# Regional Sales Analysis

# Agenda

- Problem Statement
- Approach
- Data Overview
- Project Workflow
- Exploratory Data Analysis
- Key Insights
- Recommendation

# Problem Statement

# Problem Statement

Sales teams often lack a clear, data-driven understanding of regional performance, making it difficult to identify growth opportunities and optimize resources. This project aims to analyse and visualize regional sales data to uncover trends, evaluate profitability, and support strategic decision-making.

## What's the Business Question?

- Inconsistent revenue and profit performance across U.S. regions
- Lack of visibility into seasonal swings, top SKUs, and channel profitability
- Goal: Leverage 5 years of historical data to pinpoint growth levers and optimize strategy

# Approach

# Approach – EDA

**Exploratory Data Analysis**

- Dive into historical sales, margins, products, channels, regions
- Surface trend, outliers & relationships

# Project Workflow

# Project Workflow

1

## Define Business Objective

Understand the core problem and expected business outcomes.

2

## Collect & Consolidate Data

Gather multi-source sales data (Excel sheets) and understand schema.

3

## Data Loading & Initial Exploration

Load into Jupyter Notebook for initial profiling and data understanding using Python.

4

## Pre-processing & Cleaning

Handle nulls, join tables, format dates, and normalize columns.

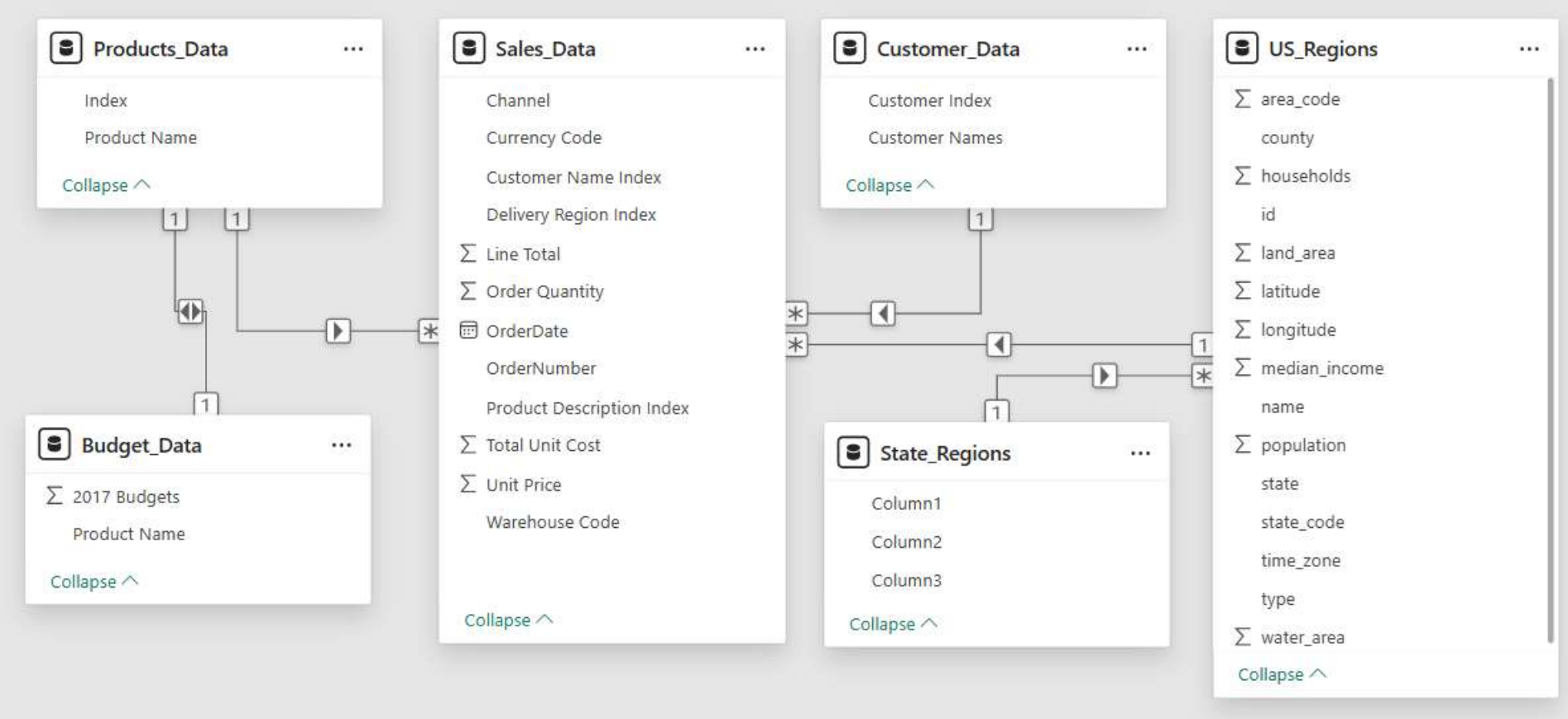
5

## Exploratory Data Analysis (EDA)

Visualize trends, compare performance, and extract key insights.

# Exploratory Data Analysis

# Raw Dataset Structure – Before Processing




Sales, products, budgets, customers, regions, and states were spread across unlinked tables. No relationships were defined initially— Pre-processing was required to clean, normalize, and join them for analysis.

# Exploratory Data Analysis (EDA)

Uncovering patterns, trends, and business insights from historical data !

Understanding the “What, Where & Why” behind the sales numbers

Exploring data through visuals, aggregations, and comparisons

Laying the groundwork for informed recommendations

Let’s Dive In

# EDA: Step-by-Step Process

- Import Libraries
- Load Data
- Initial Exploration
- Pre-processing & Cleaning
- Feature Engineering
- EDA & Visualization
- Key Insights
- Recommendations

# Pre-processing & Feature Engineering

The necessary steps applied to prepare this dataset for analysis.

- Set header row for state – region table
- Merge Sales, Customers, Products, Regions, State–Region & Budgets tables
- Drop redundant columns
- Standardize column names to lowercase
- Select key columns that are used for that analysis
- Rename columns to more sensible names
- Create profit and profit_margin_pct columns

✓*Note: No missing values or duplicate rows were found in the dataset*

# Final Dataset Structure – Ready for Analysis

| | order_number | order_date | customer_name | channel | product_name | quantity | unit_price | revenue | cost | state | state_name | us_region | lat | lon | budget | profit | profit_margin_pct |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | SO - 000225 | 2014-01-01 | Rhynoodle Ltd | Wholesale | Product 27 | 6 | 2499.1 | 14994.6 | 1824.343 | GA | Georgia | South | 32.08354 | -81.09983 | NaN | 13170.257 | 87.833333 |
| 1 | SO - 0003378 | 2014-01-01 | Thoughtmix Ltd | Distributor | Product 20 | 11 | 2351.7 | 25868.7 | 1269.918 | IN | Indiana | Midwest | 39.61366 | -86.10665 | NaN | 24598.782 | 95.090909 |
| 2 | SO - 0005126 | 2014-01-01 | Amerisourc Corp | Wholesale | Product 26 | 6 | 978.2 | 5869.2 | 684.740 | CA | California | West | 37.66243 | -121.87468 | NaN | 5184.460 | 88.333333 |
| 3 | SO - 0005614 | 2014-01-01 | Colgate-Pa Group | Export | Product 7 | 7 | 2338.3 | 16368.1 | 1028.852 | IN | Indiana | Midwest | 39.16533 | -86.52639 | NaN | 15339.248 | 93.714286 |
| 4 | SO - 0005781 | 2014-01-01 | Deseret Group | Wholesale | Product 8 | 8 | 2291.4 | 18331.2 | 1260.270 | CT | Connecticut | Northeast | 41.77524 | -72.52443 | NaN | 17070.930 | 93.125000 |

Identifiers: order_number, order_date, customer_name, channel, product_name

Financials: quantity, unit_price, revenue, cost, profit, profit_margin_pct

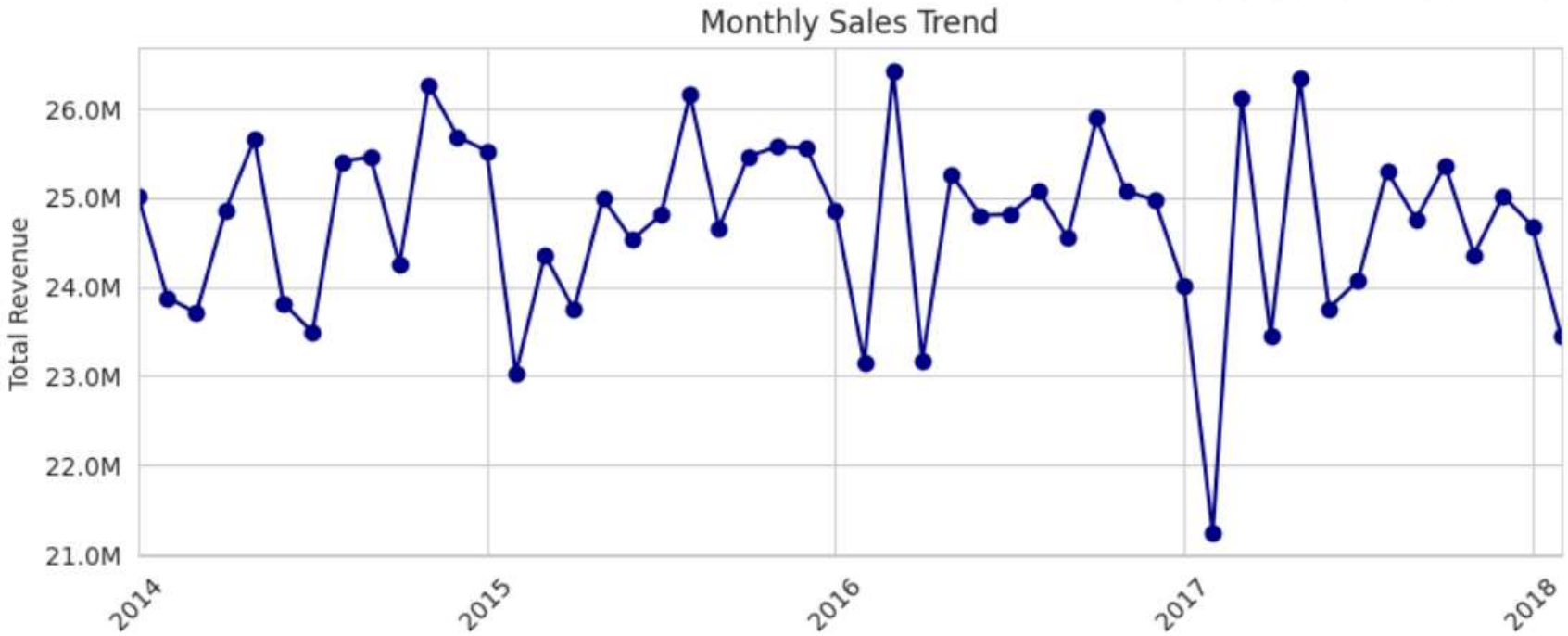
Calendar:  order_month_name, order_month_num, order_month

Geography: state (code), state_name, us_region, lat, lon

Planning: budget (2017)

# Charts & Insights

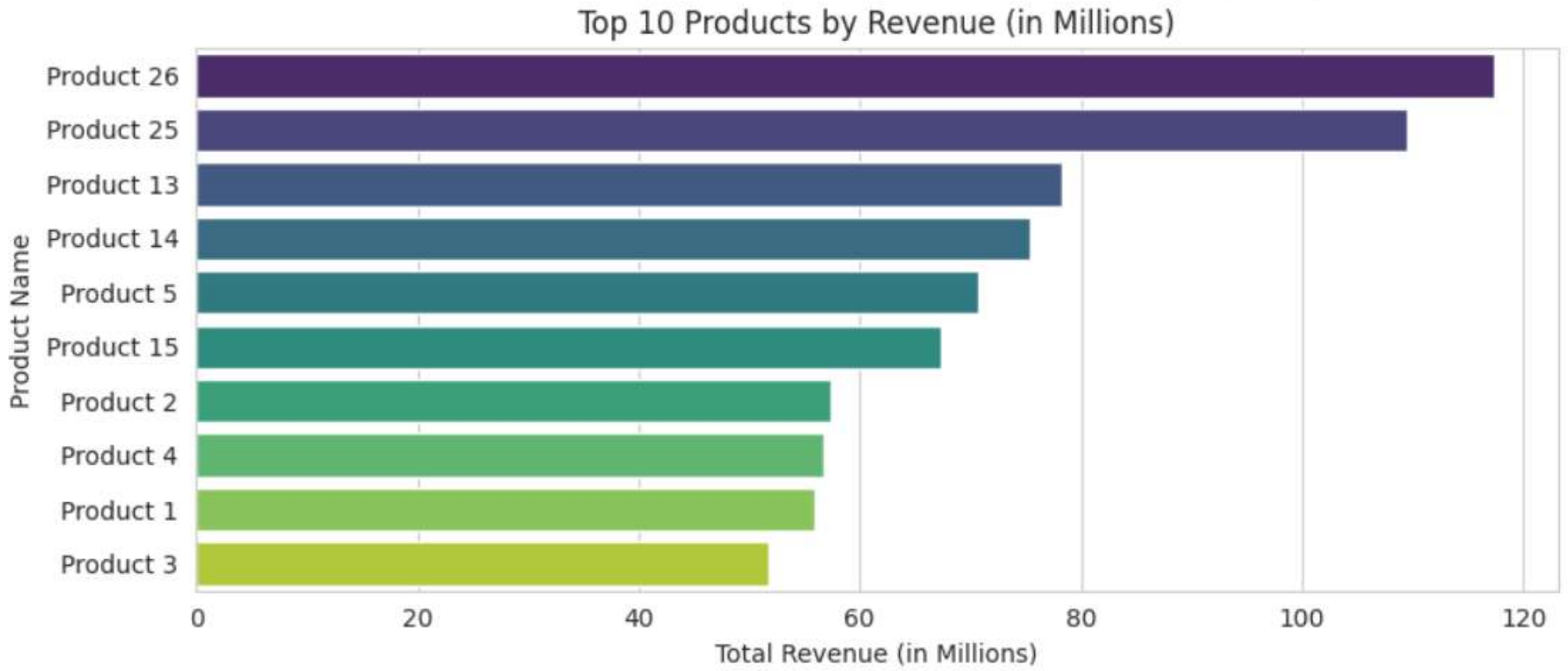
# Monthly Sales Trend Over Time




## Insights

- Consistent sales cycle: $24M to $26M.
- Seasonal peaks: Late spring/early summer (May-June).
- Annual low: January.
- Notable outlier: Sharp revenue drop in early 2017.

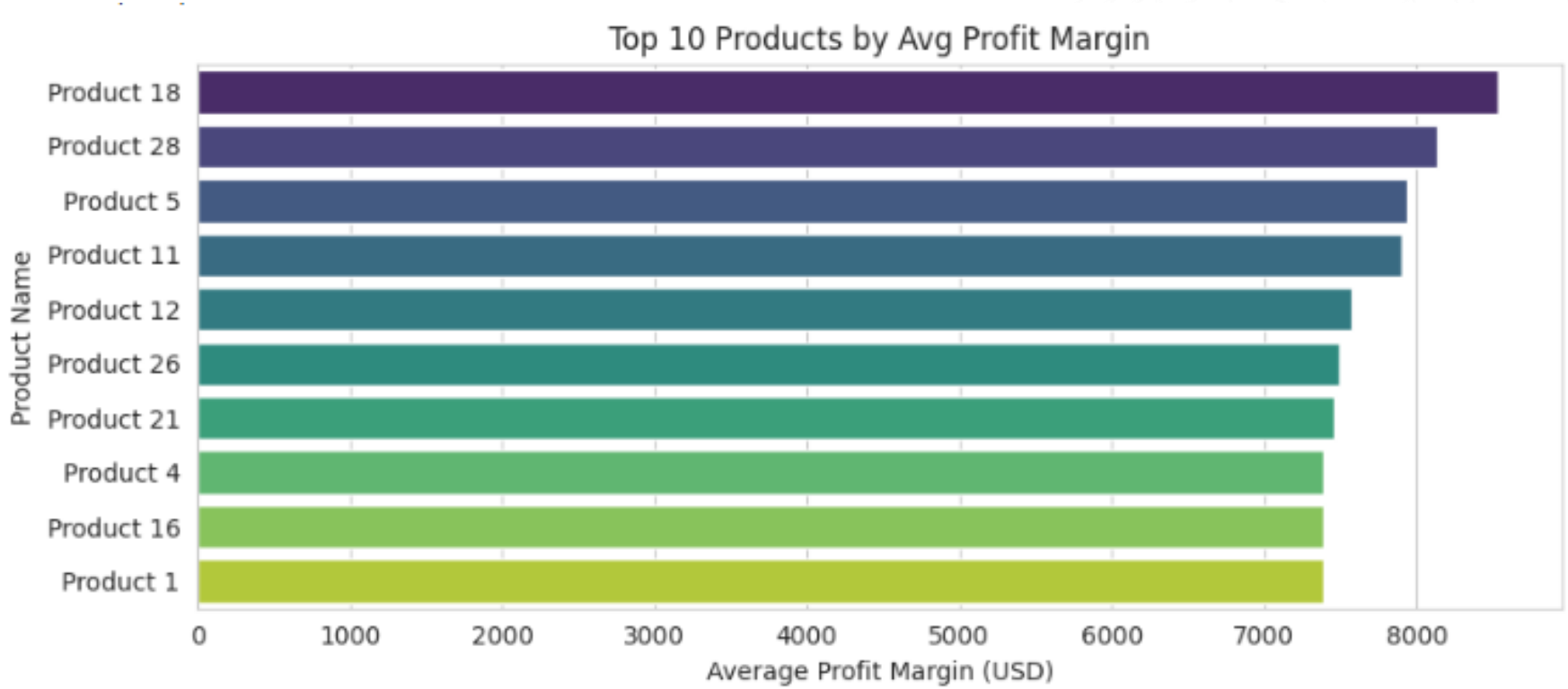
# Top 10 Products by Revenue




## Insights

- Revenue leaders: Products 26 & 25 dominate.
- Mid-range: Products 5, 13, 14, 15 show similar revenue.
- Bottom cluster: Products 1, 2, 3, 4 have the lowest revenue.
- Strategy: Grow mid-tier, improve lower performers.

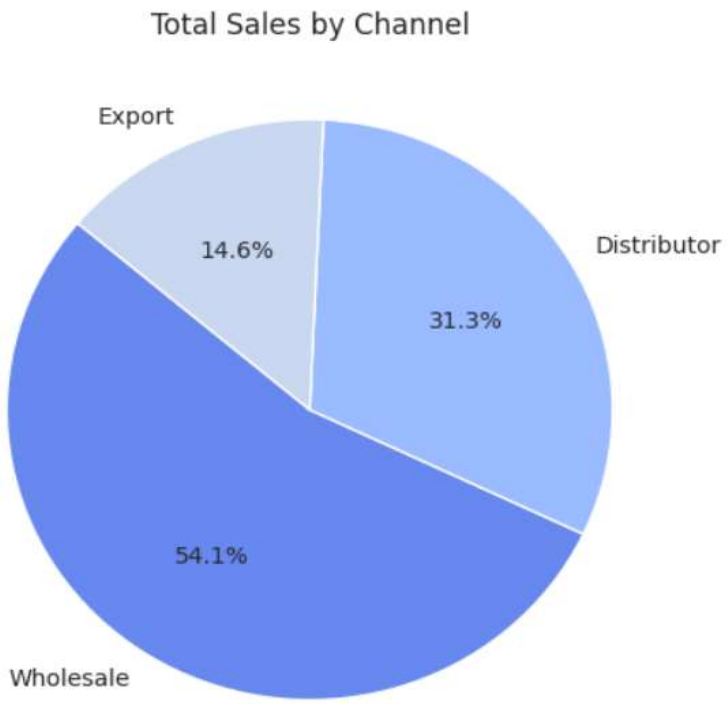
# Top 10 Products by Average Profit Margin




## Insights

- Top: Products 18 & 28 lead at ~$8.1–8.4K.
- Next tier: Products 5, 11, 12 & 26 at ~$7.5–7.8K.
- Entry-level: Products 1, 4, 16 & 21 around $7.3K.
- Takeaway: Top 10 all exceed $7.3K—consistent high margins.

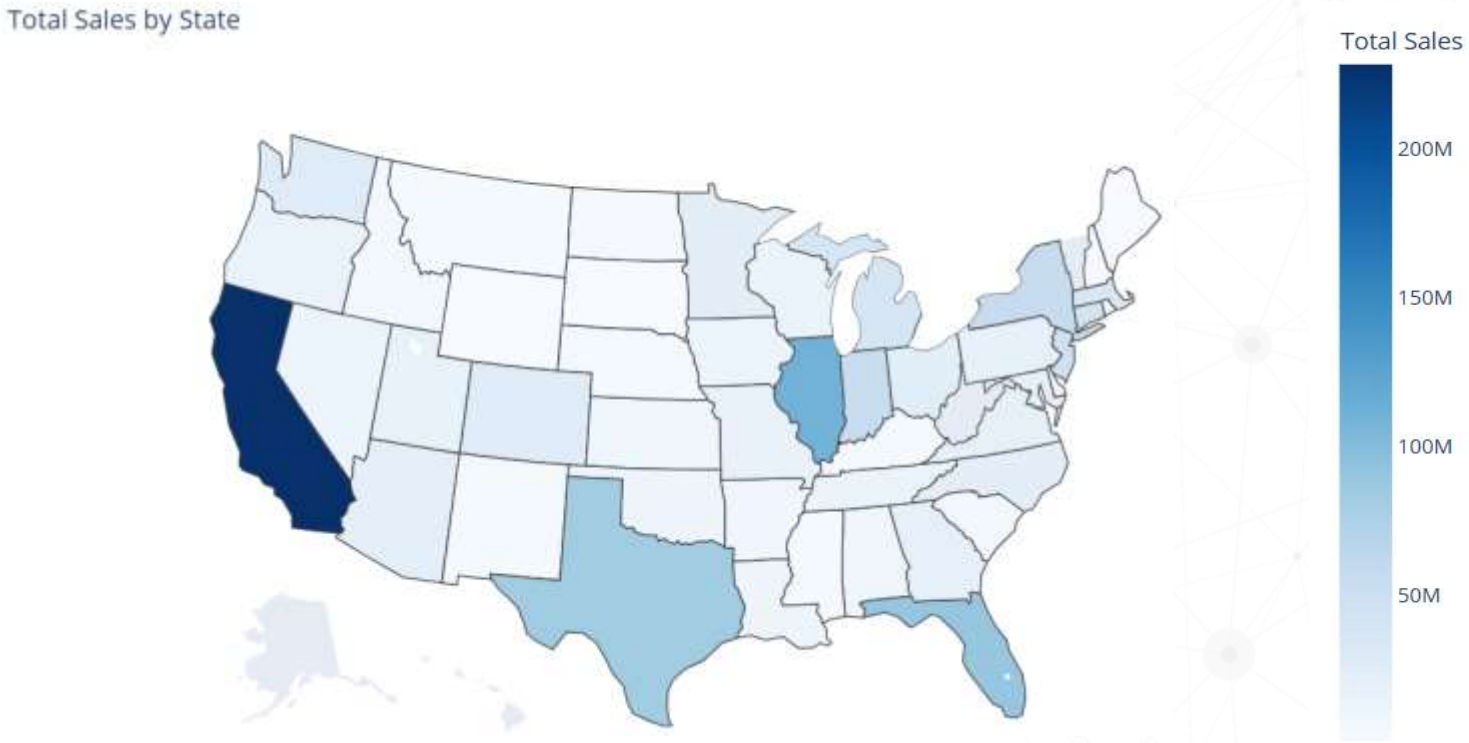
# Sales by Channel




## Insights

- Wholesale dominates: Generates the majority of total sales at 54.1%.
- Distributor is significant: Contributes a substantial 31.3% to total sales.
- Export is a smaller portion: Accounts for 14.6% of the total sales.

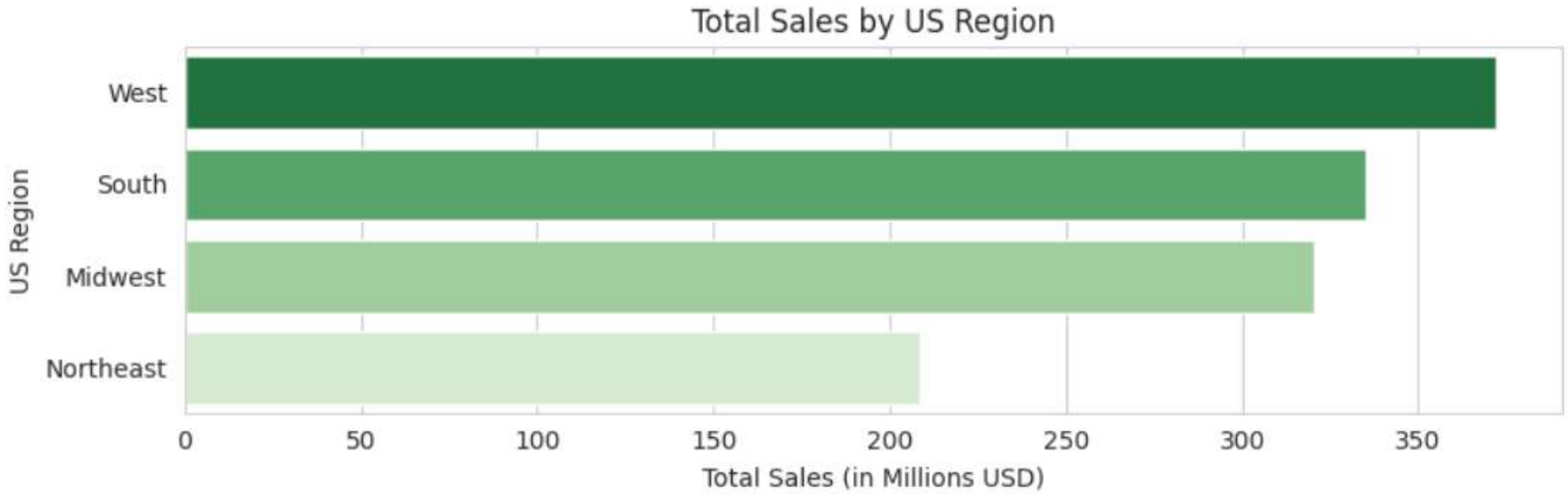
# Total Sales by State (Choropleth Map)




## Insights

- California leads: Highest total sales.
- Texas, Florida, Illinois: Significant sales.
- Varying sales: Other states show moderate to low sales.
- Visual pattern: Higher sales in larger and some coastal states.

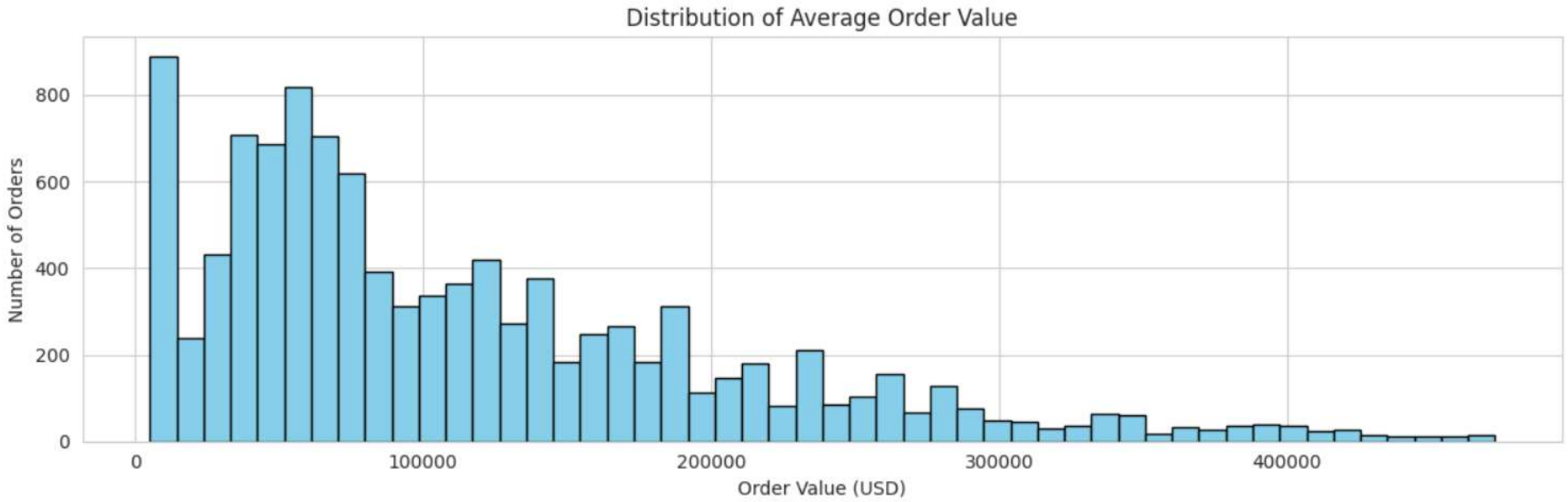
# Total Sales by US Region




## Insights

- West: Highest sales, strong market influence.
- South: Major sales contributor, key market area.
- Midwest: Steady sales performance, moderate market size.
- Northeast: Lowest sales, suggests need for deeper market understanding.

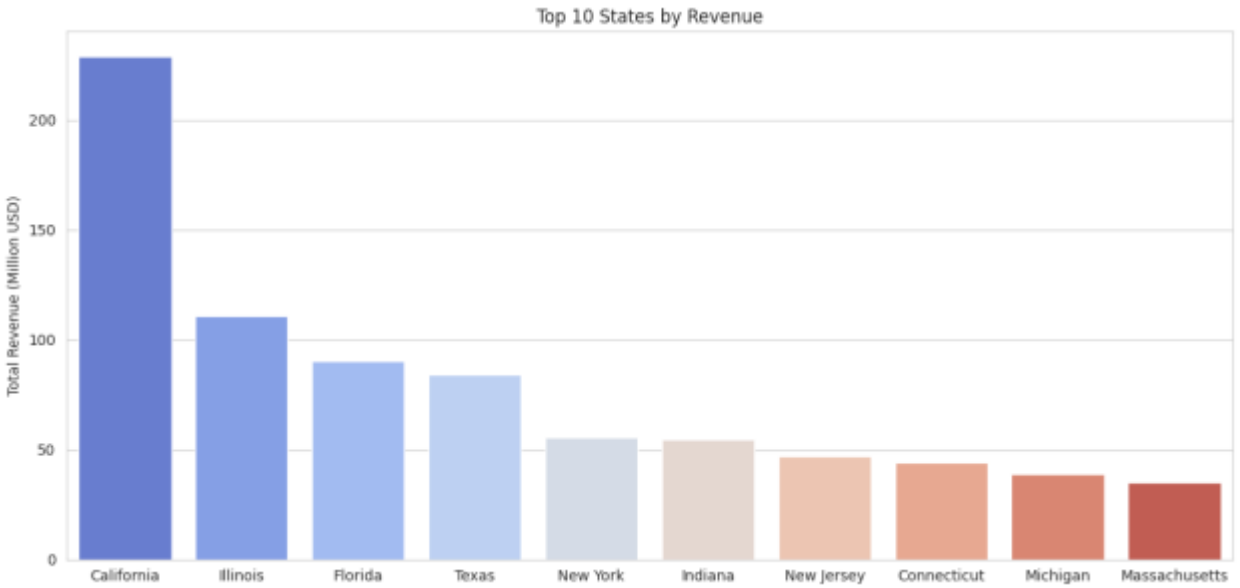
# Average Order Value (AOV) Distribution

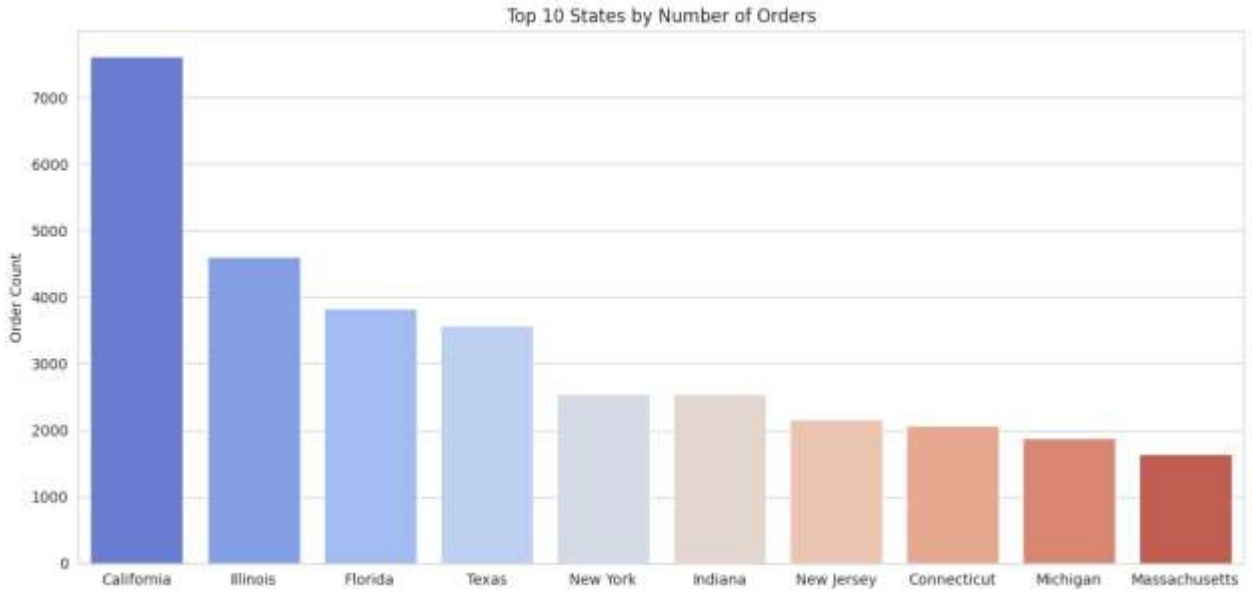



## Insights

- Low average order values are frequent.
- Distribution is right-skewed (long tail of high-value orders).
- Multiple order value clusters exist.
- Higher order values are less common.

# Top State Performance: Revenue vs. Orders

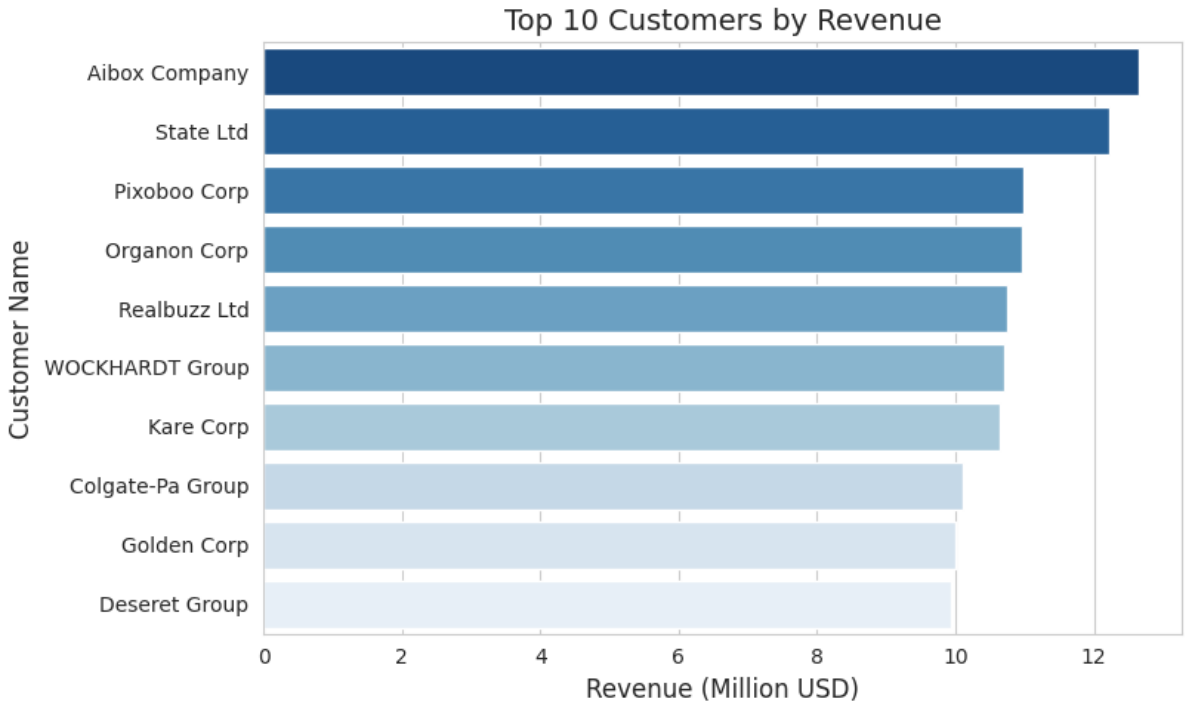


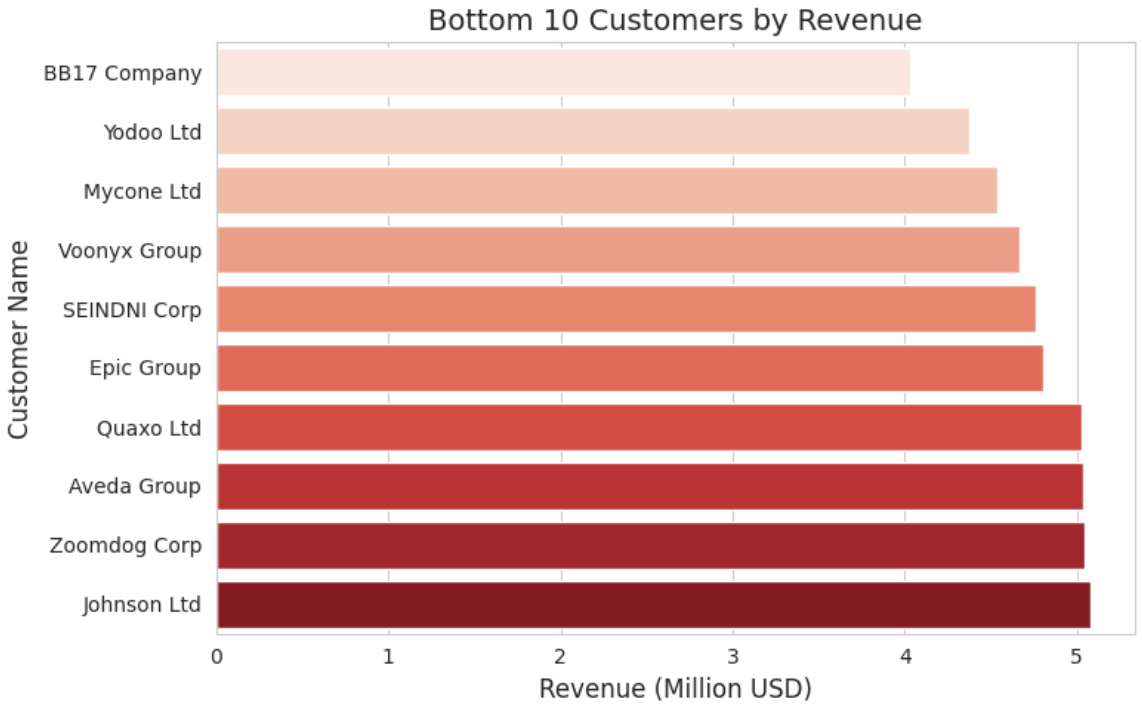




## Insights

- California tops revenue & orders.
- IL, FL, TX: High in both.
- Revenue & orders linked.
- Other top states: Lower contribution.

# Top and Bottom 10 Customers by Revenue




## Insights

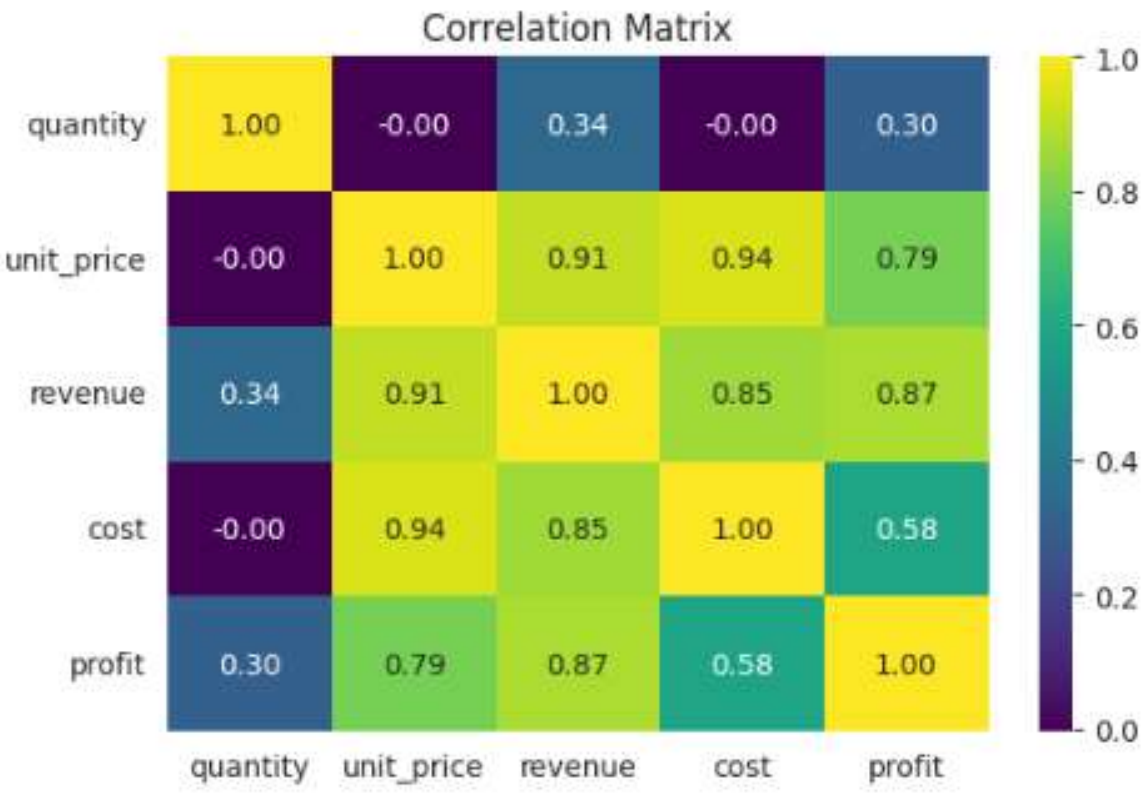
- Aibox Company leads significantly as the top revenue generator.
- Bottom 10 customers generate substantially less revenue (around $4-5M).
- Revenue concentration: Top customers drive a disproportionate share.
- Large gap: Exists between the revenue of top and bottom tier customers.

# Customer Segmentation: Revenue vs. Profit Margin


## Insights

- Those Uniform 35–40 % margins confirm strong, consistent pricing and cost control.
- >$10 M clients with <36 % margins reveal discounting hotspots—re evaluate large-account terms.
- $6–9 M clients with >40 % margins are high-value candidates for targeted upsell.

# Correlation Heatmap of Numeric Features




## Insights

- Unit price is the primary driver, showing very strong correlations with cost (0.94), revenue (0.91) and profit (0.79).
- Revenue & profit maintain a high link (0.87), underscoring direct profitability gains.
- Quantity's impact is minimal ($\leq$ 0.34 vs. financials), indicating volume plays a secondary role.
- Cost vs. profit correlation (0.58) is moderate, suggesting margin improvement focus should center on pricing.

# Key Insights

# Key Insights 📊

- Pronounced Seasonality: January revenues average $124 M, dipping to $95 M in April.
- SKU Concentration: Products 26 & 25 together drive ~25 % of total sales.
- Channel Trade-Off: Wholesale captures 54 % of volume; Export leads with ~38 % average margin.
- Geographic Dominance: California alone logs 7.6K orders ($230 M); the West region shows the largest swings.
- Aibox Company and State Ltd are the most valuable customers in terms of Revenue.

# Recommendations

# Recommendations

- Seasonal Promotions: Launch recovery campaigns in April and amplify January offers to smooth revenue swings.
- SKU Optimization: Double down on top products 26 & 25 and re-evaluate pricing or phase out low-margin SKUs.
- Channel Expansion: Incentivize Export partnerships for high margins and introduce volume deals in Wholesale.
- Regional Investment: Replicate California's success in other regions and boost marketing in the Northeast & Midwest.
- Margin Monitoring: Flag orders below 80 % margin and analyse cost drivers to uplift underperforming segments.

# Thank You