

Applied A.I. Solutions

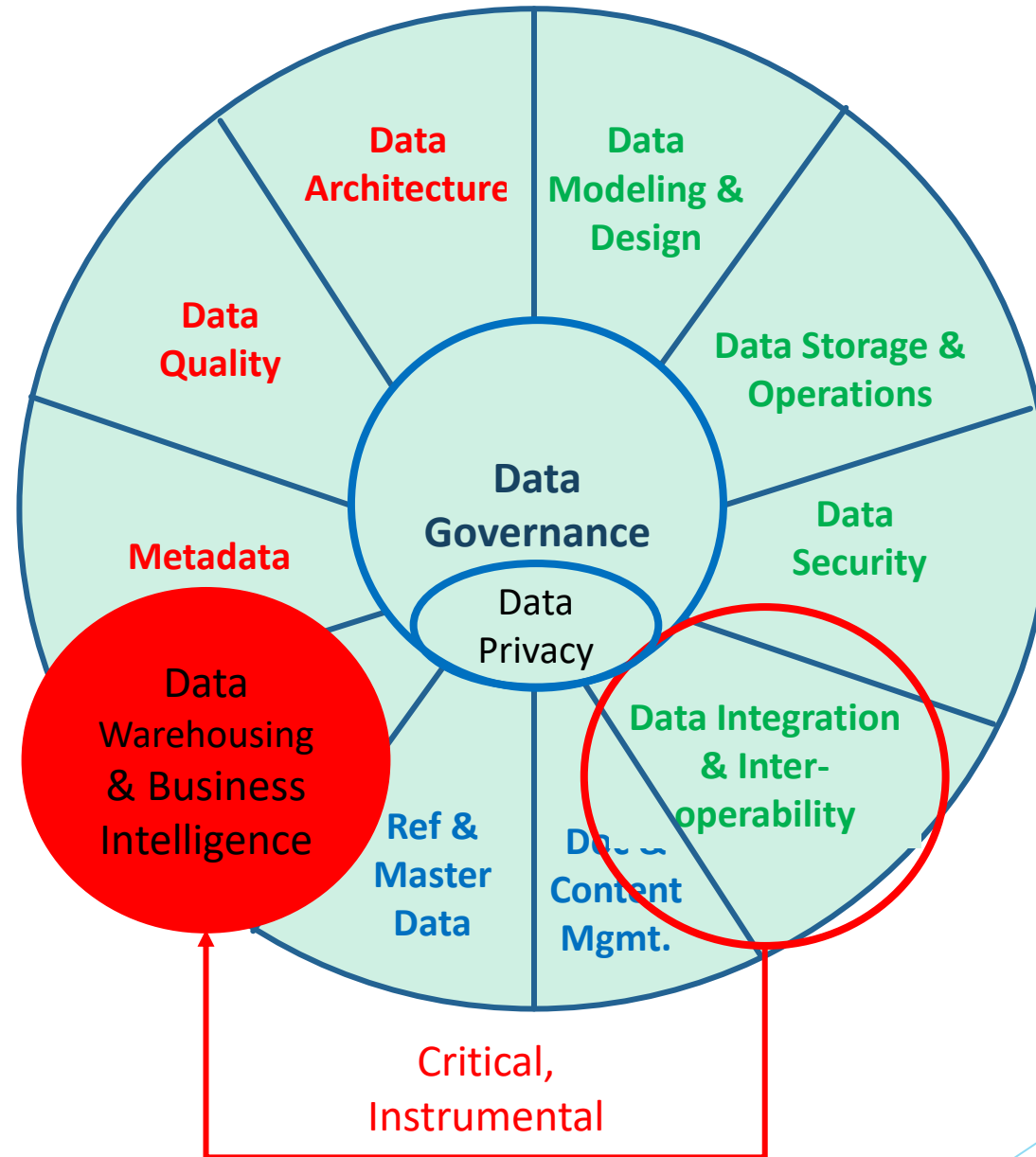
Data Visualization Techniques Business Intelligence & Analytics

**Professor
Daniel Vitaver, EMBA**

daniel.vitaver@georgebrown.ca

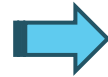
DATA WAREHOUSING AND BUSINESS INTELLIGENCE

DAMA Wheel



Business Intelligence: challenges

- You cannot see the big picture
- You cannot get the right information fast enough for decision-making
- You lack a reliable source of data
- Your personnel cannot collaborate efficiently
- Reporting is hard to pull-out from your systems



Business Intelligence: response

- **Technology** that enables to integrate data from a variety of sources into a common enterprise model, specially designed for rapid query and analysis
- **Evidence-based decision-making** organizational culture

DW & BI Management Governance Framework

Definition

Planning, implementation, and control processes to provide **decision support data** for reporting, query, and analysis.

Goals

To build a technical environment and processes needed to deliver integrated data in support to business operations, business intelligence activities and compliance requirements.

Business Intelligence Framework

-  Data Discovery (e-discovery)
-  Learning & Knowledge Management
-  Informed Decision-making
-  Performance Management & Improvement
-  Research & Development
-  Collaboration & Information Sharing
-  Monitoring and Control



Business Drivers

- Foster evidence-based decision-making
- Improve efficiency and competitive advantage
- Support operational functions, compliance requirements, BI activities
- Provide and maintain historical data

Principles:

1. Focus on business goals and priorities
2. Think corporate (architecture), build locally (incrementally)
3. Promote transparency, self-service and collaboration
4. Build Metadata with the warehouse

Main Inputs

- Business requirements
- IT Strategy, policies and standards
- Other Data Governance policies and procedures

Activities

- a. Develop the DW and Data Marts
- b. Populate the DW
- c. Implement the BI portfolio

Main Participants

- Product Owner
- Architects and Analysts
- DW/BI specialists

Technical Drivers: Tools

- Metadata repositories
- Data integration tools
- BI and Analytic applications

Main Deliverables

DW and BI architecture

- Data products
- Population process (ETL)
- Governance activities
- Lineage dictionary

Essential Concepts

1. Business Intelligence
2. Data Warehouse
3. Data Warehousing
4. DW Architecture Component / Integration
5. Approaches to Data Warehouse - Data Models
 - a. Corporate Information factory (Inmon) **vs**
 - b. Dimensional DW (Kimball)
6. Conformed Dimensions
7. DW Data Bus
8. Star Schema

1. **Business Intelligence**

- a. Type of **data analysis** aimed at understanding organizational activities and opportunities
- b. Set of **technologies** that support data analysis, and advanced analytics through the **discovery and transformation** of data into meaningful information

2. **Data Warehouse**

- Integrated decision support database and related software programs used to administer data from a variety of sources

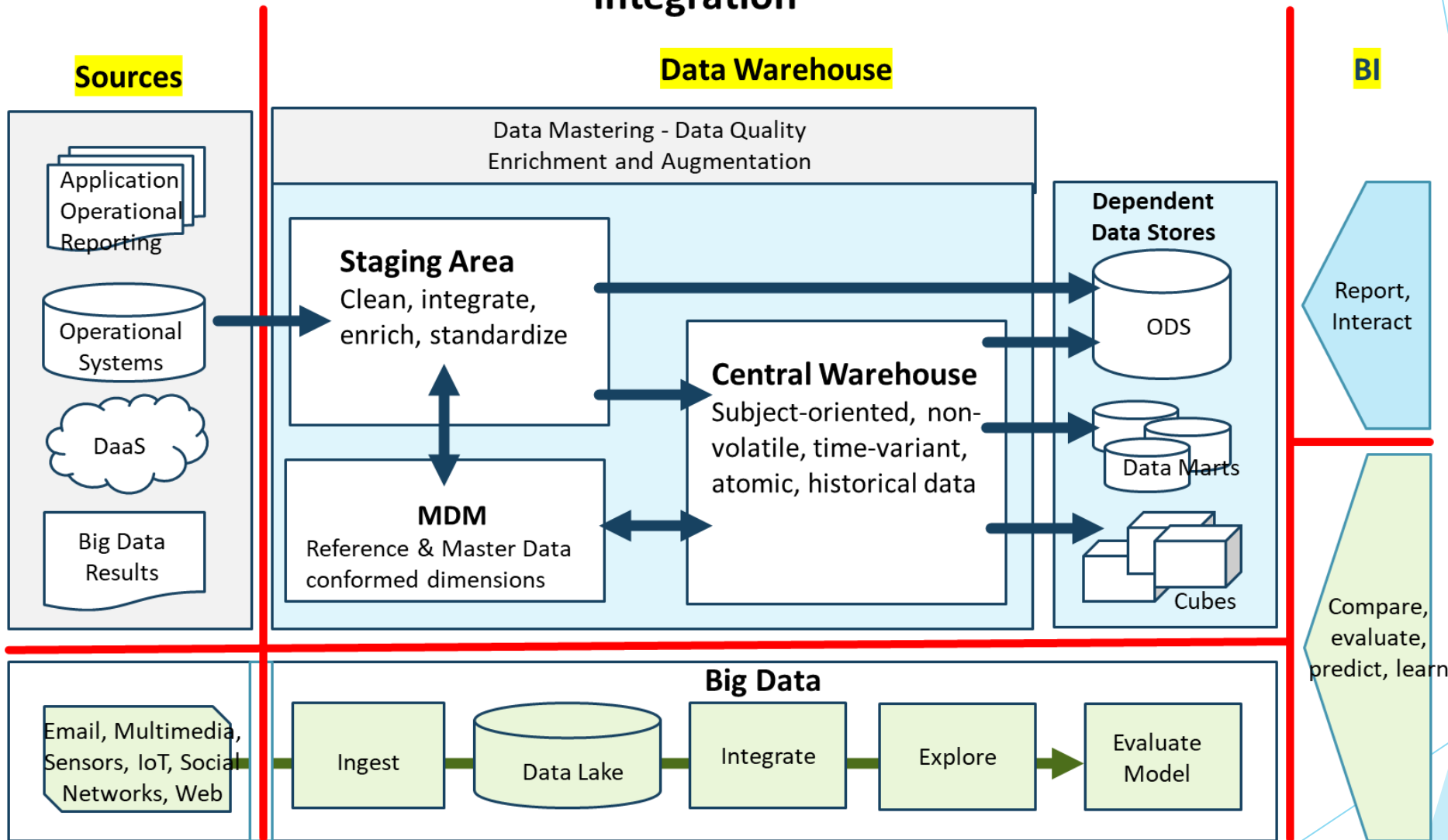
3. **Data Warehousing**

- Describes the operational extract, cleansing, transformation, control, and load processes that maintain the data in the DW
- Enforces business rules, maintains business data relationships

4. **DW Architecture Component**

- Source Systems (operational systems and external data)
- Data Integration (ETL, data visualization, and other techniques)
- Data Storage Areas
 - Staging Area
 - Reference and Master Data conformed dimensions
 - Central Warehouse
- Operational Data Storage (ODS)
- Data Marts
- Cubes for OLAP

Integration



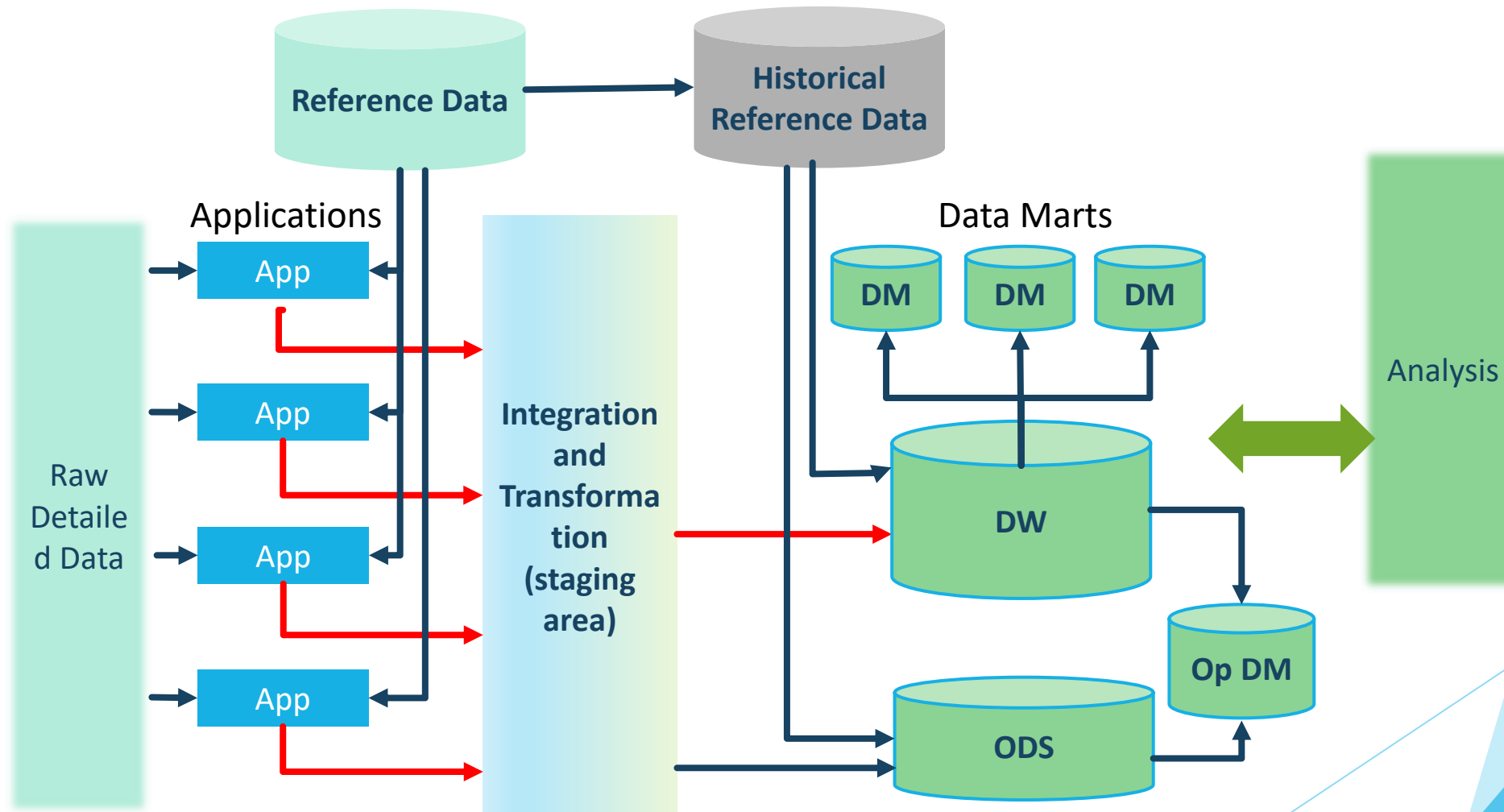
5. Approaches to Data Warehouse – Data Models

- a. **Bill Inmon:** DW is a “subject-oriented, integrated, time-variant and non-volatile collection of data (historical, detailed) in support of management’s decision-making process” (normalized relational model)
- b. **Ralph Kimball:** DW is “a copy of transaction data specially structured for query and analysis” (dimensional model)

a. Corporate Information Factory (Inmon)

- Relational Database Model
- Data Warehouse / Data Marts
 - Staging Area
 - Reference Data, Master Data
 - Integration and Transformation
 - Operational Data Store (OD)
 - Operational Data Mart (OpDM)
 - Operational Reports

5. a. The Corporate Information Factory

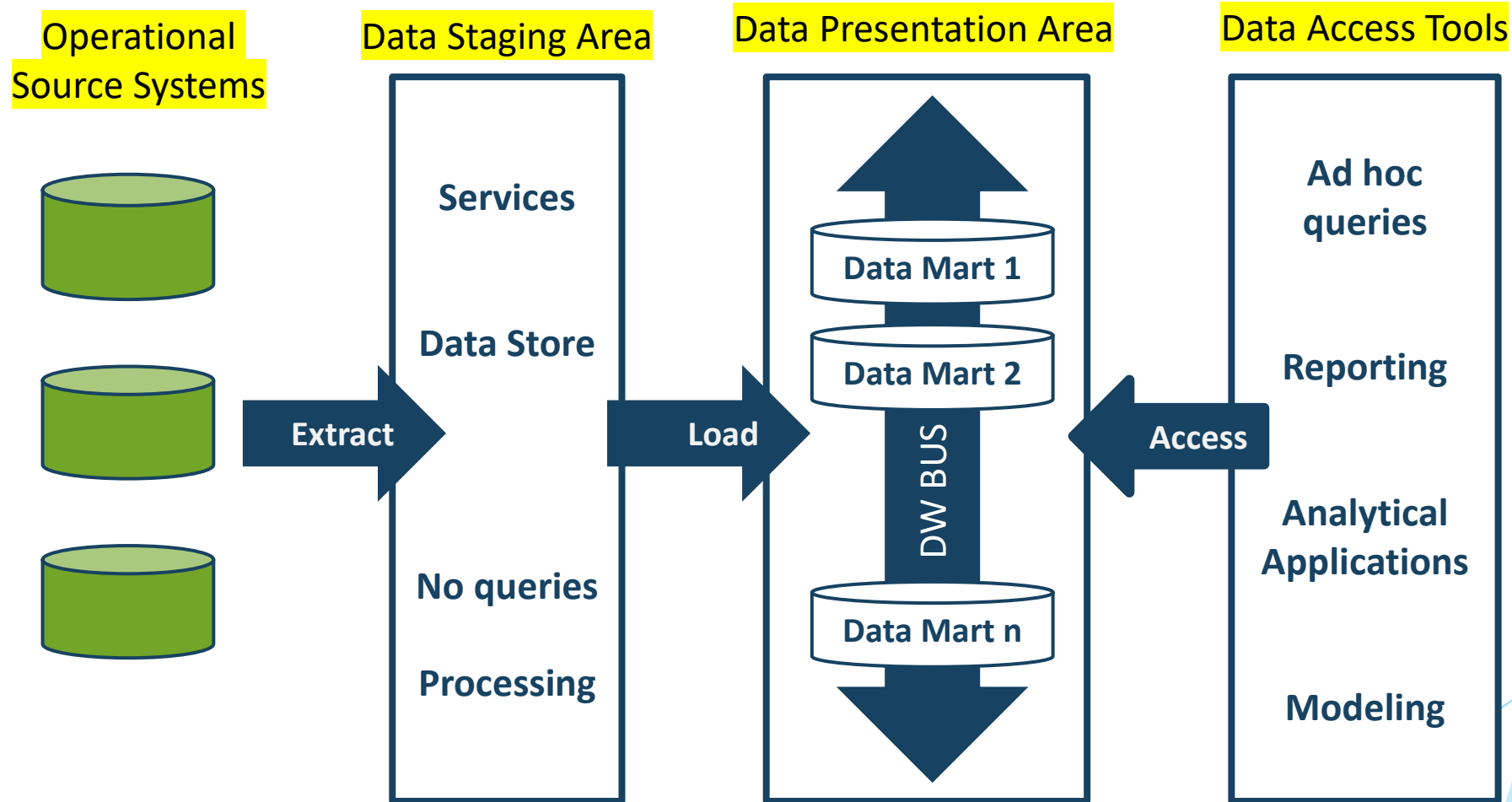


b. Dimensional Data Warehouse (Kimball)

The DW encompasses all components in the data staging and data presentation areas

- **Chess Pieces view of DW/BI architecture:**
 1. Operational source system
 2. Data staging area
 3. Data presentation area
 4. Data access tools

5. b. Kimball's DW Chess Pieces view of DW/BI



Essential concepts – cont'd

6. DW - Conformed Dimensions

- Dimension tables **conform** when **attributes** in separate dimension tables have the same column names and domain contents²
- Conformed dimensions allow facts to be **categorized** in the same way **across** multiple fact tables, and data marts, ensuring consistent reporting, analytics across the enterprise³
- When a Dimension is not shared by multiple facts is not a conformed dimension, or only dimensions that are shared for more than 1 fact may be a conformed dimension

¹ Source: Copyright © 2023 DAMA International – DMBOK2 - Technics Publications, Basking Ridge, New Jersey, USA

²<https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques/conformed-dimension/>

³<https://searchdatamanagement.techtarget.com/>

Essential concepts – cont'd

7. DW - Bus

Multiple Fact tables will share the common, or conformed dimensions via a “DW Bus”

	Processes	Dimensions				
		Subject Areas				
		Date	Product	Store	Vendor	Warehouse facility
Fact tables	Sales	X	X	X		
	Inventory	X	X	X	X	X
	Orders	X	X		X	
	Conformed Dimension Candidate	YES	YES	YES	YES	NO

Date is a common conformed dimension because its attributes (day, week, month, quarter, year, etc.) have the same meaning when joined to any fact table.

Essential concepts – cont'd

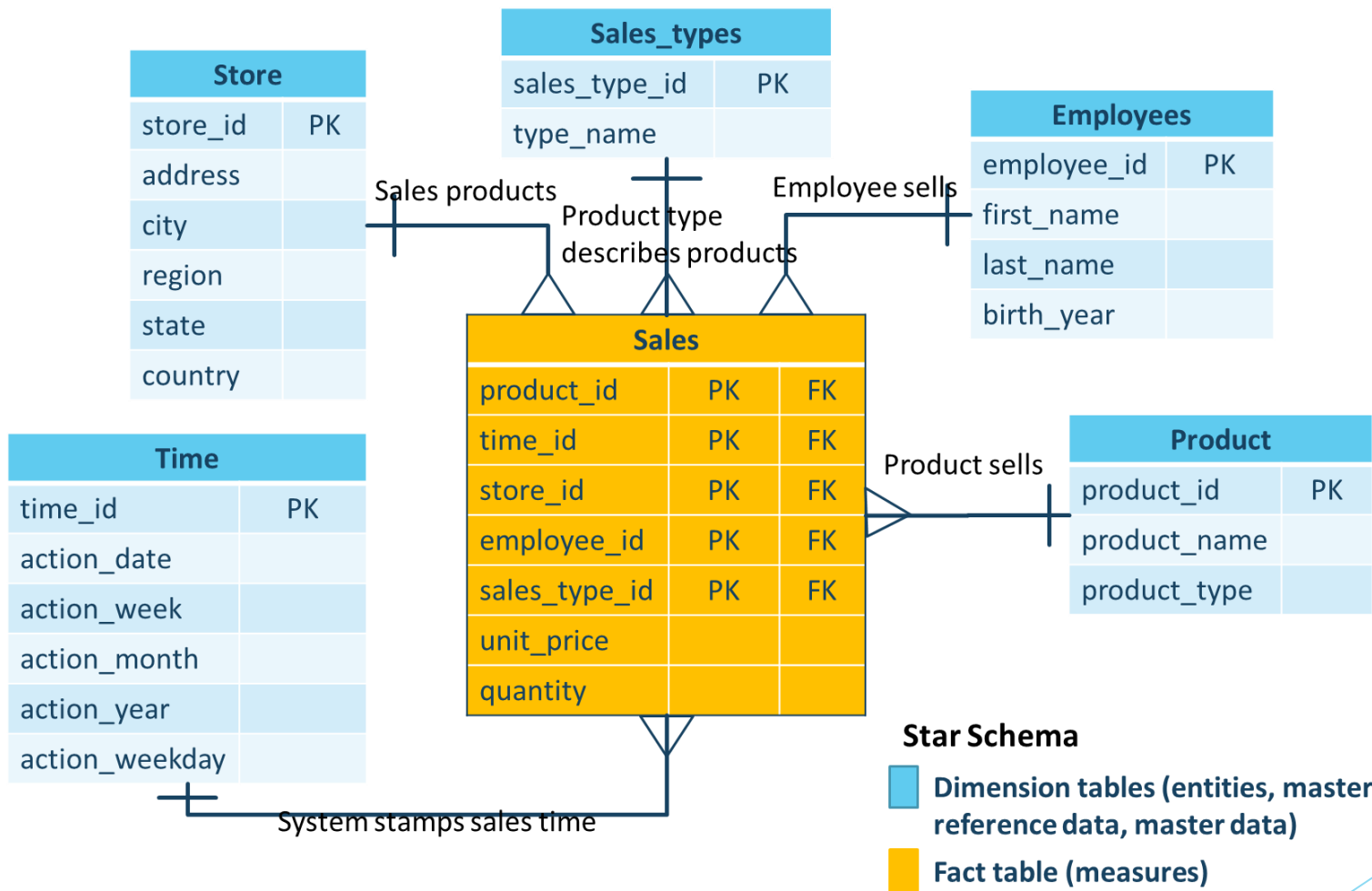
8. **Star Schema**

Star Schema comes from a model in which **one Fact table joins with many Dimension tables**, and when view as a diagram, appears as a **Star**.

Dimensional models are comprised of ...

- a) **Facts or Measures**, which contain **quantitative** data about business process
- b) **Dimensions**, store descriptive attributes (nouns) related to fact data
- Note: in a Star schema dimension tables are not normalized.

8. Star Schema



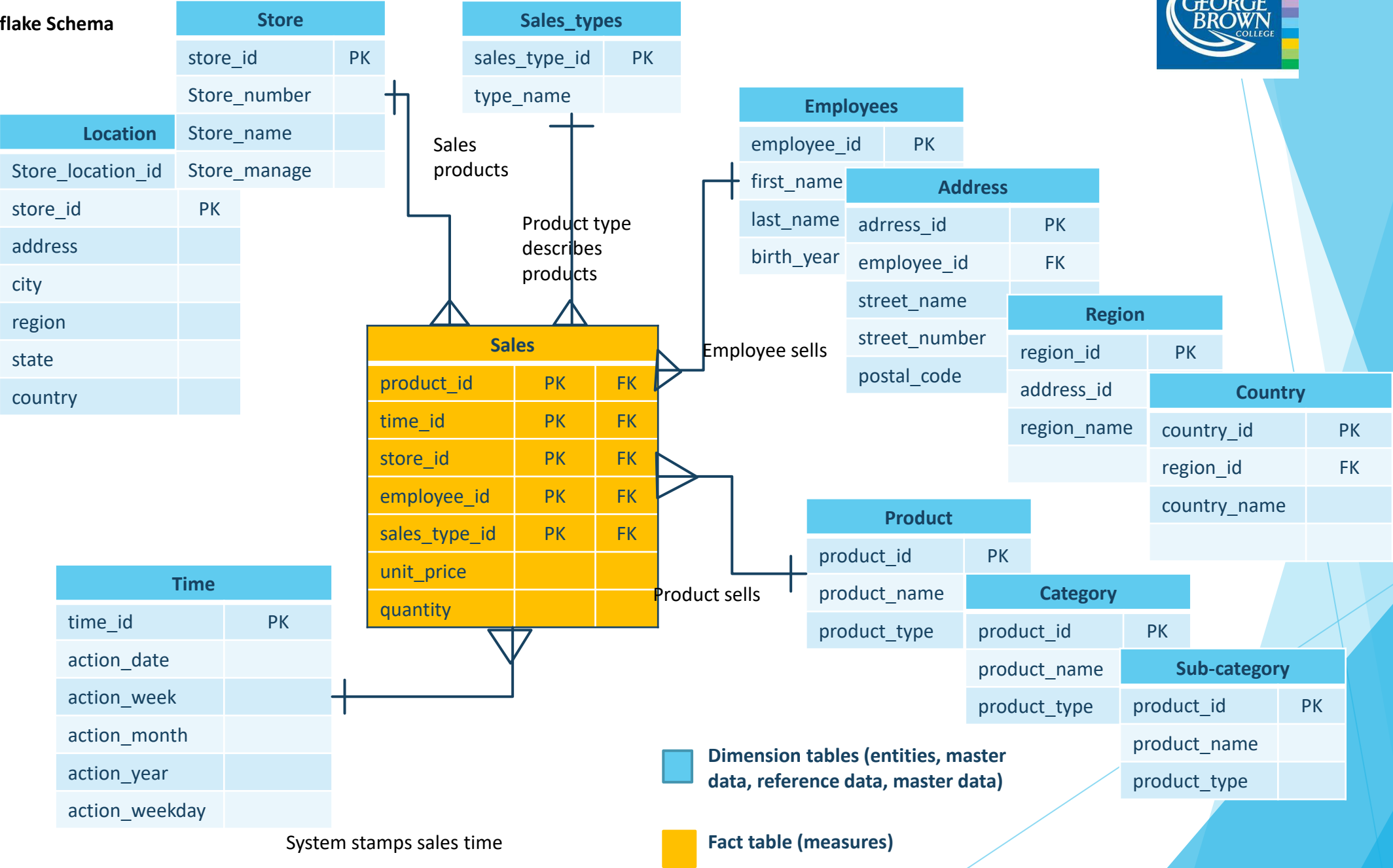
Essential concepts – cont'd

2. **Snowflake Schema**¹

- In the snowflake schema, dimension tables are completely normalized.
- Dimension tables in the snowflake schema divide themselves into more than one table. That creates the snowflake pattern.
- The snowflake schema is a “multi-dimensional” structure.

¹ <https://www.xplenty.com/blog/snowflake-schemas-vs-star-schemas-what-are-they-and-how-are-they-different/>

Snowflake Schema



Essential concepts – cont'd

Star Schema¹

- **Benefits**
 - queries are simpler
 - easier business insights reporting
 - better-performing queries
 - apps can use star schema to build cubes
- **Challenges**
 - decreased data integrity (redundancies)
 - less capable of handling complex queries

Snowflake Schema¹

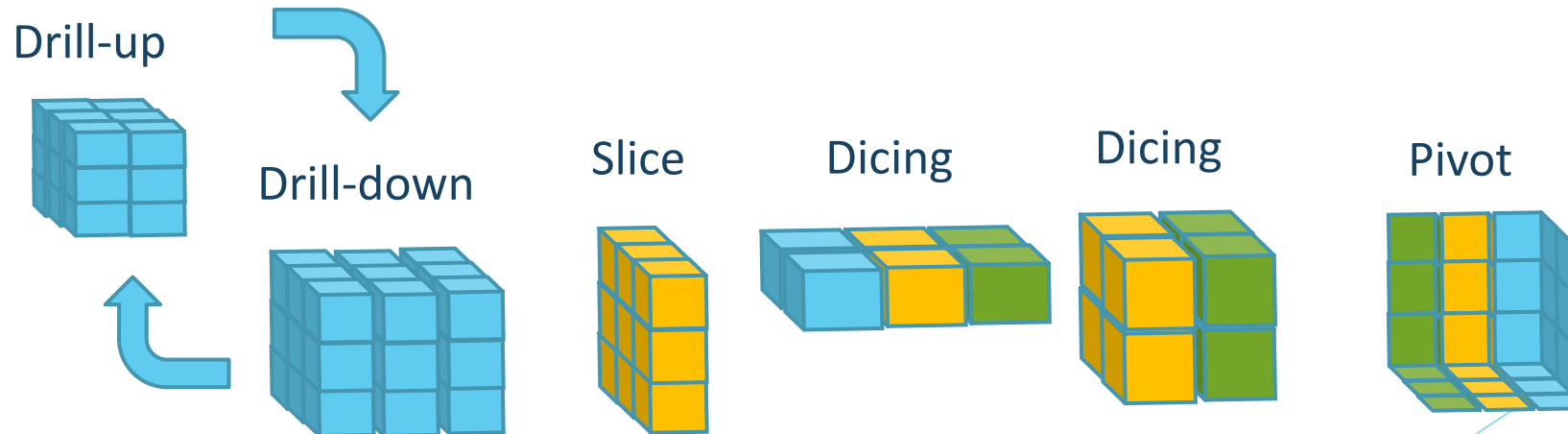
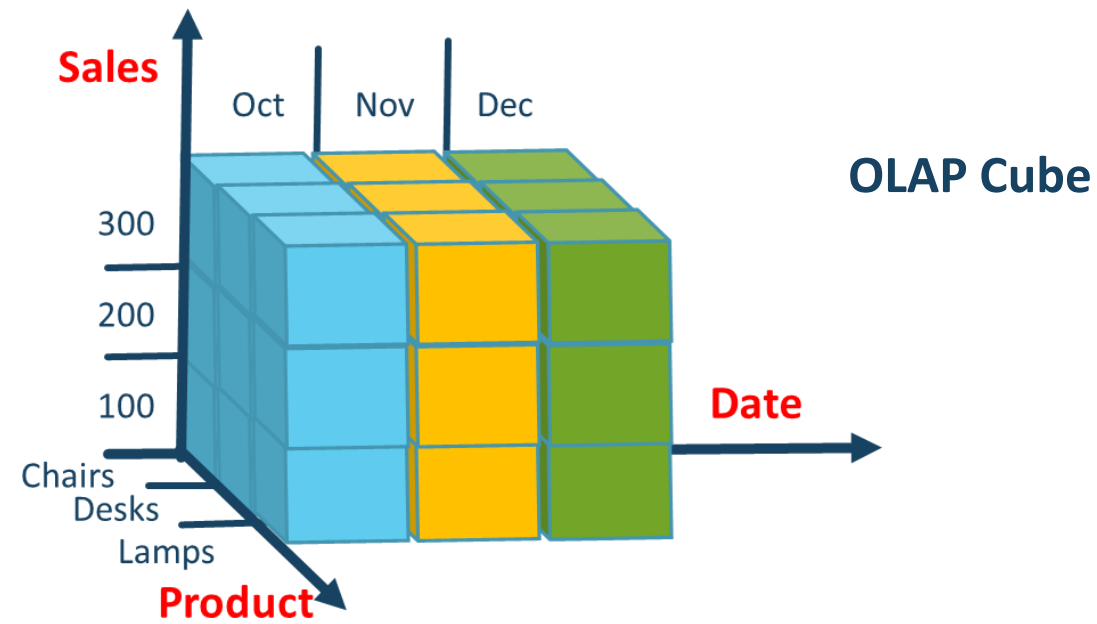
- **Benefits**
 - Compatible with many OLAP database modeling tools
 - Saves on data storage requirements
- **Challenges**
 - Complex data schemas
 - Slower at processing cube data

¹ <https://www.xplenty.com/blog/snowflake-schemas-vs-star-schemas-what-are-they-and-how-are-they-different/>

TOOLS AND TECHNIQUES

- c) Business Intelligence Technique
 - Operational Analytic Applications
 - Multi-dimensional Analysis – OLAP
 - Slice
 - Dice
 - Drill down/up
 - Pivot

Multi-dimensional Analysis



DW/BI GOVERNANCE

- Enabling Business Acceptance
 - Conceptual Model
 - Data Quality feedback loop
 - End-to-end Metadata
 - End-to-end verifiable data lineage
- Customer / User Satisfaction
- SLA
- Reporting Strategy

