Vendor Performance Data Analytic Business Report

Yufan(Krystal) Yao 9/17/2028

Executive Summary

This report presents a comprehensive vendor performance analysis aimed at improving profitability and operational efficiency within the retail and wholesale business context. Using SQL and Python for data extraction, cleaning, and exploratory data analysis (EDA), and Power BI for interactive dashboard development, this analysis delivers actionable insights into vendor and brand-level performance.

The analysis reveals that total sales reached $29.85M, with total purchases of $19.14M, generating a gross profit of $17.48M and an overall profit margin of 58.6%. A concentration risk was identified—the top 10 vendors accounted for 62.9% of total purchases, indicating dependency on a small number of suppliers. While several vendors such as Diageo North America Inc. and Martignetti Companies emerged as top sales contributors, a subset of vendors showed consistently low performance and profit margins, warranting further review and potential renegotiation of terms or promotional support. Similarly, brand-level analysis highlighted a few dominant high-performing brands driving most of the revenue, while numerous low-performing brands exhibited below-average profit margins and minimal sales contribution. This imbalance suggests an opportunity to optimize product assortments and reallocate inventory investments toward more profitable offerings. By pinpointing underperforming vendors and brands, this analysis enables decision-makers to:

•Focus marketing and promotional efforts where they will drive the highest returns

•Adjust bulk purchasing strategies to improve unit economics

•Reduce holding costs through better inventory turnover management

Overall, this project demonstrates the value of data-driven vendor and brand evaluation, providing a scalable framework to support strategic sourcing decisions and enhance profitability.

Introduction / Background

This analysis was initiated to address growing concerns about profitability, inventory efficiency, and vendor dependency within the retail and wholesale business context. Companies often struggle with challenges such as inefficient pricing structures, slow inventory turnover, and over-reliance on a limited number of suppliers.

These issues can result in higher holding costs, lower profit margins, and increased operational risk. To support data-driven decision-making and improve overall financial performance, this project focuses on evaluating vendor and brand performance through comprehensive sales and purchase data analysis. The primary objectives of this analysis are to identify underperforming brands that may require promotional or pricing adjustments, determine the top vendors contributing to overall sales and gross profit, analyze the impact of bulk purchasing on unit costs, assess inventory turnover to reduce holding costs and improve efficiency, and investigate profitability differences between high-performing and low-performing vendors.

By uncovering these insights, the analysis aims to inform strategic sourcing, pricing, and marketing decisions. This project involves several key stakeholders across the organization. Executive management will use the insights to guide strategic sourcing and vendor negotiation decisions, while the procurement team will rely on the findings to adjust purchasing allocations and contract terms. The sales and marketing teams can leverage the results to prioritize high-potential brands for promotional efforts, and the finance team will apply the profitability insights to evaluate return on investment from vendor partnerships. Additionally, the data analytics team plays a central role in maintaining data quality and enabling ongoing performance tracking.

The analysis covers one full fiscal year of sales and purchase data at both the vendor and brand levels, including metrics such as sales dollars, purchase costs, quantities, and profit margins. It assumes that the dataset is complete and accurate following the data cleaning process, though it does not account for external market fluctuations, seasonal demand shifts, or sudden supply chain disruptions. As such,

while the findings provide a strong historical benchmark, they should be interpreted within the context of these limitations.

Exploratory Data Analytic Insight

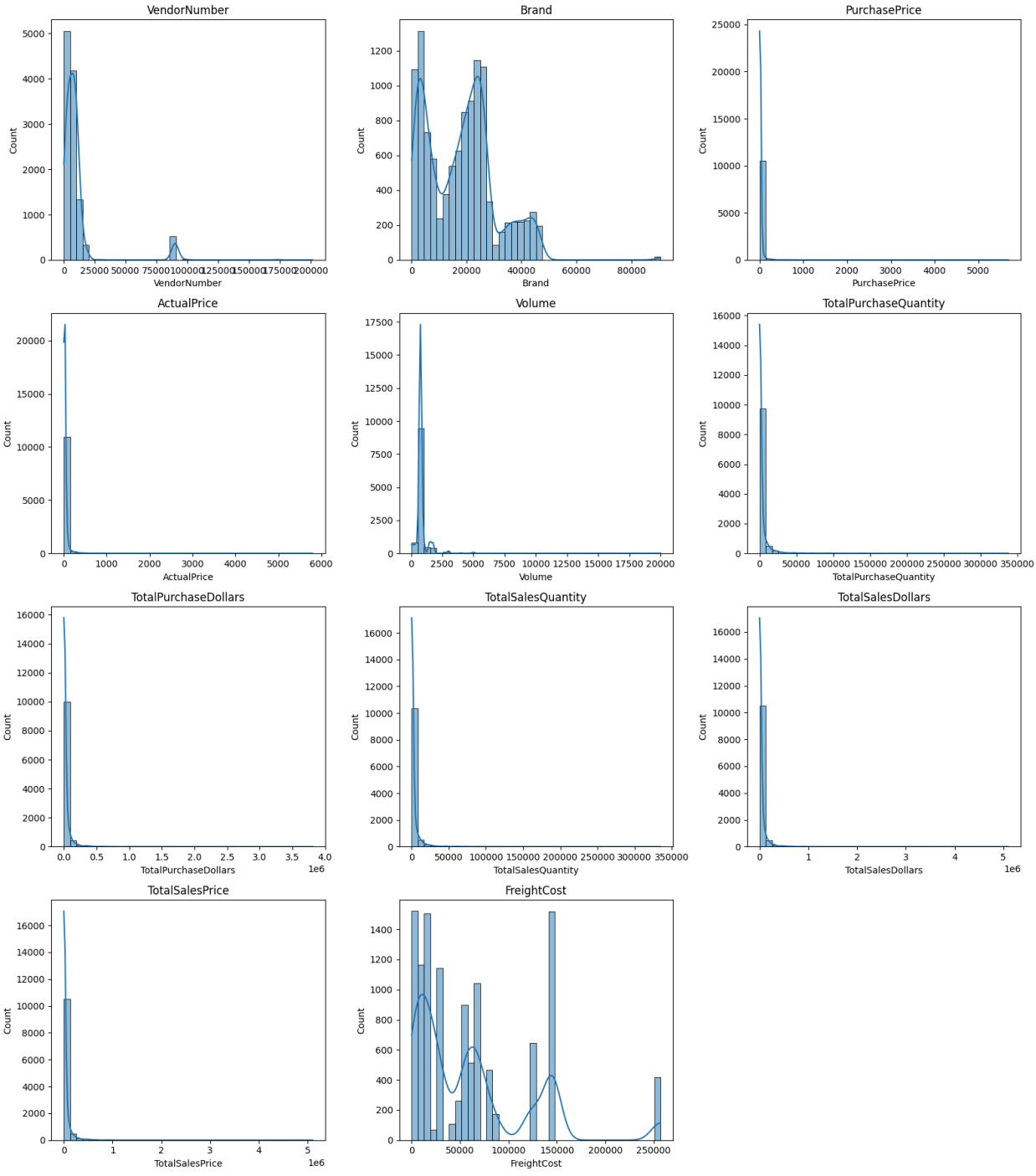
Summary Statistics

```
        VendorNumber         Brand  TotalPurchaseQuantity  \
count   11460.000000  11460.000000           10699.000000
mean    10711.366230  18070.617627            3139.020189
std     18821.779022  12626.698931           11091.078020
min         2.000000     58.000000               1.000000
25%      3960.000000   5928.000000              36.000000
50%      7153.000000  18848.500000             262.000000
75%      9552.000000  25296.500000            1976.500000
max    201359.000000  90631.000000          337660.000000

       TotalPurchaseDollars  PurchasePrice  TotalSalesQuantity  \
count          1.069900e+04   10699.000000        11277.000000
mean           3.008700e+04      24.383592         2919.027756
std            1.230286e+05     109.236900        10685.836505
min            0.000000e+00       0.000000            1.000000
25%            4.527150e+02       6.840000           26.000000
50%            3.650240e+03      10.450000          198.000000
75%            2.072748e+04      19.495000         1728.000000
max            3.811252e+06    5681.810000       334939.000000

       TotalSalesDollars  TotalSalesPrice        Volume   ActualPrice  \
count       1.127700e+04     1.127700e+04  11277.000000  11277.000000
mean        4.008716e+04     4.008716e+04    844.255210     34.727173
std         1.635082e+05     1.635082e+05    674.134671    132.071992
min         9.800000e-01     9.800000e-01     50.000000      0.490000
25%         5.756400e+02     5.756400e+02    750.000000     10.490000
50%         4.306050e+03     4.306050e+03    750.000000     15.932857
75%         2.562546e+04     2.562546e+04    750.000000     28.990000
max         5.101920e+06     5.101920e+06  20000.000000   5799.990000
```

```
         FreightCost   GrossProfit  ProfitMargin
count   11457.000000  1.146000e+04  11277.000000
mean    61260.142797  1.135796e+04     -8.009013
std     60931.477381  4.475273e+04    432.845349
min         0.090000 -5.200278e+04 -23730.638953
25%     14069.870000  5.399750e+01     17.846791
50%     50293.620000  1.069960e+03     31.838631
75%     79528.990000  7.599700e+03     44.989216
max    257032.070000  1.290668e+06    100.000000
```



Negative & Zero Values:

1.Zero sales metrics: Both Total Sales Quantity and Total Sales Dollars show heavy mass at 0. That implies many vendor–SKU combos were purchased but didn't sell in the period (slow-moving or obsolete stock), or sales weren't recorded for those lines.

2.Zero purchase metrics: A cluster at 0 for Total Purchase Quantity/Dollars indicates items that sold without a matching purchase in-window (older inventory, dropship, or data gaps).

3.Pricing zeros: If any 0 values appear for Purchase Price or Actual Price, they're data quality flags (missing price, returns netting to zero, or unit mismatch).

Outliers & Spread:

1.Prices (Purchase Price, Actual Price): Tight center around low double-digits with a long right tail and clear outliers (high-priced premium items or mis-keyed decimals). These points will dominate models unless capped or transformed.

2.Dollar totals (Total Purchase Dollars, Total Sales Dollars, TotalSalesPrice): Extremely right-skewed with rare, very large transactions—likely bulk buys, seasonal spikes, or key accounts. A log scale or historization is recommended for analysis.

3.Quantities (Total Purchase Quantity, Total Sales Quantity): Heavy right tail with many small transactions and a few very large ones. Those large points deserve review (promotions, case vs. unit counting, or one-off deals).

4.FreightCost: Very widespread with many small/zero shipments and occasional huge spikes. That pattern is typical of LTL vs. FTL shipments or consolidated loads; spikes may signal inefficiencies or surcharges.

5.Volume/Size: Mostly small pack sizes with a few very large ones. Consider normalizing monetary metrics per liter/ounce to compare fairly across products.

6.Brand / Vendor Number: Distributions are highly concentrated with a few extreme codes, suggesting a long-tail vendor and brand mix (many small vendors, a few dominant ones).

Data Filtering

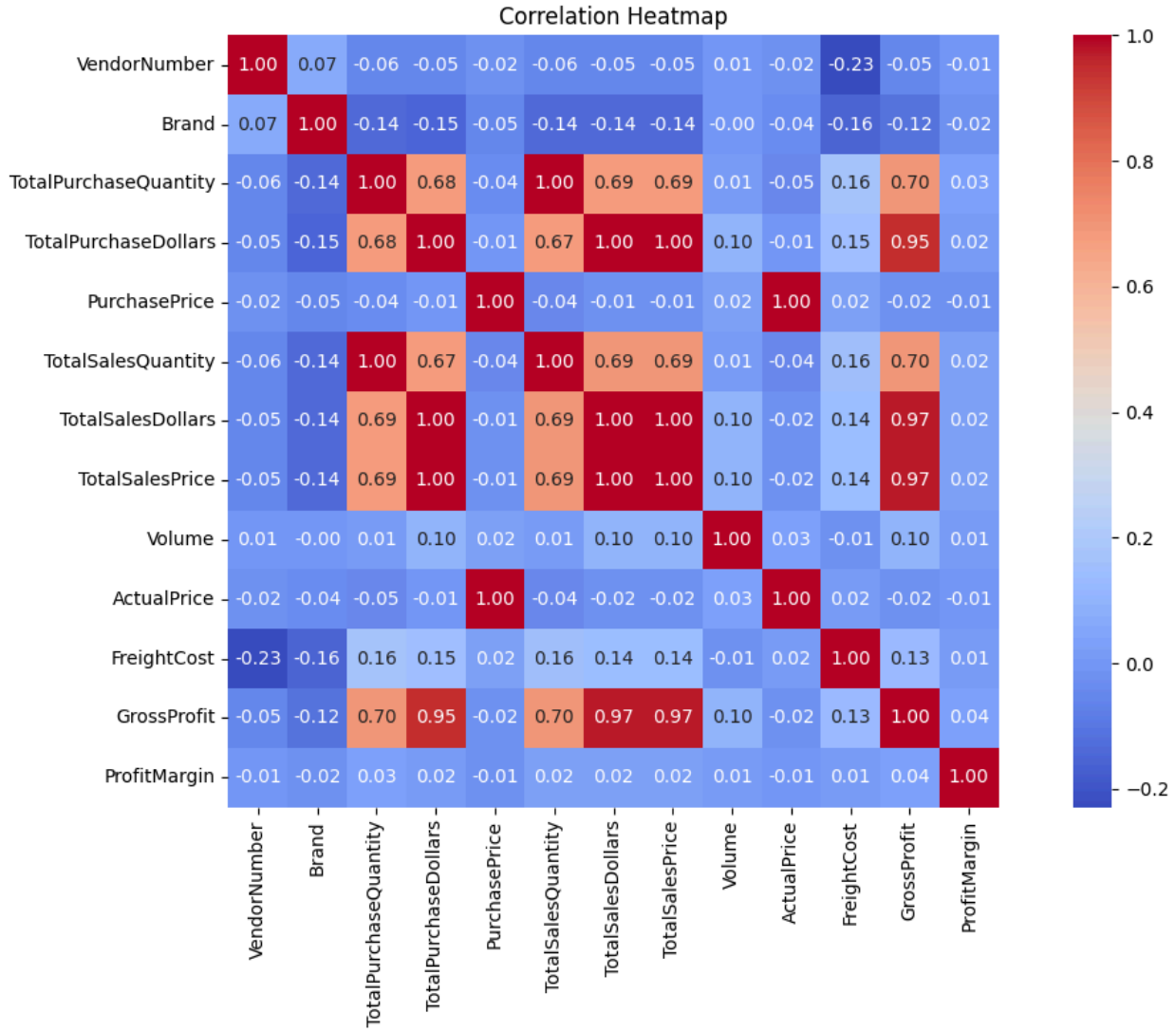
To enhance the reliability of the insights, we removed inconsistent data points where:

-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



Strong Positive Relationships

1.Total Purchase Quantity, Total Purchase Dollars, Total Sales Quantity, Total Sales Dollars, TotalSalesPrice, and Gross Profit are all highly correlated (0.67–0.98).Higher purchase volume/dollars directly translate to higher sales and profit outcomes.

2.TotalSales Dollars vs. Gross Profit (0.98) – Revenue and profit move almost in lockstep, meaning profit is strongly driven by sales volume rather than cost efficiency.

3.Total Sales Price vs. Total Sales Dollars (0.98) – Price setting is a major driver of total revenue.

Moderate Correlations Relationship

1.Purchase Price vs. Total Sales Dollars (0.69) – Higher purchase costs are moderately linked to higher sales revenue, likely because more expensive products are being sold more.

2.FreightCost shows weak-to-moderate positive correlation with sales (~0.15–0.17) – Suggests that shipping costs rise as sales volumes grow, but not proportionally.

Weak or No Relationships

1.Vendor Number and Brand have little to no correlation with financial or volume metrics. It means Vendor/brand identity alone doesn't predict sales or profitability.

2.Volume shows very weak correlations (near 0), meaning physical size/units don't explain much about financial performance.

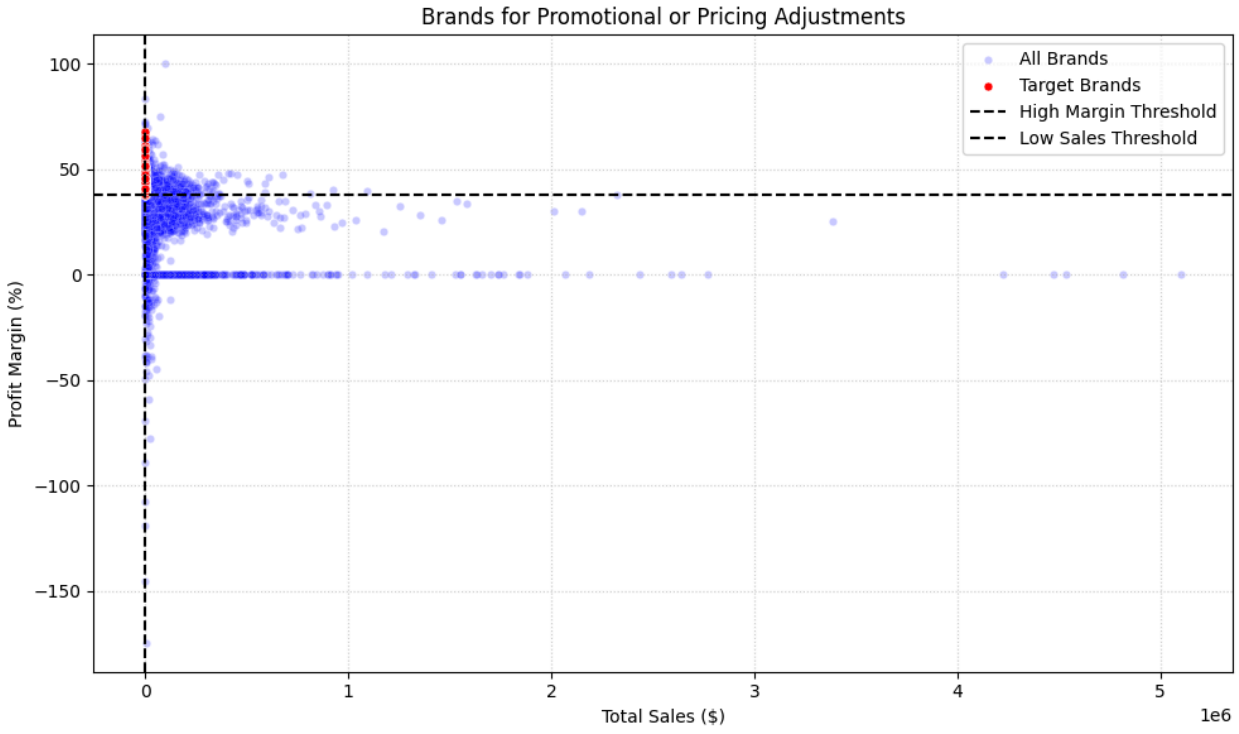
3.Profit Margin has weak negative correlations with Total Sales and Gross Profit (around -0.11 to -0.14).This Indicates that as sales increase, margin % tends to shrink (possibly due to discounting or higher costs).

Research Questions & Key Findings

1.Brands for Promotional or Pricing Adjustments

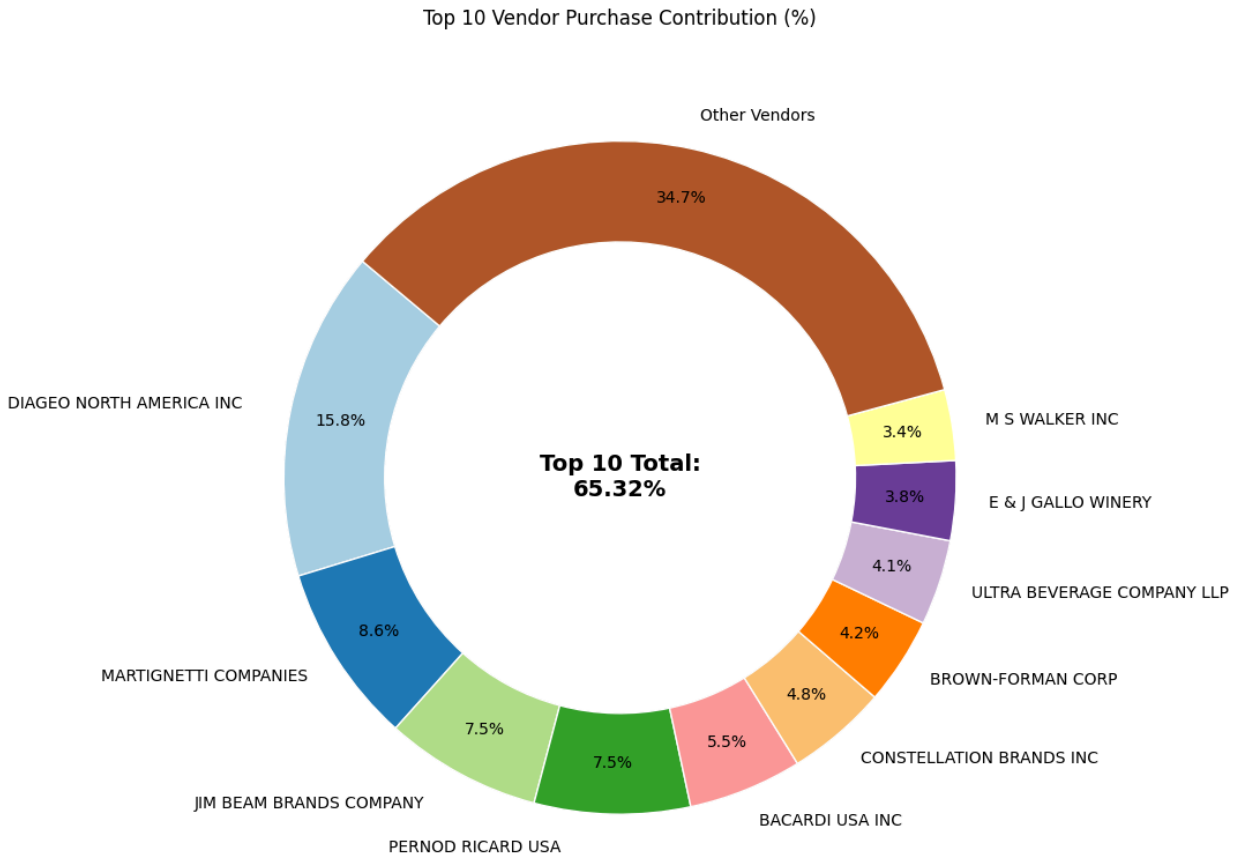
This analysis highlights brands that have high profit margins but very low total sales, making them candidates for promotional or pricing adjustments. The scatter plot shows most brands clustered at low sales with varying profit margins, while the red points (target brands) stand out above the high-margin threshold but below the low-sales threshold. The accompanying table lists these target brands, showing they each have profit margins around 40–66% yet very low total sales (under $50), suggesting untapped profitability potential if sales volume can be increased.

	Brand	Description	TotalSalesDollars	GrossProfit	ProfitMargin	IsTargetBrand
6921	22161	Santa Rita Organic Svgn Bl	9.99	6.64	66.466466	True
6115	20019	Camelot Chard	9.99	4.55	45.545546	True
9742	31972	Trinchero Main St Chard	11.99	6.11	50.959133	True
6328	20628	Clos Amador Brut Cava	13.99	6.25	44.674768	True
5754	19163	Terra Antiga Vinho Verde	14.37	5.67	39.457203	True
6787	21713	Gia Frizzante Chard	14.99	6.89	45.963976	True
4744	15672	Chatom Vyds Cuvee Red CA	14.99	5.66	37.758506	True
3510	8527	Concannon Glen Ellen Wh Zin	15.95	9.35	58.620690	True
6904	22080	Hacienda Zargon Merlot	19.98	8.96	44.844845	True
9837	33325	Jacobs Crk Merlot	25.96	10.52	40.523883	True
7766	24005	Wild Thing Chard	29.98	13.88	46.297532	True
5111	17018	Riebeek Cellars Pinotage	32.97	12.45	37.761601	True
5048	16822	Simply Naked Unoaked Moscato	33.96	15.20	44.758539	True
4594	15084	Project Paso Lonely Oak Red	35.97	17.49	48.623853	True
4730	15589	Right Hand Man Syrah	37.98	14.62	38.493944	True
4647	15289	Backhouse Pnt Grigio	39.95	20.00	50.062578	True
5655	18874	Jacobs Crk Sweet Red	41.94	17.64	42.060086	True
4240	13539	Solar De Randez Rioja	41.97	16.32	38.884918	True
3954	11873	Columbia Crest Svgn Bl Wash	43.94	16.70	38.006372	True
3821	10917	Coyam Emiliana Colchagua Red	45.98	22.46	48.847325	True



Top Vendors by sales and Contributions

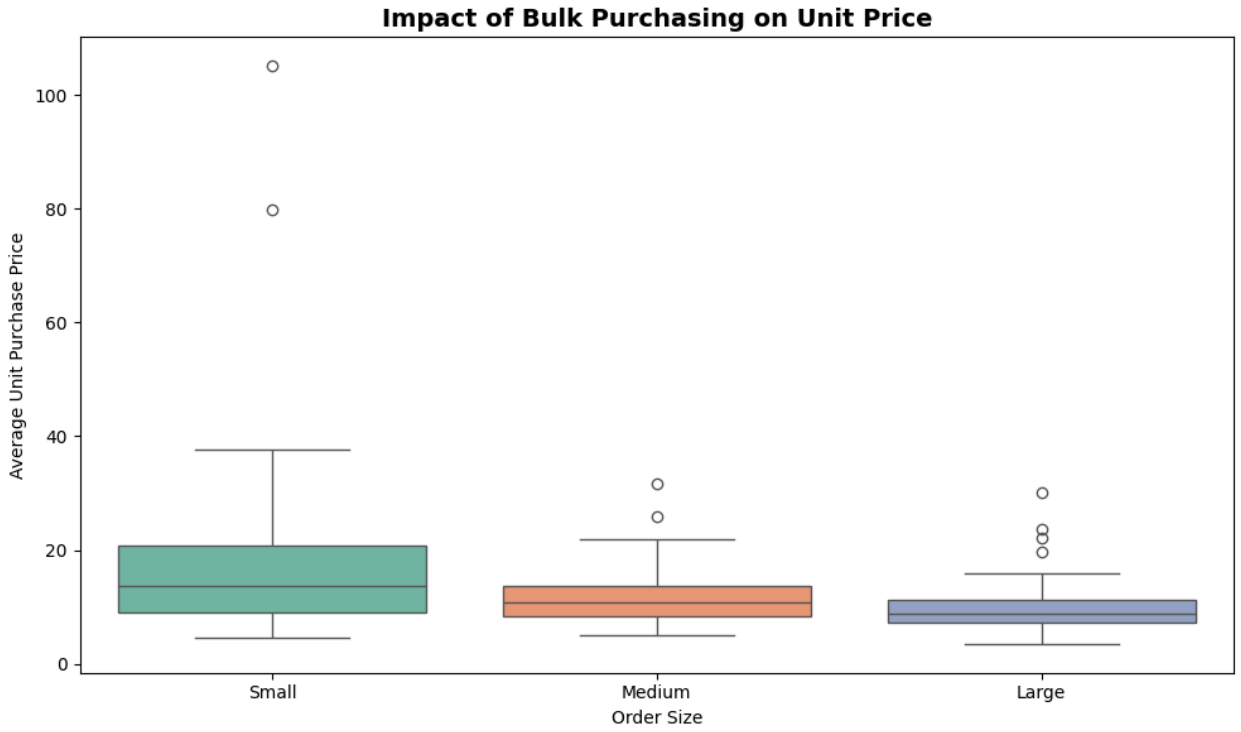
This donut chart shows that the top 10 vendors collectively contribute 65.32% of total purchases, with Diageo North America Inc alone accounting for 15.8%. The remaining 34.7% of purchases are spread across all other vendors, indicating a moderately concentrated vendor base.



Impact of Bulk Purchasing on Cost Savings

This table shows that although all 43 vendors are represented across order sizes, the vast majority of purchases come from large orders (32,855,926 units). In comparison, medium orders account for 688,529 units and small orders only 39,922 units, highlighting a heavy reliance on large-volume purchases.

OrderSize	Vendors	AvgUnitPurchasePrice	TotalPurchaseQuantity
Small	43	18.04	39922
Medium	43	12.06	688529
Large	43	10.35	32855926



Identifying Vendors with Low Inventory

Turnover Total Unsold Inventory Capital: $79.70M

Slow-moving Inventory increases storage costs, reduces cash flow efficiency, and affects overall profitability. Identifying vendors with low inventory turnover enables better stock management,

VendorName	StockTurnover
ALISA CARR BEVERAGES	0.615385
HIGHLAND WINE MERCHANTS LLC	0.708333
PARK STREET IMPORTS LLC	0.751306
Circa Wines	0.755676

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

minimizing financial strain.

Profit Margin Comparison: High vs. Low Performing Vendors

The confidence interval for low-performing vendors (40.48% to 42.62%) is significantly higher than that of top-performing vendors (30.74% to 31.61%). This suggests that vendors with lower sales tend to maintain higher profit margins, potentially due to premium pricing or lower operational costs.

For High-performing Vendors: If they aim to improve profitability, they could explore selective price adjustments, cost optimization, or building strategies. For Low-Performing Vendors: Despite higher margins, their low sales volume might indicate a need for better marketing, competitive pricing or improved distribution strategies.

Statistical Validation of Profit Margin Differences

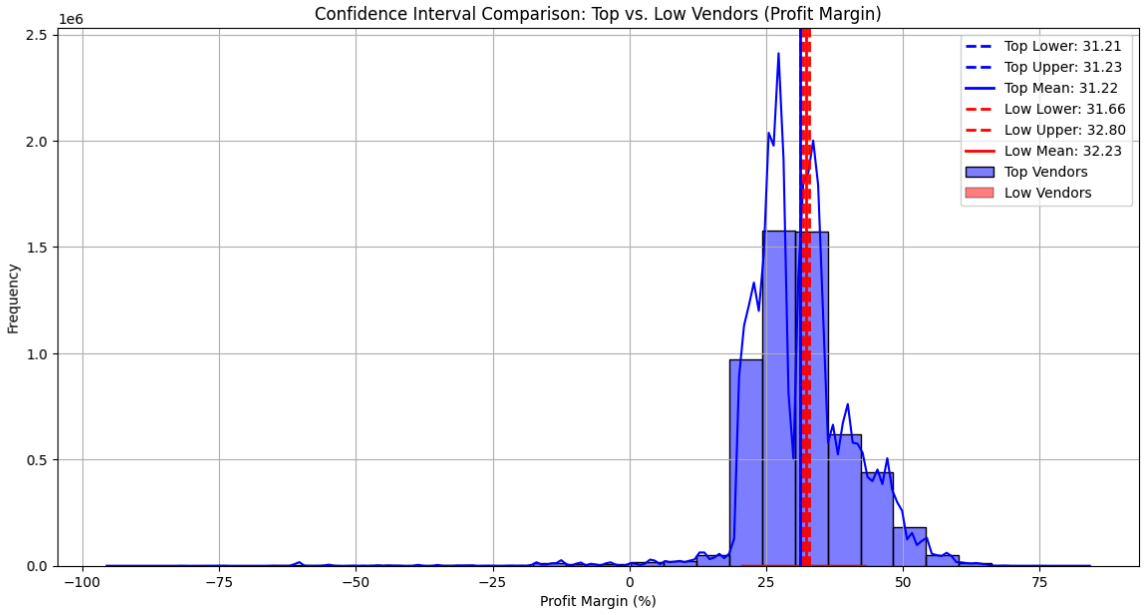
Ho(Null Hypothesis): There is no significant difference in the mean profit margins of top-performing and low-performing vendors.

H1(Alternative Hypothesis):The mean profit margins of top-performing and low-performing vendors are significantly different.

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.

Top Vendors 95% CI: (31.21, 31.23), Mean: 31.22
Low Vendors 95% CI: (31.66, 32.80), Mean: 32.23



Final Conclusion

This vendor performance analysis has provided a data-driven foundation for improving profitability, operational efficiency, and strategic decision-making within the retail and wholesale business context. By leveraging SQL and Python for data extraction, cleaning, and exploratory data analysis (EDA), and Power BI for dynamic visualization, the project transformed raw transactional data into actionable business insights.

The findings revealed significant concentration risk, with a small number of vendors driving the majority of purchases, and highlighted a long tail of underperforming vendors and brands contributing minimal sales at below-average profit margins. Conversely, top-selling vendors tended to operate on thinner margins, while low-selling vendors maintained higher margins, suggesting distinct pricing and cost structures that require tailored strategies. The discovery of slow-moving inventory and high unsold capital further underscored the need for improved inventory turnover and purchasing discipline.

Collectively, these insights enable leadership to prioritize marketing efforts on high-potential brands, renegotiate terms or offer promotional support for underperforming vendors, optimize bulk purchasing strategies to enhance unit economics, and improve cash flow efficiency through better stock management.

In conclusion, this analysis establishes a scalable and repeatable framework for vendor and brand evaluation—one that aligns operational choices with profitability goals. By continuously monitoring vendor and brand performance through this framework, the organization can proactively mitigate risk, capture new revenue opportunities, and sustain long-term competitive advantage.