Sales Data Analysis Project Report

1. Executive Summary

This project demonstrates end-to-end sales data analysis using Python. The objective was to clean raw transactional data, engineer meaningful features, perform exploratory data analysis (EDA), and generate actionable business insights. The analysis supports data-driven decision-making for revenue optimization and regional performance evaluation.

2. Business Objectives

- Analyze regional sales performance.
- Evaluate revenue and profitability across regions.
- Identify high-performing and underperforming markets.
- Calculate and interpret profit margins.
- Provide insights for revenue growth and cost optimization.

3. Data Cleaning & Processing

- Converted numerical columns using pd.to_numeric() with error handling.
- Handled missing values using median imputation.
- Ensured correct data types for analysis.
- Created derived columns: Revenue, Profit, Profit_Margin.
- Applied conditional logic to avoid division-by-zero errors.

4. Key Metrics

Metric	Formula / Description
Revenue	Quantity × Unit Price
Profit	Revenue – Cost
Profit Margin	Profit / Revenue (if Revenue > 0)

5. Exploratory Data Analysis (EDA)

- Revenue by Region comparison using bar charts.
- Profit distribution analysis.
- Profit Margin evaluation across regions.
- Detection of potential cost inefficiencies.

- Identification of highest revenue-generating region.

6. Key Business Insights

- North America recorded the highest revenue contribution.
- Some regions demonstrate strong revenue but comparatively lower margins.
- Profitability varies significantly across markets.
- Strategic focus on cost control could improve margins in mid-performing regions.

7. Tools & Technologies

- Python
- Pandas
- NumPy
- Matplotlib
- Jupyter Notebook
- Git & GitHub