

# Case Study: Building a Data Warehouse for a Retail Store

## Background

Imagine a retail company called “**SuperMart**” that sells products across multiple locations. SuperMart’s management wants a **data warehouse** to help them make better decisions about sales, inventory, and customer engagement. The goal is to have consistent, reliable, and easily accessible data for business analysis.

---

## Steps in Kimball Architecture

### 1. Identify Business Requirements:

- SuperMart’s management identifies the key questions they want to answer using the data warehouse:
  - What are the total sales per product category each month?
  - Which stores have the highest revenue?
  - Who are the most loyal customers?

### 2. Design Data Marts (Bottom-Up Approach):

- Following Kimball's approach, we start by designing specific **data marts** that serve different business needs, such as:
  - **Sales Data Mart:** Focuses on sales transactions.
  - **Customer Data Mart:** Stores customer details and purchase history.
  - **Inventory Data Mart:** Tracks stock levels and product availability.

### 3. Dimensional Modeling:

- SuperMart's data is organized into **facts** and **dimensions**.
  - **Fact Table** (central table): Holds measurable data like sales amounts, quantities sold, etc.
  - **Dimension Tables** (contextual data): Provide more details such as product information, store locations, and customer details.
- A **Star Schema** is created, with a central fact table surrounded by dimension tables.

Example:

- **Fact Table:** Sales\_Fact (columns: Sale\_ID, Product\_ID, Store\_ID, Customer\_ID, Sale\_Amount, Sale\_Date)
- **Dimension Tables:**
  - Product\_Dim (columns: Product\_ID, Product\_Name, Category)
  - Store\_Dim (columns: Store\_ID, Store\_Name, Location)
  - Customer\_Dim (columns: Customer\_ID, Customer\_Name, Email)
  - Date\_Dim (columns: Date\_ID, Year, Month, Day)

### 4. Conformed Dimensions:

- Suppose SuperMart adds another data mart for **Promotions** later on. To ensure consistent data across the different data marts, we reuse the same **Product\_Dim** and **Customer\_Dim** tables in both the Sales and Promotions data marts. These are **conformed dimensions** that provide consistency and integration.

### 5. ETL Process (Extract, Transform, Load):

- **Extract:** Data is pulled from different source systems (like point-of-sale systems, CRM, and inventory management).
  - **Transform:** Data is cleaned, formatted, and standardized (e.g., fixing inconsistent dates, handling missing data).
  - **Load:** Transformed data is loaded into the appropriate data marts (e.g., loading sales data into `Sales_Fact` and related dimensions).
6. **Access and Reporting:**
- SuperMart's business users (e.g., sales managers, marketing analysts) access the data warehouse using simple queries, dashboards, or reports. They can quickly find answers to their questions because the **star schema** design makes data easy to understand and retrieve.
- 

### Example Analysis Queries:

- To find total monthly sales, the query would sum `Sale_Amount` in the `Sales_Fact` table and group by month using the `Date_Dim`.
  - To identify top-performing stores, the query would sum sales in `Sales_Fact` and join with `Store_Dim`.
- 

## Case Study: Retail Company Data Warehouse using Kimball Architecture

**Company Background:** A retail company, *ShopNow*, operates hundreds of stores across different regions. The company collects large amounts of data related to sales, inventory, customers, and suppliers. However, this data is spread across multiple systems, making it difficult for analysts to generate reports and gain business insights quickly.

**Problem:** *ShopNow* faced challenges in:

1. **Integrating data** from various operational systems like sales, inventory management, and customer relationship management (CRM).
2. **Generating timely reports** for sales performance, inventory levels, and customer behavior.
3. **Providing analytics** for marketing and decision-making across different departments.

**Solution (Kimball Approach):** *ShopNow* decided to implement a data warehouse using Kimball's **dimensional modeling** and architecture. Here's how they did it:

1. **Identify Business Processes and Requirements:**
  - The company first met with key stakeholders to understand their reporting needs and identified key business processes, such as *sales transactions*, *inventory management*, and *customer analysis*.
2. **Design and Develop Data Marts:**
  - The **Sales Data Mart** was created first, containing a **Fact Table** with sales transactions and **Dimension Tables** for products, stores, dates, and customers.

- The **Inventory Data Mart** and **Customer Data Mart** were developed next, each structured similarly with fact and dimension tables.
- 3. **Use Dimensional Modeling:**
  - The **star schema** was applied, with each data mart having a central fact table surrounded by dimension tables, making data easy to query and analyze.
  - The **sales fact table** included measures like *sales amount*, *quantity sold*, and *discounts*, while dimension tables stored information like *product details*, *store locations*, and *sales dates*.
- 4. **ETL (Extract, Transform, Load) Process:**
  - Data from multiple sources (e.g., point-of-sale systems, customer databases) was extracted, cleaned, and transformed into the dimensional structure, and loaded into the data warehouse.
- 5. **Incremental Development:**
  - Initially, the sales data mart was deployed, allowing quick access to sales data and faster report generation.
  - Over time, additional data marts (e.g., inventory, customer) were integrated into the data warehouse, providing a comprehensive view of business data.

### **Results:**

1. **Improved Reporting Efficiency:** *ShopNow* achieved faster reporting times with consolidated data in an easy-to-query format.
2. **Better Business Insights:** Analysts could now track sales trends, monitor inventory levels, and understand customer behavior more effectively.
3. **Scalability:** The bottom-up approach allowed for incremental enhancements, adding new data marts based on evolving business needs.

**Conclusion:** The use of **Kimball's architecture** helped *ShopNow* quickly deliver business value, integrate data efficiently, and support a flexible and user-friendly reporting environment.