

# Sales Analysis Project Report

## 1. Project Overview

This project focuses on analyzing and visualizing regional sales performance to uncover key business trends, optimize product strategies, and improve profit margins. The analysis integrates Python-based Exploratory Data Analysis (EDA) and Power BI dashboards to derive actionable business insights.

The goal of this project is to:

- Identify high-performing regions, products, and sales channels
- Measure profit margins and customer contributions
- Build an interactive Power BI dashboard for real-time insights
- Support data-driven strategic decision-making

**Business Problem:** The company experienced inconsistent revenue growth, uneven profit performance, and limited visibility into channel and product-level profitability.

**Objectives:**

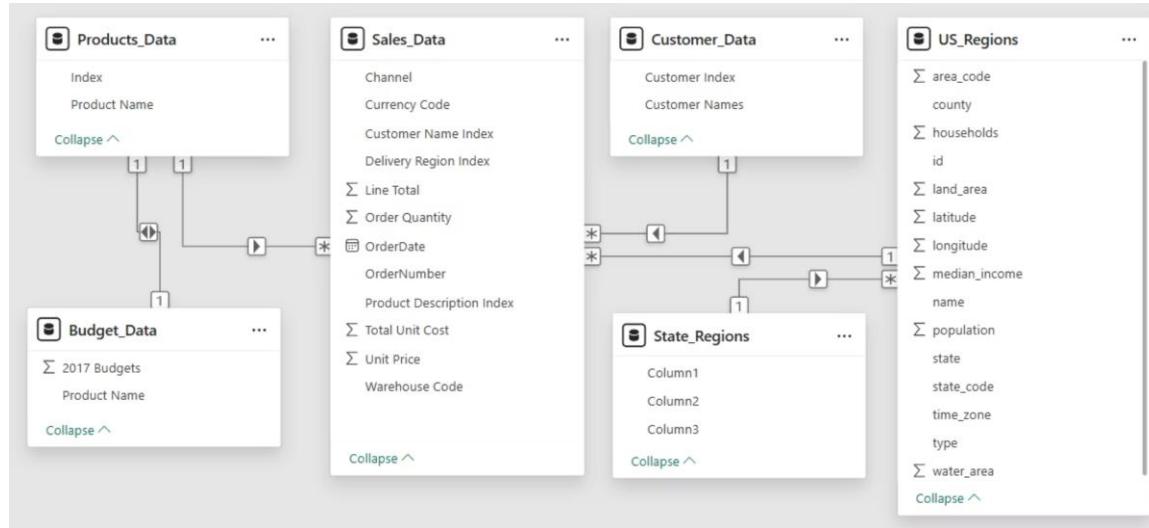
- Analyze 5 years of historical sales data
- Understand regional, product, and customer performance
- Build a live dashboard for executive decision-making

### Tech Stack

Tool	Purpose
Python (Pandas, Matplotlib, Seaborn)	Exploratory Data Analysis & Visualization
Power BI	Interactive Dashboard & KPI Visualization
Excel	Data Cleaning & Structuring
GitHub	Version Control & Project Documentation

## 2. Dataset Summary

The project dataset combines multiple tables: Sales, Products, Customers, Regions, and Budgets.



Raw Data Description:

Table	Description
Sales	Transaction-level data including order date, product, quantity, and revenue.
Products	Product details, category, and pricing information.
Customers	Customer profiles with name, channel, and region.
Regions	Mapping of states and regional sales distribution.
Budgets	Yearly financial targets for planning and comparison.

### Data Preprocessing:

- Merged datasets using keys: product\_id, customer\_id, and region\_id.
- Handled null values and standardized column names.
- Calculated derived fields: profit = revenue - cost and profit\_margin\_pct = (profit / revenue) \* 100.
- Verified no duplicates or missing records after merging.

	order_number	order_date	customer_names	channel	product_name	order_quantity	unit_price	revenue	total_unit_cost	state_code	county	state	region	latitude	longitude	budgets	total_cost	profit	profit_margin_pct
0	SO - 000225	2021-01-01	Rhynoodle Ltd	Wholesale	Product 27	6	2499.1	14994.6	1824343	GA	Chatham County	Georgia	South	32.08354	-81.09983	NaN	10946.058	4048.542	0.0027
1	SO - 0003378	2021-01-01	Thoughtmix Ltd	Distributor	Product 20	11	2351.7	25868.7	1269.918	IN	Johnson County	Indiana	Midwest	39.61366	-86.10665	NaN	13969.098	11899.602	0.0046
2	SO - 0005126	2021-01-01	Amerisource Corp	Wholesale	Product 26	6	978.2	5869.2	684740	CA	Alameda County	California	West	37.66243	-121.87468	NaN	4108.440	1760.760	0.0030
3	SO - 0005614	2021-01-01	Colgate-Palmolive Group	Export	Product 7	7	2338.3	16368.1	1026.852	IN	Monroe County	Indiana	Midwest	39.16533	-86.52639	NaN	7201.964	9166.136	0.0056
4	SO - 0005781	2021-01-01	Deseret Group	Wholesale	Product 8	8	2291.4	18331.2	1260.270	CT	Hartford County	Connecticut	Northeast	41.77524	-72.52443	NaN	10082.160	8249.040	0.0045

### 3. Exploratory Data Analysis (EDA) using Python

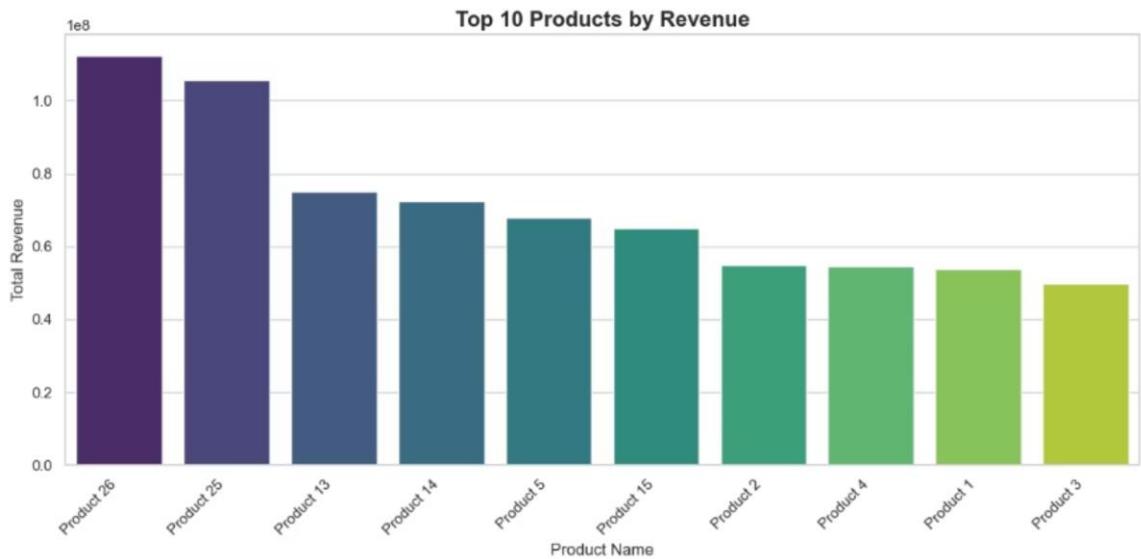
#### 1: Monthly Sales Trend



#### Insights:

- Sales show a consistent cycle with totals typically between \$24M and \$26M.
- Seasonal peaks occur in late spring/early summer (May–June), while January is typically the annual low.
- A notable outlier is a sharp revenue drop observed in early 2024 which warrants further investigation.

#### 2: Top 10 Products by Revenue



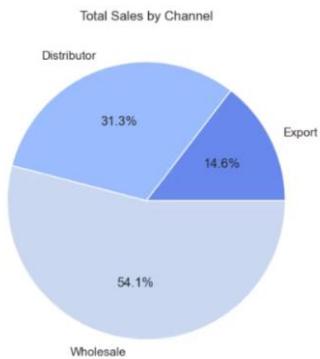
#### Insights:

- Products 26 and 25 are the dominant revenue drivers, together contributing a large portion of total sales (~25%).
- Mid-range SKUs such as Products 5, 13, 14, and 15 provide steady, mid-tier revenue and

represent opportunities for growth.

- Lower-performing SKUs (Products 1–4) contribute minimally and should be evaluated for repricing or phase-out.

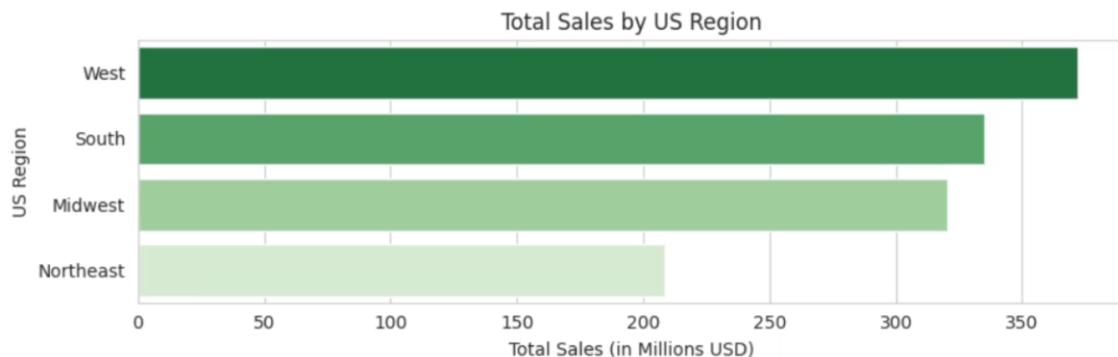
### 3: Sales by Channel



Insights:

- Wholesale is the largest revenue channel, generating approximately 54.1% of total sales.
- Distributor contributes around 31.3%, while Export accounts for about 14.6% of sales.
- Export shows higher margins on average; consider incentivizing export partnerships to improve overall profitability.

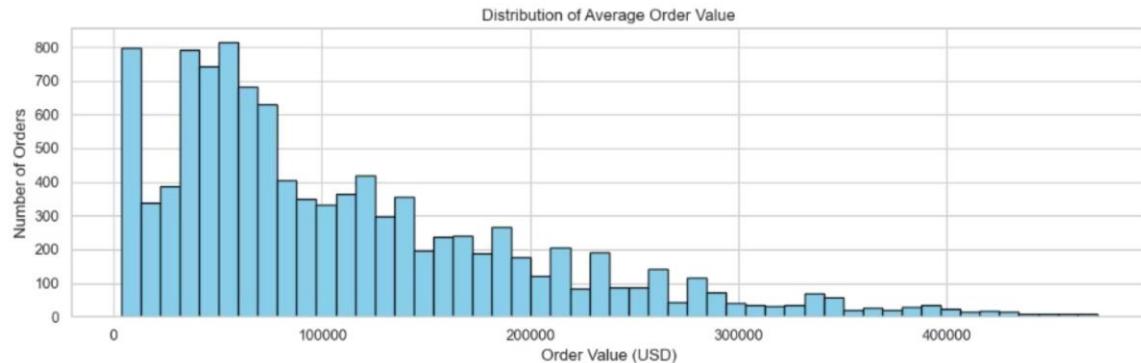
**Figure 4: Total Sales by US Region**



Insights:

- The West region leads in total sales and exhibits the largest revenue swings.
- The South region is a major contributor and the Midwest shows steady, moderate performance.
- The Northeast records the lowest sales and should be a focus for targeted growth initiatives.

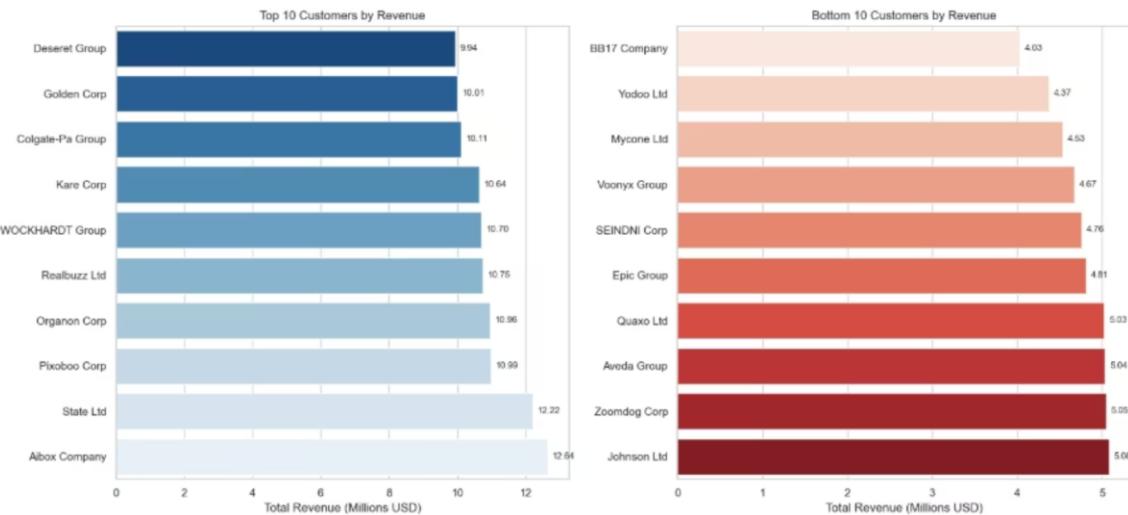
**Figure 5: Average Order Value (AOV) Distribution**



#### Insights:

- The AOV distribution is right-skewed: most orders are lower-value, with a long tail of high-value orders.
- There are multiple clusters of order values suggesting distinct customer segments (low, mid, high spenders).
- High-value orders are relatively rare, indicating potential upside in upsell or targeted promotions.

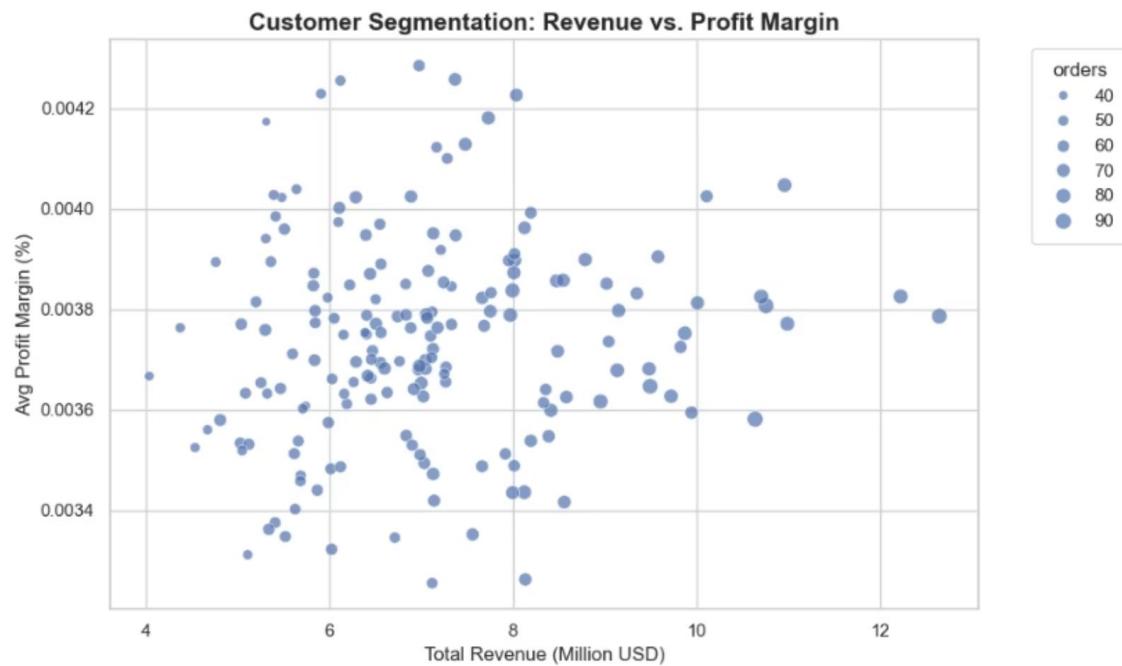
#### 6: Top & Bottom 10 Customers by Revenue



#### Insights:

- Aibox Company stands out as the top revenue-generating customer, contributing significantly more than others.
- The bottom 10 customers individually contribute substantially less (roughly \$4–5M range each).
- Revenue is concentrated among top customers, indicating a reliance on a small number of high-value clients.

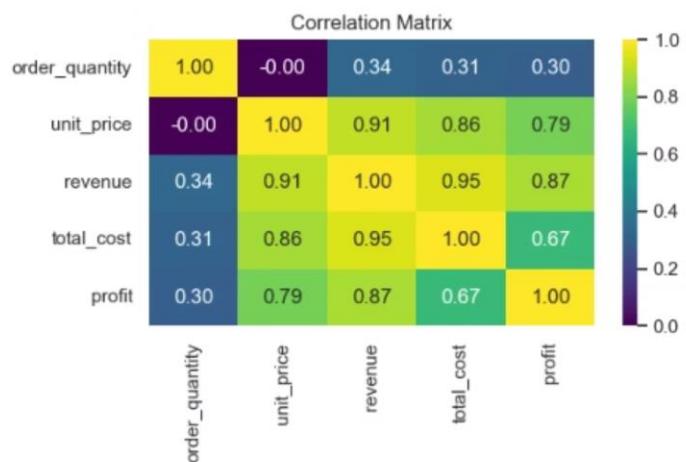
## 7: Customer Segmentation (Revenue vs Profit Margin)



**Insights:**

- Many customers maintain stable margins around 35–40%, reflecting disciplined pricing and cost control.
- Some clients with >\$10M revenue have margins below 36%, suggesting aggressive discounting for large accounts.
- Clients in the \$6–9M range with >40% margins are attractive targets for targeted upsell initiatives.

## 8: Correlation Heatmap of Numeric Features



### Insights:

- Unit price shows very strong correlations with cost (0.94), revenue (0.91), and profit (0.79), indicating pricing is a primary driver.
- Revenue and profit maintain a high correlation (~0.87), reinforcing that revenue growth translates to profitability.
- Quantity has a limited correlation with financials ( $\leq 0.34$ ), so volume plays a secondary role versus price/sku mix.

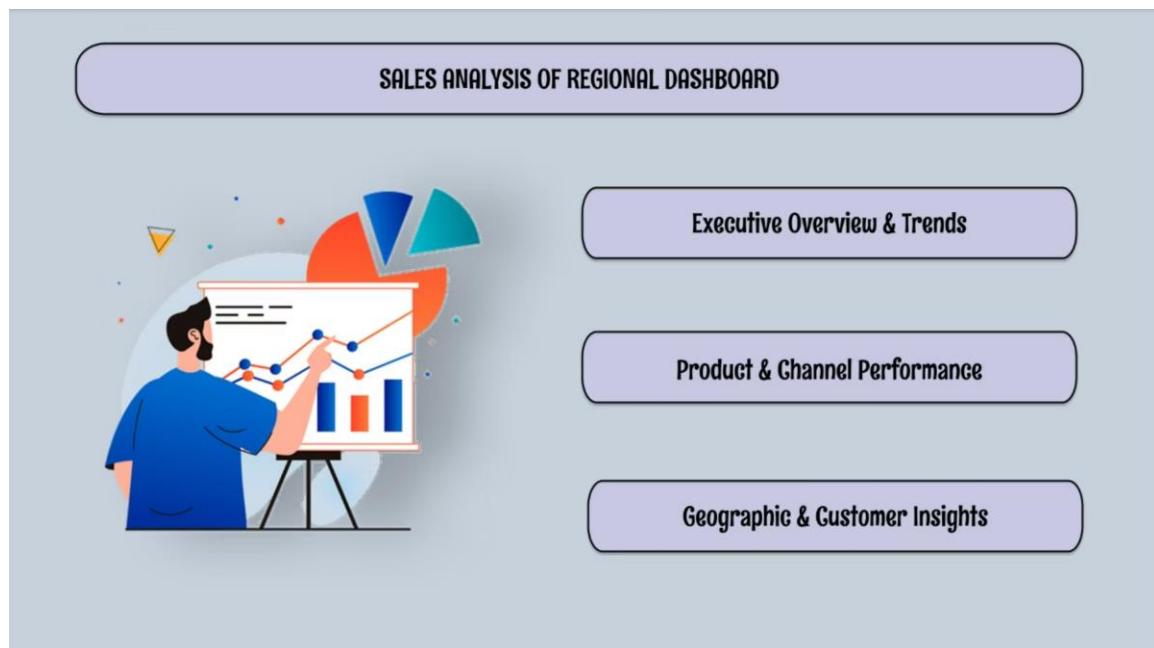
### Key Insights

- Pronounced seasonality: January revenues average \$124M, dipping to \$95M in April.
- SKU concentration: Products 26 & 25 together drive approximately 25% of total sales.
- Channel trade-off: Wholesale captures 54% of volume; Export leads with ~38% average margin (higher margins).
- Geographic dominance: California records ~7.6K orders and ~\$230M in revenue; the West region shows the largest swings.
- High-value customers (Aibox Company & State Ltd) are critical revenue drivers.

### 4. Dashboard in Power BI

Power BI Pages Included (embed screenshots in the document where appropriate):

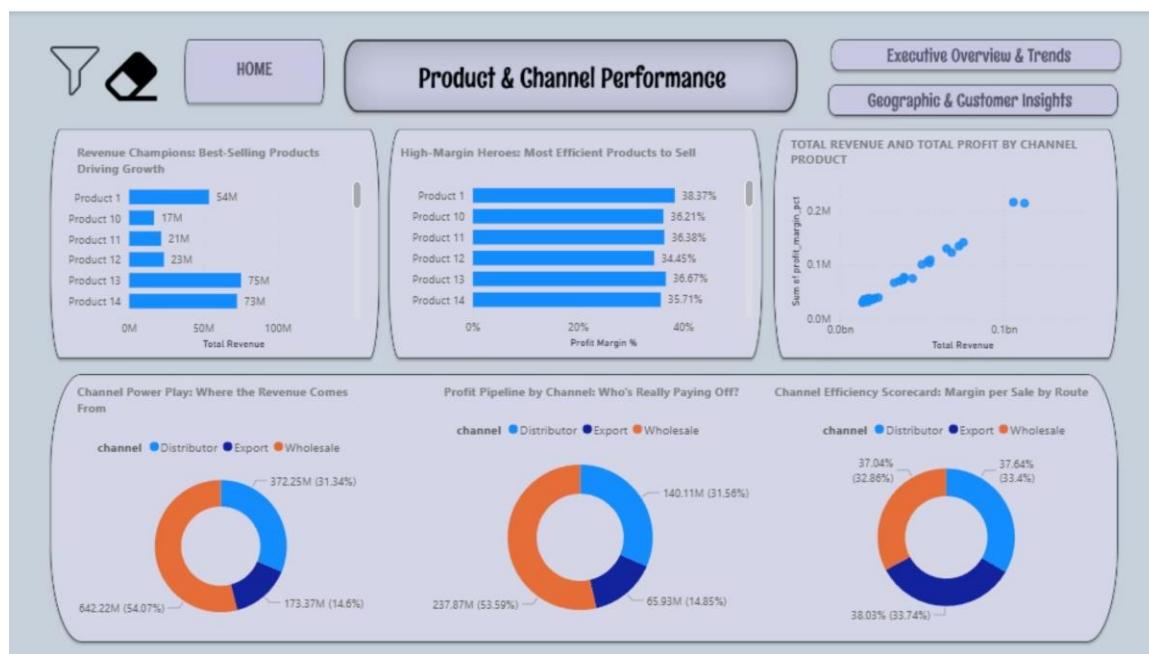
**PAGE 1**- Home Page: Interactive overview with Performance Summary, Customer Segmentation, and Revenue Scenarios:..



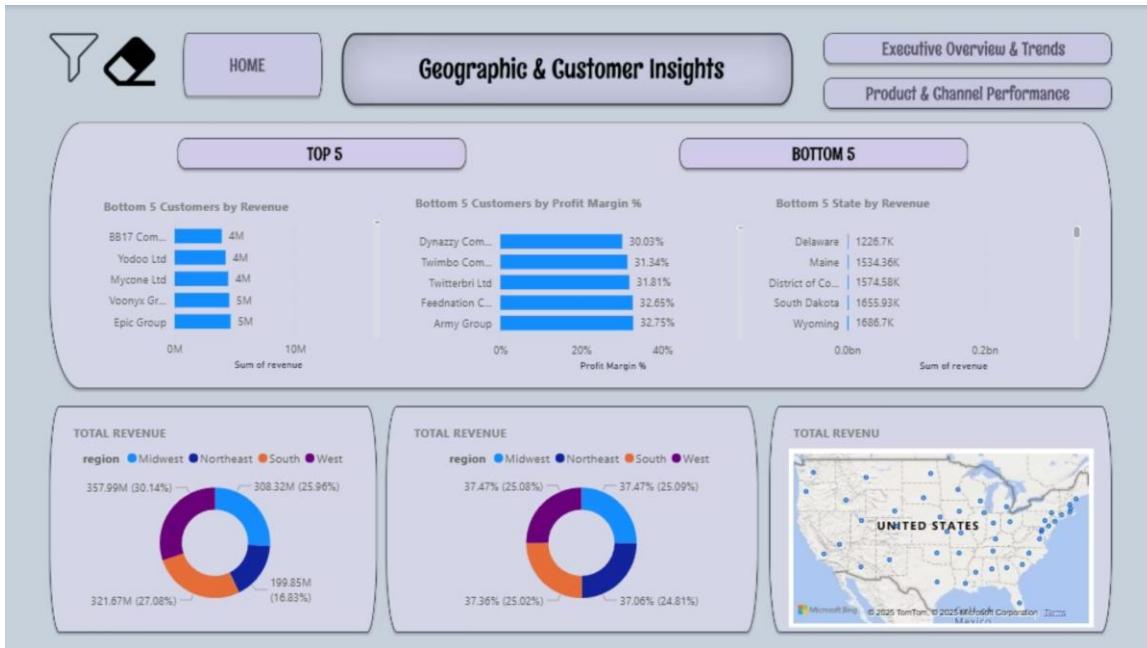
## PAGE 2 - Performance Summary: Monthly and yearly revenue trends, top contributors.



## PAGE 3- Customer Segmentation: Profit vs revenue split and target segments.



## PAGE 4: Revenue Scenarios: What-if analyses and forecasts.



## 5. Business Recommendations

Recommendations and expected business impact:

- Seasonal Promotions: Launch recovery campaigns in April and strengthen January promotions to smooth revenue swings and recover lost volume.
- SKU Optimization: Double down on Products 26 & 25; review pricing and phase out low-performing SKUs to improve assortment efficiency.
- Channel Expansion: Incentivize Export partnerships to capture higher margins and implement volume deals in Wholesale to retain share.
- Regional Investment: Replicate California/West region best-practices (marketing & distribution) in Northeast & Midwest to balance performance.
- Margin Monitoring: Flag orders below 80% target margin for cost analysis and renegotiate terms with low-margin large accounts.

## 6. Conclusions

Conclusions:

- Completed end-to-end EDA and interactive Power BI dashboard, surfacing seasonality, SKU concentration, channel performance, and regional insights.
- Insights support operational planning (inventory, promotions) and strategic decisions

(pricing, channel expansion).

### **Author**

Mukesh Gopi Nandh

Email: mukeshudatha7@gmail.com

LinkedIn: [www.linkedin.com/in/mukesh-gopi-nandh](https://www.linkedin.com/in/mukesh-gopi-nandh)

GitHub: [github.com/Mukeshgn](https://github.com/Mukeshgn)