

# **FIT3003 – Business Intelligence and Data Warehousing**

Week 1b – Introduction to Business Intelligence  
and Data Warehousing

Semester 2, 2022

Developed by:  
Dr. Agnes Haryanto  
[Agnes.Haryanto@monash.edu](mailto:Agnes.Haryanto@monash.edu)

# Database

“A **database** is a shared, integrated computer structure that stores a collection of the following:

- End-user data—that is, raw facts of interest to the end user
- Metadata, or data about data, through which the end-user data is integrated and managed” [1]

“A **database** is a collection of data, typically describing the activities of one or more related organizations” [2]

“A **database management system (DBMS)** is a collection of programs that manages the database structure and controls access to the data stored in the database.” [1]

# Database

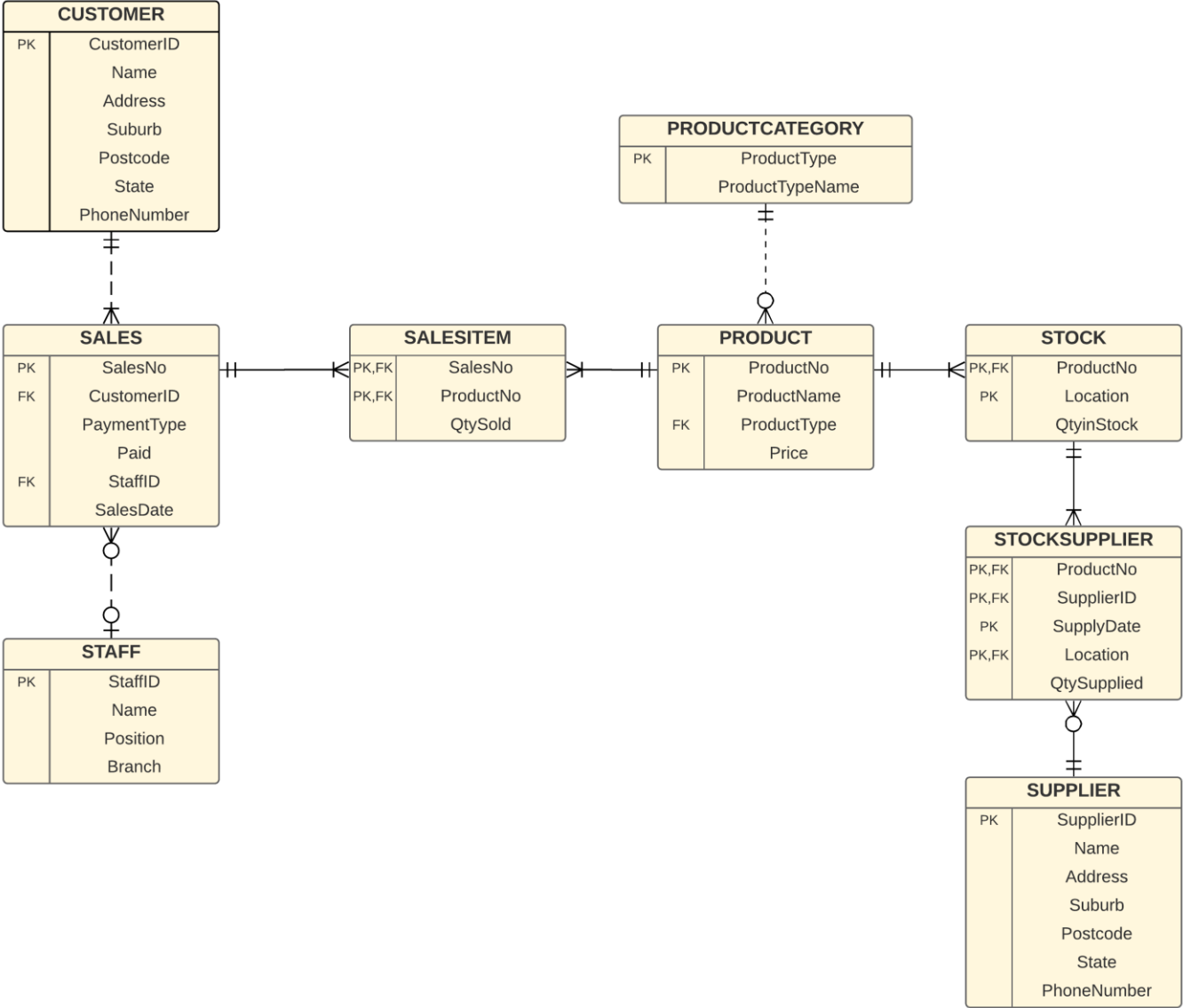
- Advantage of having a database:
  - Management can develop an understanding of how the company is performing.
  - Assess whether the existing strategies meet organizational goals.
- There is a high need for **data analysis** and having the right data at the right time to support the **decision-making process**.

# Operational Database

- Every company has a database system. This database system is used to operate the daily business activities of the company.
  - It can be sales, finance, booking systems, or any transaction events relating to the breath of the business.
- This database is then called an ***Operational Database***.
  - The operational database is used to support the operation of the business.
  - It automates the business processes.

# Operational Database

## The Product Sales System Database



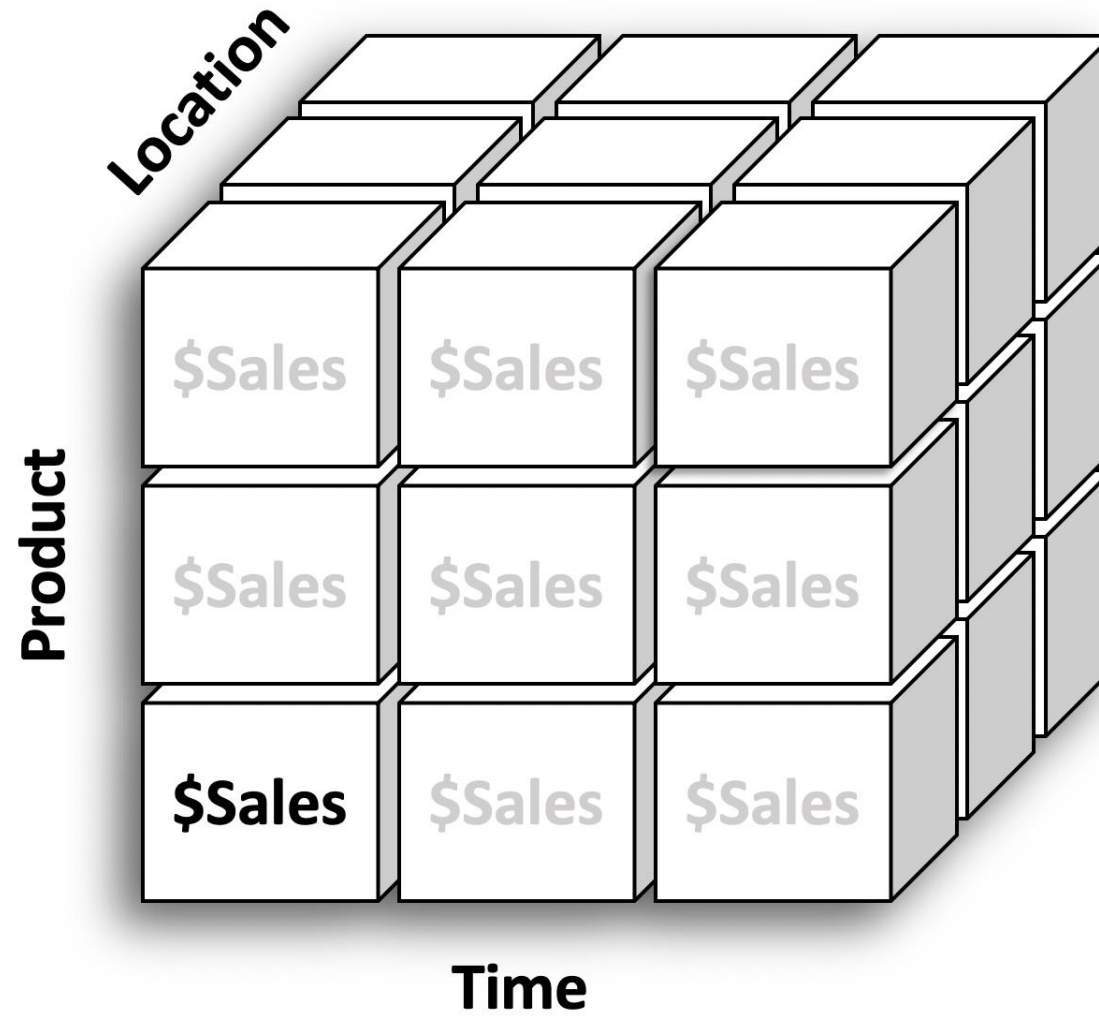
# Operational Database

- The use of an operational database is centered around transactions.
- Operational database is normally designed using a structured design, such as an Entity-Relationship (E/R) diagram.
- Operational database is suitable for decision-making.
- But to support an efficient decision-making, pre-computation of aggregation, summarization, etc becomes the main aim.
  - Unfortunately, this is **not** the **primary function** of an operational database.
  - Hence, we need a **data warehouse** to address this need.

# Data Warehouse

- To address the drawback of operational database, and a need for decision-making support data, **data warehouse is needed**.
- A **data warehouse** is a multi-dimensional view of databases, with aggregates and pre-computed summaries.
  - In many ways, it is basically doing aggregates in advance; that is exactly pre-computation done at the design level, rather than at the query level.

# Data Warehouse

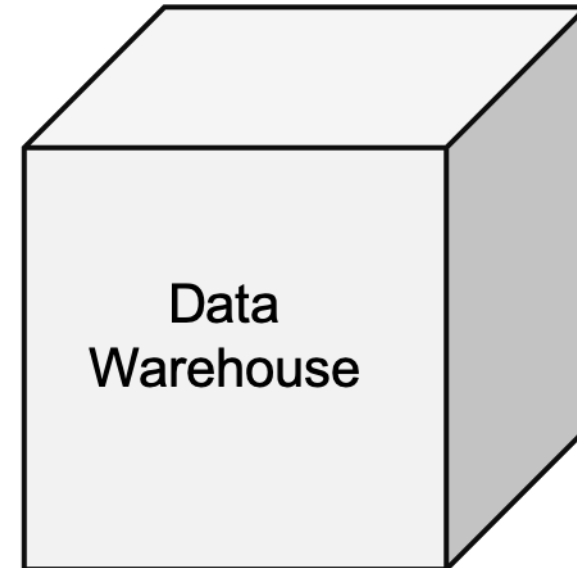
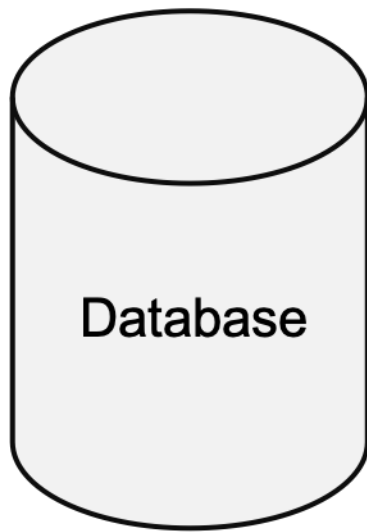




# Building Data Warehouse

- There are two important parts about data warehousing: **building it**, and **using it**.
- A data warehouse is created by transforming an operational database to a data warehouse.
- The transformation includes a series of steps of data manipulation.
  - such as extracting, cleaning, aggregating, summarizing, combining, altering, appending, etc
  - all of which are called an **Extract-Transform-Load** or an **ETL**

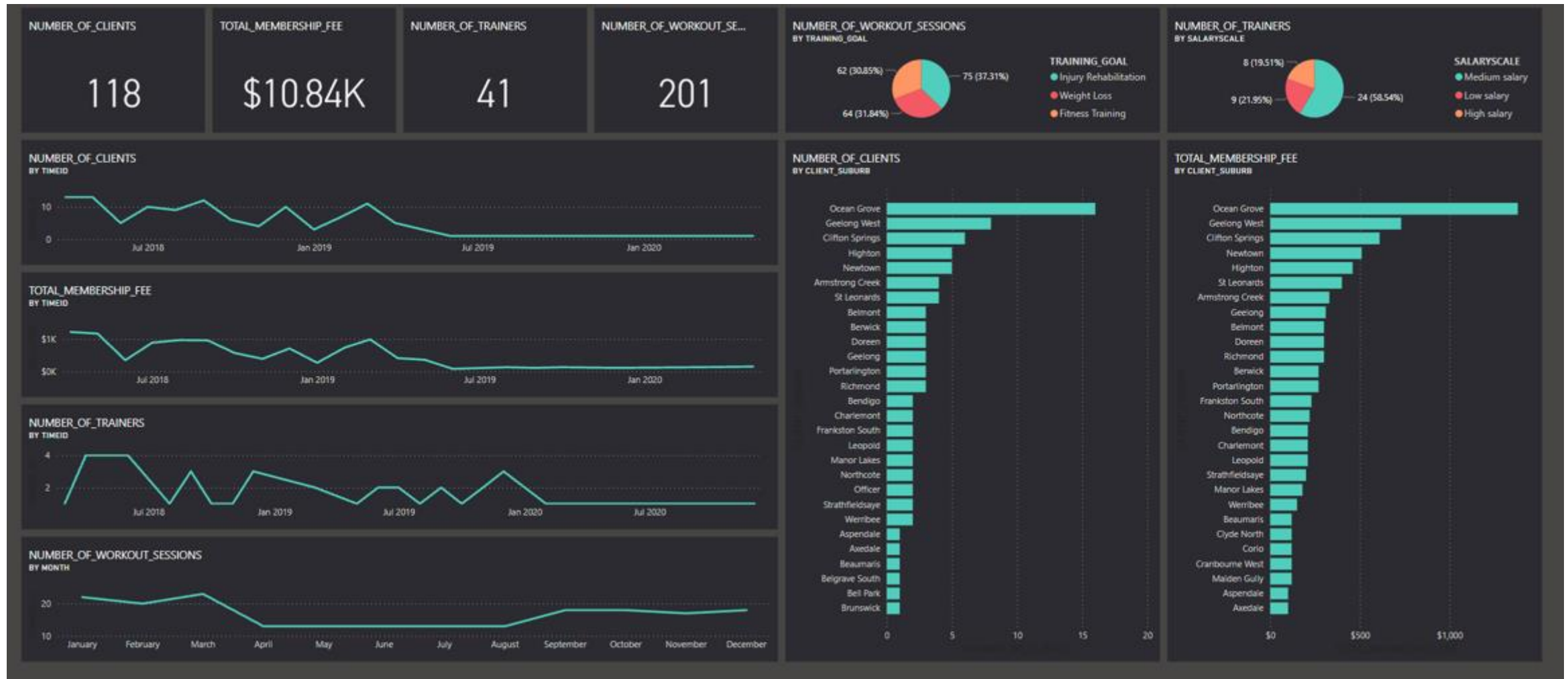
# Building Data Warehouse



# Using Data Warehouse

- Using a data warehouse means to extract data from the data warehouse for further data analysis.
- The query to extract data from the data warehouse is an ***Online Analytical Tool*** or ***OLAP***.
- OLAP retrieves raw data, which can then be later formatted using any Business Intelligence (BI) tools.
  - In OLAP, the focus is on the data – the retrieved data is the most important.
  - The BI tool is for further presentation and visualization.
  - BI tools which receive the raw data can present the data in any forms: reports, graphs, dashboards, etc.

# Business Intelligence



# The Need of Data Analysis

- Analyzing company data can provide insightful information about short-term tactical evaluations and strategic questions, such as:
  - Are our sales promotions working?
  - What market percentage are we controlling?
  - Are we attracting new customers?
- Tactical and strategic decisions are also shaped by constant pressure from external and internal forces, including globalization, the cultural and legal environment, and technology.

# Business Intelligence

- Business intelligence (BI) is a term that describes a comprehensive, cohesive, and integrated set of tools and processes used to capture, collect, integrate, store, and analyze data with the purpose of generating and presenting information to support business decision making.
- BI is a framework that allows a business to transform data into information, information into knowledge, and knowledge into wisdom.

# Business Intelligence

- In general, BI provides a framework for:
  - Collecting and storing operational data
  - Aggregating the operational data into decision support data
  - Analyzing decision support data to generate information
  - Presenting such information to the end user to support business decisions
  - Making business decisions, which in turn generate more data that is collected, stored, and so on (restarting the process)
  - Monitoring results to evaluate outcomes of the business decisions, which again provides more data to be collected, stored, and so on
  - Predicting future behaviors and outcomes with a high degree of accuracy

# Business Intelligence

- In practice, the first point, collecting and storing operational data, does not fall into the realm of a BI system per se; rather, it is the function of an operational system.
- However, the BI system will use the operational data as input material from which information will be derived.
- The rest of the processes and outcomes explained in the preceding points are oriented toward generating knowledge, and they are the focus of the BI system.



# The Big Picture

Operational  
Database



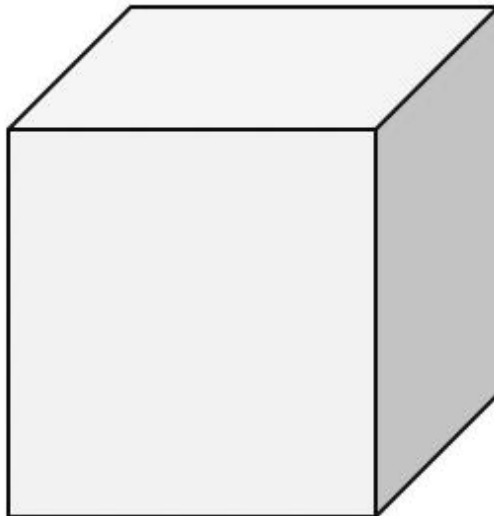
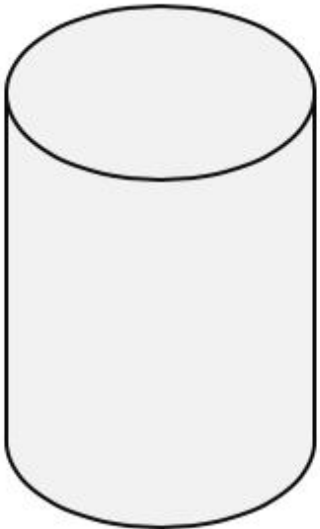
Data  
Warehouse



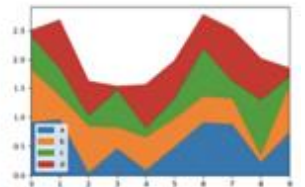
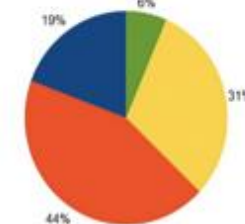
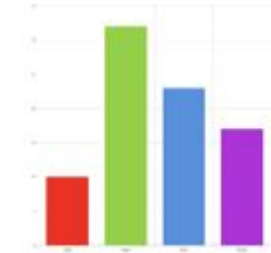
OLAP



Business  
Intelligence

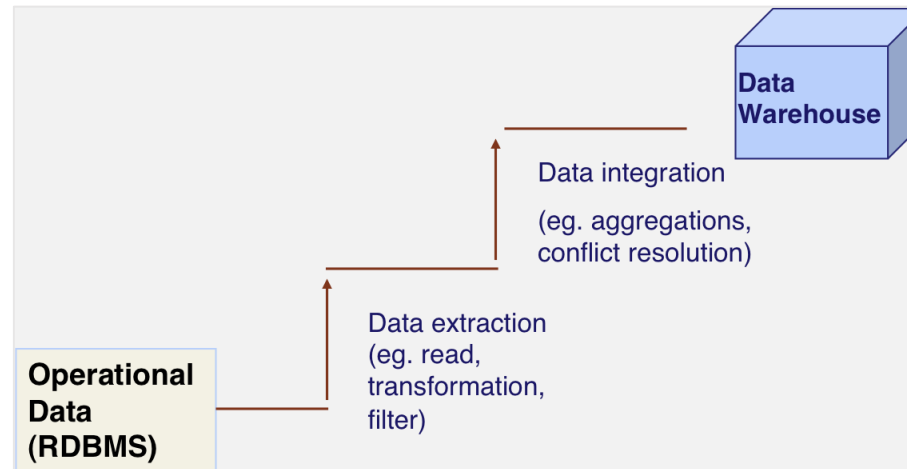


	...	...	...	...
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...



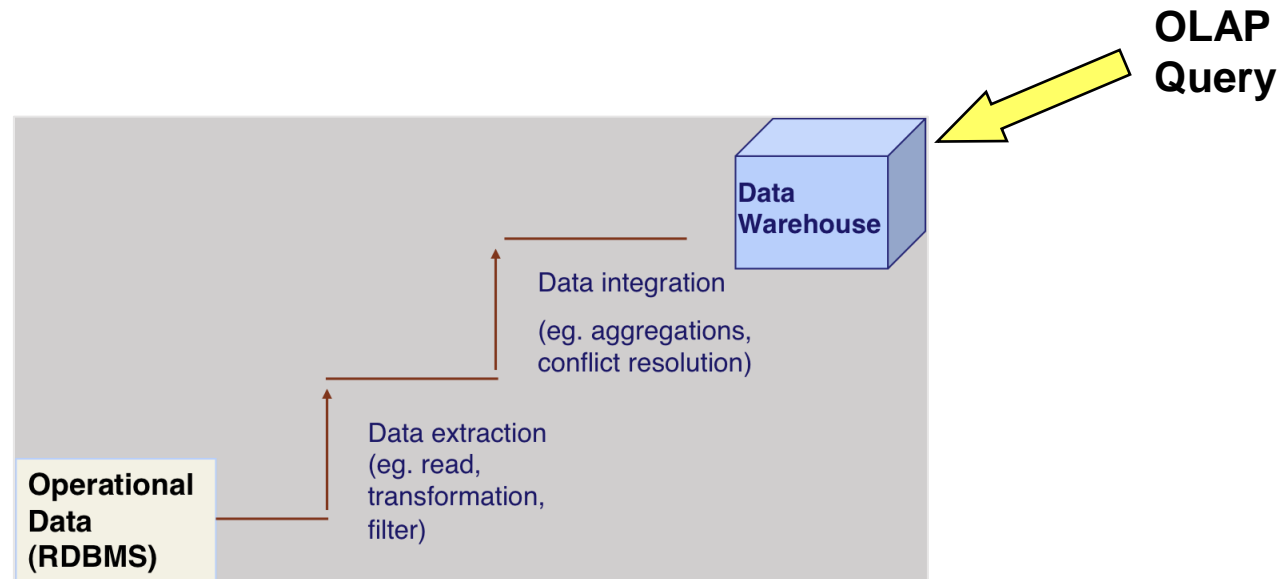
# Summary

- The need for an effective decision support system has motivated the emergence of the new data storage facility called **Data Warehouse**.



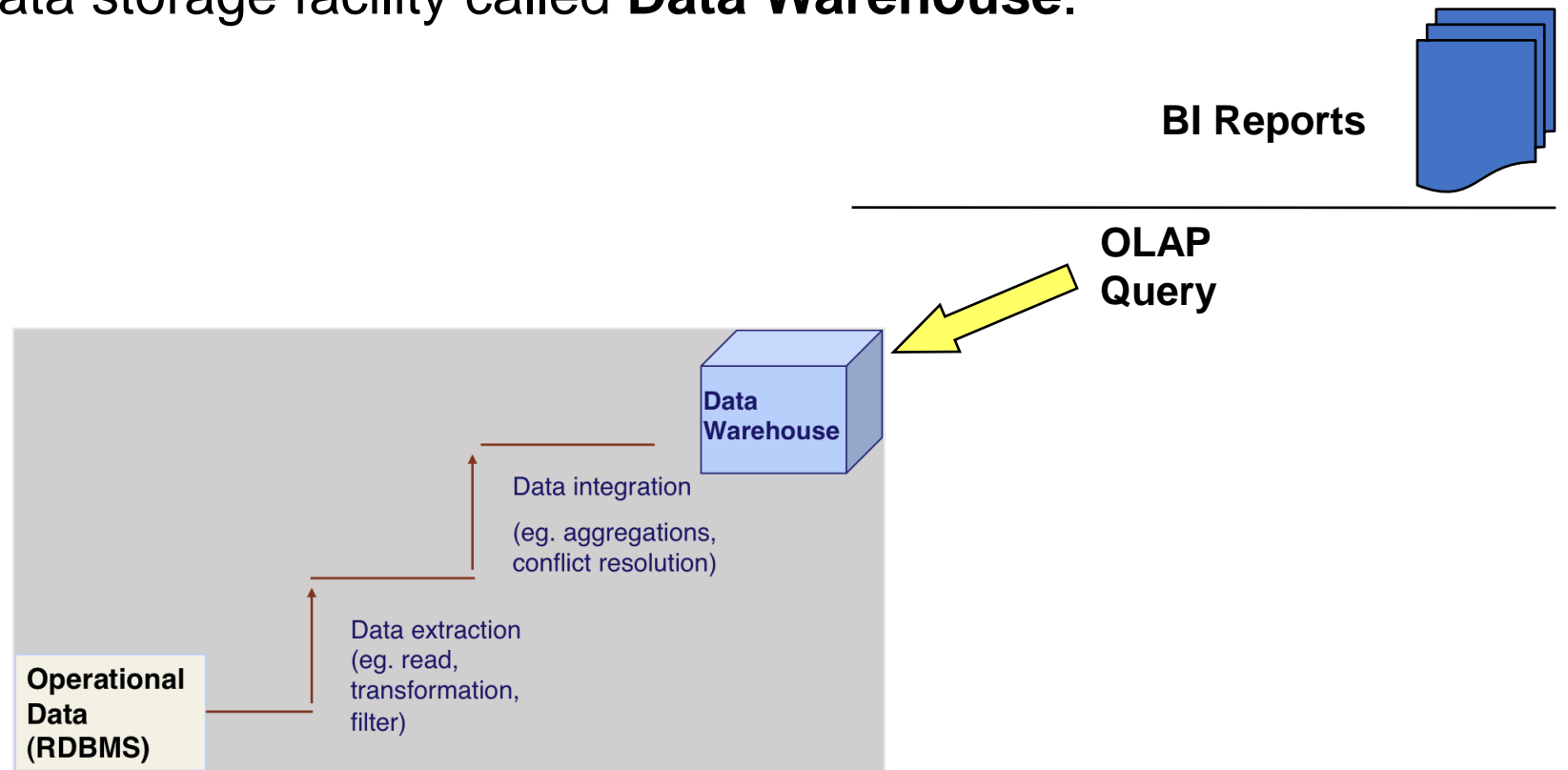
# Summary

- The need for an effective decision support system has motivated the emergence of the new data storage facility called **Data Warehouse**.



# Summary

- The need for an effective decision support system has motivated the emergence of the new data storage facility called **Data Warehouse**.



# Main Questions...

- How do we design the data warehouse model?
- What methods are available to implement it?
- Case studies
  
- However, before we do this, we need to revise SQL

# References

- [1] C. Carlos and M. Steven, Database systems: *Design, implementation, & management*. Cengage Learning, 12<sup>th</sup> ed., 2016.
- [2] R. Ramakrishnan and J. Gehrke, Database Management Systems. New York, NY, USA: McGraw-Hill, Inc., 2<sup>nd</sup> ed., 2000.
- [3] Taniar, D., & Rahayu, Wenny. (2022). *Data Warehousing and Analytics : Fueling the Data Engine*. Chapter 1 ("Introduction").