

VENDOR PERFORMANCE ANALYSIS

End-to-End Data Analytics Project

Dipanshu Kumar

15/09/2025

PROJECT OVERVIEW

This project demonstrates a complete data analytics pipeline from raw CSV files to actionable business insights. The analysis identifies vendor optimization opportunities through comprehensive performance evaluation.

TECHNICAL IMPLEMENTATION

- DATA INGESTION: Automated CSV to PostgreSQL pipeline with chunked processing
- DATA EXPLORATION: SQL-based analysis of vendor relationships and business rules
- DATA CLEANING: Python implementation of data quality checks and feature engineering
- ANALYSIS: Statistical testing and performance metric calculation
- VISUALIZATION: Interactive Tableau dashboards for business intelligence

TOOLS & TECHNOLOGIES

PostgreSQL | Python | Tableau | Pandas | SQLAlchemy | Scipy

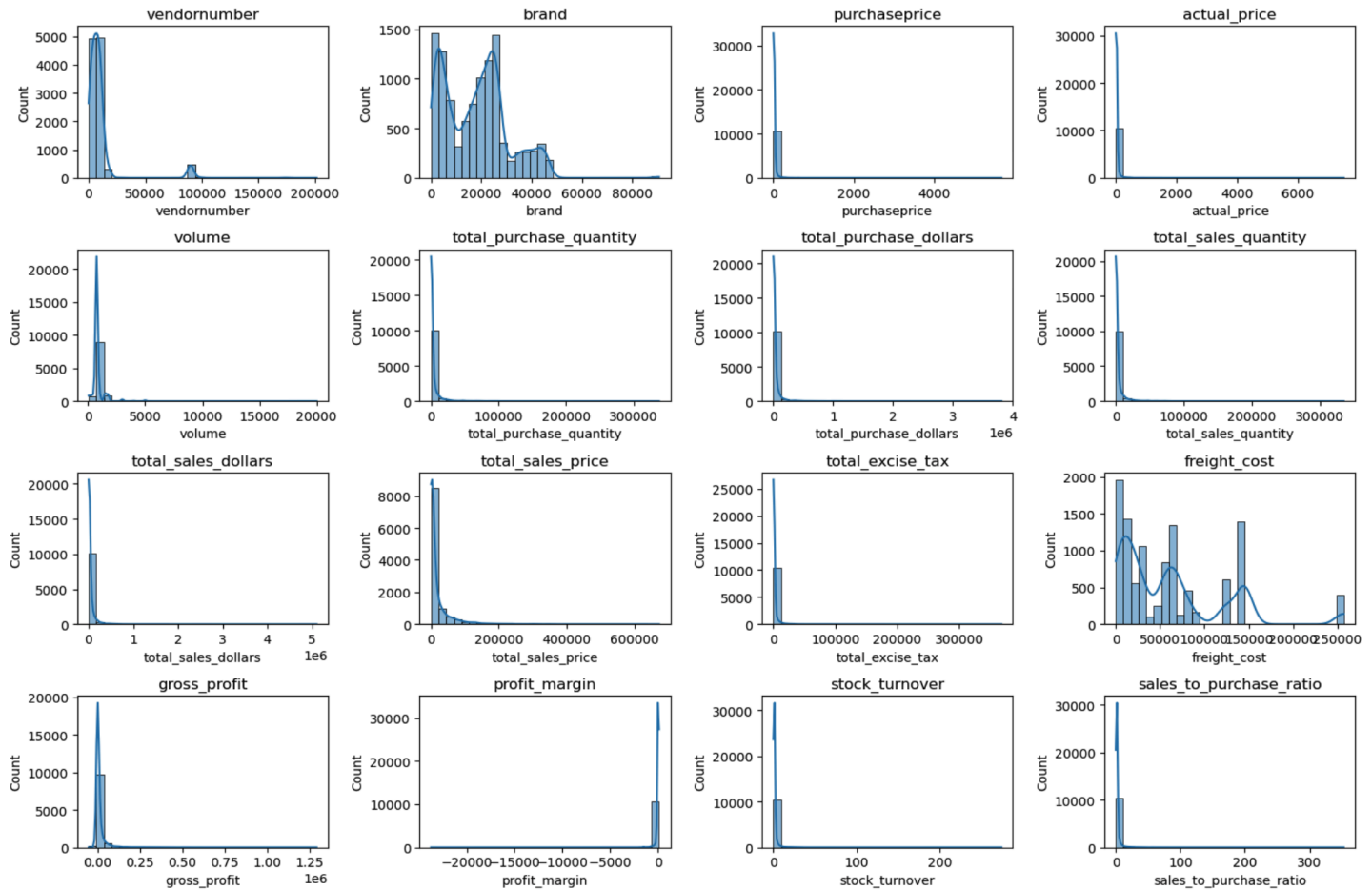
DELIVERABLES

- Automated ETL pipeline scripts
- Comprehensive SQL exploration documentation
- Vendor performance summary dataset
- Interactive Tableau dashboard
- Actionable business recommendations

EXPLORATORY DATA ANALYSIS INSIGHTS

SUMMARY STATISTICS

	count	mean	std	min	25%	50%	75%	max
vendornumber	10692.0	10650.649458	18753.519148	2.000000	3951.000000	7153.000000	9552.000000	2.013590e+05
brand	10692.0	18039.228769	12662.187074	58.000000	5793.500000	18761.500000	25514.250000	9.063100e+04
purchaseprice	10692.0	24.385303	109.269375	0.360000	6.840000	10.455000	19.482500	5.681810e+03
actual_price	10692.0	35.643671	148.246016	0.490000	10.990000	15.990000	28.990000	7.499990e+03
volume	10692.0	847.360550	664.309212	50.000000	750.000000	750.000000	750.000000	2.000000e+04
total_purchase_quantity	10692.0	3140.886831	11095.086769	1.000000	36.000000	262.000000	1975.750000	3.376600e+05
total_purchase_dollars	10692.0	30106.693372	123067.799627	0.710000	453.457500	3655.465000	20738.245000	3.811252e+06
total_sales_quantity	10692.0	3077.482136	10952.851391	0.000000	33.000000	261.000000	1929.250000	3.349390e+05
total_sales_dollars	10692.0	42239.074419	167655.265984	0.000000	729.220000	5298.045000	28396.915000	5.101920e+06
total_sales_price	10692.0	18793.783627	44952.773386	0.000000	289.710000	2857.800000	16059.562500	6.728193e+05
total_excise_tax	10692.0	1774.226259	10975.582240	0.000000	4.800000	46.570000	418.650000	3.682428e+05
freight_cost	10692.0	61433.763214	60938.458032	0.090000	14069.870000	50293.620000	79528.990000	2.570321e+05
gross_profit	10692.0	12132.381048	46224.337964	-52002.780000	52.920000	1399.640000	8660.200000	1.290668e+06
profit_margin	10692.0	-15.620770	443.555329	-23730.638953	13.324515	30.405457	39.956135	9.971666e+01
stock_turnover	10692.0	1.706793	6.020460	0.000000	0.807229	0.981529	1.039342	2.745000e+02
sales_to_purchase_ratio	10692.0	2.504390	8.459067	0.000000	1.153729	1.436894	1.665449	3.529286e+02



Negative & Zero Values:

- **Gross Profit:** Minimum value is -52,002.78, indicating losses. Some products or transactions may be selling at a loss due to high costs or selling at discounts lower than the purchase cost.
- **Profit Margin:** Has a minimum of -23730.64%, which suggests cases where costs significantly exceed revenue, indicating severe pricing issues or potential data outliers that require investigation.
- **Total Sales Quantity & Sales Dollars:** Minimum values are 0, meaning some products were purchased but never sold. These could be slow-moving or obsolete stock, leading to inventory inefficiencies.

Outliers Indicated by High Standard Deviations:

- **Purchase & Actual Prices:** The max values (5,681.81 & 7,499.99) are significantly higher than the mean (24.39 & 35.64), indicating potential premium products.

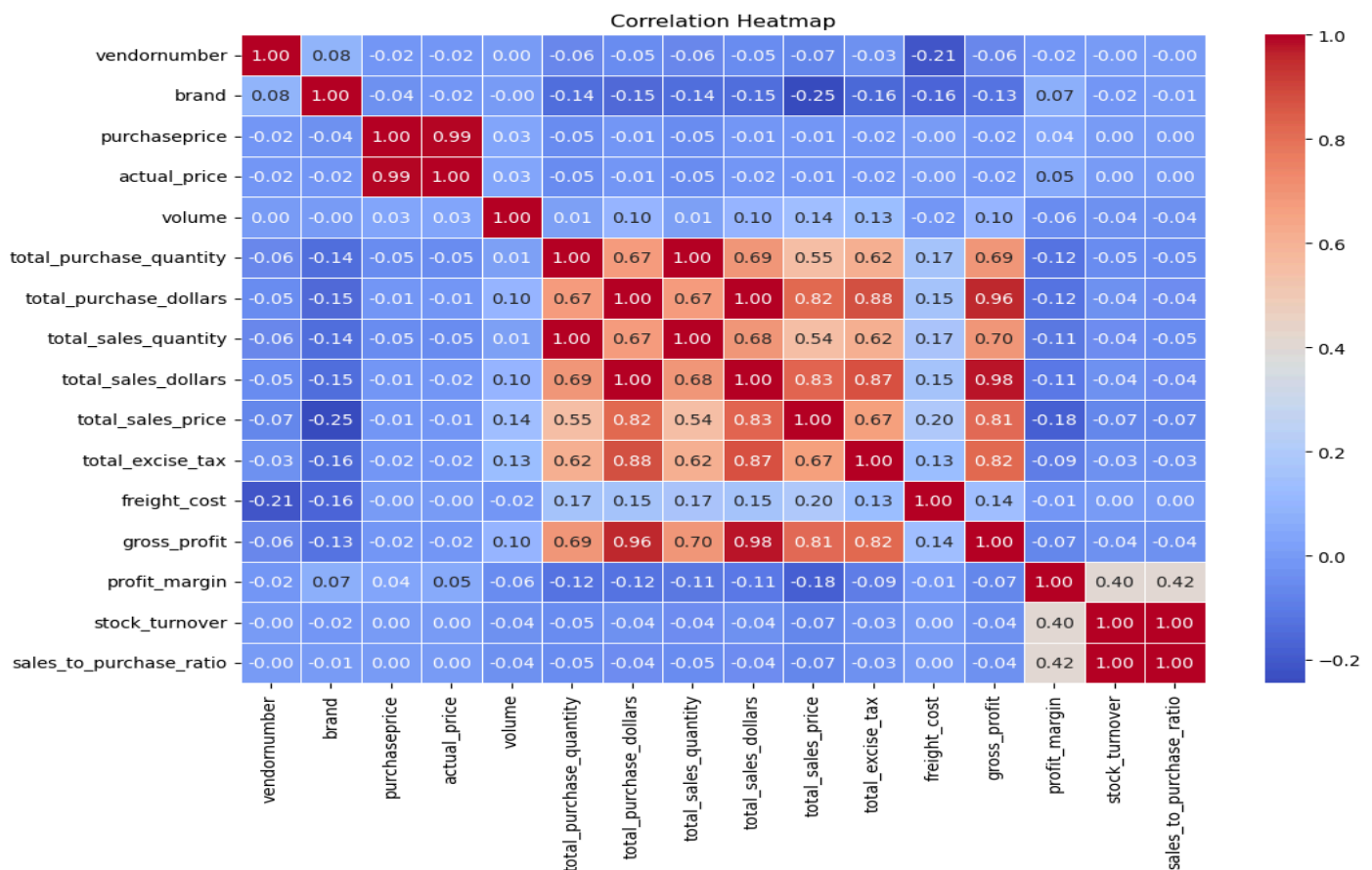
- **Freight Cost:** Huge variation, from 0.09 to 257,032.07, suggests logistics inefficiencies, bulk shipments, or erratic costs across different products.
- **Stock Turnover:** Ranges from 0 to 274.5, implying some products sell extremely fast while others remain in stock indefinitely. Value more than 1 indicates that sales quantity for that product is higher than purchased quantity due to either sales being fulfilled from older stock.

DATA FILTERING

To enhance the reliability of the insights, we removed inconsistent data points where:

- Gross Profit ≤ 0 (to exclude transactions leading to losses).
- Profit Margin ≤ 0 (to ensure analysis focuses on profitable transactions).
- Total Sales Quantity = 0 (to eliminate inventory that was never sold).

CORRELATION INSIGHTS



- **Purchase Price vs. Total Sales Dollars & Gross Profit:** Weak correlation (-0.012 and -0.016), indicating that price variations do not significantly impact sales revenue or profit.
- **Total Purchase Quantity vs. Total Sales Quantity:** Strong correlation (0.999), confirming efficient inventory turnover.
- **Profit Margin vs. Total Sales Price:** Negative correlation (-0.179), correlation (-0.179), suggesting increasing sales prices may lead to reduced margins, possibly due to competitive pricing pressures.
- **Stock Turnover vs. Gross Profit & Profit Margin:** Weak negative correlation (-0.038 & -0.055), indicating that faster stock turnover does not necessarily equate to higher profitability.

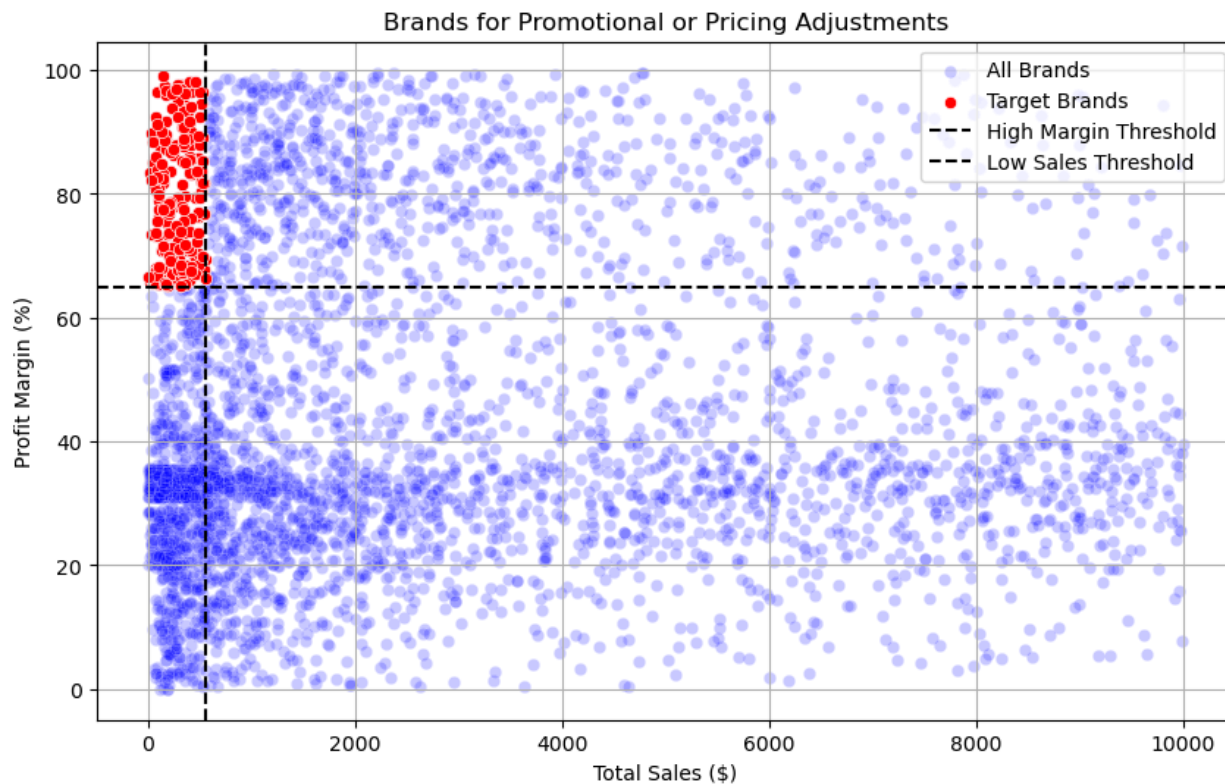
RESEARCH QUESTIONS AND KEY FINDINGS

1) BRANDS FOR PROMOTIONAL OR PRICING ADJUSTMENTS

Brands with Low Sales but High Profit Margins:			
	description	total_sales_dollars	profit_margin
6199	Santa Rita Organic Svgn Bl	9.99	66.466466
2369	Debauchery Pnt Nr	11.58	65.975820
2070	Concannon Glen Ellen Wh Zin	15.95	83.448276
2188	Crown Royal Apple	27.86	89.806174
6237	Sauza Sprklg Wild Berry Marg	27.96	82.153076
...
5074	Nanbu Bijin Southern Beauty	535.68	76.747312
2271	Dad's Hat Rye Whiskey	538.89	81.851584
57	A Bichot Clos Marechaudes	539.94	67.740860
6245	Sbragia Home Ranch Merlot	549.75	66.444748
3326	Goulee Cos d'Estournel 10	558.87	69.434752

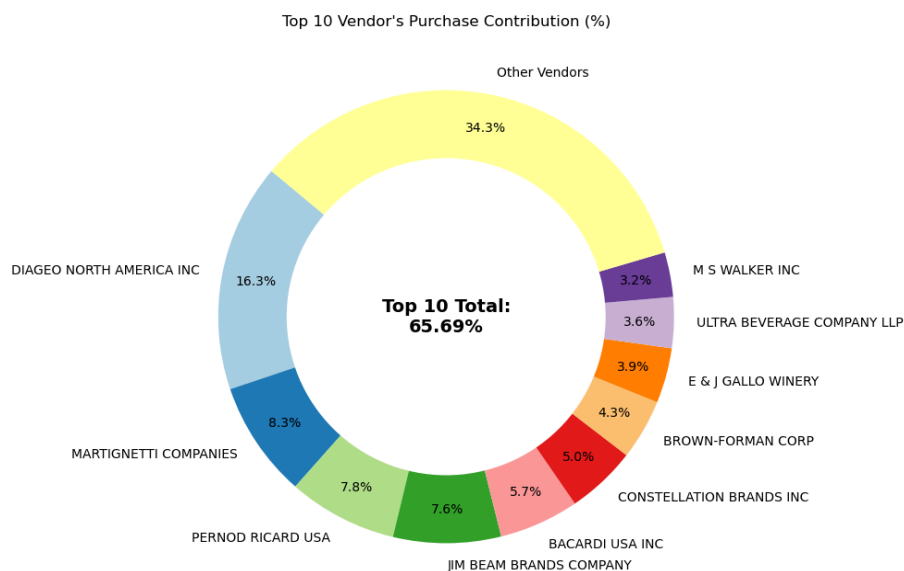
198 rows × 3 columns

198 brands exhibit lower sales but higher profit margins, which could benefit from targeted marketing, promotions, or price optimizations to increase volume without compromising profitability.



2) TOP VENDORS BY SALES & PURCHASE CONTRIBUTION

The top 10 vendors contribute 65.69% of total purchases, while the remaining vendors contribute only 34.31%. This over-reliance on a few vendors may introduce risks such as supply chain disruptions, indicating a need for diversification.



3) IMPACT OF BULK PURCHASING ON COST SAVINGS

Vendors buying in large quantities receive a 72% lower unit cost (\$10.78 per unit vs. higher unit costs in smaller orders).

Bulk pricing strategies encourage larger orders, increasing total sales while maintaining profitability.

UnitPurchasePrice	
OrderSize	
Small	39.068186
Medium	15.486414
Large	10.777625

4) IDENTIFYING VENDORS WITH LOW INVENTORY TURNOVER

Total Unsold Inventory Capital: \$2.71M

Slow-moving inventory increases storage costs, reduces cash flow efficiency, and affects overall profitability.

Identifying vendors with low inventory turnover enables better stock management, minimizing financial strain.

stock_turnover	
vendorname	
ALISA CARR BEVERAGES	0.615385
HIGHLAND WINE MERCHANTS LLC	0.708333
PARK STREET IMPORTS LLC	0.751306
Circa Wines	0.755676
Dunn Wine Brokers	0.766022
CENTEUR IMPORTS LLC	0.773953
SMOKY QUARTZ DISTILLERY LLC	0.783835
TAMWORTH DISTILLING	0.797078
THE IMPORTED GRAPE LLC	0.807569
WALPOLE MTN VIEW WINERY	0.820548

vendorname	UnsoldInventoryValue
DIAGEO NORTH AMERICA INC	722.21K
JIM BEAM BRANDS COMPANY	554.67K
PERNOD RICARD USA	470.63K
WILLIAM GRANT & SONS INC	401.96K
E & J GALLO WINERY	228.28K
SAZERAC CO INC	198.44K
BROWN-FORMAN CORP	177.73K
CONSTELLATION BRANDS INC	133.62K
MOET HENNESSY USA INC	126.48K
REMY COINTREAU USA INC	118.60K

5) PROFIT MARGIN COMPARISON: HIGH VS LOW

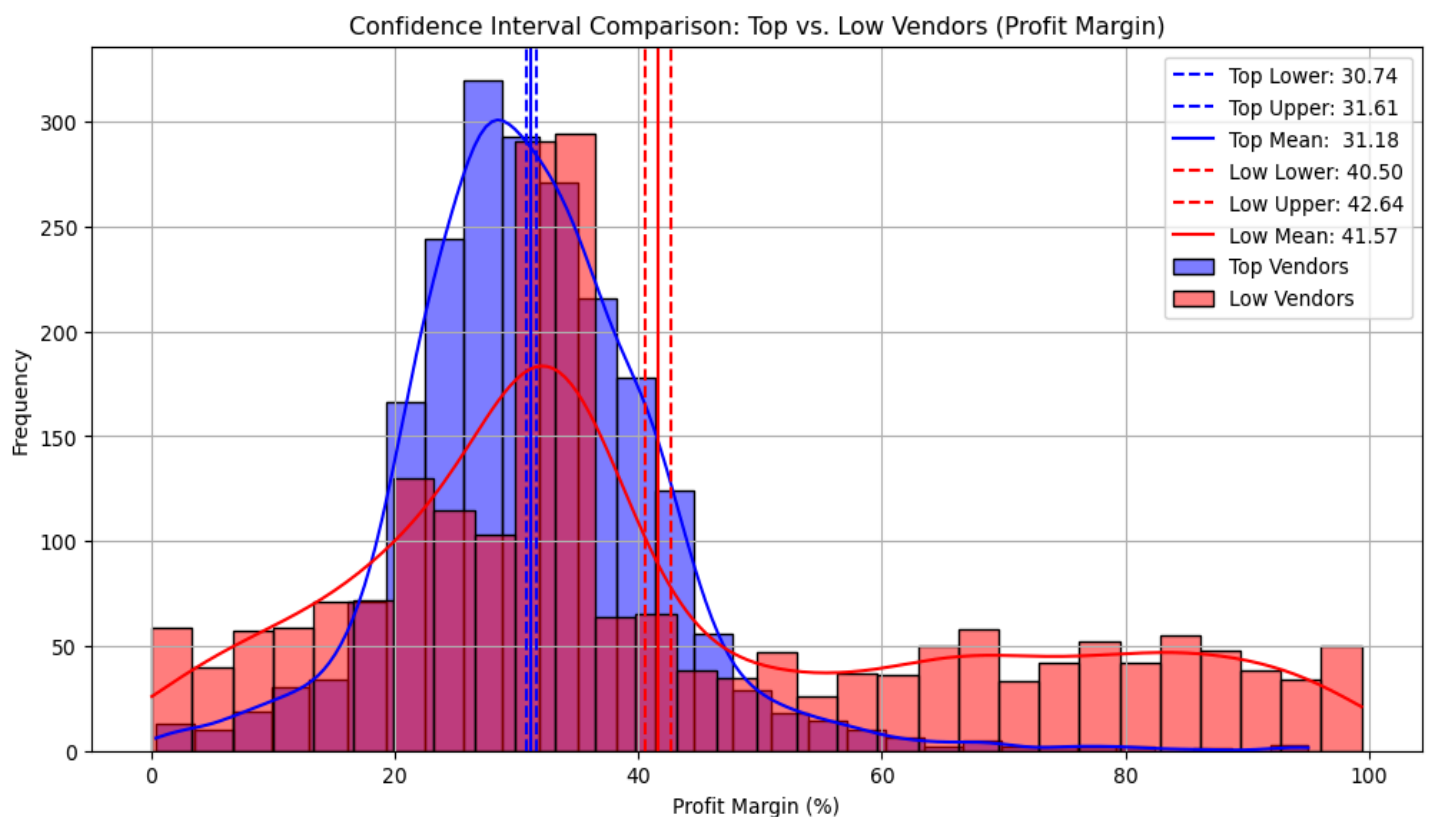
Top Vendors' Profit Margin (95% CI): (30.74%, 31.61%), Mean: 31.17%

Low Vendors' Profit Margin (95% CI): (40.48%, 42.62%), Mean: 41.55%

Low-performing vendors maintain higher margins but struggle with sales volumes, indicating potential pricing inefficiencies or market reach issues.

Actionable Insights:

- Top-performing vendors: Optimize profitability by adjusting pricing, reducing operational costs, or offering bundled promotions.
- Low-performing vendors: Improve marketing efforts, optimize pricing strategies, and enhance distribution networks.



6) Statistical Validation of Profit Margin Differences

Hypothesis Testing:

H_0 (Null Hypothesis): No significant difference in profit margins between top and low-performing vendors.

H_1 (Alternative Hypothesis): A significant difference exists in profit margins between the two vendor groups.

Result: The null hypothesis is rejected, confirming that the two groups operate under distinctly different profitability models.

Implication: High-margin vendors may benefit from better pricing strategies, while top-selling vendors could focus on cost efficiency.

FINAL RECOMMENDATION

- Re-evaluate pricing for low-sales, high-margin brands to boost sales volume without sacrificing profitability.
- Diversify vendor partnerships to reduce dependency on a few suppliers and mitigate supply chain risks.
- Leverage bulk purchasing advantages to maintain competitive pricing while optimizing inventory management.
- Optimize slow-moving inventory by adjusting purchase quantities, launching clearance sales, or revising storage strategies.
- Enhance marketing and distribution strategies for low-performing vendors to drive higher sales volumes without compromising profit margins.
- By implementing these recommendations, the company can achieve sustainable profitability, mitigate risks, and enhance overall operational efficiency.