

# Sales Data Pipeline and Dashboard

---

## Introduction

The **Sales Data Pipeline and Dashboard** project was developed to provide actionable insights into sales performance for a computer hardware business operating across multiple regions.

The business faced challenges in monitoring revenue, identifying top customers, and tracking regional sales performance in a timely manner. This project addresses those challenges by building an **end-to-end data pipeline** and an **interactive Power BI dashboard** that connects directly to sales records stored in a **MySQL database**.

By integrating SQL-based data processing with real-time dashboards, the project enhances reporting accuracy, timeliness, and decision-making capabilities.

## Tech Stack

- **SQL (MySQL):** Data extraction, transformation, and query building.
- **Power BI:** Data modeling, KPI tracking, DAX calculations, and interactive dashboards.
- **Excel:** Quick data validation and exploratory analysis.

## Project Workflow

### 1. Data Discovery

- Sales records obtained from a MySQL database containing 150K+ transactions.
- Data included customers, products, markets, and sales details.

### 2. ETL Process (Extract, Transform, Load)

- Performed in SQL and Power BI.
- Tasks included cleaning, handling invalid values, and normalizing currencies.

### 3. Data Modeling

- Relationships defined between sales, products, customers, and markets.
- Applied DAX calculations to generate KPIs and time-based insights.

### 4. Dashboard Development

- Built an interactive dashboard in Power BI to provide real-time insights.
- Visuals included revenue by region, sales quantities, top customers, and revenue trends.

## Dashboard Features

- **Revenue by Region:** Compare total revenue contributions across markets.
- **Sales Quantity by Region:** Monitor demand distribution.
- **Top Customers:** Identify high-value customers and their revenue contributions.
- **Top Performing Markets:** Pinpoint growth opportunities by region.
- **Revenue Trend:** Track monthly fluctuations and seasonal peaks.

## Key Insights

- **Delhi NCR** emerged as the top-performing region, generating **0.17bn revenue**.
- **Electricalsara Stores** was the leading customer, contributing **~139M revenue**.
- The business experienced **seasonal revenue spikes in July 2019**, followed by a sharp decline in Q4.
- Certain regions (e.g., Patna, Lucknow) contributed minimal revenue, suggesting potential areas for growth or re-allocation of resources.

## Business Impact

- Improved **reporting timeliness and accuracy** by automating data pipelines.
- Enabled **data-driven decision-making** with real-time dashboards.
- Provided **visibility into customer and market performance**, helping prioritize sales strategies.
- Created a foundation for future **scalability** with cloud or enterprise-level BI integration.

## Learning Outcomes

- Hands-on experience in **end-to-end BI project execution** (data discovery → ETL → visualization).
- Practical application of **SQL pipelines** for data processing.
- Advanced skills in **Power BI modeling, DAX calculations, and dashboard design**.

## Future Enhancements

- **Cloud Migration:** Host the database on AWS RDS or Azure SQL for scalability.
- **Scheduled ETL:** Automate refreshes with tools like Apache Airflow or Power Automate.
- **Advanced Analytics:** Incorporate predictive modeling to forecast sales.

- **User Access Control:** Deploy dashboards with role-based access for business units.

## Project Deliverables

- **SQL Dataset** (sales records in .sql format).
- **SQL Queries** used for extraction and validation.
- **Power BI Dashboard** (.pbix file).
- **Dashboard Screenshot** for quick reference.

## Dashboard Preview

