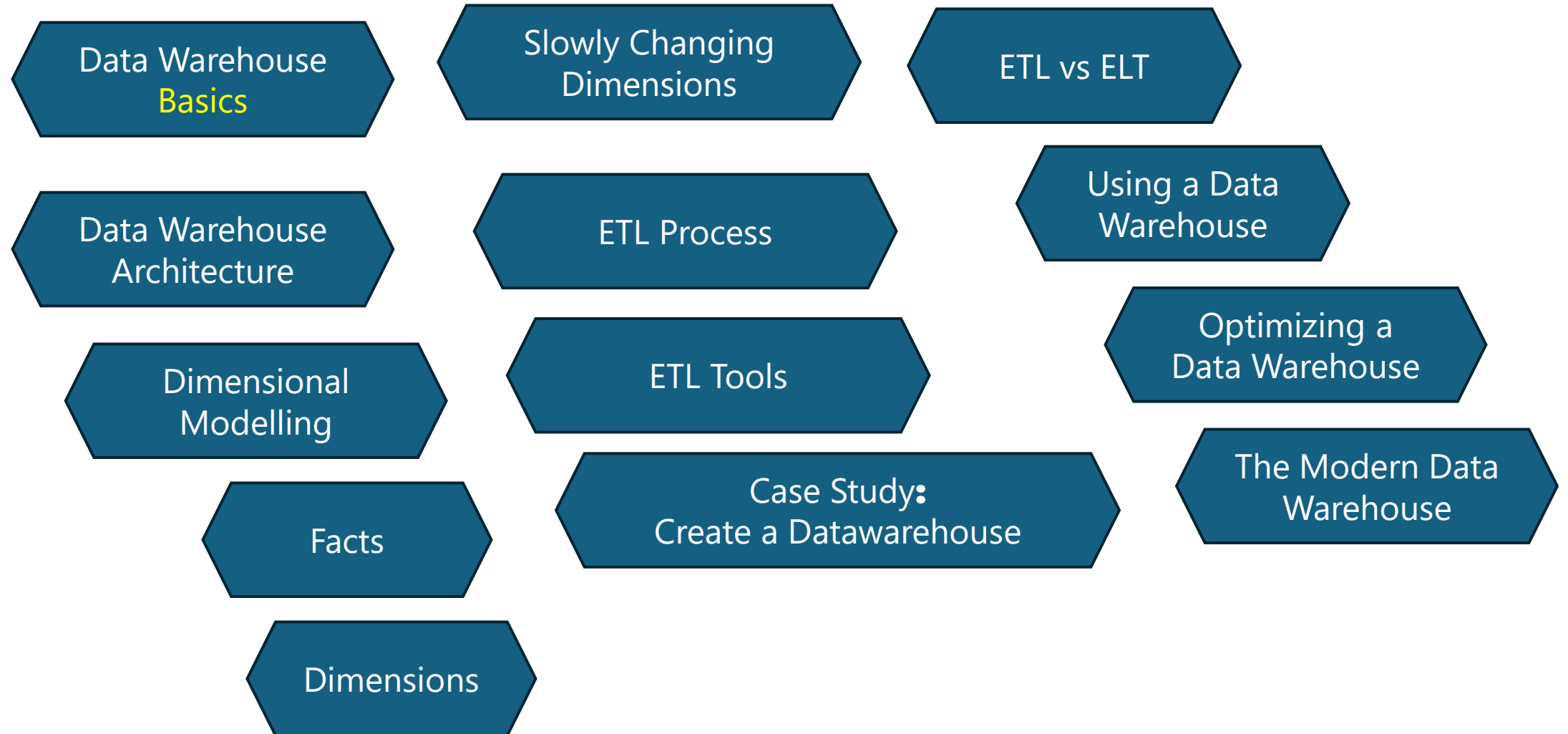


Data Warehouse The Complete Guide

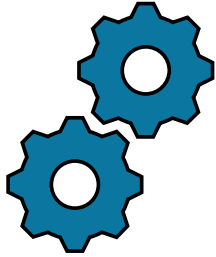
What you will learn?



Why we need Data Warehouse?

Two Purposes

Operational Data keeping

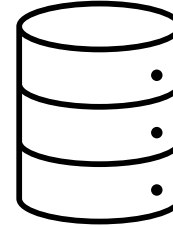


- Receive orders
- React to complaints
- Fill up Stocks

Run the
wheel

Keep the wheel running

Analytical Decision making



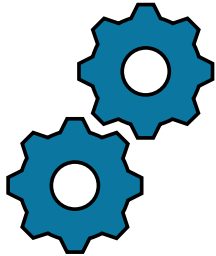
- What's the best category?
- How many sales compared to last month?
- What can be improved?

Watch the
wheel

- ☐ Evaluate Performance
- ☐ Decision-making

Two Purposes

Operational Data keeping



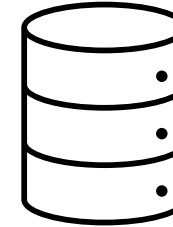
- Receive orders
- React to complaints
- Fill up Stocks

Run the wheel

Keep the wheel running

OLTP = Online Transactional processing

Analytical Decision making



- What's the best category?
- How many sales compared to last month?
- What can be improved?

Watch the wheel

- ☐ Evaluate Performance
- ☐ Decision-making

OLAP = Online Analytical processing

“Yes, we have lot of data, but we don’t use it”

“Our data is very complicated and difficult to analyze ”

“It’s spread all over different systems and difficult to access ”

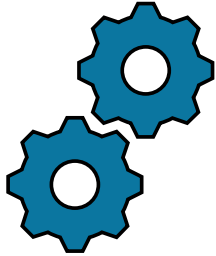
“I just want to see what is relevant ”

“We need to access data quick and easily ”

“We want to make fact-based decisions ”

Two Requirements

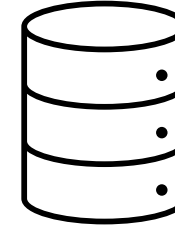
Operational Data keeping



- One record at a time
- Data input
- No long history

OLTP = Online Transactional processing

Analytical Decision making



- Thousands of record at a time
- Fast query performance
- Historical context
- usability

OLAP = Online Analytical processing

Two Requirements

Operational Data keeping



Analytical Decision making



DWH is there to address those analytical needs

- Data input
- No long history

- Fast query performance
- Historical context

Used for reporting and data analysis

OLTP = Online Transactional processing

OLAP = Online Analytical processing

What is a Data Warehouse?

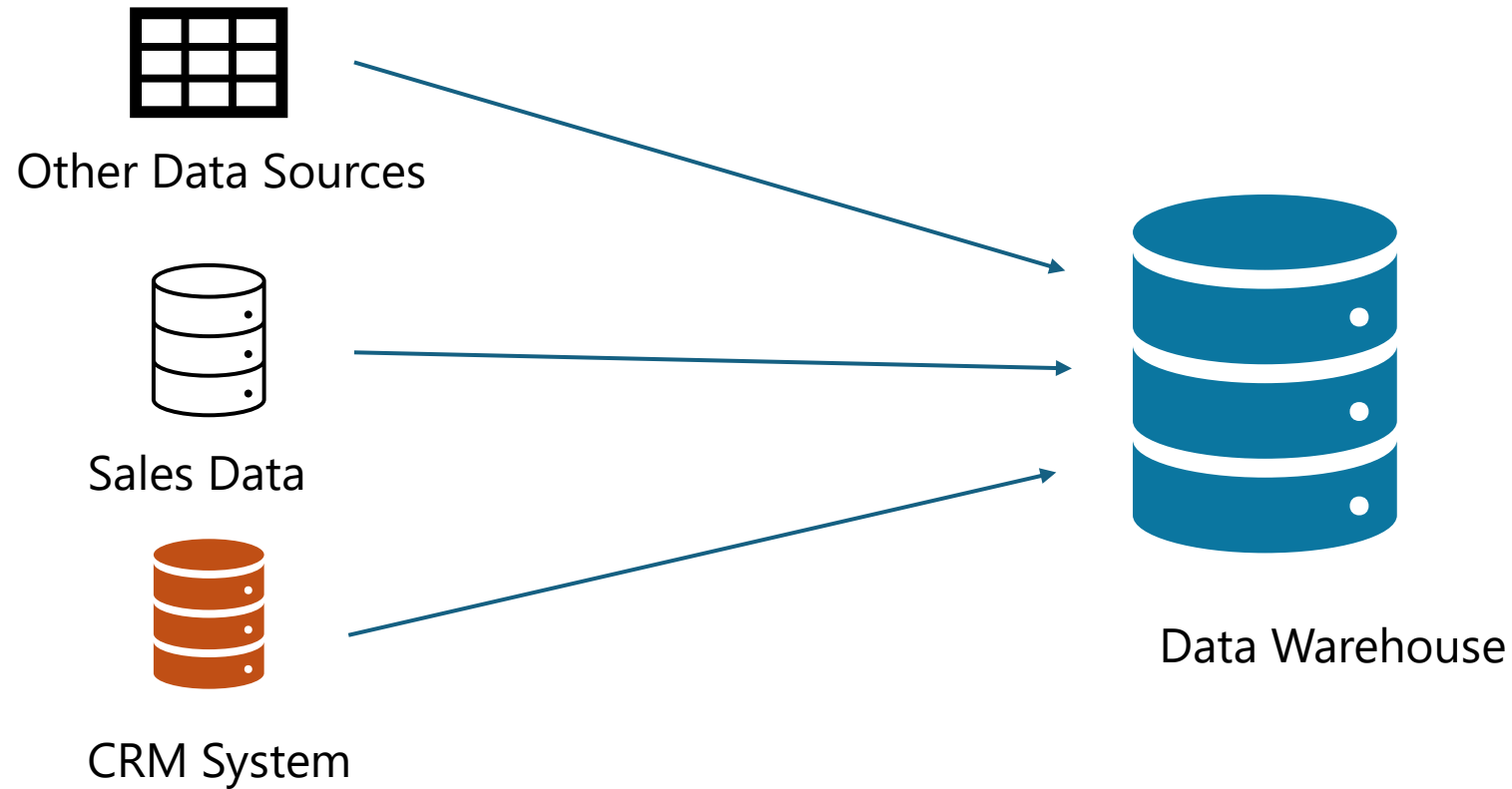
What is a Data Warehouse?

Data Warehouse:

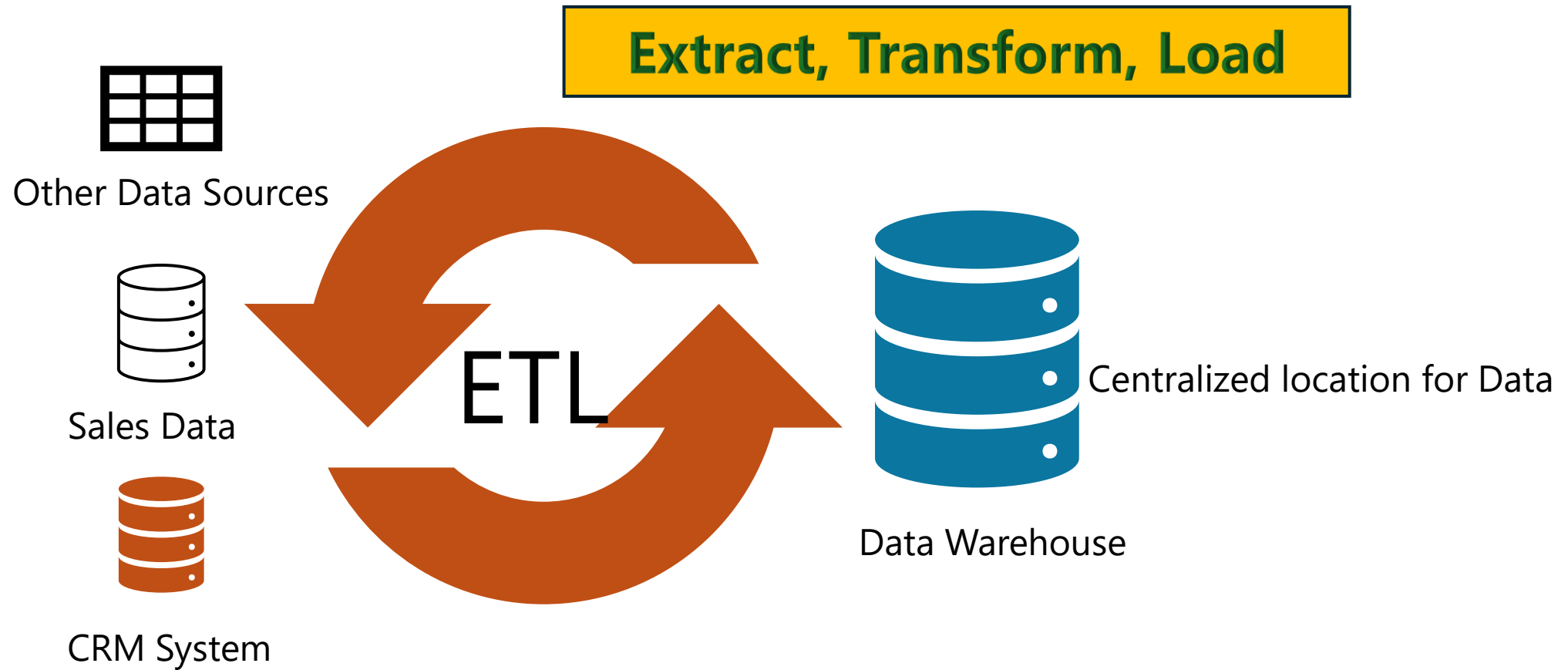
A Database used & optimized for analytical purpose

- ✓ User friendly
- ✓ Fast query performance
- ✓ Enabling data analysis

Understanding Data Warehouse



Understanding Data Warehouse



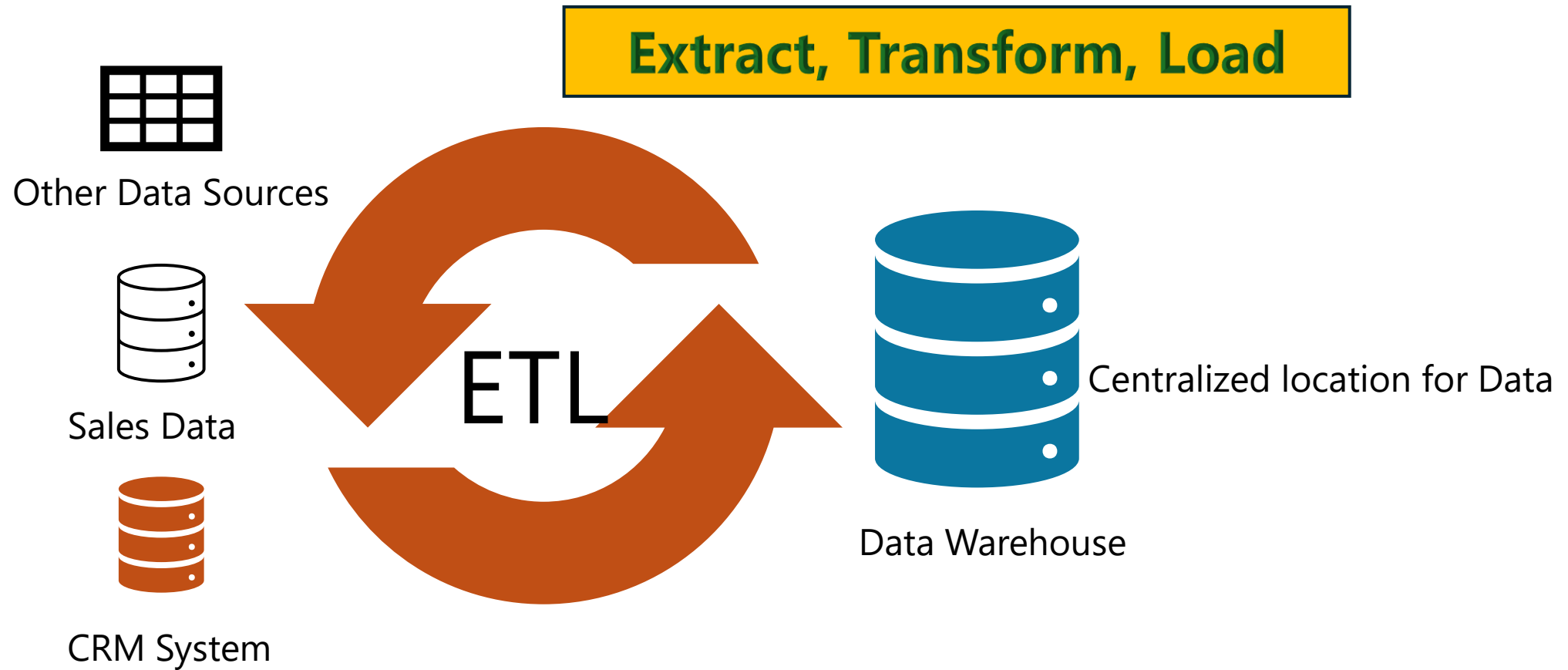
Goals of a Data Warehouse

Layers of Data Warehouse

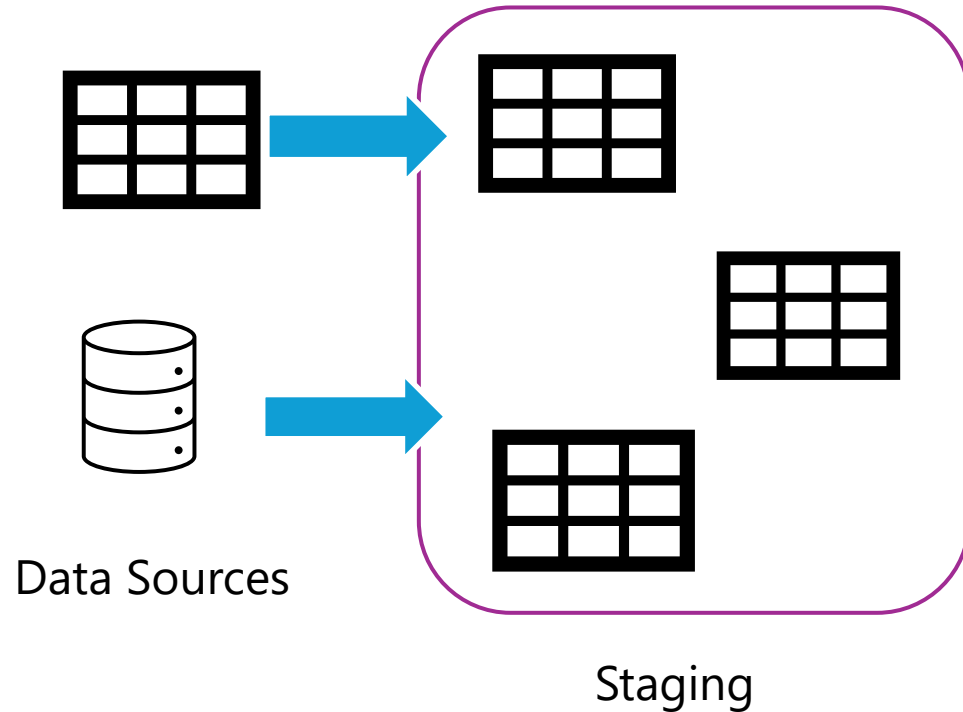
- ❑ Centralized and consistent location for data
- ❑ Data must be accessible fast (query performance)
- ❑ User-friendly (Easy to understand)
- ❑ Must load data consistently and repeatedly (ETL)
- ❑ Reporting and data visualization built on top

Data Warehouse Architecture

Data Warehouse Layers



Data Warehouse Layers



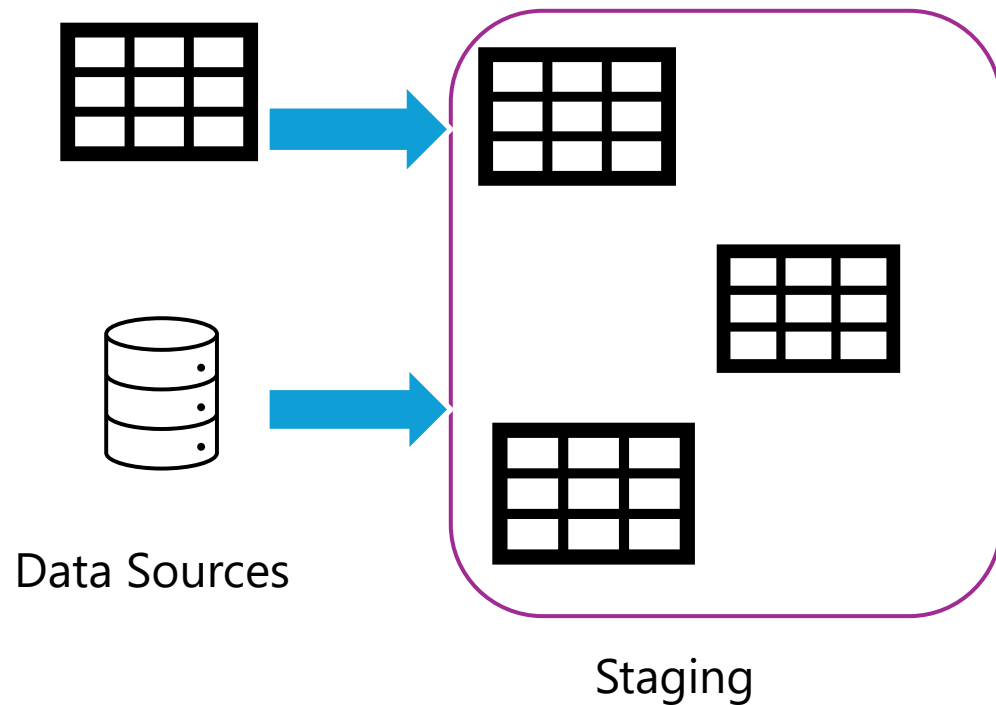
Department 1

employee_id	entry_date	position_level
1	01-01-2025	HR
2	02-01-2025	IT
3	04-01-2025	IT
4	04-01-2025	UM
5	05-01-2025	PM

Department 2

employee_id	entry_date	position_level
6	05-01-2025	HR
7	08-01-2025	HR
8	10-01-2025	PM
9	11-01-2025	UM
10	12-01-2025	PM

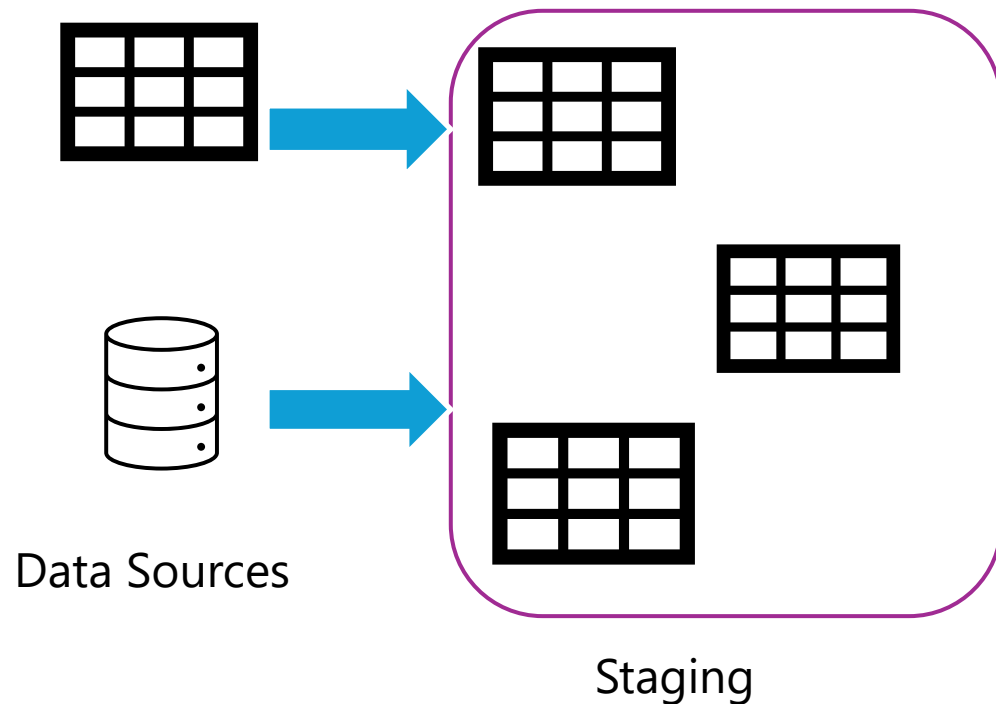
Data Warehouse Layers



Departments

employee_id	entry_date	position_level
1	01-01-2025	HR
2	02-01-2025	IT
3	04-01-2025	IT
4	04-01-2025	UM
5	05-01-2025	PM
6	05-01-2025	HR
7	08-01-2025	HR
8	10-01-2025	PM
9	11-01-2025	UM
10	12-01-2025	PM

Data Warehouse Layers



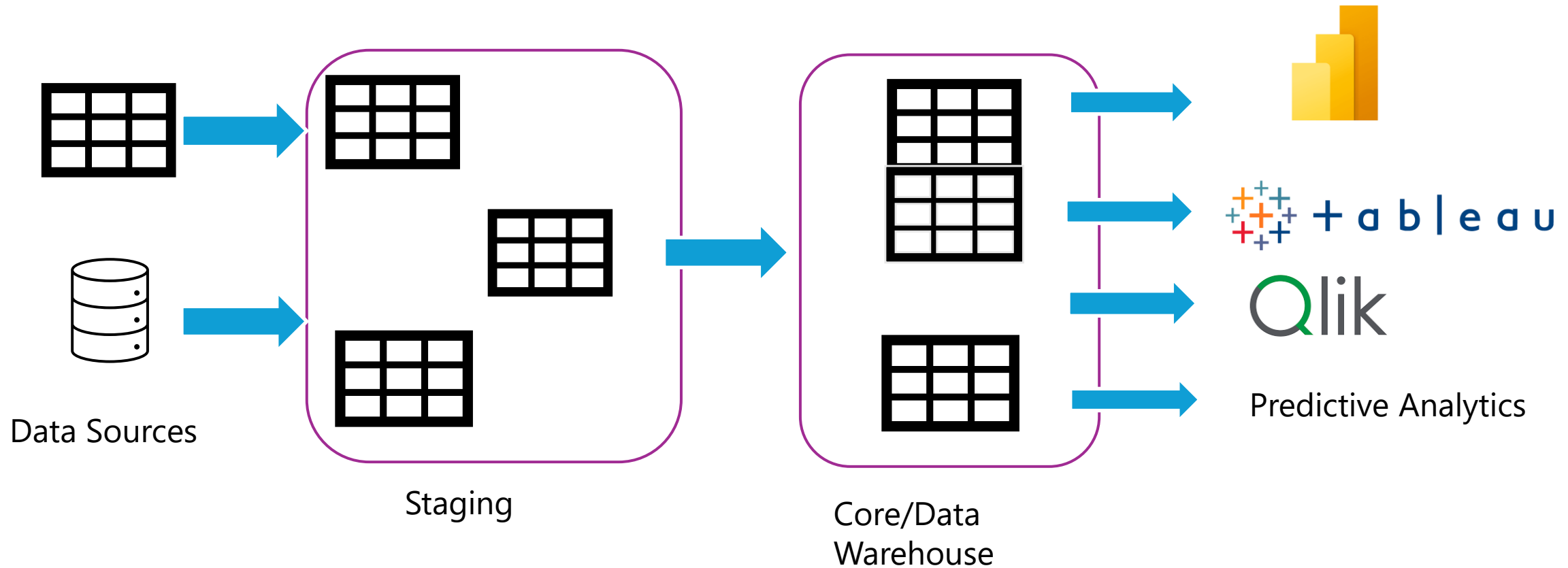
Department 1

employee_id	entry_date	position_level
1	01-01-2025	HR
2	02-01-2025	IT
3	04-01-2025	IT
4	04-01-2025	UM
5	05-01-2025	PM

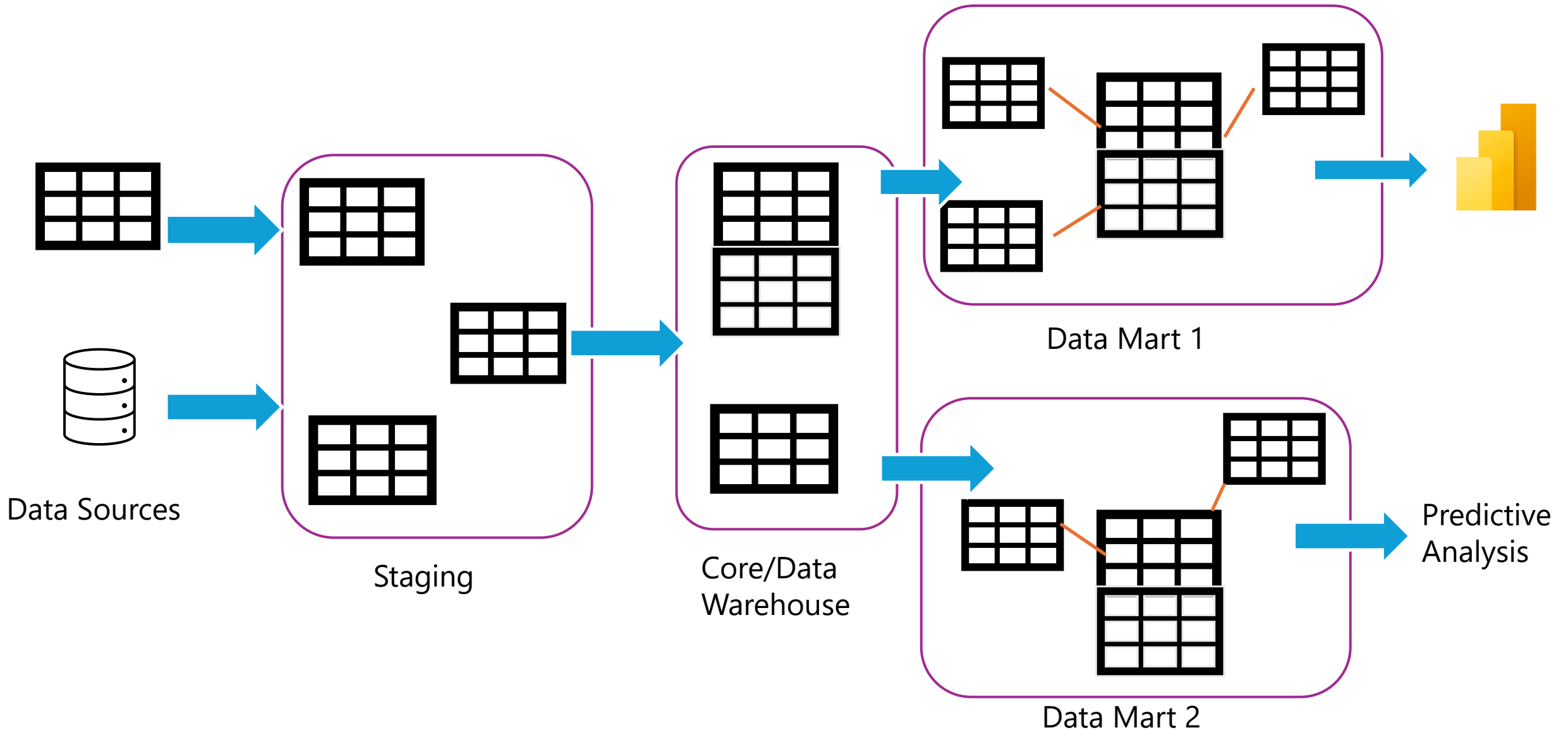
Department 2

employee_id	entry_date	position_level
6	05-01-2025	Human Resources
7	08-01-2025	Human Resources
8	10-01-2025	Project Management
9	11-01-2025	Upper Management
10	12-01-2025	Project Management

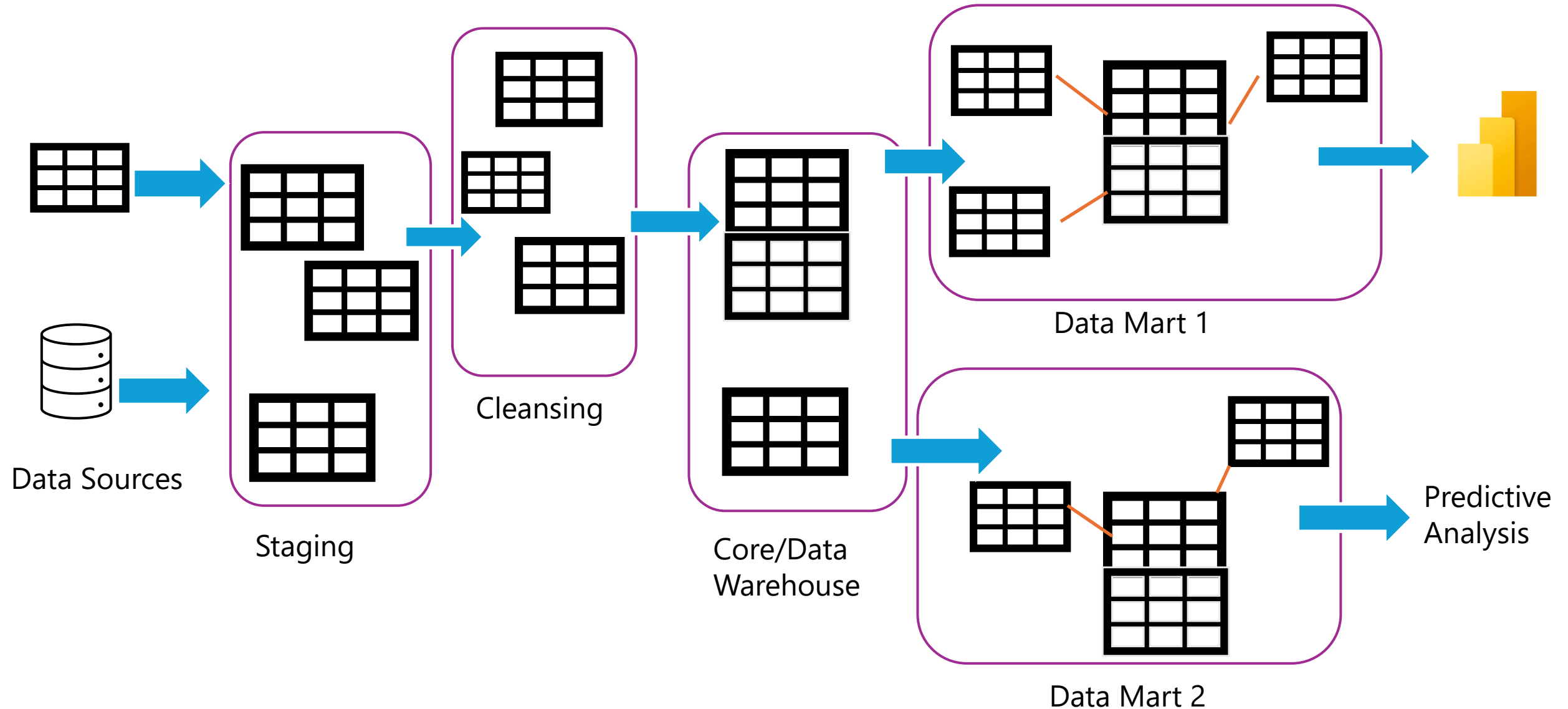
Data Warehouse Layers



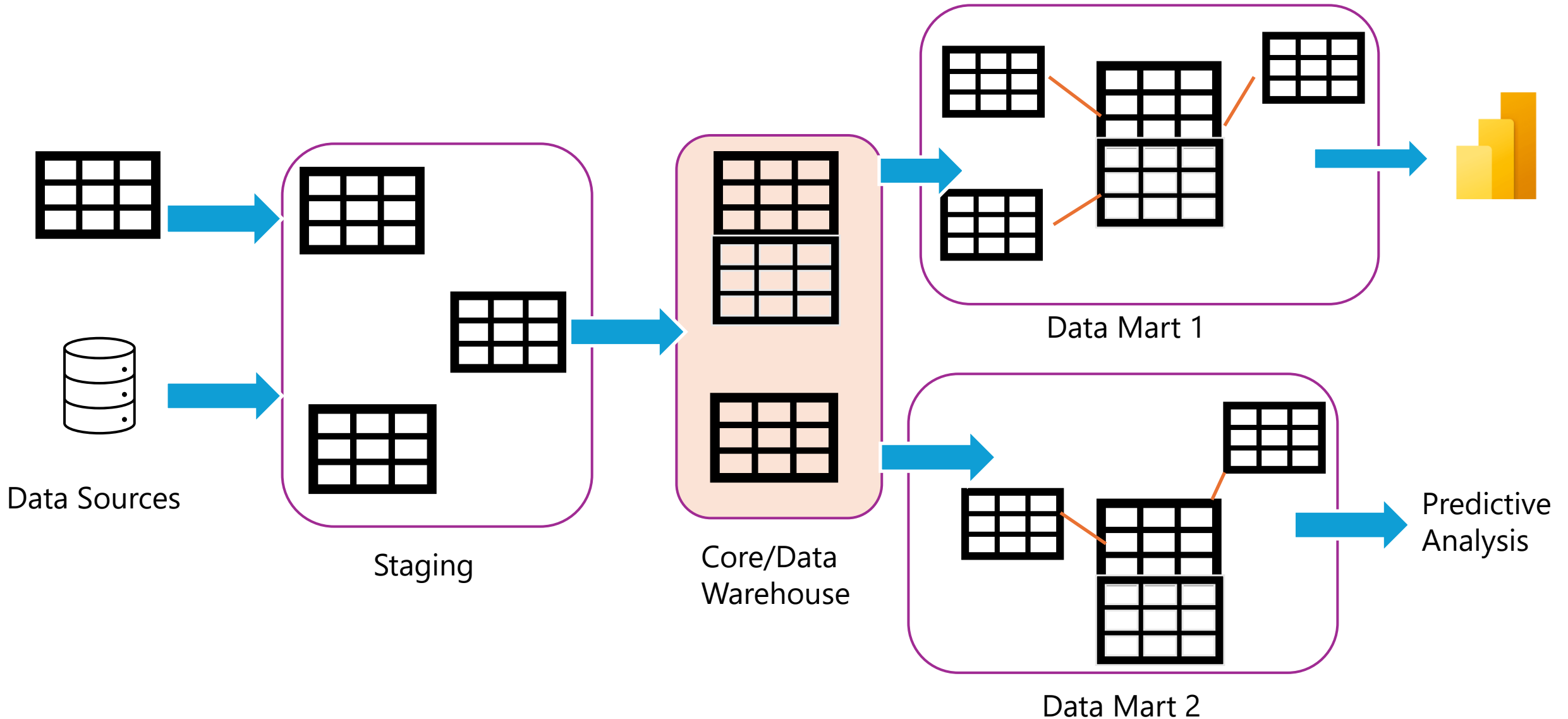
Data Warehouse Layers



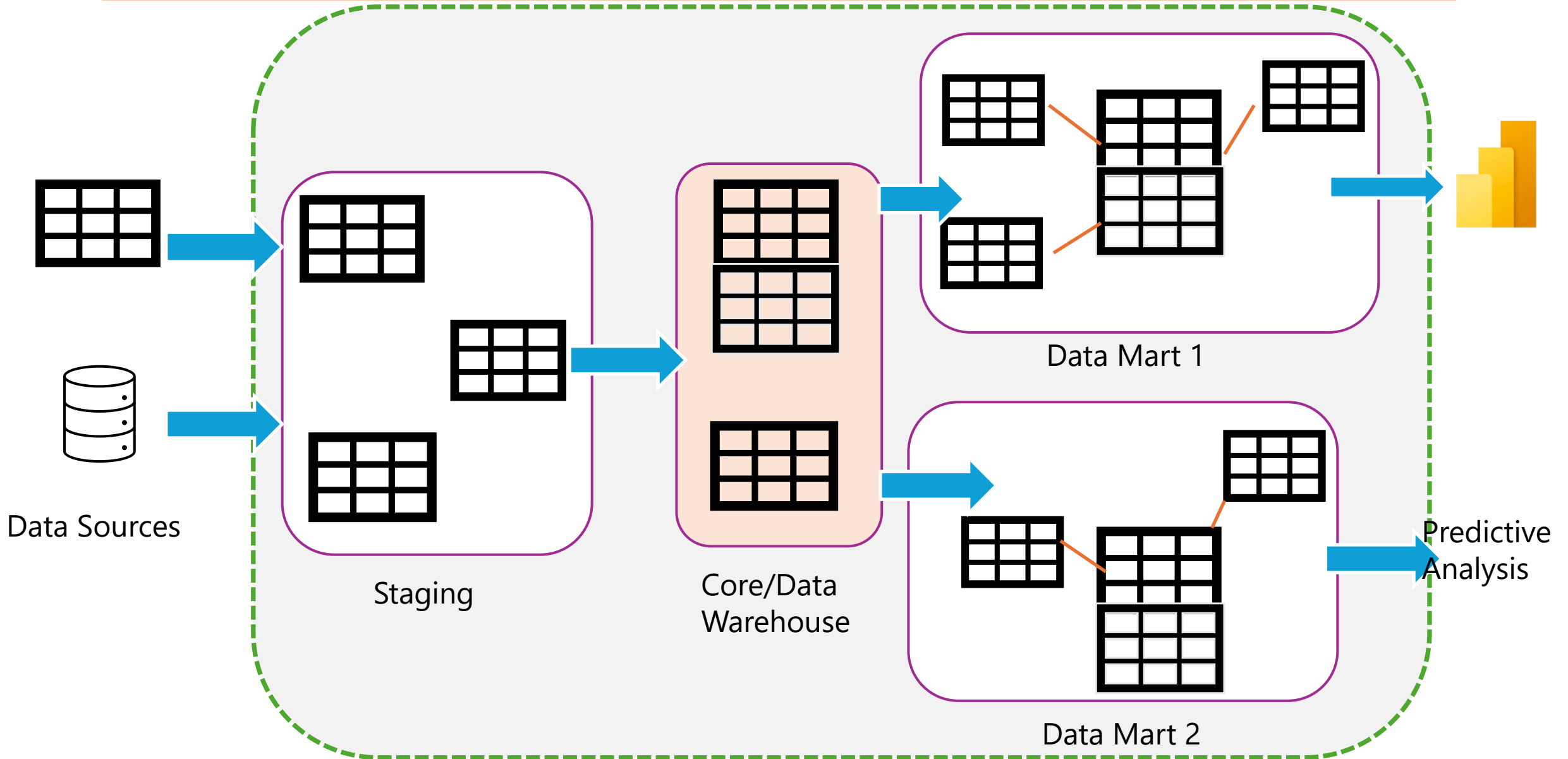
Data Warehouse Layers



Data Warehouse Layers

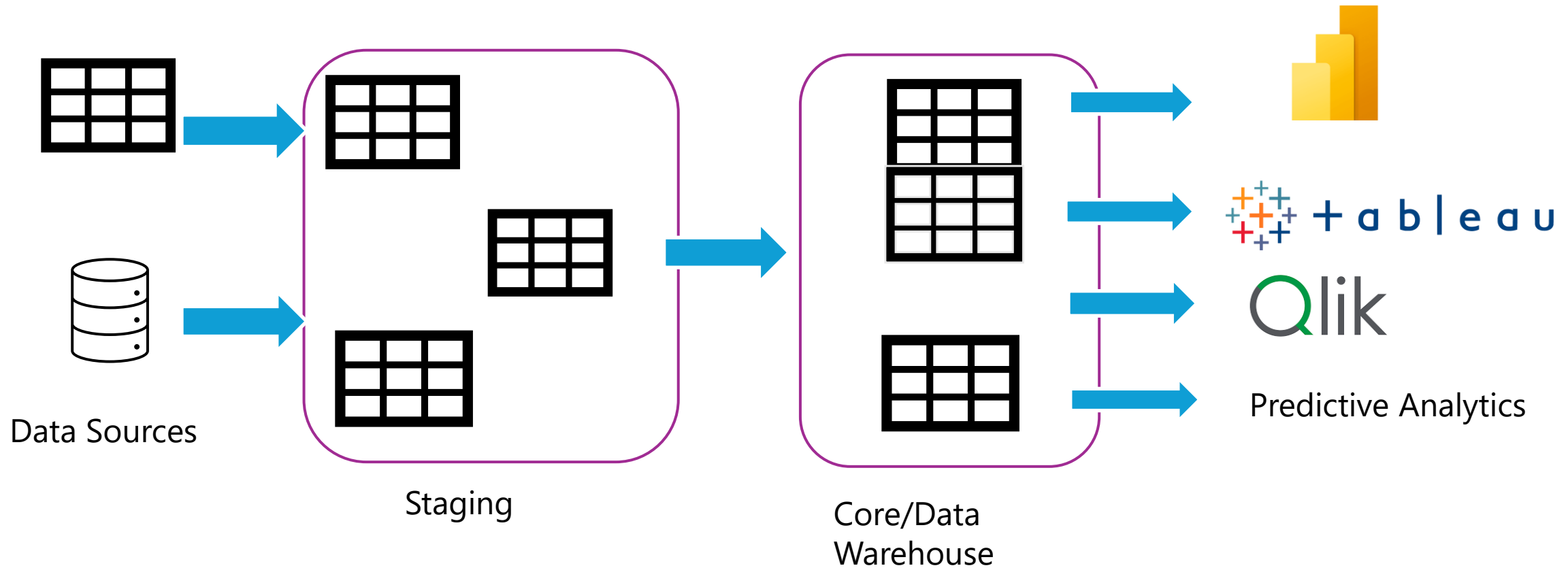


Data Warehouse Layers

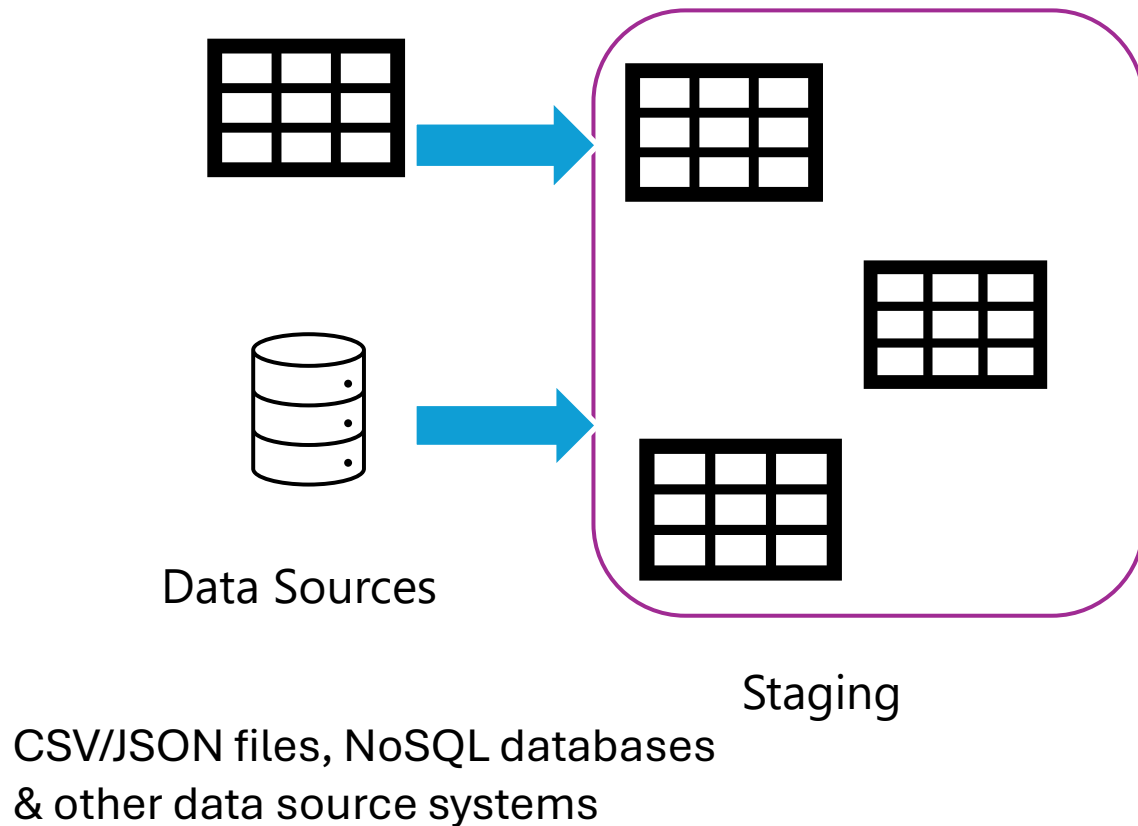


Staging Area

Staging Layer



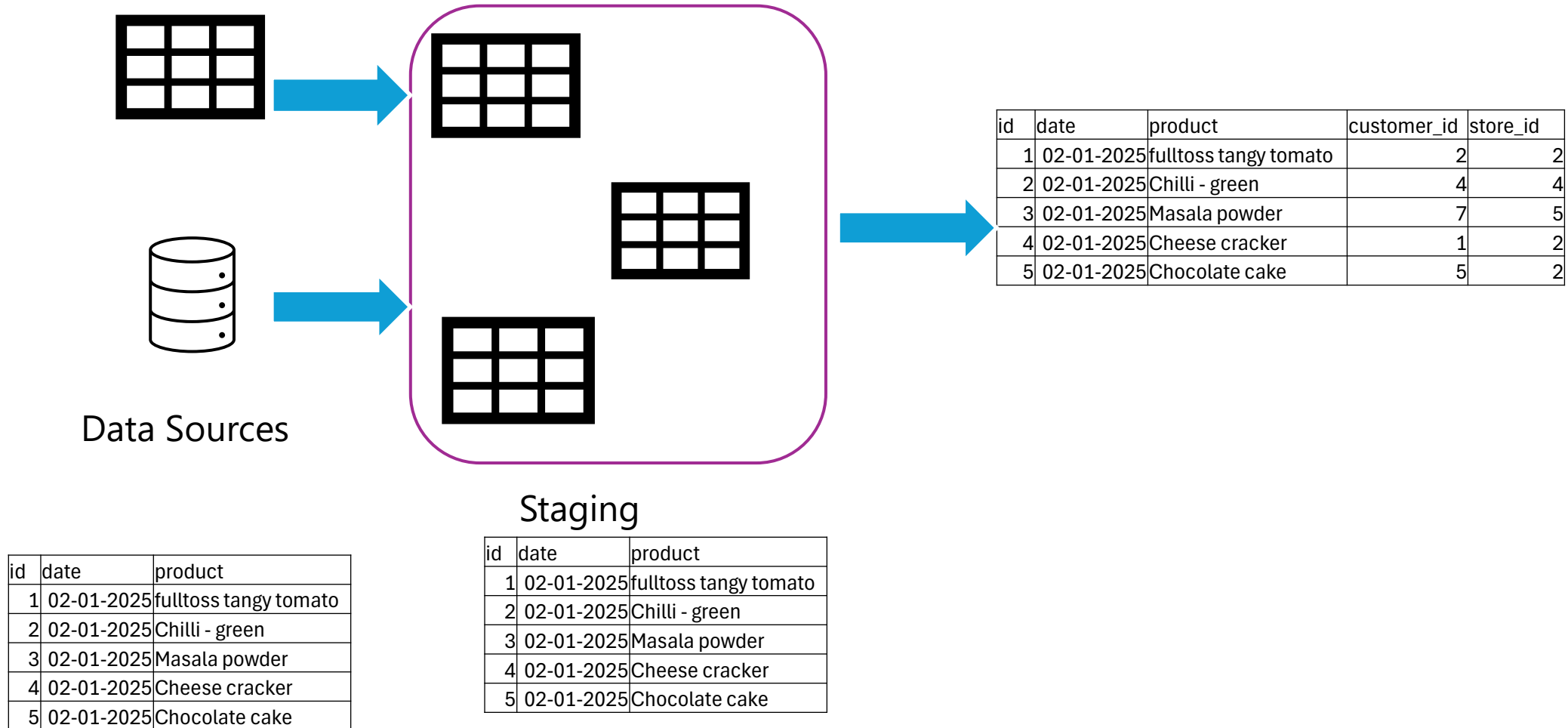
Staging Layer



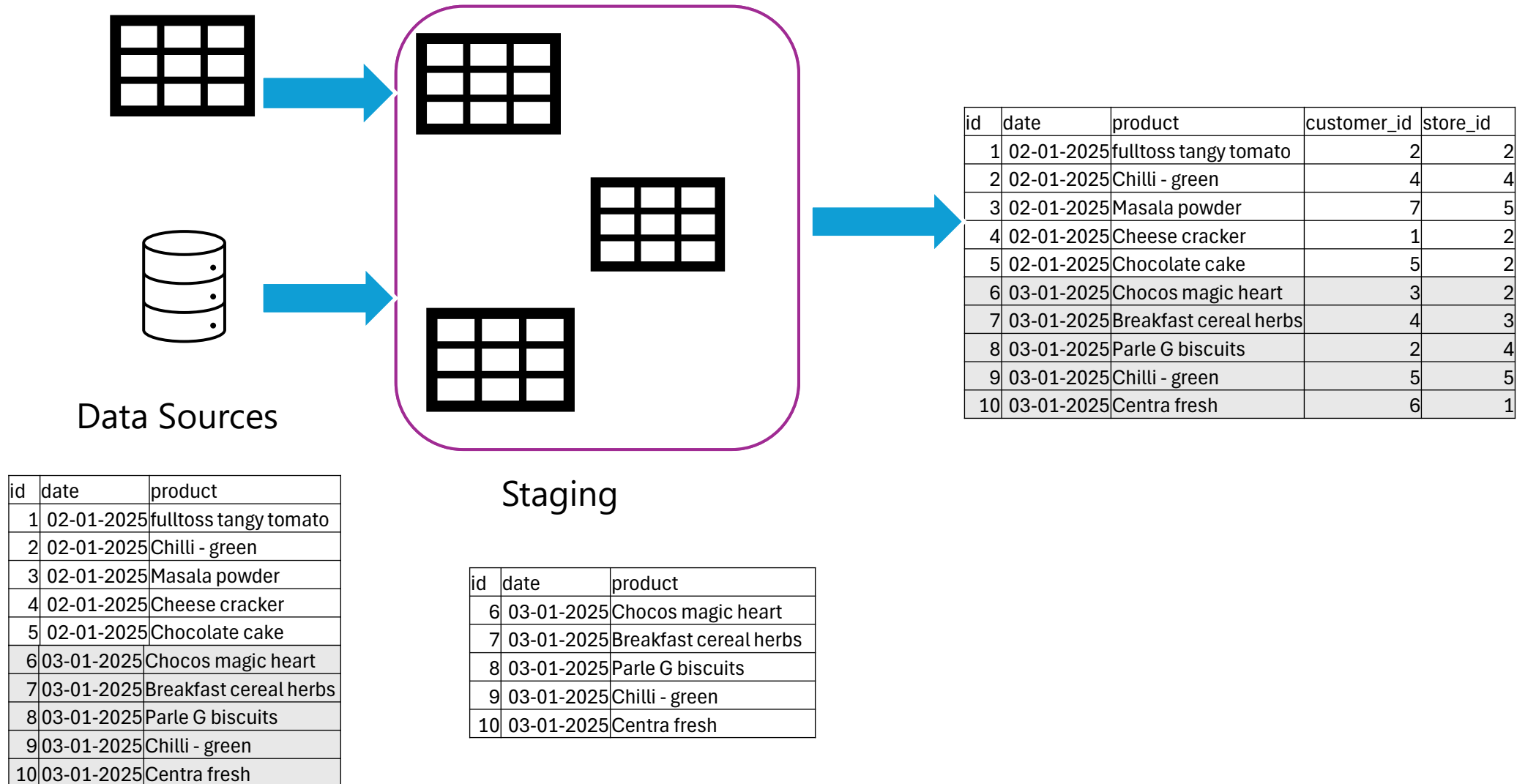
- Short time on the source systems
- Quickly extract
- Move the data in the relational database

- Start transformations from there

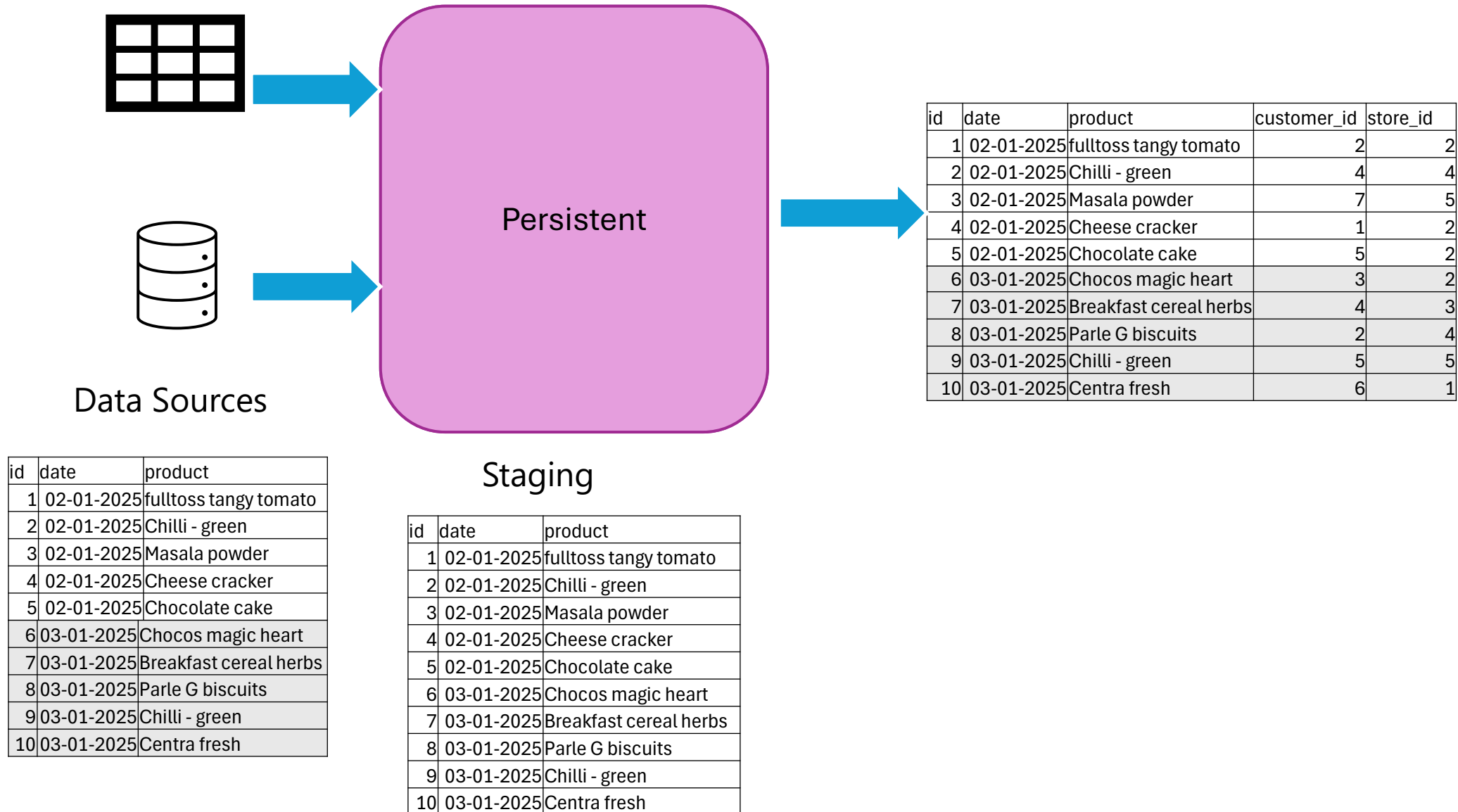
Staging Layer



Staging Layer



Staging Layer



Staging Layer

- Staging Layer is landing zone extracted data
- Data in tables and on a separate database
- As little “touching” as possible to source system
- We don’t change source systems
- Temporary or persistent staging layers