

# E-COMMERCE VENDOR PERFORMANCE ANALYSIS

## INTRODUCTION

---

The retail and wholesale industry operates in a competitive environment where **effective inventory management, vendor performance tracking, and sales optimization** are essential for sustaining profitability. Businesses often face challenges such as **high capital lock-up in unsold inventory, dependency on a few key vendors, inefficient pricing strategies, and inconsistent procurement practices**.

This project, "**E-Commerce Vendor Performance Analysis**," aims to address these challenges by leveraging data analytics to:

- *Uncover vendor and brand performance trends.*
- *Identify opportunities for cost savings through bulk purchasing.*
- *Improve inventory turnover and reduce capital locked in unsold stock.*
- *Highlight underperforming brands for targeted promotional or pricing adjustments.*

By combining **Exploratory Data Analysis (EDA)** in Python with an **interactive Power BI dashboard**, this project provides a comprehensive view of vendor, sales, and inventory performance, enabling **data-driven decision-making for procurement and sales strategies**.

## BUSINESS STATEMENT

---

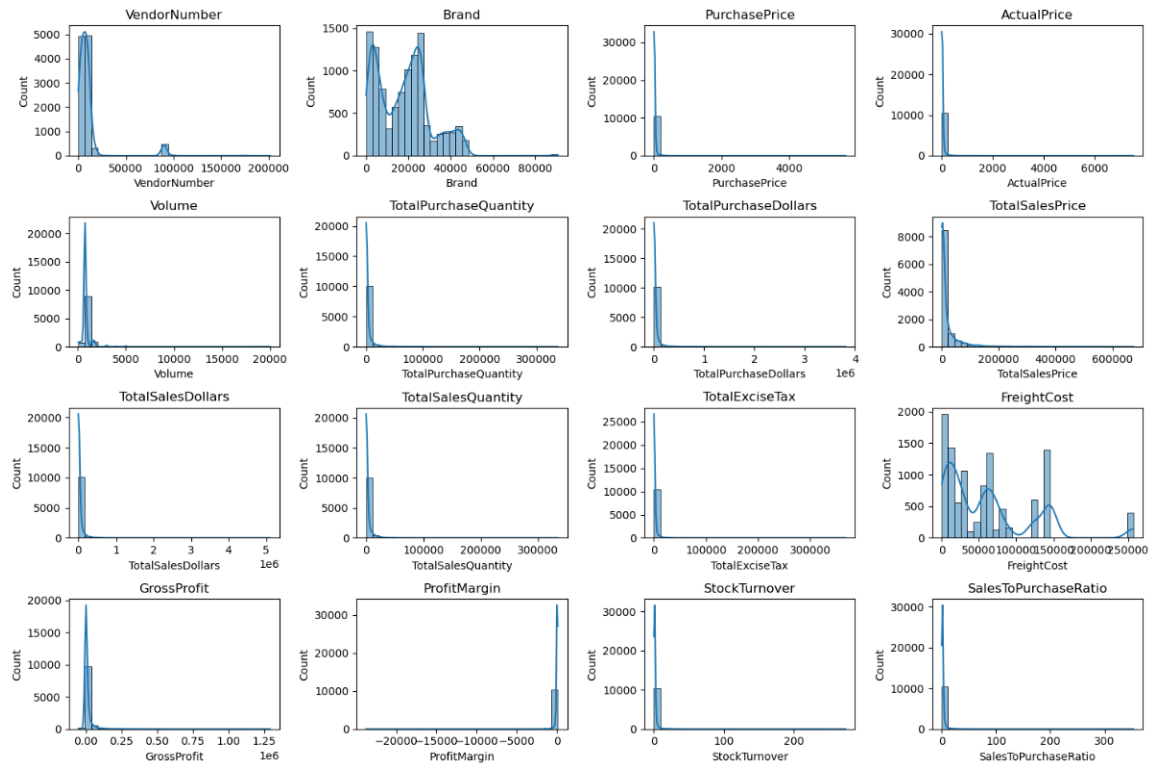
Effective inventory and sales management are critical for optimizing profitability in the retail and wholesale industry. Companies must minimize losses caused by inefficient pricing, excess inventory, poor vendor performance, and dependency on a limited vendor base.

This analysis leverages **Exploratory Data Analysis (EDA), advanced visualizations, and Power BI dashboard** to:

- *Identify underperforming brands requiring promotional or pricing adjustments.*
- *Determine top vendors driving sales and gross profit.*
- *Analyze the impact of bulk purchasing on reducing unit costs.*
- *Assess inventory turnover to reduce holding costs and improve efficiency.*
- *Investigate profitability variance between high-performing and low-performing vendors.*

# EXPLORATORY DATA ANALYSIS INSIGHTS

## I) Distribution Plots of Numerical Columns



### Analysis Performed:

- Distribution plots were generated for all key numerical variables to assess their spread, identify skewness, and detect outliers. This helps in understanding the underlying data patterns, variability, and concentration of values across vendors, sales, purchases, and profitability metrics.

### Key Insights:

#### - Vendor & Brand Distribution:

- > Majority of vendors have lower vendor numbers, while a few dominate the dataset.
- > Certain brands appear significantly more frequently, indicating brand concentration.

#### - Pricing (PurchasePrice & ActualPrice):

- > Both prices are right-skewed, with most transactions at lower price points.
- > A small number of high-value purchases contribute to a long-tailed distribution.

#### - Volume & Quantities:

- > Purchase and sales quantities are skewed toward smaller orders, with occasional bulk transactions.

**- Monetary Values (Purchase & Sales Dollars):**

-> Both show long-tailed distributions, where a few high-value transactions contribute disproportionately to overall revenue.

**- Profitability (GrossProfit & ProfitMargin):**

-> Most profits are at lower levels, but a few vendors contribute significantly higher profits.

-> Profit margins vary widely, suggesting differences in pricing strategies or cost structures.

**- Operational Metrics (StockTurnover & SalesToPurchaseRatio):**

-> Low to moderate stock turnover dominates, with only a few vendors showing high efficiency.

-> Sales-to-purchase ratio indicates that most sales closely align with purchases, but a few outliers exhibit very high ratios.

**Conclusion**

- The distributions reveal that most variables are highly right-skewed, highlighting that a small number of vendors, transactions, or products drive a majority of the value.

- This concentration emphasizes the importance of focusing on top-performing vendors, high-value transactions, and outlier analysis for targeted strategic decision-making.

## II) Summary Statistics

	count	mean	std	min	25%	50%	75%	max
VendorNumber	10692.0	1.065065e+04	18753.519148	2.00	3951.000000	7153.000000	9552.000000	2.013590e+05
Brand	10692.0	1.803923e+04	12662.187074	58.00	5793.500000	18761.500000	25514.250000	9.063100e+04
PurchasePrice	10692.0	2.438530e+01	109.269375	0.36	6.840000	10.455000	19.482500	5.681810e+03
ActualPrice	10692.0	3.564367e+01	148.246016	0.49	10.990000	15.990000	28.990000	7.499990e+03
Volume	10692.0	8.473605e+02	664.309212	50.00	750.000000	750.000000	750.000000	2.000000e+04
TotalPurchaseQuantity	10692.0	3.140887e+03	11095.086769	1.00	36.000000	262.000000	1975.750000	3.376600e+05
TotalPurchaseDollars	10692.0	3.010669e+04	123067.799627	0.71	453.457500	3655.465000	20738.245000	3.811252e+06
TotalSalesPrice	10692.0	1.879378e+04	44952.773386	0.00	289.710000	2857.800000	16059.562500	6.728193e+05
TotalSalesDollars	10692.0	4.223907e+04	167655.265984	0.00	729.220000	5298.045000	28396.915000	5.101920e+06
TotalSalesQuantity	10692.0	3.077482e+03	10952.851391	0.00	33.000000	261.000000	1929.250000	3.349390e+05
TotalExciseTax	10692.0	1.774226e+03	10975.582240	0.00	4.800000	46.570000	418.650000	3.682428e+05
FreightCost	10692.0	6.143376e+04	60938.458032	0.09	14069.870000	50293.620000	79528.990000	2.570321e+05
GrossProfit	10692.0	1.213238e+04	46224.337964	-52002.78	52.920000	1399.640000	8660.200000	1.290668e+06
ProfitMargin	10692.0	-inf	NaN	-inf	13.324515	30.405457	39.956135	9.971666e+01
StockTurnover	10692.0	1.706793e+00	6.020460	0.00	0.807229	0.981529	1.039342	2.745000e+02
SalesToPurchaseRatio	10692.0	2.504390e+00	8.459067	0.00	1.153729	1.436894	1.665449	3.529286e+02

**Analysis Performed**

- Summary statistics were calculated for all key numerical variables to evaluate central

tendency, spread, and outliers. Metrics such as mean, standard deviation, minimum, maximum, and quartiles were analyzed to assess variability, detect extreme values, and identify potential data quality issues.

## Key Insights

### - *VendorNumber & Brand:*

-> Wide ranges observed (VendorNumber: 2 to 213M; Brand: up to 90K), indicating skewed distribution and a few dominant vendors/brands driving large portions of activity.

### - *Pricing (PurchasePrice & ActualPrice):*

-> Average purchase price: \$23.4, Actual price: \$36.4, with outliers reaching up to \$7.5K, suggesting occasional high-value deals.

### - *Volume & Quantities:*

-> Average purchase and sales quantities (~3,100 each) with maximums exceeding 33K-34K, pointing to bulk transactions.

### - *Financial Metrics (Purchase, Sales, Profit):*

-> Total Purchase Dollars: Mean \$30K, max \$2.08M.

-> Total Sales Dollars: Mean \$16.7K, max \$2.8M.

-> Gross Profit: Mean \$21K, max \$1.29M, indicating a few large transactions disproportionately impacting profitability.

### - *Profit Margin:*

-> Includes NaN and -inf values, likely due to division errors or missing sales data. Ranges from -52% to +40%, requiring data cleaning.

### - *Operational Metrics (StockTurnover & Sales-to-Purchase Ratio):*

-> Stock Turnover: Mean 1.7, with peaks at 2.75, showing moderate movement of inventory.

-> Sales-to-Purchase Ratio: Mean 2.5, with extreme outliers (max 352), indicating anomalies or exceptional cases of high sales relative to purchases.

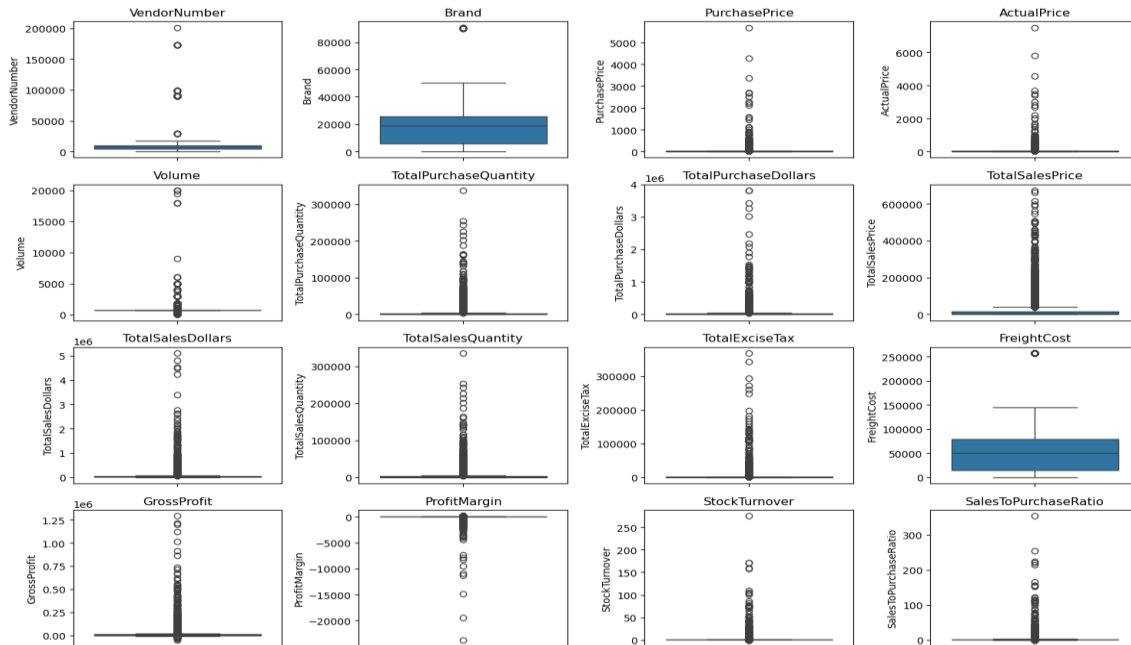
## Conclusion

- The dataset exhibits significant skewness and high-value outliers, typical of transactional financial data. While most values cluster around lower ranges, a small number of extreme values drive large portions of revenue and profit.

- Additionally, Profit Margin anomalies (NaN/-inf) highlight the need for data cleaning before applying advanced analytics.

- These insights underscore the importance of focusing on data normalization, outlier treatment, and quality checks to ensure accurate business decisions.

### III) Identifying Outliers using Boxplots



#### Analysis Performed

- Boxplots were used to detect outliers across key numerical variables. Outliers represent extreme data points beyond the typical range, potentially indicating anomalies, exceptional transactions, or data quality issues.

#### Key Insights

##### - **High-Value Outliers:**

-> Significant outliers observed in PurchasePrice, ActualPrice, TotalPurchaseDollars, and TotalSalesDollars, representing rare high-value transactions.

##### - **Operational Outliers:**

-> Metrics like Volume, StockTurnover, and Sales-to-Purchase Ratio showed extreme values, likely tied to bulk transactions or rapid inventory cycles.

##### - **Profit Margin Anomalies:**

-> Presence of extreme negative and positive margins, including invalid (-inf) values, signals calculation or data quality issues.

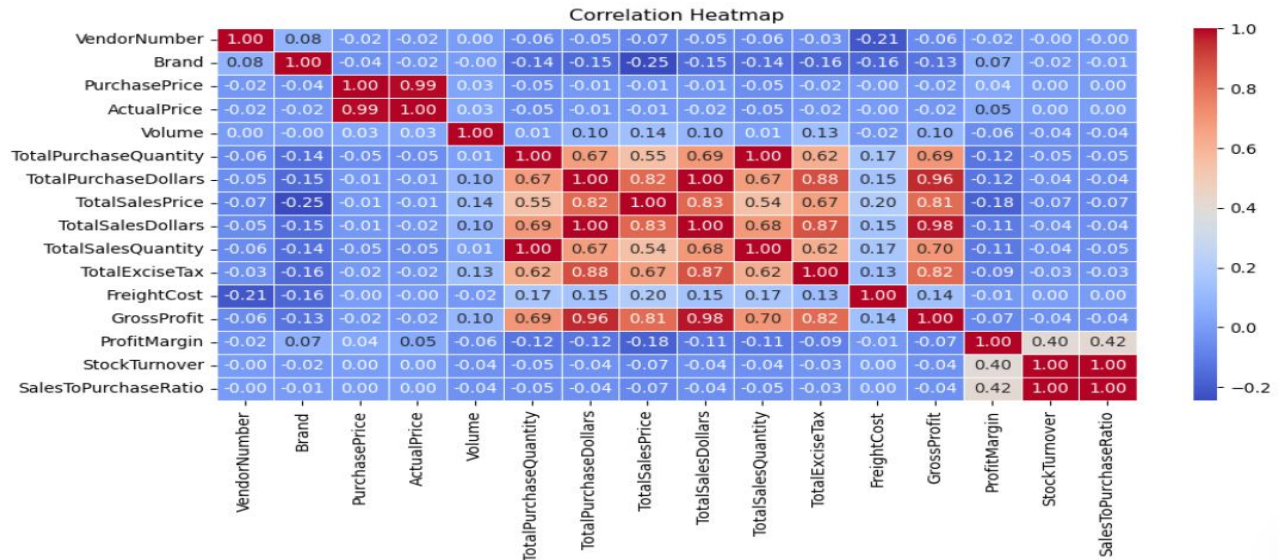
##### - **FreightCost & GrossProfit:**

-> Few transactions incurred unusually high costs or profits, indicating special cases or premium orders.

#### Conclusion

- Outliers are prominent across financial and operational metrics. While some represent valid high-impact business events, others require further validation and potential treatment (e.g., normalization or exclusion) to prevent skewed analysis results. Addressing these ensures more reliable insights for decision-making.

## IV) Correlation Analysis of Variables



### Analysis Performed

- A correlation heatmap was used to identify relationships between key numerical variables, highlighting dependencies and potential multicollinearity risks.

### Key Insights

#### - **Strong Positive Correlations:**

- > PurchasePrice & ActualPrice (0.99): Nearly perfect correlation, confirming they move together.
- > GrossProfit & ProfitMargin (0.96): High correlation as profit margins directly drive gross profit.
- > TotalSalesQuantity & TotalSalesDollars (0.83): Higher sales volume leads to greater revenue.

#### - **Moderate Correlations:**

- > TotalPurchaseDollars & TotalPurchaseQuantity (0.67): Larger purchases often involve higher quantities.
- > FreightCost & TotalExciseTax (~0.87): High-tax transactions often incur higher freight costs.

#### - **Weak or Negligible Correlations:**

- > VendorNumber: Minimal correlation with other variables; serves only as an identifier.
- > ProfitMargin: Lower correlations with most metrics, indicating influence from external factors.

### Conclusion

- Key financial and operational metrics show strong interdependence (e.g., prices, quantities, and revenue), validating their role as performance indicators. Vendor identifiers and low correlated metrics have limited predictive value. These insights help focus modeling and decision-making on the most impactful variables.

# RESEARCH & BUSINESS PROBLEMS

## Problem 1: Identifying Brands that Need Promotional or Pricing Adjustments

In this step, I focused on finding brands that are not performing well in sales but have a high profit margin.

### I used percentile-based filtering:

- Selected brands in the bottom 15% for sales (low sales).
- Selected brands in the top 15% for profit margin (high profitability).

### After filtering, I plotted these brands on a scatter plot:

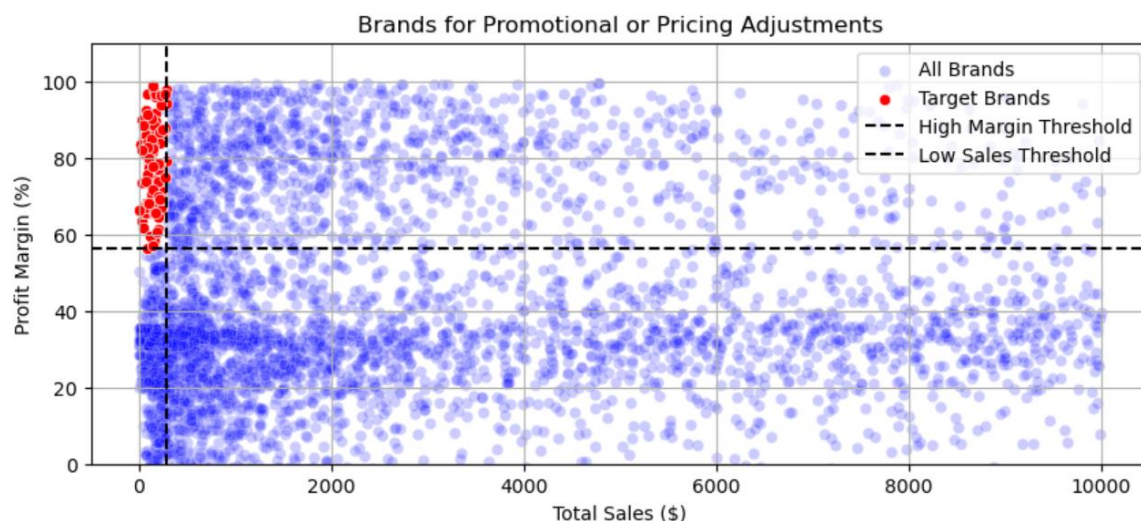
- X-axis: Total Sales (to show sales performance).
- Y-axis: Profit Margin (to show profitability).
- Each point represents a brand, and the size of the points is based on profit margin.

### Key Insights:

- Most brands are clustered in the middle range, meaning they have average sales and margins.
- The highlighted brands clearly stand out in the low sales but high profit margin zone.
- These brands have profitability potential, but their sales volumes are low.

### Conclusion:

These brands should be the focus of promotions, discounts, or pricing adjustments to increase sales while retaining their high profitability. Targeting these specific brands could help boost overall revenue without reducing margins significantly.





## Problem 2: Which Vendors and Brands demonstrate the highest sales performance?

### Analysis Performed:

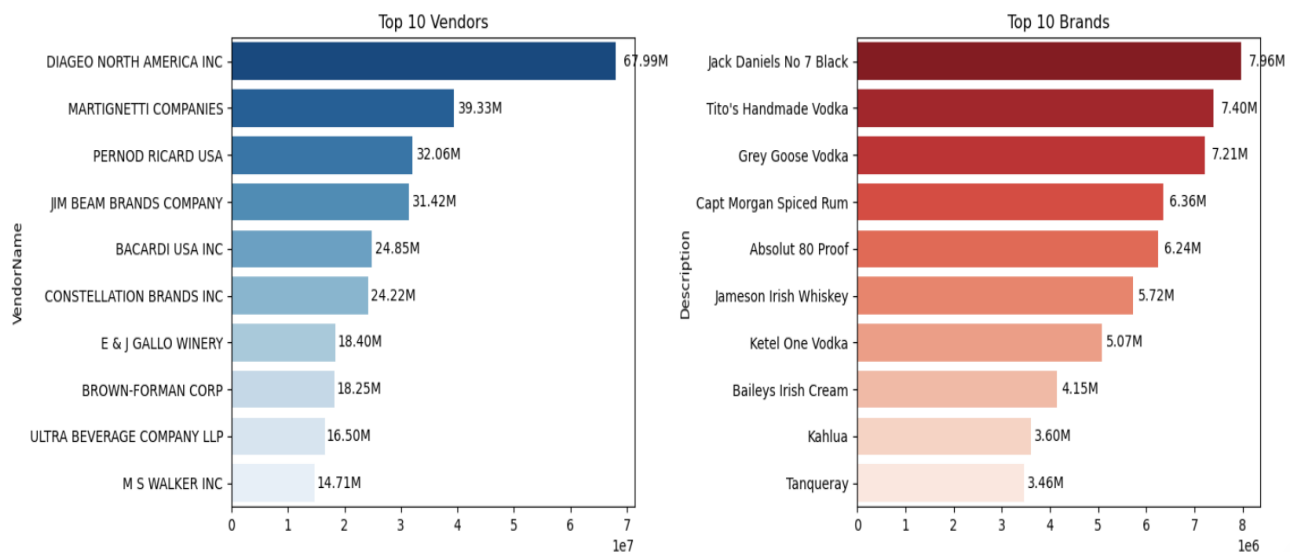
- Grouped sales data by VendorName and Brand to rank them based on Total Sales Dollars.
- Identified the Top 10 Vendors and Top 10 Brands contributing most significantly to overall sales.
- Sorted the results in descending order to highlight the leading contributors clearly.

### Key Insights:

- The sales performance is highly concentrated among a few top vendors and brands, showing heavy reliance on these key players.
- The Top 10 vendors collectively contribute a substantial portion of total sales, and a similar trend is observed for brands.
- This distribution indicates that while some vendors are dominant, mid- and low-tier vendors contribute minimally.

### Conclusion:

Sales are driven predominantly by a select group of high-performing vendors and brands. For long-term business stability, it is crucial to strengthen relationships with top contributors while also focusing on strategies to improve performance of underperforming vendors and brands to reduce dependency risk.





### Problem 3: Which Vendors Contribute the Most to Total Purchase Dollars?

#### Analysis Performed:

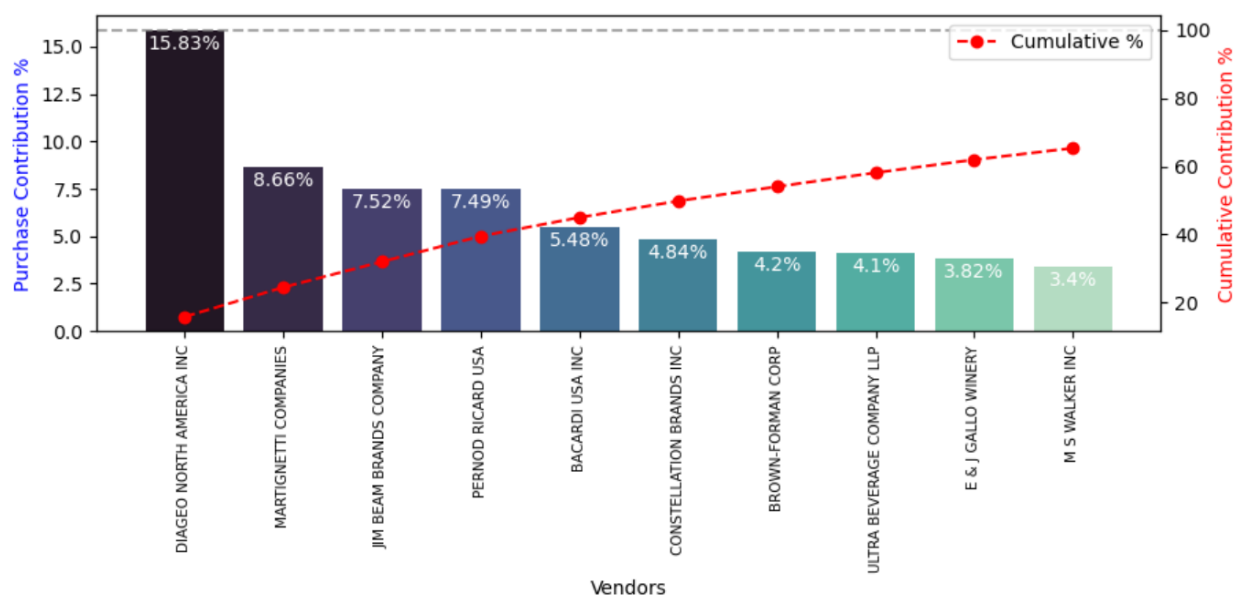
- A Pareto Chart was used to evaluate vendor contributions to total purchase dollars.
- The analysis plotted:
  - > Individual vendor purchase contribution (%) using bars.
  - > Cumulative purchase contribution (%) using a line graph.
- This approach identifies key vendors driving procurement costs and validates the Pareto principle (80/20 rule) in vendor spend.

#### Key Insights:

- Top Contributing Vendors:
  - > Diageo North America Inc contributes the highest at 15.83%.
  - > Followed by Martignetti II Companies (8.66%) and Jim Beam Brands Company (7.52%).
- Cumulative Contribution:
  - > The top 3-4 vendors contribute ~37-40% of total purchases, showing high vendor concentration.
- Spend Distribution:
  - > Smaller vendors like MS Walker Inc contribute as low as 3.4%, reinforcing that procurement is concentrated among few high-value vendors.

#### Conclusion:

- Procurement spend is heavily reliant on a few vendors, aligning with the Pareto principle.
- Strategic focus should be on cost optimization and negotiation with top vendors while assessing risks of over-dependency.
- This insight also opens opportunities to diversify the vendor base or leverage high-volume vendors for better terms.



## Problem 4: How Much Total Procurement is Dependent on the Top Vendors?

### Analysis Performed:

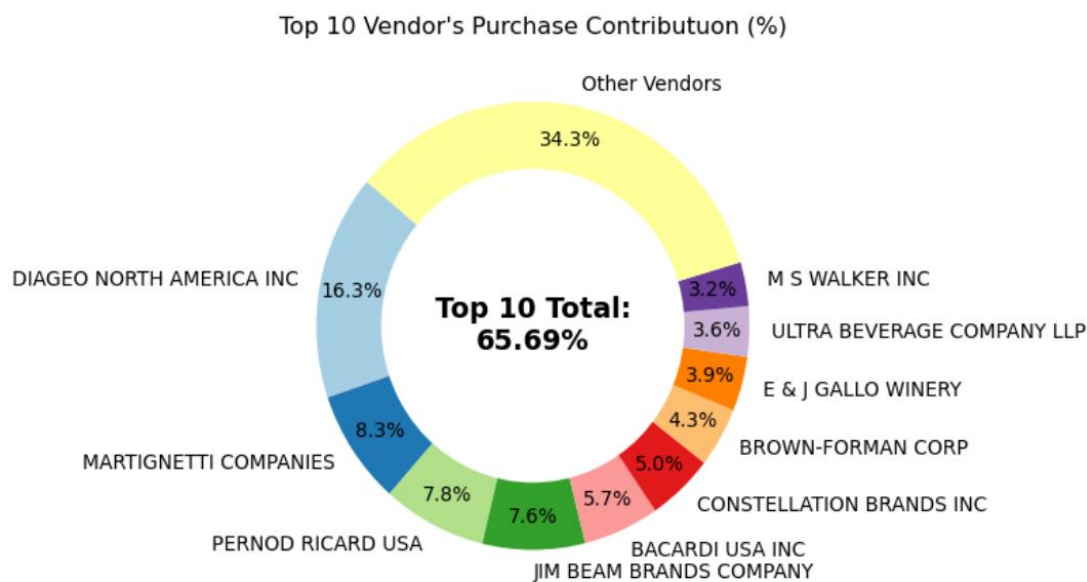
- A Donut Chart was used to assess procurement dependency on the top 10 vendors, with the remaining vendors grouped under "Others."
- The contribution of each vendor to total procurement was calculated, highlighting reliance on a few key suppliers.

### Key Insights:

- High Dependency on Top Vendors:
  - > The top 10 vendors account for 65.69% of total procurement.
  - > Diageo North America Inc leads with 16.3%, followed by Martignetti Companies (8.3%), Bacardi USA Inc (7.8%), and Jim Beam Brands Company (7.6%).
- "Others" Category:
  - > Vendors outside the top 10 collectively contribute 34.3%, representing a diversified but fragmented vendor base.
- Risk Factor:
  - > Heavy reliance on a few vendors increases supply chain risk if disruptions occur among these top suppliers.

### Conclusion:

- Procurement is concentrated among a few key vendors, making them strategically significant.
- To reduce risk, the organization should diversify supply sources, negotiate contingency contracts, and continuously monitor vendor dependency to maintain supply chain resilience.



## Problem 5: Did Purchasing in Bulk Reduce the Unit Price, and What is the Optimal Purchase Volume for Cost Savings?

### Analysis Performed:

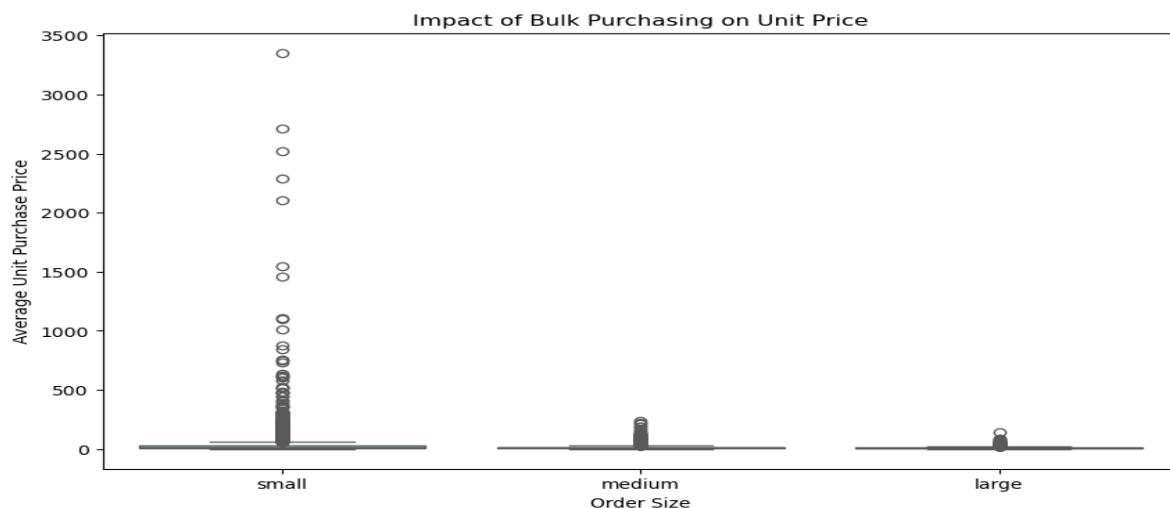
- A Boxplot was used to analyze the relationship between Order Size (small, medium, large) and Average Unit Purchase Price.
- Order sizes were categorized into three groups using purchase quantity percentiles, and the price distribution was studied to assess bulk purchase cost savings.

### Key Insights:

- Small Orders:
  - > Show high price variability, with several outliers reaching up to ~\$3,400/unit, indicating inconsistent pricing and higher costs for low-volume purchases.
- Medium Orders:
  - > Prices are more stable and lower than small orders, with fewer outliers, suggesting moderate cost efficiency.
- Large Orders:
  - > Display the lowest and most consistent prices, confirming that bulk purchasing provides significant cost savings and pricing stability.
- Trend Identified:
  - > There is a clear inverse relationship between order size and unit price — larger orders result in reduced unit costs and improved pricing stability.

### Conclusion:

- Bulk purchasing significantly lowers the average unit purchase price and reduces pricing variability.
- The "large" order category is identified as the optimal purchase volume for maximizing cost savings.
- To capitalize on these benefits, organizations should optimize inventory planning and procurement strategies to enable larger order volumes without overstocking.



## Problem 6: Which Vendors Have Low Inventory Turnover, Indicating Excess Stock and Slow-Moving Products?

### Analysis Performed:

- A table ranking vendors by Stock Turnover ratio was analyzed to identify vendors with slow-moving inventory.
- Stock Turnover is calculated as  $\text{Sales Quantity} \div \text{Purchase Quantity}$ , indicating how frequently stock is sold and replaced. Lower ratios signal excess inventory and slower sales cycles.

### Key Insights:

- Lowest Turnover Vendor:
  - > ALISA CARR BEVERAGES has the lowest turnover ratio (~0.615), indicating inventory is replenished less than once every two cycles, suggesting weak sales and overstocking.
- Other Low-Turnover Vendors:
  - > HIGHLAND WINE MERCHANTS LLC (~0.708) and PARK STREET IMPORTS LLC (~0.751) also show slow stock movement, reflecting inefficient inventory utilization.
- Business Implications:
  - > Low turnover can lead to:
  - > Increased storage and holding costs.
  - > Risk of product obsolescence or wastage.
  - > Capital being tied up in unsold stock, reducing cash flow.

### Conclusion:

- ALISA CARR BEVERAGES and similar vendors with low turnover require immediate attention to optimize inventory and improve sales performance.
- Strategies such as better demand forecasting, promotional campaigns, or revising order volumes can mitigate overstocking risks.
- Continuous monitoring of turnover ratios is critical to reduce holding costs, improve liquidity, and ensure healthier inventory cycles.

StockTurnover	
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

## Problem 7: How Much Capital is Locked in Unsold Inventory per Vendor, and Which Vendors Contribute Most to It?

### Analysis Performed:

- The analysis calculated unsold inventory capital (in \$K) for each vendor to identify the amount of financial resources tied up in slow-moving or unsold stock.
- This metric highlights vendors causing significant capital lock-up, impacting cash flow and working capital efficiency.

### Key Insights:

- Top Vendors by Unsold Inventory Capital:
  - > DIAGEO NORTH AMERICA INC: Highest unsold inventory value at \$722.21K, indicating substantial financial resources tied in unsold stock.
  - > JIM BEAM BRANDS COMPANY: Significant capital lock-up of \$554.67K.
  - > PERNOD RICARD USA: Close behind with \$470.63K in unsold inventory.
- Moderate Capital Locks:
  - > Vendors like WILLIAM GRANT & SONS INC (\$401.96K) and E & J GALLO WINERY (\$228.28K) also contribute but at moderate levels.
- Efficient Vendors:
  - > Vendors such as CONSTELLATION BRANDS INC and MOET HENNESSY USA INC show lower unsold inventory, suggesting effective inventory turnover or better sales performance.
- Business Implications:
  - > Vendors with high unsold inventory represent financial risk and poor inventory efficiency, requiring targeted actions like better demand forecasting, inventory control, and vendor collaboration.

### Conclusion:

- DIAGEO NORTH AMERICA INC, JIM BEAM BRANDS COMPANY, and PERNOD RICARD USA are the top contributors to unsold inventory capital, tying up significant funds.
- Addressing these vendors through optimized stock planning, improved sales strategies, and stronger vendor negotiations can reduce capital lock-up and improve cash flow.
- Continuous monitoring of unsold inventory values is crucial for minimizing financial risks and enhancing operational efficiency.

	VendorName	UnsoldInventoryValue
25	DIAGEO NORTH AMERICA INC	722.21K
46	JIM BEAM BRANDS COMPANY	554.67K
68	PERNOD RICARD USA	470.63K
116	WILLIAM GRANT & SONS INC	401.96K
30	E & J GALLO WINERY	228.28K
79	SAZERAC CO INC	198.44K
11	BROWN-FORMAN CORP	177.73K
20	CONSTELLATION BRANDS INC	133.62K
61	MOET HENNESSY USA INC	126.48K
77	REMY COINTREAU USA INC	118.60K

# Final Recommendations

---

## 1. Vendor Management & Procurement Optimization

- **Focus on Key Vendors:** Prioritize negotiation and collaboration with top vendors like Diageo North America Inc and Martignetti Companies, who collectively contribute significantly to procurement and sales.
- **Diversify Vendor Base:** Reduce dependency on top vendors (65% procurement concentration) by onboarding new suppliers to mitigate supply chain risks.
- **Performance-Based Contracts:** Implement SLAs (Service Level Agreements) and performance-based contracts for vendors to ensure consistent supply and quality.

## 2. Inventory & Capital Management

- **Reduce Unsold Inventory:** Work closely with vendors with high unsold inventory (e.g., Diageo North America Inc) to improve demand forecasting and reduce capital lock-up.
- **Promote Low-Turnover Stock:** Introduce promotions or bundle offers for slow-moving products (e.g., ALISA CARR BEVERAGES) to improve stock turnover.
- **Adopt Inventory Thresholds:** Implement minimum and maximum stock thresholds to avoid overstocking and improve cash flow.

## 3. Pricing & Sales Strategies

- **Leverage Bulk Discounts:** The analysis confirms bulk purchasing reduces unit cost. Optimize procurement by consolidating orders to leverage economies of scale.
- **Target High-Margin, Low-Sales Brands:** Increase marketing or promotional focus on profitable but underperforming brands to boost revenue.
- **Dynamic Pricing:** Introduce competitive pricing strategies for products with low sales-to-purchase ratios to increase turnover and reduce holding costs.

## 4. Data Quality & Monitoring

- **Address Data Issues:** Resolve identified issues like negative or infinite profit margins and inconsistent vendor identifiers to improve analysis accuracy.
- **Continuous Monitoring Dashboards:** Implement Power BI dashboards for real-time monitoring of vendor performance, inventory turnover, and profitability metrics.

## 5. Strategic Growth Initiatives

- **Strengthen Vendor Relationships:** Engage in quarterly business reviews (QBRs) with top vendors to align procurement strategies and secure better terms.
- **Diversify Product Mix:** Analyze customer demand trends to diversify offerings and reduce reliance on slow-moving SKUs.
- **Automated Forecasting:** Integrate predictive analytics to forecast demand and optimize procurement cycles, reducing excess stock and costs.