RETAIL & WAREHOUSE SALES ANALYSIS

Project Overview

This project analyzes sales performance from warehouse and retail channels for the year 2020. Using Python for data preparation and Power BI for visualization, the project uncovers:

- Monthly sales trends
- Best-selling products
- Category-wise performance
- Supplier contribution
- Retail vs warehouse sales patterns
- Sales forecasting

The final output is a fully interactive Power BI dashboard and a clean dataset ready for business decisions.

Problem Statement

The client/company wants to understand:

- Which products generate the most revenue?
- How do retail sales differ from warehouse sales?
- Which suppliers contribute the most?
- What are the monthly sales patterns?
- Can future sales be predicted?

The goal is to support decision-making in inventory, procurement, and business planning.

Objectives

Primary Objectives:

- Clean and preprocess raw sales data
- Generate meaningful KPIs
- Identify best-selling items and suppliers
- Compare retail vs warehouse sales
- Build an interactive dashboard
- Provide actionable business insights

Secondary Objectives:

- Use forecasting to predict future trends
- Automate cleaning using Python
- Create documentation and project deliverables

Tools & Technologies

Tool	Purpose
Python (Pandas, Matplotlib, Seaborn)	Data cleaning, transformation, EDA
Power BI	Dashboard & visualization
Power Query	Data shaping & modeling
Excel	Intermediate checks
GitHub	Project hosting

Dataset Description

Dataset Name: Warehouse_and_Retail_Sales
Rows: 900+
Columns: 13

Key Columns

- YEAR
- MONTH
- SUPPLIER
- ITEM CODE
- ITEM DESCRIPTION
- ITEM TYPE
- RETAIL SALE
- RETAIL TRANSACTIONS
- WAREHOUSE SALES

Created/Engineered Columns

Column	Description
TotalSales	RetailSale + WarehouseSales
Profit	Estimated 20% profit margin
MonthYear	Combined month & year
SalesCategory	Low / Medium / High

Data Cleaning & Preprocessing (Python)

✓ Steps Performed:

1. Loaded raw Excel data
2. Removed duplicates
3. Standardized column names
4. Converted Year/Month to numeric
5. Filled missing values
6. Created TotalSales and Profit columns
7. Exported clean data for Power BI

✓ Sample Code:

```
df = df.drop_duplicates()
df["TotalSales"] = df["RetailSale"] + df["WarehouseSales"]
df["Profit"] = df["RetailSale"] * 0.20
df["MonthYear"] = df["Year"].astype(str) + "-" + df["Month"].astype(str)
```

Data Modeling (Power BI)

✓ Tables Used

- clean_warehouse_retail
- monthly_trend

✓ Relationship

clean_warehouse_retail[MonthYear]

---->

monthly_trend[MonthYear]

Type: **Many-to-One (* : 1)**
Cross Filter Direction: **Single**

Dashboard Visualizations

1. KPI Cards

- Total Sales
- Total Profit
- Retail Transactions
- Average Monthly Sales

2. Line Chart

- Monthly Sales Trend + Forecast

3. Bar Charts

- Sales by Category
- Top 10 Items

4. Clustered Column Chart

- Retail vs Warehouse Performance

5. Donut Chart

- Supplier Contribution

6. Slicers

- Month
- Supplier
- Item Type

Conclusion

The project provides a clear, data-driven overview of warehouse and retail sales performance. Key findings assist in: Better inventory planning

- Supplier performance evaluation
- Decision-making for stock and promotions
- Forecasting future demand
- The dashboard enables interactive analysis for stakeholders.