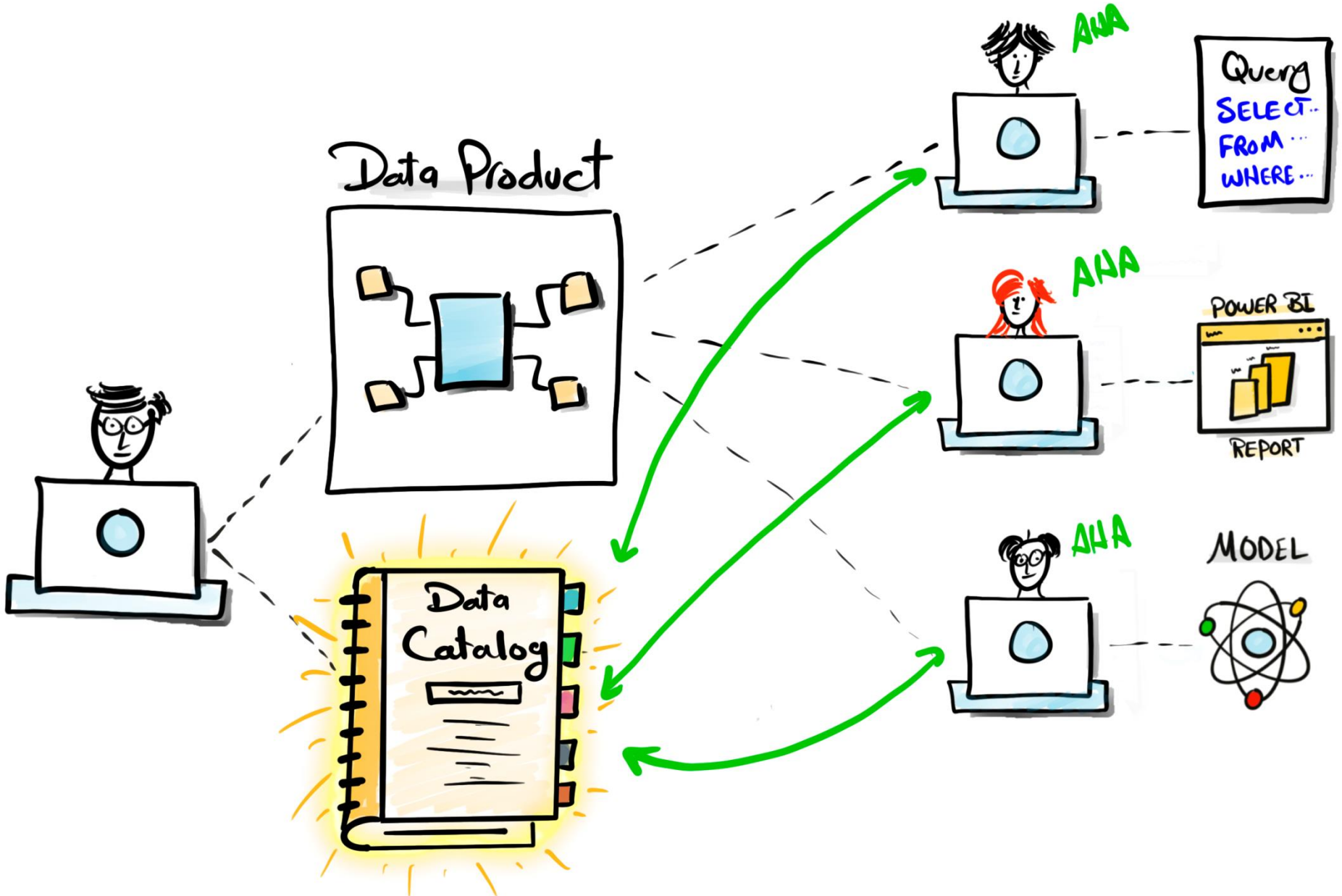
ASHESH SAMRAJYA

ASHESH SAMRAJYA

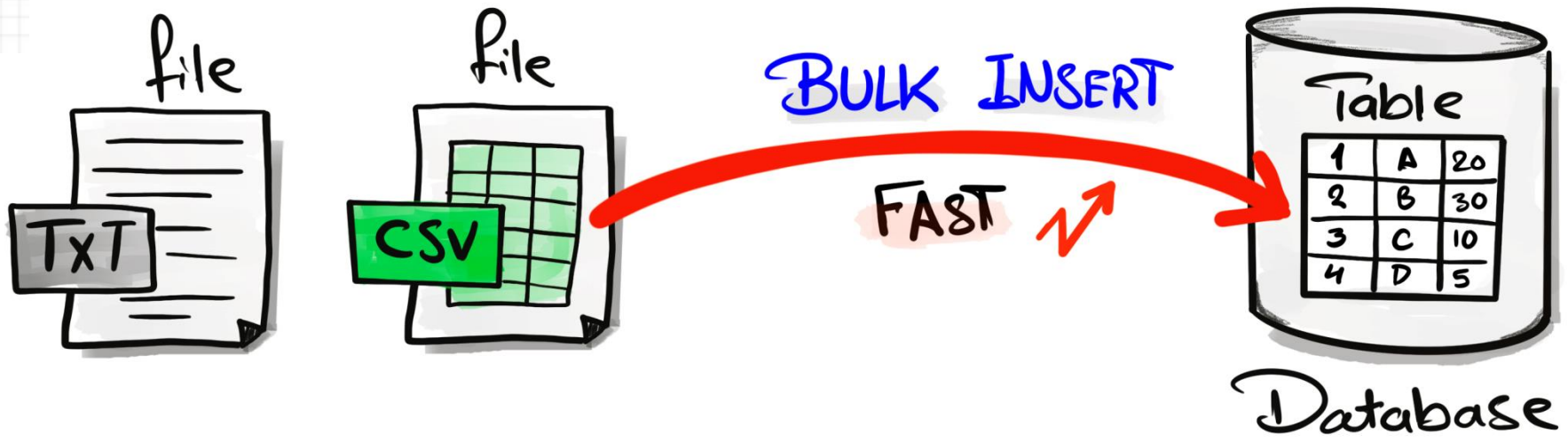
Data Integration

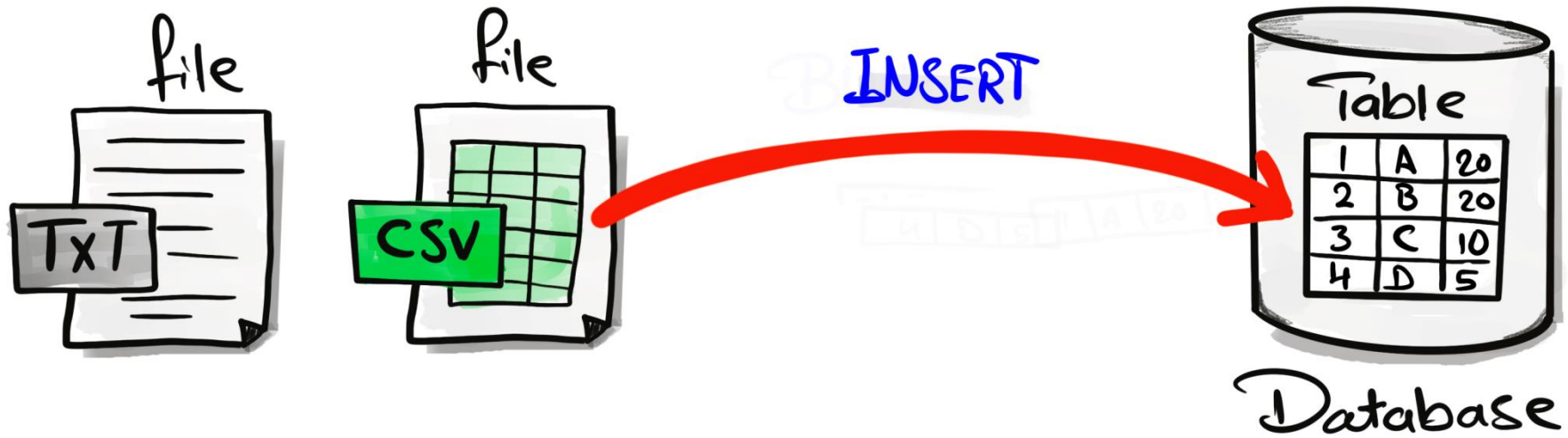


ASHESH SAMRAJYA

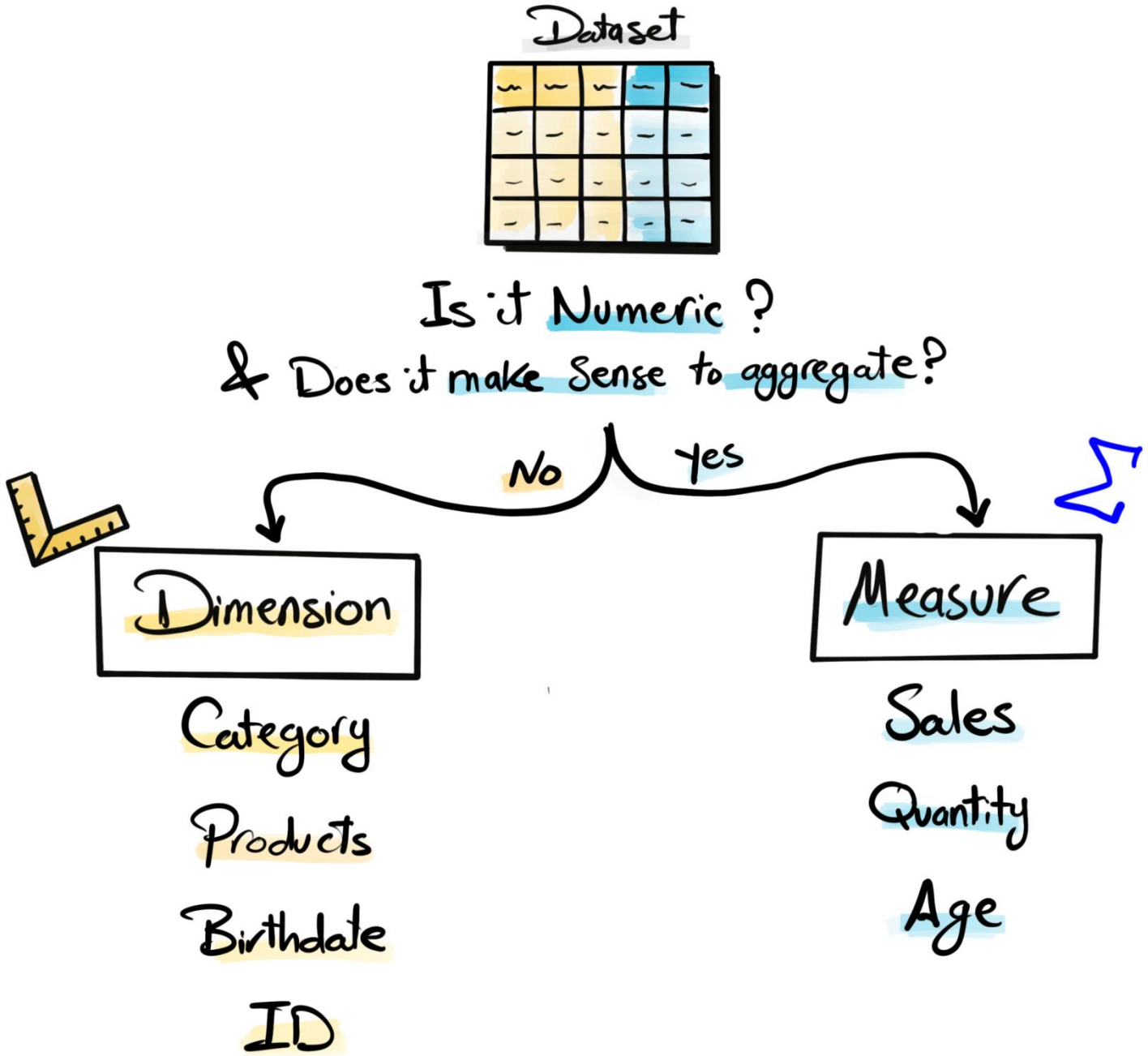


ASHESH SAMRAJYA





ASHESH SAMRAJYA



ASHESH SAMRAJYA

Dimensions Exploration

DISTINCT [Dimension]

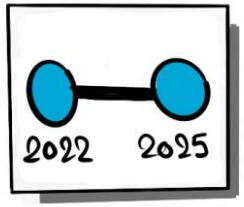
DISTINCT Country
DISTINCT Category
DISTINCT Product

⋮
A
B
B
C
A
B
C
A
⋮

⇒

A
B
C

ASHESH SAMRAJYA



Date Exploration

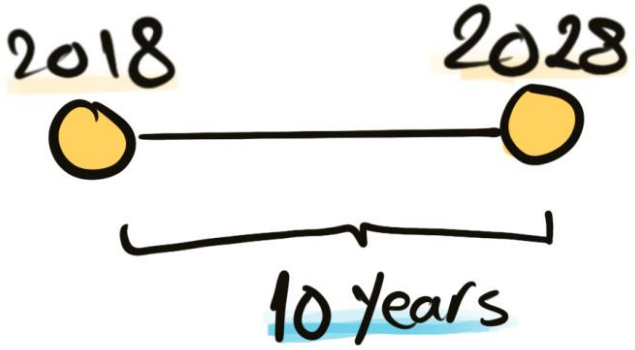
MIN/MAX [Date Dimension]

MIN Order_date

MAX Create_date

MIN Birthdate

2019
2020
2018
2018
2022
2023
2023
2023
2022



DATEDIFF

ASHESH SAMRAJYA

999

Measures Exploration

Σ [Measure]

SUM (Sales)

AVG (Price)

SUM (Quantity)

10
20
50
30
10
80
30
10

$\Rightarrow$ **240**

BIG Number

Key Metric

ASHESH SAMRAJYA

Magnitude

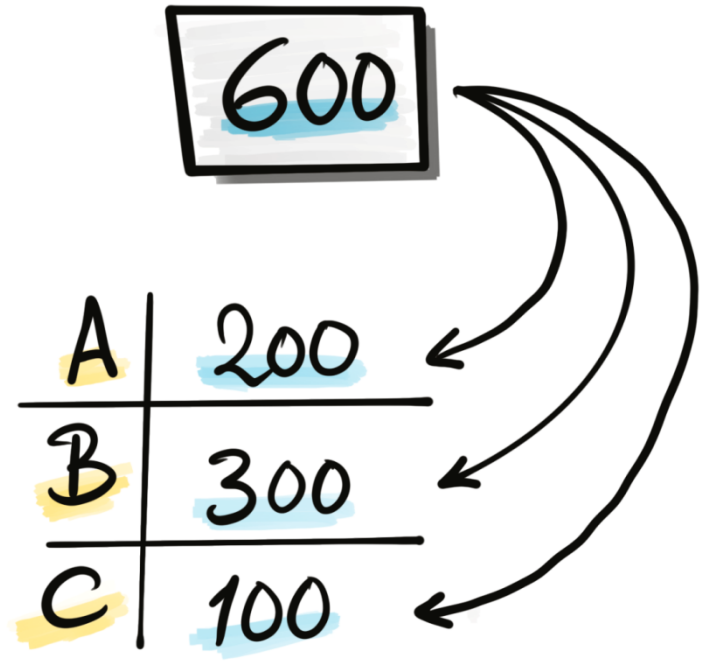
Σ [Measure] By [Dimension]

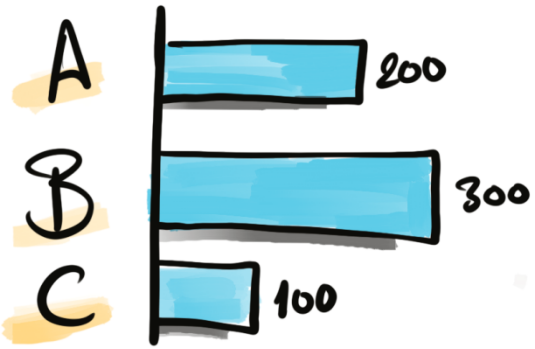
Total Sales By Country

Total Quantity By Category

Average Price By Product

Total Orders By Customer





ASHESH SAMRAJYA

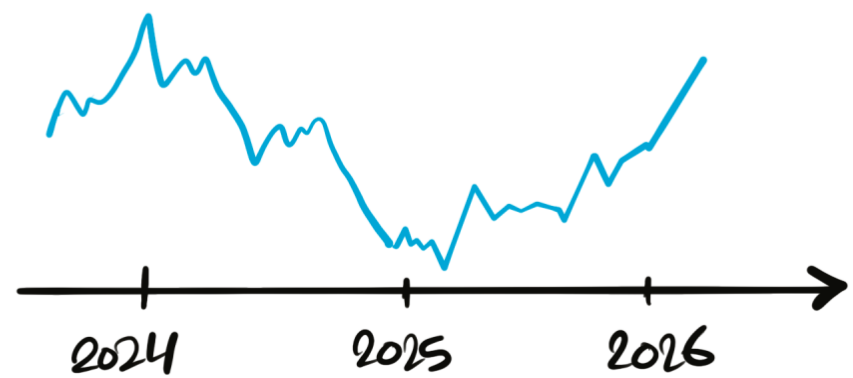
Change-Over-Time "Trends"

Σ [Measure] By [Date Dimension]

Total Sales By Year

Average Cost By Month

2024	300
2025	100
2026	200



ASHESH SAMRAJYA

Cumulative Analysis

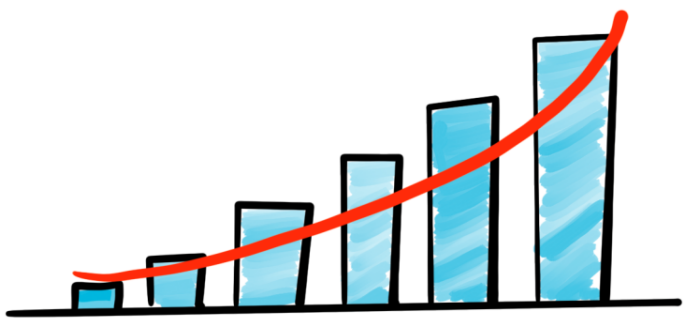
Σ [Cumulative Measure] By [Date Dimension]

Running Total Sales By Year

Moving Average of Sales By Month

Cumulative

2024	300	300
2025	100	400
2026	200	600



WINDOW FUNCTIONS

ASHESH SAMRAJYA

Performance Analysis

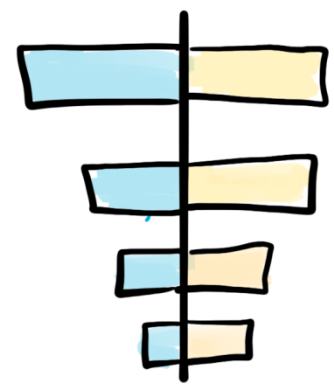
Current [Measure] – Target [Measure]

Current Sales – Average Sales

Current Year Sales – Previous Year Sales

Current Sales – lowest Sales

	Current	− Target (AVG)	Performance
A	200	200	0
B	300	200	100
C	100	200	-100



WINDOW FUNCTIONS

ASHESH SAMRAJYA

Part-to-Whole

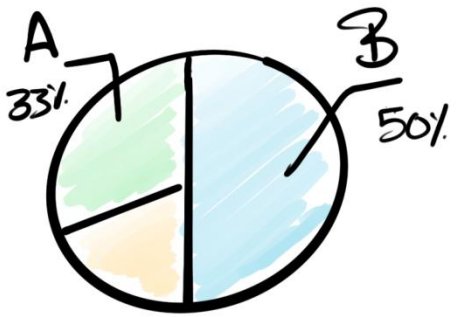
Proportional Analysis

([Measure] / Total [Measure]) * 100 By [Dimension]

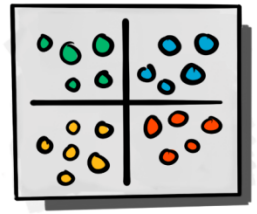
(Sales / Total Sales) * 100 By Category

(Quantity / Total Quantity) * 100 By Country

A	200	33%
B	300	50%
C	100	17%



ASHESH SAMRAJYA



Data Segmentation

[Measure] By [Measure]

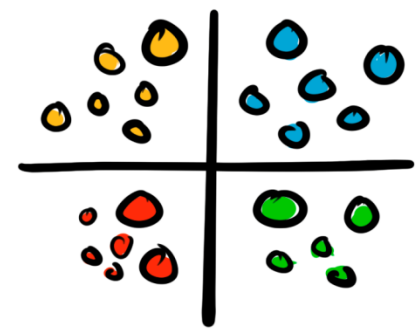
Total Products By Sales Range

Total Customers By Age

Σ

Categorize

3	50
4	100
5	150
1	200
10	250
5	300

Low	7
Medium	6
Large	15

CASE WHEN STATEMENT