Goal

To build a complete, end-to-end data analytics solution using:

- Excel for raw data input
- Python for data cleaning and transformation
- MySQL for storage and querying
- Power BI for interactive dashboard and business insights

The primary objective: Provide actionable insights into sales performance, customer behavior, and marketing effectiveness.

What We Did

Data Preparation:

- Simulated 4 datasets in Excel: sales, customers, products, and marketing spend

Data Cleaning (Python):

- Removed missing values and duplicates
- Standardized dates and categories
- Created derived columns like total_price and SaleMonth

Data Loading (MySQL):

- Created retail_db schema
- Loaded cleaned data using SQLAlchemy and Pandas

Dashboarding (Power BI):

- Connected to MySQL
- Built visualizations and interactive insights

Challenges Faced

Challenge: Mismatched date granularity

Solution: Created a Month-Year key to link tables

Challenge: Missing data and duplicates

Solution: Used Pandas to clean and merge

Challenge: Complex relationships

Solution: Normalized tables and performed pre-joins

Challenge: Comparing marketing and sales

Solution: Unified date dimension and DAX calculations

Insights Uncovered

- Revenue spikes align with marketing campaigns
- Top 5 products contribute about 40 percent of revenue
- Certain regions outperform others
- Steady customer growth and repeat purchases

Bottom Line

Successfully built a scalable, end-to-end analytics pipeline using:

- Excel to Python to MySQL to Power BI
- Automated data cleaning and integration
- Insightful KPIs for real-time decision making

Dashboard empowers teams to make data-driven decisions across departments.