# Sales Analysis Project Report

## 1. Project Overview

This project aims to analyze and visualize a large-scale sales dataset using PostgreSQL and Power BI. The data pipeline involves importing data from text files into PostgreSQL, creating relationships among multiple tables, and visualizing insights through a detailed dashboard.

## 2. Tools & Technologies Used

- PostgreSQL: For data storage and querying

- Power BI: For data modeling, visualization, and reporting

- Power Query: For data transformation

## 3. Data Schema

- customers: Stores customer information (ID, name, email, country)

- orders: Contains order-level details (order ID, date, customer ID, total amount)

- order_items: Line items for each order (order ID, product ID, quantity, unit price)

- products: Product catalog (ID, name, category)

- sales_person: Information about the sales representatives (ID, name, region)

## 4. Data Import Process

- Data was imported using PostgreSQL's COPY functionality from `.txt` files.

- Adjustments included correcting data types (e.g., extending varchar lengths, removing currency symbols from numeric fields, converting ID formats).

- Relationships and foreign key constraints were validated post-import to ensure referential integrity.

## 5. Data Cleaning & Transformation

- In Power Query, calculated columns were added for `Month Name`, `Year`, and clean numeric values.

- Date fields were standardized to support time-based slicing and aggregation.

- Redundant and helper columns were removed or hidden for cleaner modeling.

## 6. Power BI Dashboard

- KPI Cards with Icons: Average Order Value, Total Sales, Total Orders, Total Customers

- Donut Chart: Total Sales by Year

- Bar Chart: Top 10 Countries by Sales

- Bar Chart: Top 5 Customers

- Bar Chart: Top 5 Products by Sales

- Month Slicer: Enables dynamic month-based filtering

## 7. Insights Extracted

- 2025 had the highest sales performance, with over 113.89M in revenue.

- Congo was the top-performing country by sales volume.

- Robert Smith led customer contributions with over 108K in purchases.

- The product “Street” led all product categories in revenue.

## 8. Challenges Faced

- Resolving data type mismatches during import

- Handling duplicate keys and refining constraints

- Cleaning irregular values in the `total_amount` and `country` fields

- Ensuring correct date formatting for time-based slicers

## 9. Conclusion

The project demonstrates the construction of an end-to-end analytics pipeline using PostgreSQL and Power BI. By transforming raw sales data into a highly visual and interactive dashboard, the project empowers stakeholders with actionable business insights across regions, customers, and product categories.

## Appendix: Dashboard Preview

*Final Power BI Dashboard Screenshot:*