**📊 Vendor Performance Analysis Report**

**📁 Project Overview**

This project analyzes vendor performance using data stored in multiple relational database tables such as purchases, sales, vendor invoices, and product prices. The goal is to create a data-driven approach to:

* Assess vendor profitability and cost-efficiency.
* Identify underperforming vendors.
* Enable informed procurement and pricing decisions.

**🛠️ Data Sources and Preprocessing**

The data was pulled from a MySQL database using SQLAlchemy connectors and integrated into a unified DataFrame for further processing. Key tables involved:

1. purchases: Actual purchase data by vendors including cost, quantity.
2. sales: Sales transaction data including selling price, sales quantity, and total revenue.
3. purchase\_prices: Vendor-wise actual vs. purchase price for each brand.
4. vendor\_invoice: Aggregated purchase quantity, purchase dollars, and freight charges per vendor.

The final summary table combined these for enriched analysis, with engineered columns like:

* GrossProfit = Sales – PurchaseCost – Freight
* ProfitMargin = GrossProfit / Sales
* StockTurnover = SalesQty / PurchaseQty

**🔍 Exploratory Data Analysis (EDA)**

**Key Observations:**

* **Gross Profit** showed extreme negative values (min: -52,002.78), indicating some products were sold at a loss.
* **Profit Margin** had entries of -∞, caused by zero sales (denominator issue).
* **High Freight Costs** (up to 257,032) hint at inefficient logistics or large vendor contracts.
* **Stock Turnover** ranged from 0 to 274.5:
  + 0: Unsold stock.
  + 1: Sales possibly coming from older inventory.

**Data Quality Checks:**

* Identified presence of outliers in price columns.
* Null and zero values were handled to avoid bias during analysis.
* Log transformations and visualizations (histograms, box plots) helped to assess skewness and normalize distributions.

**📈 Analysis Highlights**

**1. Vendor Performance Metrics**

* Each vendor was evaluated based on:
  + **Average Purchase Price**
  + **Average Profit Margin**
  + **Gross Profit Contribution**
  + **Stock Turnover Efficiency**

**2. Hypothesis Testing**

* A two-sample **T-Test** was conducted to compare:
  + Vendors with **high** vs. **low** Stock Turnover
  + Vendors with **above** vs. **below-average** Profit Margins
* Result: There was a **statistically significant** difference in profitability across high-performing and low-performing vendors.

**3. Visualizations**

* **Heatmaps** to identify correlation between pricing, profit margin, freight cost, and sales.
* **Box plots** to detect pricing outliers.
* **Bar charts** for gross profit per vendor.
* **Scatter plots** to visualize stock turnover vs profit margin.

**📌 Key Insights**

* Vendors offering **low purchase prices** don’t always yield high profits — freight and sales volume are important influencers.
* **Stock Turnover > 1** is ideal but can mask deeper issues like relying on old stock.
* A few vendors drive **majority of profit** — hinting at Pareto Principle (80/20 rule).
* Certain **high-freight vendors** consistently reduce overall margin.

**✅ Conclusion**

This vendor performance analysis successfully transforms raw transactional data into actionable insights. It empowers procurement and business teams to:

* Make vendor decisions based on **quantifiable performance**.
* Identify key levers for **profitability optimization**.
* Drive a more **data-informed sourcing strategy**.