



Debby Rofiko Malik

Data Analysis Portfolio

2025



Tahun
2021 - 2025

Education

Biomedical Eng. - Bachelor Degree
Universitas Indonesia



Tahun
2023 - 2024

Work Experience

3D Design Engineer
Covent Indonesia

Tools



Hard Skills

Math and Statistics **Data Cleaning**
Data Visualization **Machine Learning**

Soft Skills

Critical Thinking **Problem Solving**
Communication **Teamwork** **Adaptability**
Leadership **Time Management** **Creativity**

Past Projects

- ★ Hotel Booking Analysis
- ★ Vendor Performance Analysis

Certifications

- ★ Data Science & Data Analysis (MySkill)
- ★ Microsoft Excel, Word, and PowerPoint (MySkill)

Contact Me

[Debby Rofiko Malik](#)

debbymalik21@gmail.com

[github.com/debbyrofikomalik](#)

Vendor Performance Analysis

Python (Google Colab)

SQL

Power BI

`</>` [Access Code Here](#)

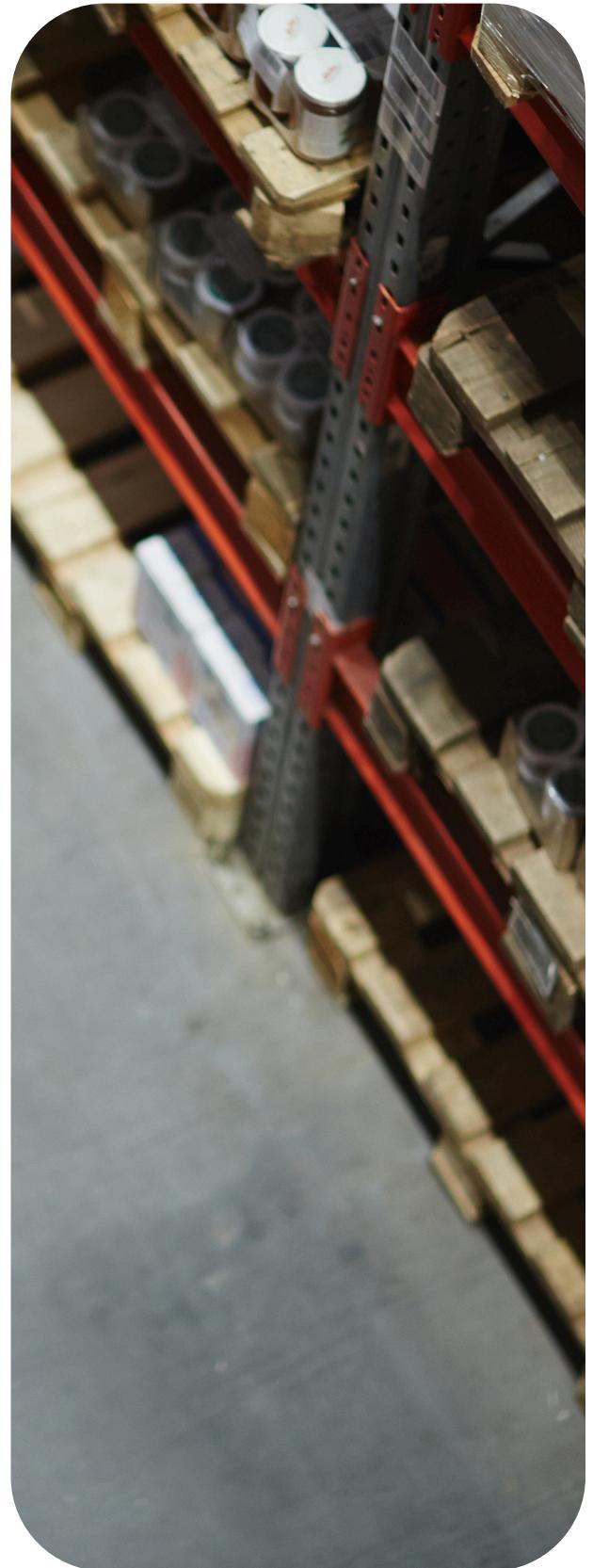


Problem

Efficient management of inventory and sales is essential for maximizing profitability in both retail and wholesale sectors. Businesses must ensure that losses do not occur as a result of poor pricing strategies, low inventory turnover, or excessive dependence on certain vendors.

This analysis aims to:

- **Identify brands with low performance** that may need pricing or promotional strategies.
- **Highlight top-performing vendors** driving sales and gross profit.
- **Analyze the impact of bulk purchasing** on unit costs.
- **Assess inventory turnover** to reduce holding costs and improve efficiency.
- **Investigate the profitability variance** between high-performing and low-performing vendors.



About the Dataset

This dataset provides a complete overview of vendor performance across key operational aspects such as purchases, sales, inventory, pricing, and invoices. It captures essential business interactions and financial details that allow for assessing vendor efficiency, profitability, and overall supply chain performance. It is available on [Kaggle](#).

Breakdown of Dataset Files

purchases.csv

→ Detailed vendor purchase transactions

sales.csv

→ Records sales performance data linked to vendors

**inventory.csv
(begin & end)**

→ Represents stock levels at different periods

purchase_prices.csv

→ Provides historical vendor pricing information

vendor_invoice.csv

→ Invoice details for reconciliation

vendor_sales_summary.csv

→ Aggregated vendor-wise sales insights

SQL Data Analysis and Cleaning

Ingesting Dataset Files into SQLite Database

The script imports the `begin_inventory.csv` file into the `inventory.db` database stored in Google Drive. It reads the data in 500-row chunks for efficient processing, prevents duplicate table creation, and inserts the records into the database. Once finished, it verifies the process by listing all tables and their row counts.

These steps are also applied to `end_inventory.csv`, `purchase_prices.csv`, `purchases.csv`, `sales.csv`, and `vendor_invoice.csv`.

```
from google.colab import drive
drive.mount('/content/drive')

import pandas as pd
from sqlalchemy import create_engine, inspect
import sqlite3

db_path = '/content/drive/My Drive/Vendor Performance Analysis/inventory.db'
data_path = '/content/drive/My Drive/Vendor Performance Analysis/data/begin_inventory.csv'

engine = create_engine(f'sqlite:///{{db_path}}')

def ingest_vendor_invoice(chunksize=500):
    table_name = 'begin_inventory'
    inspector = inspect(engine)
    existing_tables = inspector.get_table_names()
    if table_name in existing_tables:
        print(f'Tabel {table_name} sudah ada. Proses ingest dihentikan untuk mencegah duplikasi.')
        return
    chunk_num = 0
    for chunk in pd.read_csv(data_path, chunksize=chunksize, low_memory=True):
        chunk.to_sql(
            table_name,
            con=engine,
            if_exists='append',
            index=False,
            method='multi'
        )
        chunk_num += 1
        print(f"[{table_name}] Inserted chunk {chunk_num} ({len(chunk)} rows)")
    print(f"Ingession of {table_name} complete!")

ingest_vendor_invoice()
conn = sqlite3.connect(db_path)
cursor = conn.cursor()

cursor.execute("SELECT name FROM sqlite_master WHERE type='table';")
tables = cursor.fetchall()
print("Tables in DB:", tables)

for table in tables:
    table_name = table[0]
    cursor.execute(f"SELECT COUNT(*) FROM {table_name}")
    count = cursor.fetchone()[0]
    print(f"{table_name}: {count} rows")

conn.close()
```

Check Tables in the Database

Check the Tables' Structure and Column Data Types

	name
0	begin_inventory
1	end_inventory
2	purchase_prices
3	purchases
4	sales
5	vendor_invoice

begin_inventory									
Count of records: 206529									
InventoryId	Store	City	Brand	Description	Size	onHand	Price	startDate	
0	1_HARDERSFIELD_58	1	HARDERSFIELD	58 Gekkeikan Black & Gold Sake	750mL	8	12.99	2024-01-01	
1	1_HARDERSFIELD_60	1	HARDERSFIELD	60 Canadian Club 1858 VAP	750mL	7	10.99	2024-01-01	
2	1_HARDERSFIELD_62	1	HARDERSFIELD	62 Herradura Silver Tequila	750mL	6	36.99	2024-01-01	
3	1_HARDERSFIELD_63	1	HARDERSFIELD	63 Herradura Reposado Tequila	750mL	3	38.99	2024-01-01	
4	1_HARDERSFIELD_72	1	HARDERSFIELD	72 No. 3 London Dry Gin	750mL	6	34.99	2024-01-01	

end_inventory									
Count of records: 224489									
InventoryId	Store	City	Brand	Description	Size	onHand	Price	endDate	
0	1_HARDERSFIELD_58	1	HARDERSFIELD	58 Gekkeikan Black & Gold Sake	750mL	11	12.99	2024-12-31	
1	1_HARDERSFIELD_62	1	HARDERSFIELD	62 Herradura Silver Tequila	750mL	7	36.99	2024-12-31	
2	1_HARDERSFIELD_63	1	HARDERSFIELD	63 Herradura Reposado Tequila	750mL	7	38.99	2024-12-31	
3	1_HARDERSFIELD_72	1	HARDERSFIELD	72 No. 3 London Dry Gin	750mL	4	34.99	2024-12-31	
4	1_HARDERSFIELD_75	1	HARDERSFIELD	75 Three Olives Tomato Vodka	750mL	7	14.99	2024-12-31	

purchase_prices									
Count of records: 12261									
Brand	Description	Price	Size	Volume	Classification	PurchasePrice	VendorNumber	VendorName	
0	58 Gekkeikan Black & Gold Sake	12.99	750mL	750	1	9.28	8320	SHAW ROSS INT'L IMP LTD	
1	62 Herradura Silver Tequila	36.99	750mL	750	1	28.67	1128	BROWN-FORMAN CORP	
2	63 Herradura Reposado Tequila	38.99	750mL	750	1	30.46	1128	BROWN-FORMAN CORP	
3	72 No. 3 London Dry Gin	34.99	750mL	750	1	26.11	9165	ULTRA BEVERAGE COMPANY LLP	
4	75 Three Olives Tomato Vodka	14.99	750mL	750	1	10.94	7245	PROXIMO SPIRITS INC.	

purchases																
Count of records: 2372474																
InventoryId	Store	Brand	Description	Size	VendorNumber	VendorName	PONumber	PoDate	ReceivingDate	InvoiceDate	PayDate	PurchasePrice	Quantity	Dollars	Classification	
0	69_MOUNTMEND_8412	69	8412	Tequila Ocho Plata Fresno	750mL	105	ALTAMAR BRANDS LLC	8124	2023-12-21	2024-01-02	2024-01-04	2024-02-16	35.71	6	214.26	1
1	30_CULCHETH_5255	30	5255	TGI Fridays Ultimate Mudslide	1.75L	4466	AMERICAN VINTAGE BEVERAGE	8137	2023-12-22	2024-01-01	2024-01-07	2024-02-21	9.35	4	37.40	1
2	34_PITMERDEN_5215	34	5215	TGI Fridays Long Island Iced	1.75L	4466	AMERICAN VINTAGE BEVERAGE	8137	2023-12-22	2024-01-02	2024-01-07	2024-02-21	9.41	5	47.05	1
3	1_HARDERSFIELD_5255	1	5255	TGI Fridays Ultimate Mudslide	1.75L	4466	AMERICAN VINTAGE BEVERAGE	8137	2023-12-22	2024-01-01	2024-01-07	2024-02-21	9.35	6	56.10	1
4	76_DONCASTER_2034	76	2034	Glendalough Double Barrel	750mL	388	ATLANTIC IMPORTING COMPANY	8169	2023-12-24	2024-01-02	2024-01-09	2024-02-16	21.32	5	106.60	1

sales														
Count of records: 12825363														
InventoryId	Store	Brand	Description	Size	SalesQuantity	SalesDollars	SalesPrice	SalesDate	Volume	Classification	ExciseTax	VendorNo	VendorName	
0	1_HARDERSFIELD_1004	1	1004	Jim Beam w/2 Rocks Glasses	750mL	1	16.49	16.49	2024-01-01	750.0	1	0.79	12546	JIM BEAM BRANDS COMPANY
1	1_HARDERSFIELD_1004	1	1004	Jim Beam w/2 Rocks Glasses	750mL	2	32.98	16.49	2024-01-02	750.0	1	1.57	12546	JIM BEAM BRANDS COMPANY
2	1_HARDERSFIELD_1004	1	1004	Jim Beam w/2 Rocks Glasses	750mL	1	16.49	16.49	2024-01-03	750.0	1	0.79	12546	JIM BEAM BRANDS COMPANY
3	1_HARDERSFIELD_1004	1	1004	Jim Beam w/2 Rocks Glasses	750mL	1	14.49	14.49	2024-01-08	750.0	1	0.79	12546	JIM BEAM BRANDS COMPANY
4	1_HARDERSFIELD_1005	1	1005	Maker's Mark Combo Pack	375mL 2 Pk	2	69.98	34.99	2024-01-09	375.0	1	0.79	12546	JIM BEAM BRANDS COMPANY

vendor_invoice										
Count of records: 5543										
VendorNumber	VendorName	InvoiceDate	PONumber	PoDate	PayDate	Quantity	Dollars	Freight	Approval	
0	105	ALTAMAR BRANDS LLC	2024-01-04	8124	2023-12-21	2024-02-16	6	214.26	3.47	None
1	4466	AMERICAN VINTAGE BEVERAGE	2024-01-07	8137	2023-12-22	2024-02-21	15	140.55	8.57	None
2	388	ATLANTIC IMPORTING COMPANY	2024-01-09	8169	2023-12-24	2024-02-16	5	106.60	4.61	None
3	480	BACARDI USA INC	2024-01-12	8106	2023-12-20	202				

The next steps focus on preparing the essential transactional data. The immediate goal is to establish data integrity and create summary metrics for the target analysis by focusing on Vendor Number 4466 as a prototype.

Data Isolation and Validation

All purchasing (purchases and purchase_prices), invoicing (vendor_invoice), and sales (sales) records specifically for Vendor 4466 are extracted. This step confirms data relationships across the tables are sound.

# Retrieve all records for VendorNumber 4466 from the purchase_prices table									
purchase_prices = pd.read_sql_query("SELECT * FROM purchase_prices WHERE VendorNumber = 4466", conn)									
Brand	Description	Price	Size	Volume	Classification	PurchasePrice	VendorNumber	VendorName	
0	TGI Fridays Long Island Iced	12.99	1750mL	1750	1	9.41	4466	AMERICAN VINTAGE BEVERAGE	
1	TGI Fridays Ultieme Mudslide	12.99	1750mL	1750	1	9.35	4466	AMERICAN VINTAGE BEVERAGE	
2	TGI Fridays Orange Dream	14.99	1750mL	1750	1	11.19	4466	AMERICAN VINTAGE BEVERAGE	

# Fetch all records for VendorNumber 4466 from the vendor_invoice table									
vendor_invoice = pd.read_sql_query("SELECT * FROM vendor_invoice WHERE VendorNumber = 4466", conn)									
VendorNumber	VendorName	InvoiceDate	PONumber	PODate	PayDate	Quantity	Dollars	Freight	Approval
0	AMERICAN VINTAGE BEVERAGE	2024-01-07	8137	2023-12-22	2024-02-21	15	140.55	8.57	None
1	AMERICAN VINTAGE BEVERAGE	2024-01-19	8207	2023-12-27	2024-02-26	335	3142.33	16.97	None
2	AMERICAN VINTAGE BEVERAGE	2024-01-18	8307	2024-01-03	2024-02-18	41	383.35	1.99	None

# Retrieve all sales records for VendorNo 4466 from the sales table									
sales = pd.read_sql_query("SELECT * FROM sales WHERE VendorNo = 4466", conn)									
InventoryId	Store	Brand	Description	Size	SalesQuantity	SalesDollars	SalesPrice	SalesDate	
0	1_HARDERSFIELD_5215	1	TGI Fridays Long Island Iced	1.75L	1	12.99	12.99	2024-01-09	
1	1_HARDERSFIELD_5215	1	TGI Fridays Long Island Iced	1.75L	1	12.99	12.99	2024-01-12	
2	1_HARDERSFIELD_5215	1	TGI Fridays Long Island Iced	1.75L	1	12.99	12.99	2024-01-15	

Volume	Classification	ExciseTax	VendorNo	VendorName
1750.0	1	1.84	4466	AMERICAN VINTAGE BEVERAGE
1750.0	1	1.84	4466	AMERICAN VINTAGE BEVERAGE
1750.0	1	1.84	4466	AMERICAN VINTAGE BEVERAGE

Purchasing Aggregation

Purchase volume (Quantity and Dollars) is summarized at the Brand and Purchase Order (PO) Number levels. This allows for **verification that the detailed purchase line items sum up correctly to the total transaction amounts found in the vendor_invoice table**.

```
# Aggregate total quantity and dollars spent by Brand and PurchasePrice
purchases.groupby(["Brand", "PurchasePrice"])[["Quantity", "Dollars"]].sum()
```

Brand	PurchasePrice	Quantity	Dollars
3140	11.19	4640	51921.60
5215	9.41	4923	46325.43
5255	9.35	6215	58110.25

```
# Aggregate total quantity and dollars spent by purchase order number
purchases.groupby(["PONumber"])[["Quantity", "Dollars"]].sum()
```

PONumber	Quantity	Dollars
8137	15	140.55
8207	335	3142.33
8307	41	383.35

Sales Aggregation

Sales metrics (SalesQuantity, SalesDollars, ExciseTax) are **aggregated by Brand**. This establishes the revenue and demand profile for the products supplied by this specific vendor.

```
# Aggregate total sales metrics by brand
sales.groupby("Brand")[['SalesQuantity', 'SalesDollars', 'SalesPrice', 'ExciseTax']].sum()
```

Brand	SalesQuantity	SalesDollars	SalesPrice	ExciseTax
3140	3890	50531.10	30071.85	7149.25
5215	4651	60416.49	41542.02	8548.96
5255	6096	79187.04	51180.60	11204.28

These summarized results will then be used as the building blocks to construct the final Consolidated Summary Table (vendor_sales_summary) for all vendors.

The next step was to **create a consolidated summary table combining data from purchases, sales, and vendor invoices**. Three subqueries (FreightSummary, PurchaseSummary, and SalesSummary) were built to calculate total purchase amounts, sales revenue, product prices, and freight costs for each vendor. These datasets were then joined to form the `vendor_sales_summary` table, providing a complete view of every vendor's activity across brands and transactions.

1

```
freight_summary = pd.read_sql_query("""SELECT
    VendorNumber,
    SUM(Freight) as FreightCost
FROM vendor_invoice
GROUP BY VendorNumber""", conn)
freight_summary
```

	VendorNumber	FreightCost
0	2	27.08
1	54	0.48
2	60	367.52
3	105	62.39
4	200	6.19

2

```
purchase_summary = pd.read_sql_query("""SELECT
    p.VendorNumber,
    p.VendorName,
    p.Brand,
    p.Description,
    p.PurchasePrice,
    pp.Price as ActualPrice,
    pp.Volume,
    SUM(p.Quantity) AS TotalPurchaseQuantity,
    SUM(p.Dollars) AS TotalPurchaseDollars
FROM purchases p
JOIN purchase_prices pp
    ON p.Brand = pp.Brand
WHERE p.PurchasePrice > 0
GROUP BY p.VendorNumber,p.VendorName, p.Brand, p.Description, p.PurchasePrice, pp.Price, pp.Volume
ORDER BY TotalPurchaseDollars
""",conn)
purchase_summary
```

	VendorNumber	VendorName	Brand	Description	PurchasePrice	ActualPrice	Volume	TotalPurchaseQuantity	TotalPurchaseDollars
0	7245	PROXIMO SPIRITS INC.	3065	Three Olives Grape Vodka	0.71	0.99	50.0	1	0.71
1	3960	DIAGEO NORTH AMERICA INC	6127	The Club Strawbry Margarita	1.47	1.99	200.0	1	1.47
2	3924	HEAVEN HILL DISTILLERIES	9123	Deep Eddy Vodka	0.74	0.99	50.0	2	1.48

3

```
sales_summary = pd.read_sql_query("""SELECT
    VendorNo,
    Brand,
    SUM(SalesQuantity) AS TotalSalesQuantity,
    SUM(SalesDollars) AS TotalSalesDollars,
    SUM(SalesPrice) AS TotalSalesPrice,
    SUM(ExciseTax) AS TotalExciseTax
FROM sales
GROUP BY VendorNo, Brand""",conn)
sales_summary
```

	VendorNo	Brand	TotalsalesQuantity	TotalsalesDollars	TotalsalesPrice	TotalExciseTax
0	2	90085	18	665.82	295.92	2.00
1	2	90609	24	599.76	449.82	0.52
2	60	771	47	704.53	494.67	37.01

4

```
vendor_sales_summary.columns
Index(['VendorNumber', 'VendorName', 'Brand', 'Description', 'PurchasePrice',
       'ActualPrice', 'Volume', 'TotalPurchaseQuantity',
       'TotalPurchaseDollars', 'TotalSalesQuantity', 'TotalSalesDollars',
       'TotalSalesPrice', 'TotalExciseTax', 'FreightCost'],
      dtype='object')
```

This pre-aggregated summary table allows faster and more efficient analysis. Instead of running multiple heavy queries on large tables each time, dashboards and reports can directly access summarized results from `vendor_sales_summary`, enabling quick insights into vendor performance, profitability, and pricing trends.

Data Cleaning

Vendor Sales Summary Dtypes	
VendorNumber	int64
VendorName	object
Brand	int64
Description	object
PurchasePrice	float64
ActualPrice	float64
Volume	float64
TotalPurchaseQuantity	int64
TotalPurchaseDollars	float64
TotalSalesQuantity	float64
TotalSalesDollars	float64
TotalSalesPrice	float64
TotalExciseTax	float64
FreightCost	float64

Vendor Sales Summary Isnull().sum()	
VendorNumber	0
VendorName	0
Brand	0
Description	0
PurchasePrice	0
ActualPrice	0
Volume	0
TotalPurchaseQuantity	0
TotalPurchaseDollars	0
TotalSalesQuantity	178
TotalSalesDollars	178
TotalSalesPrice	178
TotalExciseTax	178
FreightCost	0

Vendor Sales Summary VendorName Unique	
'PALM BAY INTERNATIONAL INC'	'REMY COINTREAU USA INC'
'SIDNEY FRANK IMPORTING CO'	'E & J GALLO WINERY'
'WILLIAM GRANT & SONS INC'	'HEAVEN HILL DISTILLERIES'
'DISARONNO INTERNATIONAL LLC'	'EDRINGTON AMERICAS'
'CASTLE BRANDS CORP.'	'SOUTHERN WINE & SPIRITS NE'
'STE MICHELLE WINE ESTATES'	'TRINCHERO FAMILY ESTATES'
'MHW LTD'	'WINE GROUP INC'
'PERFECTA WINES'	'LUXCO INC'
'TREASURY WINE ESTATES'	'DIAGEO CHATEAU ESTATE WINES'

Identified Issues

- The Volume is a numerical column but having object datatype.
- There are some products that are not sold, so having missing values.
- There are white spaces in the categorical columns.

Action Taken

- Convert 'Volume' column to float for numerical analysis
- Replace missing values with 0 to ensure clean calculations
- Remove leading and trailing spaces from categorical columns

```
# Convert 'Volume' column to float for numerical analysis
vendor_sales_summary['Volume'] = vendor_sales_summary['Volume'].astype('float')
# Replace missing values with 0 to ensure clean calculations
vendor_sales_summary.fillna(0, inplace=True)
# Remove leading and trailing spaces from categorical columns
vendor_sales_summary['VendorName'] = vendor_sales_summary['VendorName'].str.strip()
vendor_sales_summary['Description'] = vendor_sales_summary['Description'].str.strip()
```

```
# Add new calculated columns to enhance analysis
vendor_sales_summary['GrossProfit'] = vendor_sales_summary['TotalSalesDollars'] - vendor_sales_summary['TotalPurchaseDollars']
vendor_sales_summary['ProfitMargin'] = (vendor_sales_summary['GrossProfit'] / vendor_sales_summary['TotalSalesDollars']) * 100
vendor_sales_summary['StockTurnover'] = vendor_sales_summary['TotalSalesQuantity'] / vendor_sales_summary['TotalPurchaseQuantity']
vendor_sales_summary['SalesToPurchaseRatio'] = vendor_sales_summary['TotalSalesDollars'] / vendor_sales_summary['TotalPurchaseDollars']
```

New columns were added to the vendor_sales_summary DataFrame to provide deeper insights into vendor performance.

- **GrossProfit** shows the profit by subtracting purchase costs from sales revenue
- **ProfitMargin** expresses this profit as a percentage of sales
- **StockTurnover** measures how efficiently inventory is sold
- **SalesToPurchaseRatio** indicates how effectively purchases generate sales.

After enhancing the vendor_sales_summary DataFrame, the dataset was saved into the database by creating a table named vendor_sales_summary with appropriate data types and a composite primary key on VendorNumber and Brand to ensure data integrity and support further analysis.

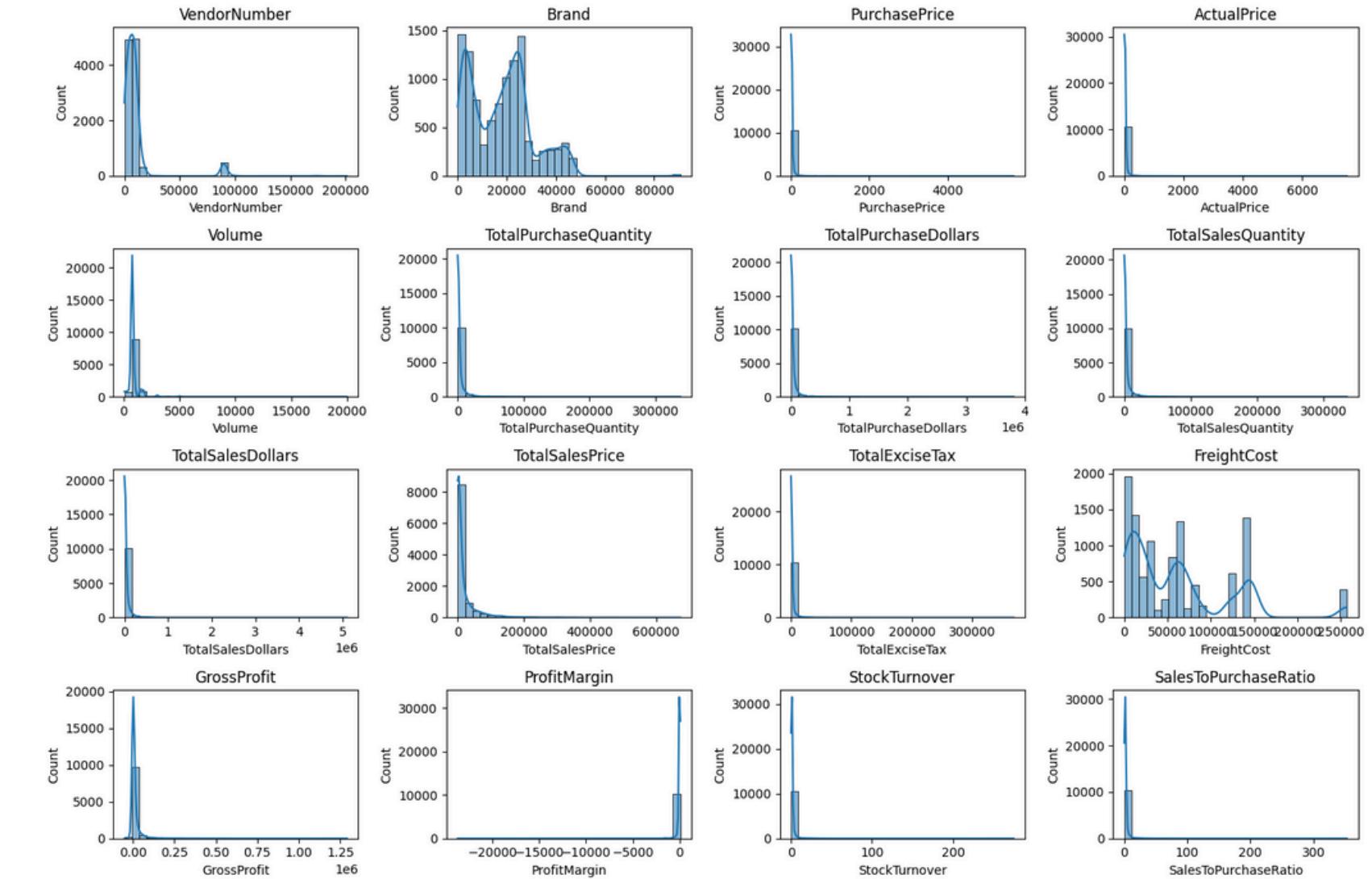
```
cursor = conn.cursor()

cursor.execute("""CREATE TABLE vendor_sales_summary (
    VendorNumber INT,
    VendorName VARCHAR(100),
    Brand INT,
    Description VARCHAR(100),
    PurchasePrice DECIMAL(10,2),
    ActualPrice DECIMAL(10,2),
    Volume INT,
    TotalPurchaseQuantity INT,
    TotalPurchaseDollars DECIMAL(15,2),
    TotalSalesQuantity INT,
    TotalSalesDollars DECIMAL(15,2),
    TotalSalesPrice DECIMAL(15,2),
    TotalExciseTax DECIMAL(15,2),
    FreightCost DECIMAL(15,2),
    GrossProfit DECIMAL(15,2),
    ProfitMargin DECIMAL(15,2),
    StockTurnover DECIMAL(15,2),
    SalesToPurchaseRatio DECIMAL(15,2),
    PRIMARY KEY (VendorNumber, Brand)
);""")
```

Exploratory Data Analysis with Python (Insights)

Summary Statistics

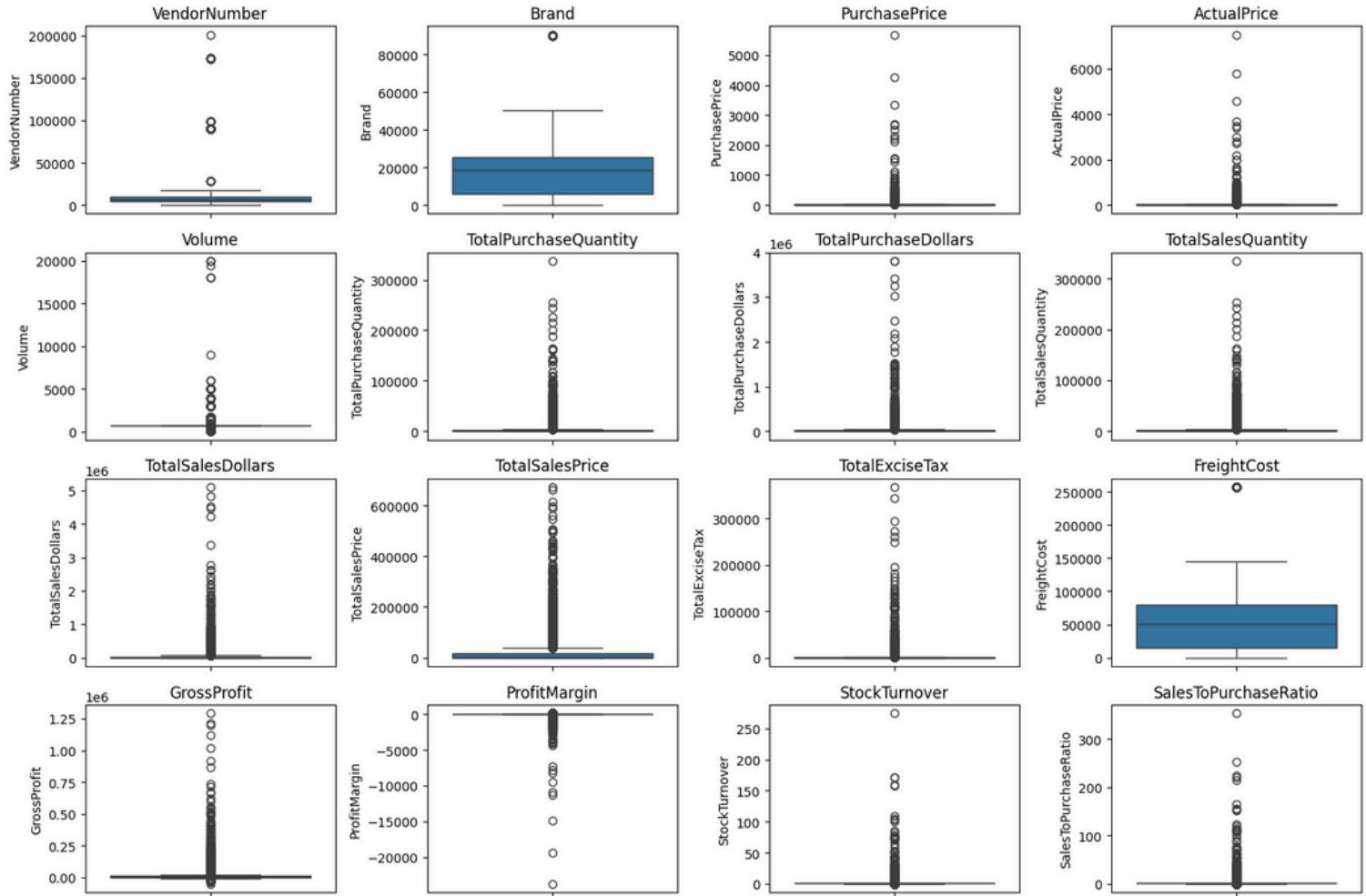
	count	mean	std	min	25%	50%	75%	max
VendorNumber	10692.0	1.065065e+04	18753.519148	2.00	3951.000000	7153.000000	9552.000000	2.013590e+05
Brand	10692.0	1.803923e+04	12662.187074	58.00	5793.500000	18761.500000	25514.250000	9.063100e+04
PurchasePrice	10692.0	2.438530e+01	109.269375	0.36	6.840000	10.455000	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
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
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



Negative & Zero Values

- Gross Profit:** The minimum value of -52,002.78 suggests **possible losses, likely caused by high purchase costs or significant discounts**. This may occur when **products are sold below their original purchase price**.
- Profit Margin:** The minimum value of $-\infty$ indicates cases where **revenue is zero or falls below total costs**, resulting in **extremely negative profit margins**.
- Total Sales Quantity & Sales Dollars:** Some products show **zero sales**, indicating they were purchased but never sold. These may be **slow-moving or obsolete stock, leading to inventory inefficiencies**.

Summary Statistics



Outliers Detected by High Standard Deviations

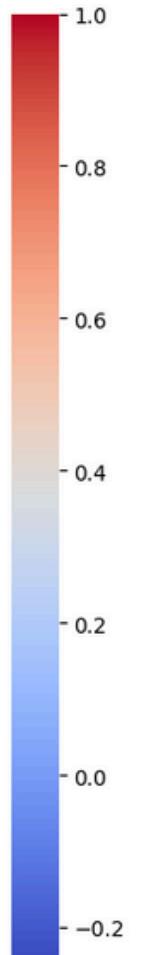
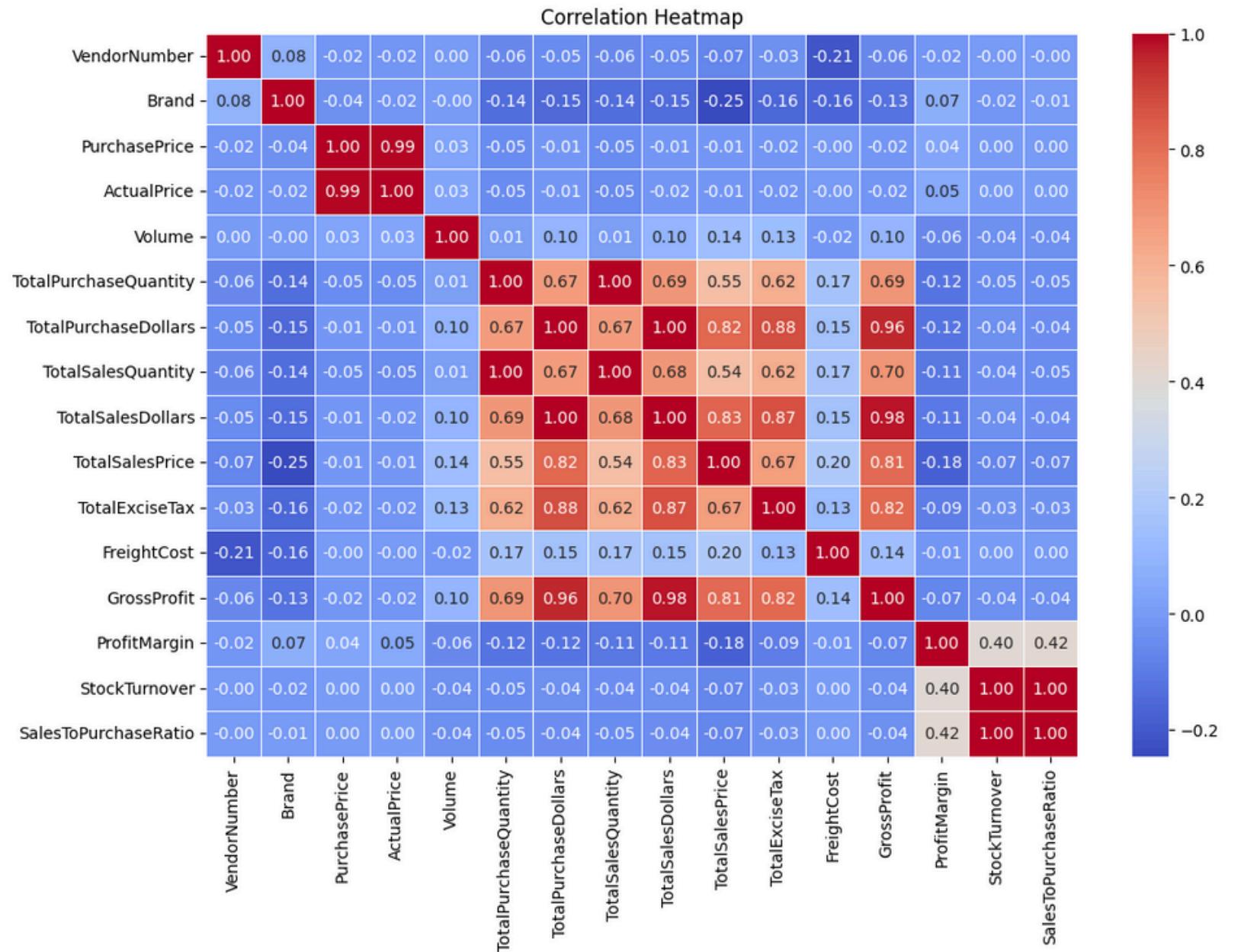
- **Purchase & Actual Prices:** The maximum values (5,681.81 & 7,499.99) are significantly higher than the mean (24.39 & 35.64), **indicating premium product offerings**.
- **Freight Cost:** Extreme variation from 0.09 to 257,032.07 **suggests logistics inefficiencies, bulk shipments, or erratic shipping costs across different products**.
- **Stock Turnover:** Ranges from 0 to 274.5, suggesting **some products sell rapidly while others remain unsold for long periods**. A value greater than 1 indicates that sales for a product exceed the purchased quantity due to older stock fulfilling orders.

Data Filtering

To improve the reliability of the insights, inconsistent data points were **removed** where:

- **Gross Profit ≤ 0 ,** excluding transactions that resulted in losses.
- **Profit Margin ≤ 0 ,** focusing the analysis on profitable transactions.
- **Total Sales Quantity = 0,** eliminating inventory that was never sold.

Correlation Insights



- Purchase Price vs. Total Sales Dollars & Gross Profit:** Exhibits a weak correlation (-0.012 and -0.016), suggesting that **fluctuations in purchase price have minimal influence** on overall sales revenue and profit levels.
- Total Purchase Quantity vs. Total Sales Quantity:** Shows a strong correlation (0.999), indicating a **highly efficient relationship between purchasing and sales activities, reflecting effective inventory management**.
- Profit Margin vs. Total Sales Price:** Displays a negative correlation (-0.179), implying that **higher sales prices may be associated with reduced profit margins, potentially due to competitive market pressures**.
- Stock Turnover vs. Gross Profit & Profit Margin:** Demonstrates a weak negative correlation (-0.038 and -0.055), suggesting that **rapid stock turnover does not necessarily result in increased profitability**.

Research Questions & Key Findings

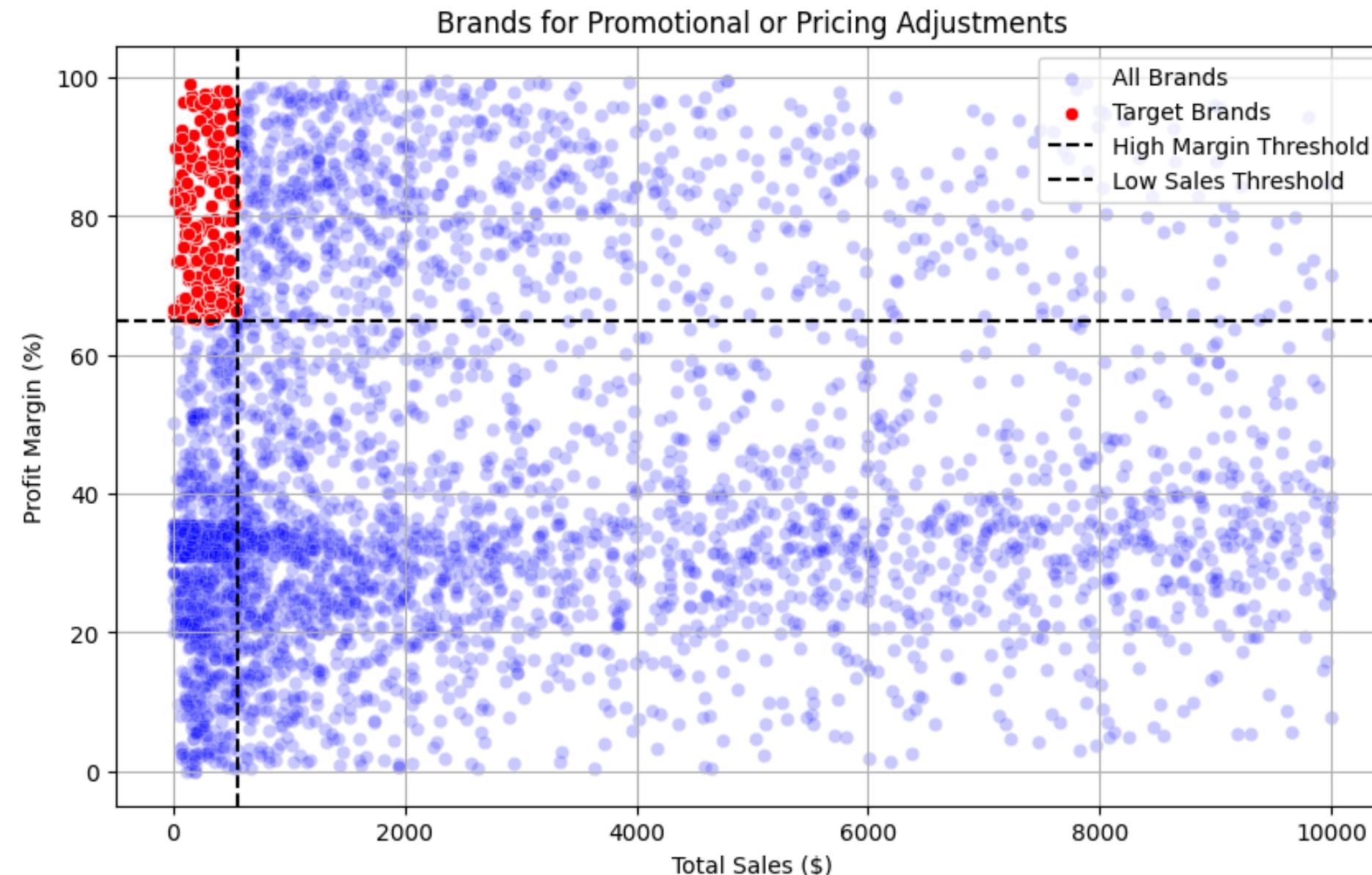
Identify Brands That Require Promotional or Pricing Adjustments

Brands with Low Sales Performance but High Profit Margins

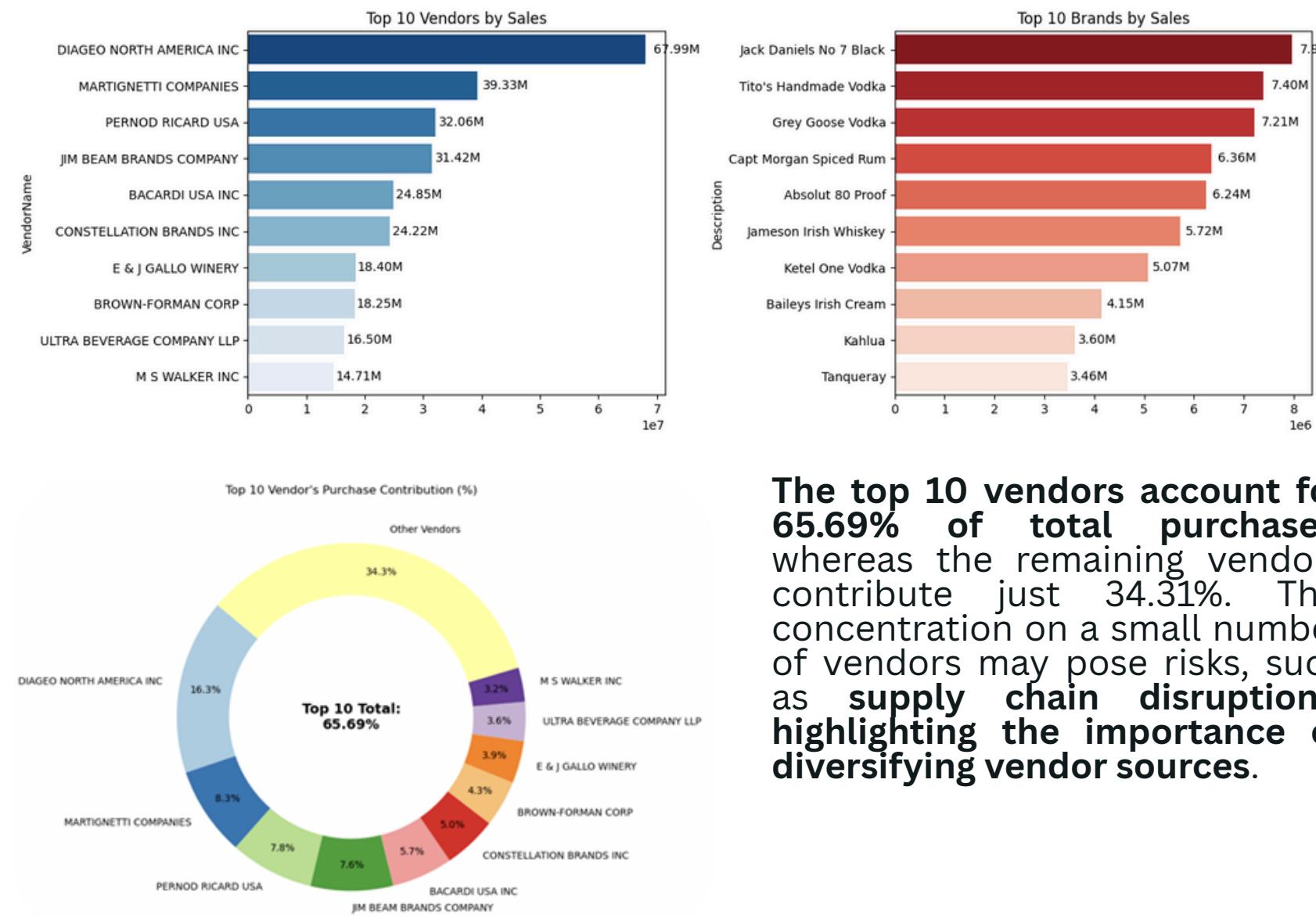
	Description	TotalSalesDollars	ProfitMargin
6199	Santa Rita Organic Svgn Bl	9.99	66.466466
2369	Debauchery Pnt Nr	11.58	65.975820
2070	Concannon Glen Ellen Wh Zin	15.95	83.448276
2188	Crown Royal Apple	27.86	89.806174
6237	Sauza Sprklg Wild Berry Marg	27.96	82.153076
...
5074	Nanbu Bijin Southern Beauty	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 x 3 columns

A total of **198 brands** show low sales despite high profit margins, indicating **opportunities for targeted marketing, promotional campaigns, or pricing adjustments** to boost sales volume while maintaining profitability.



Top Vendors by Sales and Purchase Contribution



Impact of Bulk Purchasing on Cost Savings

OrderSize	UnitPurchasePrice
0	Small
1	Medium
2	Large

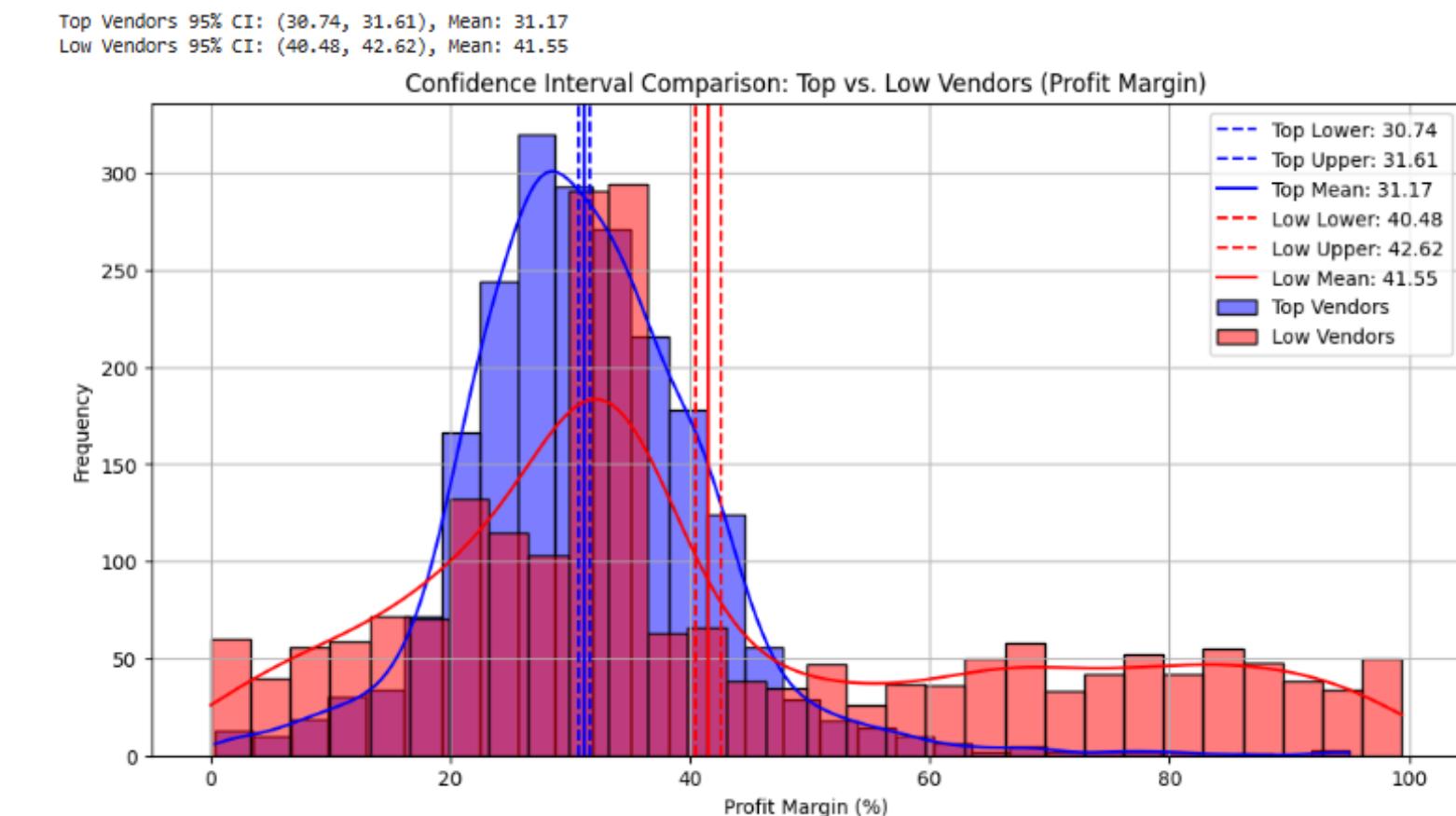
Vendors purchasing in **large quantities benefit from a 72% lower unit cost** (\$10.78 per unit compared to higher costs for smaller orders). This bulk pricing strategy **incentivizes larger purchases, boosting total sales while preserving profitability**.

Identify Vendors with Low Inventory Turnover

	VendorName	StockTurnover	Total Unsold Capital: 2.71M	VendorName	UnsoldInventoryValue
0	ALISA CARR BEVERAGES	0.615385		25	DIAGEO NORTH AMERICA INC 722.21K
36	HIGHLAND WINE MERCHANTS LLC	0.708333		46	JIM BEAM BRANDS COMPANY 554.67K
60	PARK STREET IMPORTS LLC	0.751306		68	PERNOD RICARD USA 470.63K
19	Circa Wines	0.755676		116	WILLIAM GRANT & SONS INC 401.96K
26	Dunn Wine Brokers	0.766022		30	E & J GALLO WINERY 228.28K
15	CENTEUR IMPORTS LLC	0.773953		79	SAZERAC CO INC 198.44K
78	SMOKY QUARTZ DISTILLERY LLC	0.783835		11	BROWN-FORMAN CORP 177.73K
90	TAMWORTH DISTILLING	0.797078		20	CONSTELLATION BRANDS INC 133.62K
91	THE IMPORTED GRAPE LLC	0.807569		61	MOET HENNESSY USA INC 126.48K
101	WALPOLE MTN VIEW WINERY	0.820548		77	REMY COINTREAU USA INC 118.60K

- Total Unsold Inventory Capital: \$2.71M**, indicating a substantial amount of capital tied up in unsold goods.
- The presence of **slow-moving inventory** contributes to **higher storage costs, reduced cash flow efficiency**, and a **negative impact on overall profitability**.
- Identifying vendors with lower inventory turnover supports more strategic stock management, helping to optimize inventory levels and alleviate financial pressure.

Profit Margin Comparison: High vs. Low-Performing Vendors



Low-performing vendors have higher margins but lower sales volumes, indicating **possible pricing inefficiencies or limited market reach**.

Recommendation Actions:

- Top Vendors:** Adjust pricing, reduce operational costs, and offer bundled promotions to maximize profit.
- Low Vendors:** Strengthen marketing strategies, optimize pricing, and improve distribution to boost sales performance.

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.

```
top_threshold = df["TotalSalesDollars"].quantile(0.75)
low_threshold = df["TotalSalesDollars"].quantile(0.25)

top_vendors = df[df["TotalSalesDollars"] >= top_threshold][["ProfitMargin"]].dropna()
low_vendors = df[df["TotalSalesDollars"] <= low_threshold][["ProfitMargin"]].dropna()

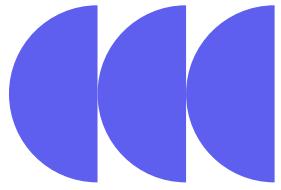
# Perform Two-Sample T-Test
t_stat, p_value = ttest_ind(top_vendors, low_vendors, equal_var=False)

# Print results
print(f"T-Statistic: {t_stat:.4f}, P-Value: {p_value:.4f}")
if p_value < 0.05:
    print("Reject  $H_0$ : There is a significant difference in profit margins between top and low-performing vendors.")
else:
    print("Fail to Reject  $H_0$ : No significant difference in profit margins.")

T-Statistic: -17.6440, P-Value: 0.0000
Reject  $H_0$ : There is a significant difference in profit margins between top and low-performing vendors.
```

The analysis indicates that the null hypothesis is rejected, confirming that the two vendor groups follow distinct profitability models. This suggests that high-margin vendors could enhance performance by refining their pricing strategies, whereas top-selling vendors may achieve greater profitability by focusing on operational cost efficiency.

Final Recommendation



Review pricing for low-sales, high-margin brands to increase sales volume while preserving profitability.



Expand vendor partnerships to reduce reliance on a few suppliers and lower supply chain risks.



Utilize bulk purchasing to maintain competitive pricing and improve inventory management.



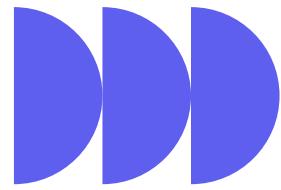
Manage **slow-moving inventory through adjusted purchase quantities**, clearance promotions, or optimized storage strategies.



Strengthen marketing and distribution for low-performing vendors to boost sales without eroding profit margins.



Implementing these measures can help the company achieve **sustainable profitability, risk mitigation, and enhanced operational efficiency**.



Power BI Dashboard

Vendor Performance Dashboard

441.41M

Total Sales (\$)

307.34M

Total Purchase (\$)

134.07M

Gross Profit (\$)

38.7

Profit Margin (%)

2.71M

Unsold Capital (\$)

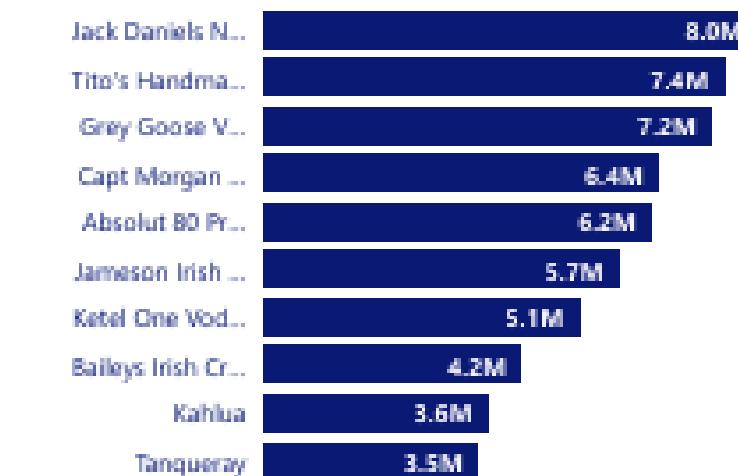
Purchase Contribution (%)

VendorName	Sum of PurchaseContribution%
M S WALKER INC	3.18
ULTRA BEVERAGE COMPANY LLP	3.63
E & J GALLO WINERY	3.93
BROWN-FORMAN CORP	4.31
CONSTELLATION BRANDS INC	4.97
BACARDI USA INC	5.67
JIM BEAM BRANDS COMPANY	7.64
PERNOD RICARD USA	7.76
MARTIGNETTI COMPANIES	8.30
Total	65.69

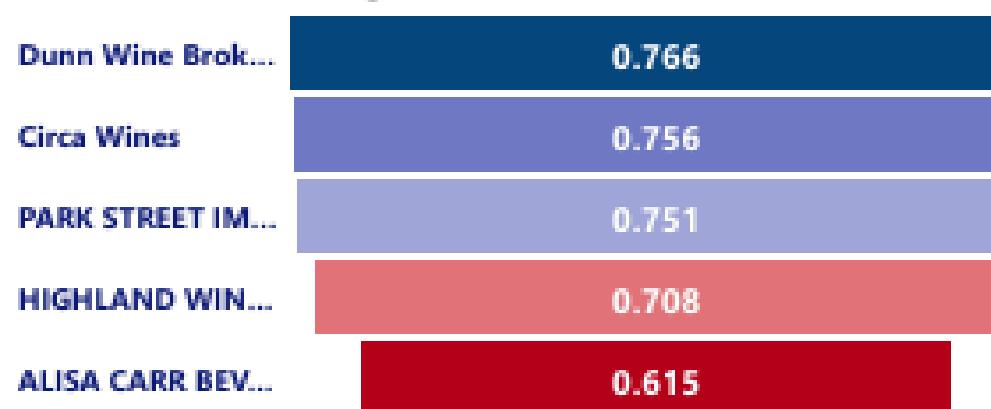
Top Vendors By Sales



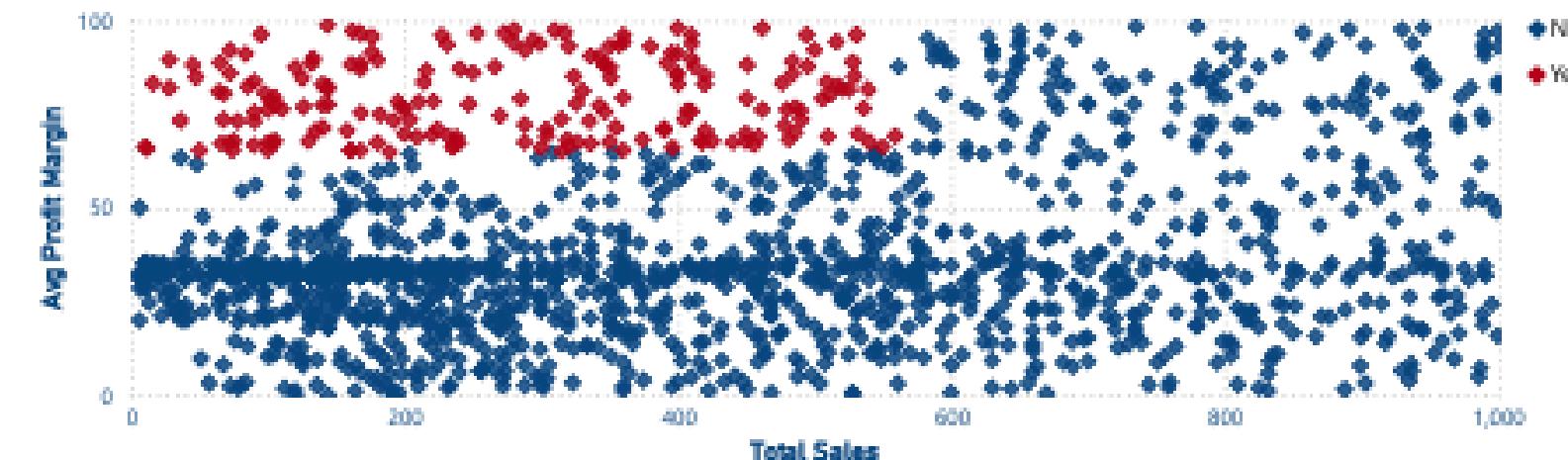
Top Brands By Sales



Low Performing Vendors



Low Performing Brands



THANK YOU