

MANAGED BY  
RUBY THALIB AI & DATA CONSULTANT

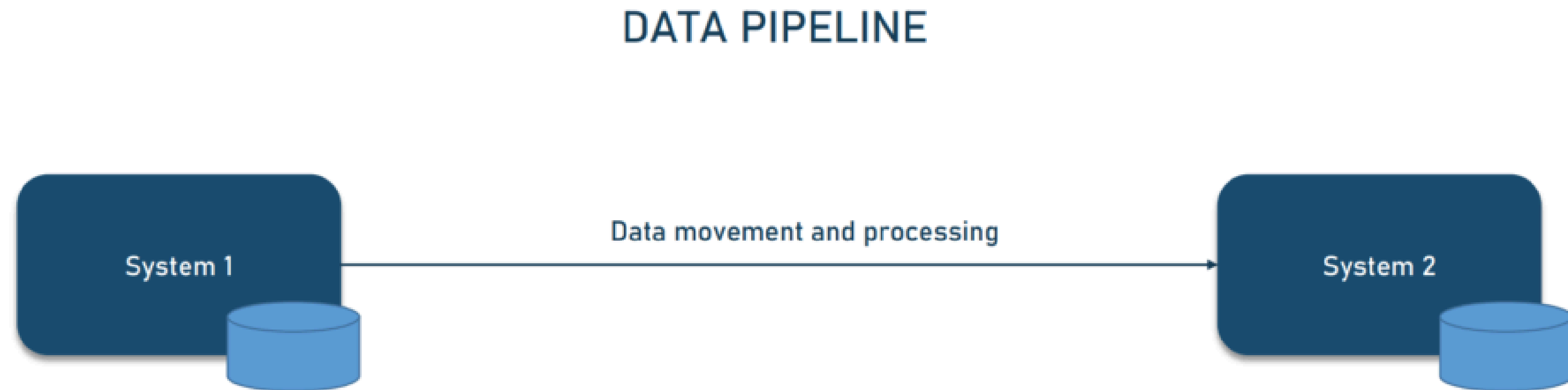
# DATA ENGINEERING

MOST SCOPE

INTERACTIVE CLASS, 2024

- **Data Pipeline**
- **ETL Concept**

- **DATA PIPELINE**



A **DATA PIPELINE** IS A SET OF TOOLS AND ACTIVITIES FOR MOVING DATA FROM ONE SYSTEM WITH ITS METHOD OF DATA STORAGE AND PROCESSING TO ANOTHER SYSTEM IN WHICH IT CAN BE STORED AND MANAGED DIFFERENTLY.

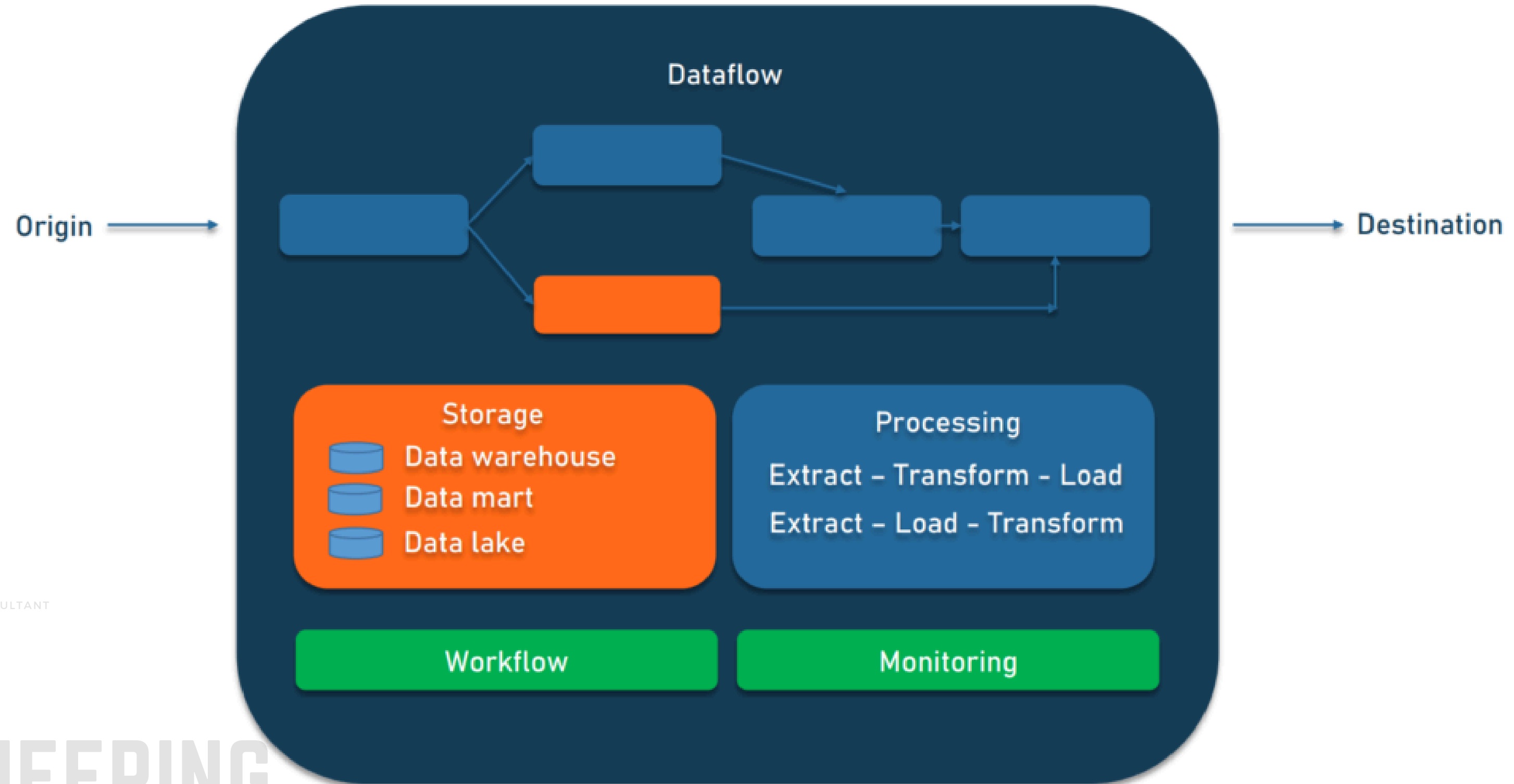
MANAGED BY  
RUBY THALIB AI & DATA CONSULTANT

**DATA  
ENGINEERING**

MOST SCOPE

INTERACTIVE CLASS, 2024

# • DATA PIPELINE



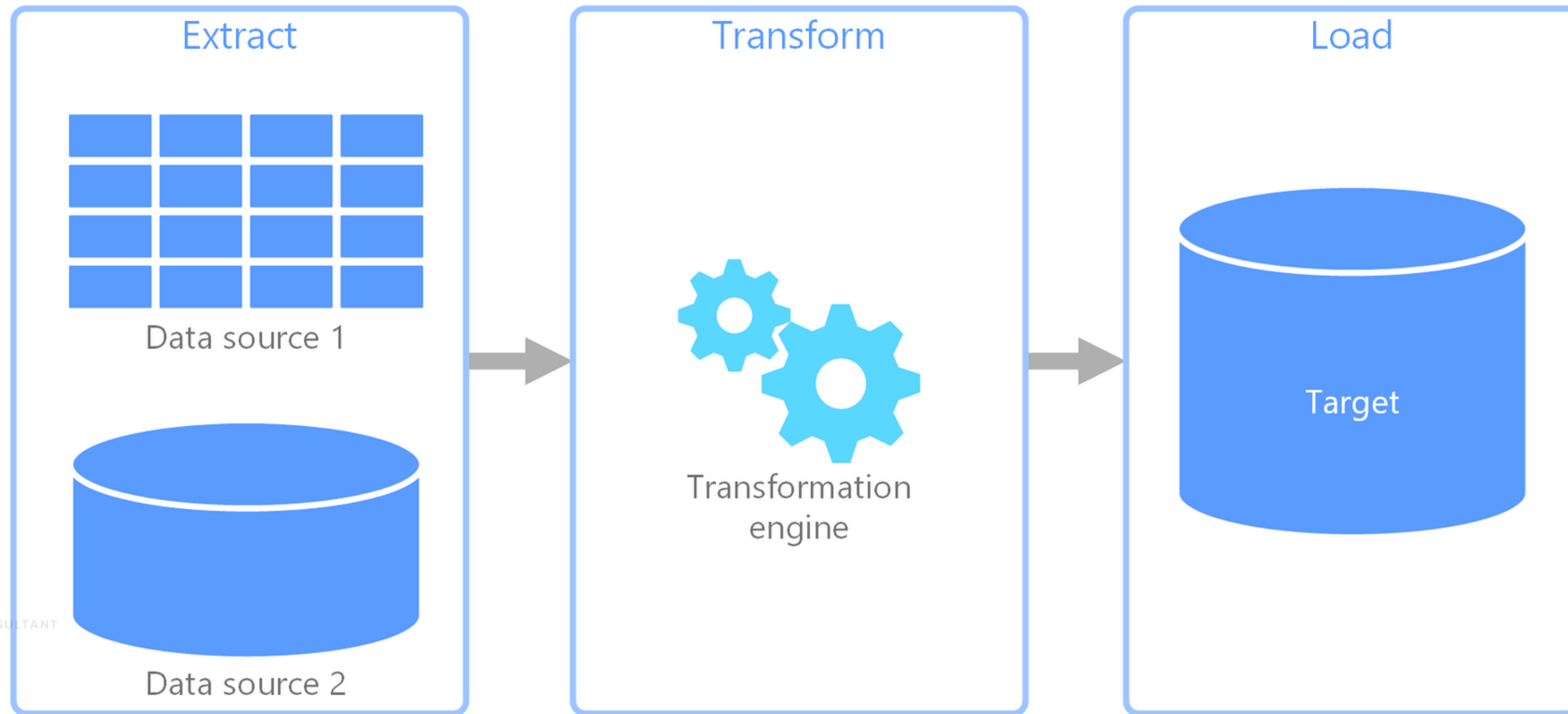
MANAGED BY  
RUBY THALIB AI & DATA CONSULTANT

# DATA ENGINEERING

MOST SCOPE

INTERACTIVE CLASS, 2024

- **ETL (Extract Transform Load)**



MANAGED BY  
RUBY THALIB AI & DATA CONSULTANT

DATA  
ENGINEERING

MOST SCOPE

INTERACTIVE CLASS, 2024

# • ETL (Extract Transform Load)



VS.



Aggregate  
Function in  
SQL

Count()

Sum()

AVG()

Max()

Min()

SUBSTRING – Return characters within another string

S	Q	L		S	e	r	v	e	r
1	2	3	4	5	6	7	8	9	10

SUBSTRING('SQL Server',5,3)

MANAGED BY  
RUBY THALIB AI & DATA CONSULTANT

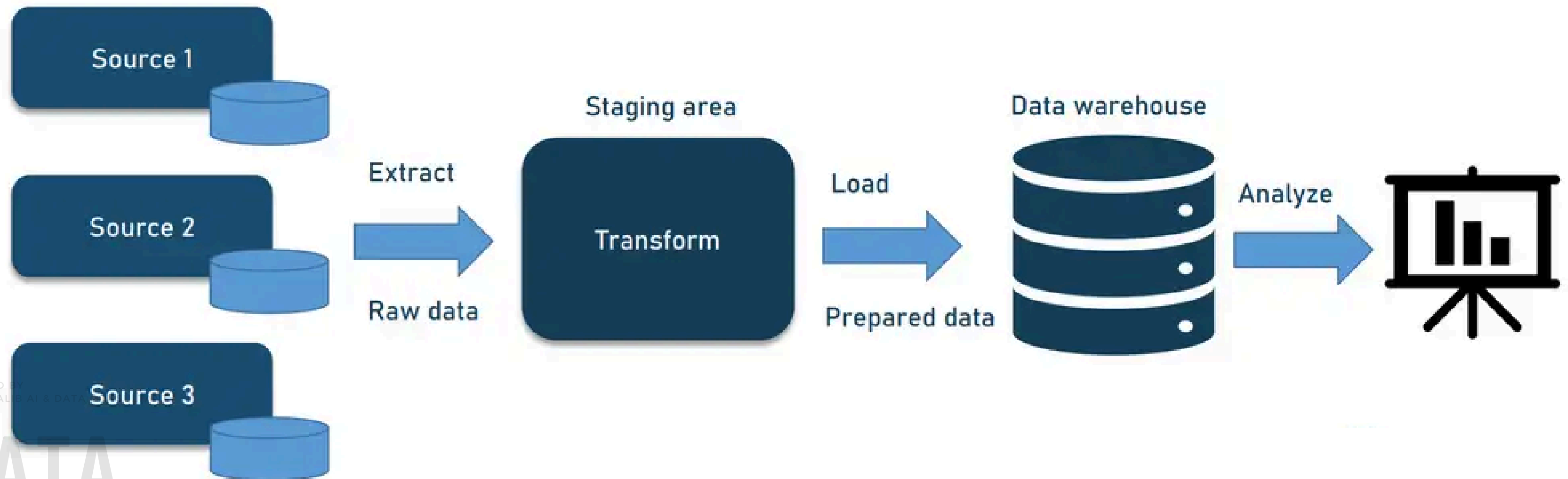
DATA  
ENGINEERING

MOST SCOPE

INTERACTIVE CLASS, 2024

- ETL Pipelines

## ETL PIPELINE



MANAGED BY  
RUBY THAI & AI & DATA

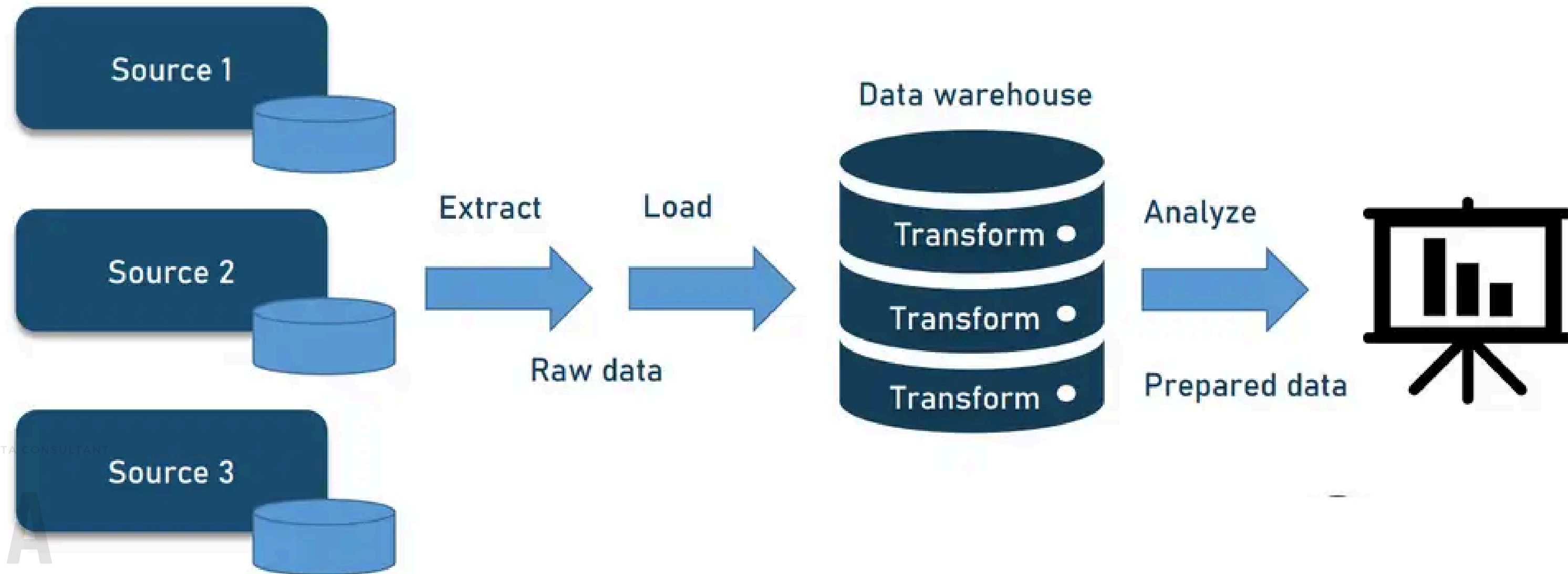
DATA  
ENGINEERING

MOST SCOPE

INTERACTIVE CLASS, 2024

- ETL Pipelines

## ELT PIPELINE



MANAGED BY  
RUBY THALIB AI & DATA CONSULTANT

DATA  
ENGINEERING

MOST SCOPE

INTERACTIVE CLASS, 2024

MANAGED BY  
RUBY THALIB AI & DATA CONSULTANT

# DATA ENGINEERING

MOST SCOPE

INTERACTIVE CLASS, 2024

# DAY 5 TRIAL :D