

Exploratory Data Analysis Insights

1. Previously, we examined the various tables in the database to identify key variables, understand their relationships, and determine which ones should be included in the final analysis.
2. In this phase of EDA, we will analyse the resultant table to gain insights into the distribution of each column. This will help us understand data patterns, identify anomalies, and ensure data quality before proceeding with further analysis.

Summary Statistics:

Out[91]:

	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
TotalSalesQuantity	10692.0	3.077482e+03	10952.851391	0.00	33.000000	261.000000	1929.250000	3.349390e+05
TotalSalesDollars	10692.0	4.223907e+04	167655.265984	0.00	729.220000	5298.045000	28396.915000	5.101920e+06
TotalSalesPrice	10692.0	1.879378e+04	44952.773386	0.00	289.710000	2857.800000	16059.562500	6.728193e+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
GrossProduct	10692.0	1.213238e+04	46224.337964	-52002.78	52.920000	1399.640000	8660.200000	1.290668e+06
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

Fig 1. Summary Statistics

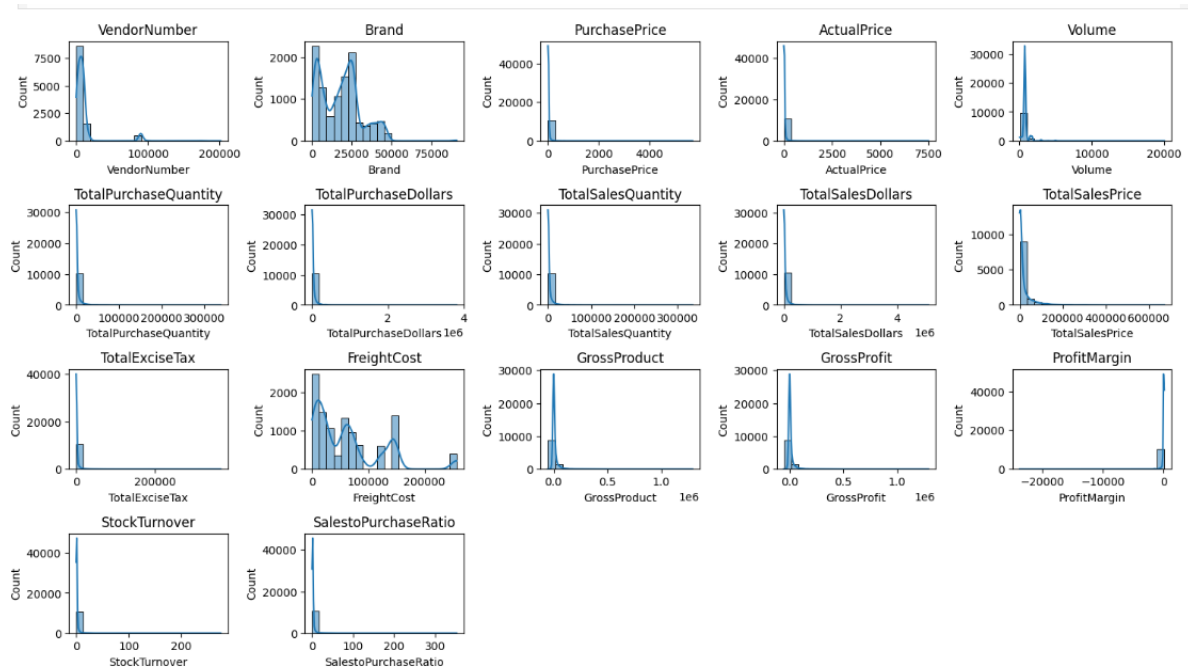


Fig. 2 Histogram of all the columns in the vendor summary dataset

Summary Statistics Insights

Negative & Zero values:

1. **Gross Profit:** Minimum value is -52002.78, indicating losses. Some products or transaction may be selling at a loss due to high cost or selling at discounts lower than the purchase price.
2. **Profit Margin:** has a minimum of -Infinite which suggests cases where revenue is zero or even lower than costs.
3. **Total Sales Quality & sales Dollars:** Minimum values are 0, meaning some products were purchased but never sold. These could be slow-moving or obsolete stock.

Outliers Indicated by High Standard Deviation:

1. **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.
2. **Freight Cost:** Huge variation, from 0.09 to 257,032.07, suggests logistics inefficiencies or bulk shipments
3. **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 Sold quantity for that product is higher than purchased quantity due to either sales are 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

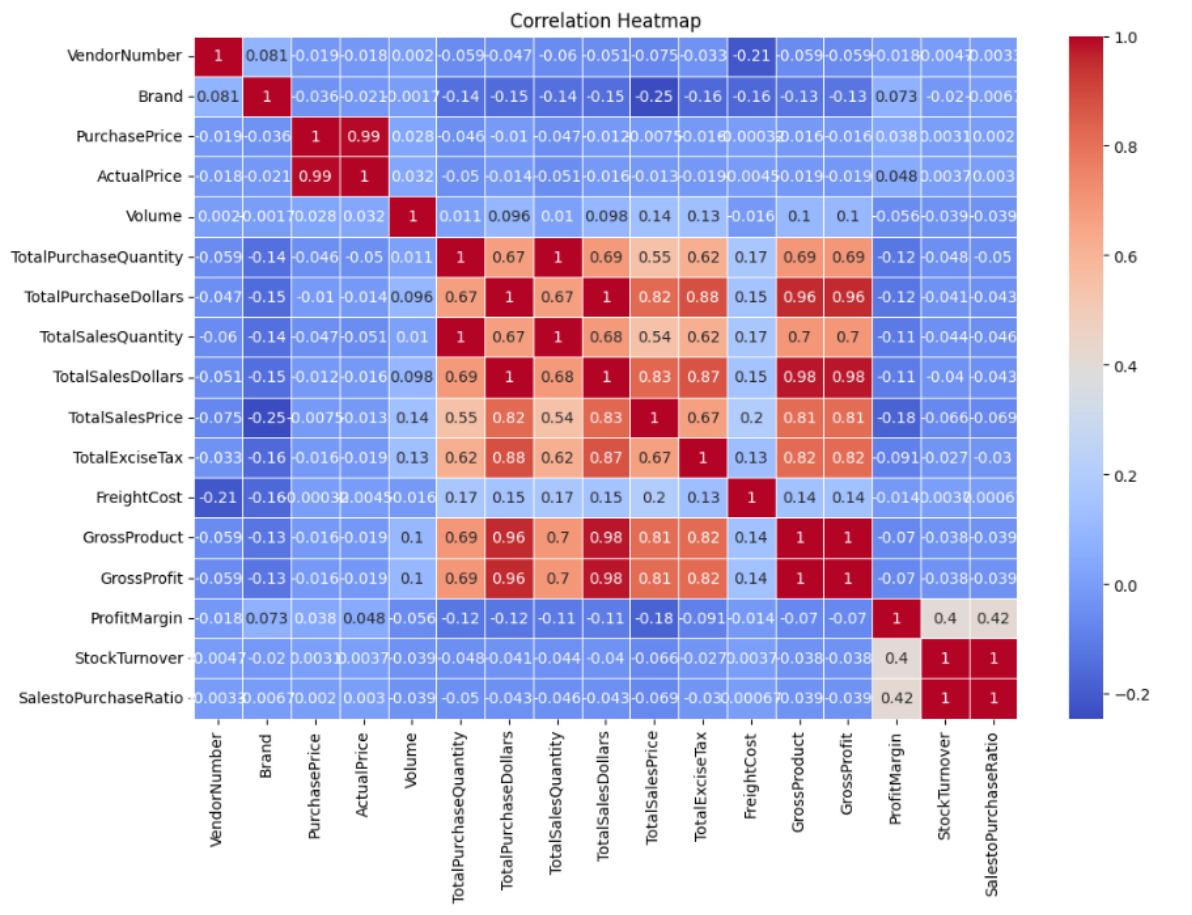


Fig.3 Correlation heatmap

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), 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. Brand Promotional and Pricing Adjustments

Brands with Low Sales but High Profit Margins:

	Description	TotalSalesDollars	ProfitMargin
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

Fig. 4 Brands with low sales but high profit margins

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.

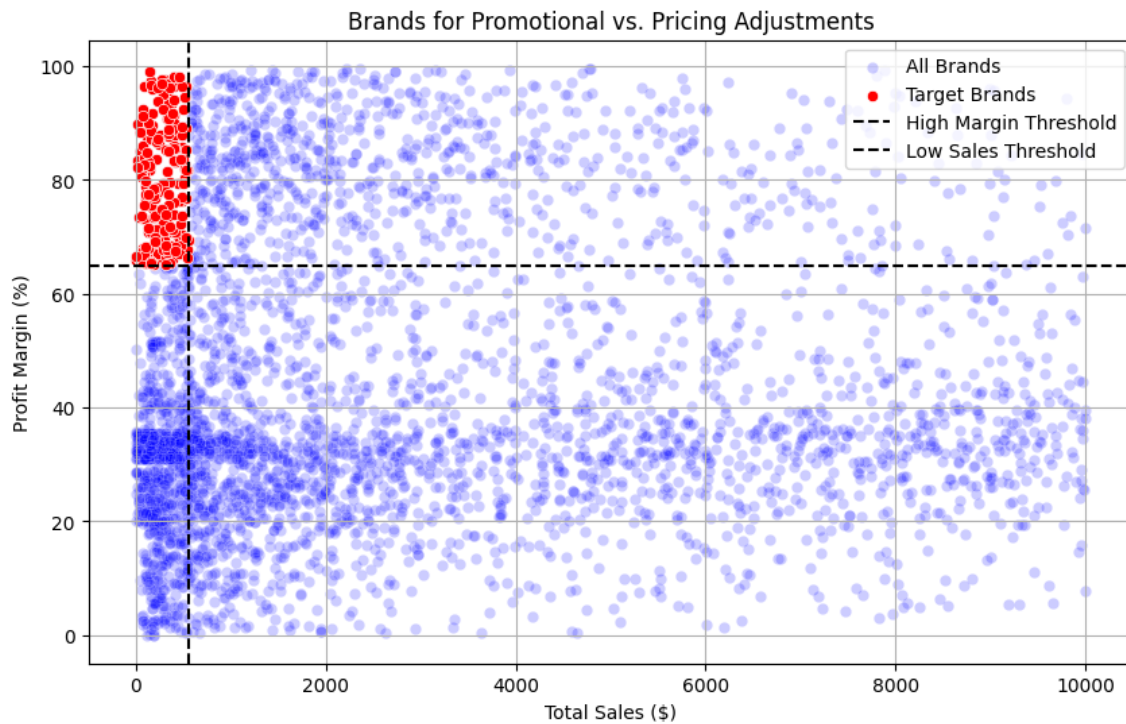


Fig. 5 Brands for Promotion vs Price Adjustments

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.

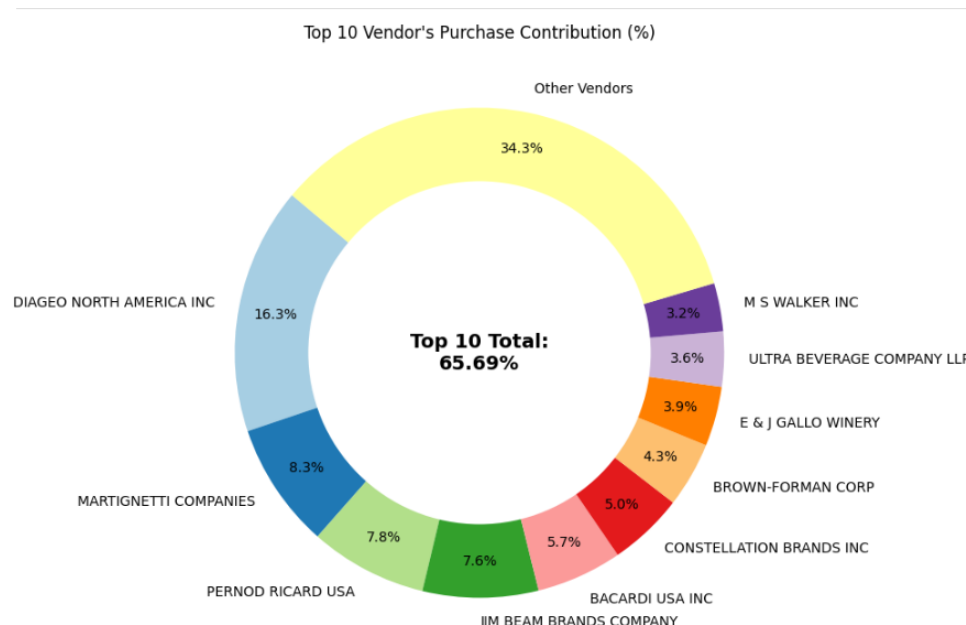


Fig. 7 Top 10 Vendor's Purchase Contribution in Percentage

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	39.057543
Medium	15.486414
large	10.777625

Fig. 8 Order size and its unit purchase price

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.

StockTurnover			
VendorName		VendorName	UnsoldInventoryValue
ALISA CARR BEVERAGES	0.615385	25 DIAGEO NORTH AMERICA INC	722.21K
HIGHLAND WINE MERCHANTS LLC	0.708333	46 JIM BEAM BRANDS COMPANY	554.67K
PARK STREET IMPORTS LLC	0.751306	68 PERNOD RICARD USA	470.63K
Circa Wines	0.755676	116 WILLIAM GRANT & SONS INC	401.96K
Dunn Wine Brokers	0.766022	30 E & J GALLO WINERY	228.28K
CENTEUR IMPORTS LLC	0.773953	79 SAZERAC CO INC	198.44K
SMOKY QUARTZ DISTILLERY LLC	0.783835	11 BROWN-FORMAN CORP	177.73K
TAMWORTH DISTILLING	0.797078	20 CONSTELLATION BRANDS INC	133.62K
THE IMPORTED GRAPE LLC	0.807569	61 MOET HENNESSY USA INC	126.48K
WALPOLE MTN VIEW WINERY	0.820548	77 REMY COINTREAU USA INC	118.60K

Fig. 9 Vendor name & its Stock-Turnover

Fig.10 Vendor name & it unsold inventory value

5. Profit Margin Comparison: High vs. Low-Performing Vendors

- **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: Optimise profitability by adjusting pricing, reducing operational costs, or offering bundled promotions.

Low-performing vendors: Improve marketing efforts, optimize pricing strategies, and enhance distribution networks.

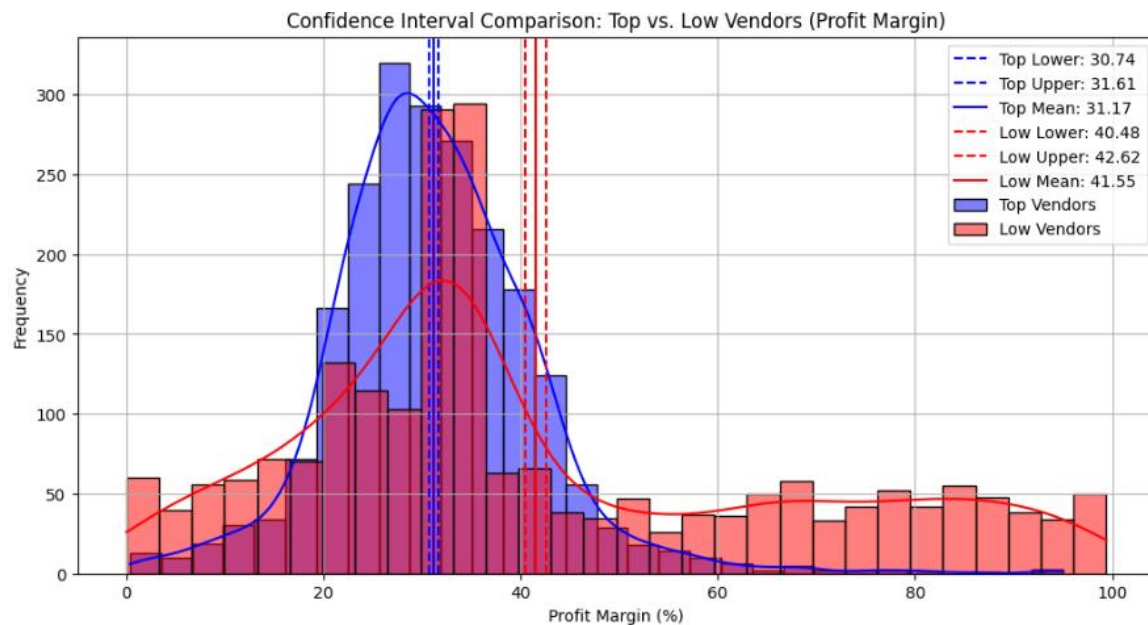


Fig. 11 Confidence Interval Comparison: Top vs Low Vendors (Profit Margins)

6. Statistical Validation of Profit Margin Differences

Hypothesis Testing:

- **H0 (Null Hypothesis):** No significant difference in profit margins between top and low-performing vendors.
- **H1 (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 Recommendations

- 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.