**BUSINESS TEMPLATE**

**DATA WAREHOUSE DESIGN AND IMPLEMENTATION**

**LUMEN ELECTRONICS**



# **1. BUSINESS DESCRIPTION**

# **1.1 BUSINESS BACKGROUND**

**LUMEN ELECTRONICS** is a growing **consumer electronics brand** specializing in smartphones, laptops, wearables, and smart home products. The company operates through two major sales channels:

1. **Offline retail stores** across the **United States**, and
2. **An online e-commerce platform** serving customers **worldwide**.

Over the past few years, LUMEN ELECTRONICS has experienced rapid expansion in both physical and digital markets. This growth has led to the generation of large volumes of data across multiple systems — including sales transactions, product inventories, customer demographics, and logistics records.

However, the data is stored in **separate systems** for offline and online operations, making it difficult for the company to monitor overall performance effectively. As a result, management faces challenges in generating unified reports, understanding customer behavior, and forecasting future trends.

To address these issues and build a data-driven culture, LUMEN ELECTRONICS decided to develop a **centralized data warehouse system** that integrates all sales data into a single analytical environment.

# 1.2 PROBLEM STATEMENT

The existing data management process at **LUMEN ELECTRONICS** lacks integration, efficiency, and analytical capability. The company’s major challenges are:

* **Data fragmentation:** Offline and online data are kept in different systems, preventing holistic analysis of sales performance.
* **Manual data handling:** Reports are generated manually by extracting and merging CSV files, which is time-consuming and error-prone.
* **Inconsistent formats:** Data differences such as currency types, date formats, and product naming conventions complicate analysis.
* **Limited analytics:** Without a unified database, advanced business intelligence tools cannot be effectively used.
* **Delayed decision-making:** Executives lack quick access to key performance indicators (KPIs) across both sales channels.

These issues result in **incomplete insights**, **redundant reporting efforts**, and **missed business opportunities**. A well-designed data warehouse will solve these problems by providing a unified, accurate, and scalable data foundation.

# **1.3 PROJECT VISION & OBJECTIVES**

## **Project Vision**

To create a **comprehensive enterprise data warehouse** for **LUMEN ELECTRONICS** that consolidates data from all business channels — both online and offline — enabling accurate reporting, advanced analytics, and data-driven decision-making.

## **Project Objectives**

* Integrate offline and online sales data into a unified data warehouse.
* Apply **Inmon’s Top-Down approach** to design a normalized 3NF warehouse.
* Ensure data consistency and integrity through cleaning, validation, and standardization.
* Enable executives to visualize key business metrics using **Power BI dashboards**.
* Simulate real-world enterprise data management practices using **Python**, **PostgreSQL**, and **DBeaver**.
* Provide clear, data-driven insights that support long-term strategic planning.

# **1.4 THE BENEFITS OF IMPLEMENTING A DATABASE / DATA WAREHOUSE**

By implementing a **data warehouse**, **LUMEN ELECTRONICS** will gain several business and technical advantages:

* **Centralized data storage:** Combines data from offline and online operations into a single repository.
* **Reliable and consistent reporting:** Eliminates data duplication and inconsistencies.
* **Improved decision-making:** Accurate data allows management to track performance, trends, and profitability with confidence.
* **Data quality assurance:** Structured cleaning and transformation processes enhance accuracy and reliability.
* **Enhanced scalability:** The data warehouse can grow with future business expansion and new data sources.
* **Time efficiency:** Automated ETL processes minimize manual reporting work.
* **Business intelligence integration:** Enables advanced visual analytics and KPI dashboards through Power BI.

Ultimately, this system will serve as **the single source of truth** for LUMEN ELECTRONICS, improving both operational efficiency and strategic insight.

# 

# 1.**5 METHODOLOGY / APPROACH**

The data warehouse is developed following the **Inmon Approach (Top-Down)** — which focuses on building an **enterprise-wide, normalized (3NF)** data warehouse before developing smaller **data marts** for department-specific analysis.

### **Development Steps**

1. **Data Generation and Extraction:**Two realistic datasets were generated using **Python**, representing offline and online sales operations.
2. **Data Transformation:**Data was cleaned and standardized — including date formatting, currency conversion, and handling missing or inconsistent values.
3. **Data Loading:**The transformed data was imported into a **PostgreSQL** database using **DBeaver** as the primary management tool.
4. **Data Modeling:**The data warehouse schema was designed using **Entity–Relationship (ER) modeling** and normalized to **3rd Normal Form (3NF)** to ensure data integrity and eliminate redundancy.
5. **Data Visualization:**After successful data integration, **Power BI** dashboards were built to visualize performance indicators such as revenue by region, best-selling products, and sales growth trends.

This structured approach ensures that LUMEN ELECTRONICS’ data warehouse is scalable, accurate, and suitable for real-world business analysis.

# 1.6 DATASET DESCRIPTIONS

## **Dataset 1: Offline Sales (Retail Stores in the USA)**

* **Concept:** Represents transactional data from LUMEN ELECTRONICS’ physical stores across the United States.
* **Size:** Approximately **550,000 rows** and **23 columns**.
* **Generated by:** Python script simulating daily store transactions.
* **Key Attributes:**
  + Transaction ID
  + Store ID
  + Product ID
  + Customer ID
  + Date
  + Region
  + Quantity Sold
  + Payment Type
  + Total Amount
  + Salesperson ID
* **Purpose:** To analyze store-level performance, customer behavior, and regional sales distribution within the U.S. market.

## **Dataset 2: Online Sales (E-Commerce Platform)**

* **Concept:** Represents global online transactions made through LUMEN ELECTRONICS’ e-commerce website.
* **Size:** Approximately **550,000 rows** and **30 columns**.
* **Generated by:** Python script with random yet realistic data distributions.
* **Key Attributes:**
  + Order ID
  + Product ID
  + Customer Country
  + Currency
  + Price in Local Currency
  + Converted Price (USD)
  + Warehouse ID
  + Seller ID
  + Order Date
  + Delivery Method
* **Special Features:**
  + Multi-currency support with conversions to USD for unified financial analysis.
  + Global customer coverage enabling cross-country comparisons.
* **Purpose:** To evaluate online sales trends, global demand, and e-commerce performance metrics.

Both datasets combined total over **1.1 million records** and **53 attributes**, forming a comprehensive foundation for LUMEN ELECTRONICS’ enterprise data warehouse.