Regional Sales Analysis Report

Agenda

Problem Statement | Approach | Data Overview | Project Workflow | Exploratory Data Analysis | Key Insights | Recommendation | Dashboard Preview

Exploratory Data Analysis

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

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

- 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

Final Dataset Structure – Ready for Analysis

```python
df.head()
```

Python

|	ordernumber	orderdate	customer names	channel	product name	order quantity	unit price	line total	total unit cost	state_code	...	state	region	latitude	longitude	2017 budgets	tot
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	109
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	139
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	41
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	72
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	100

5 rows × 21 columns

- 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)

1. Monthly Sales Trend Over Time
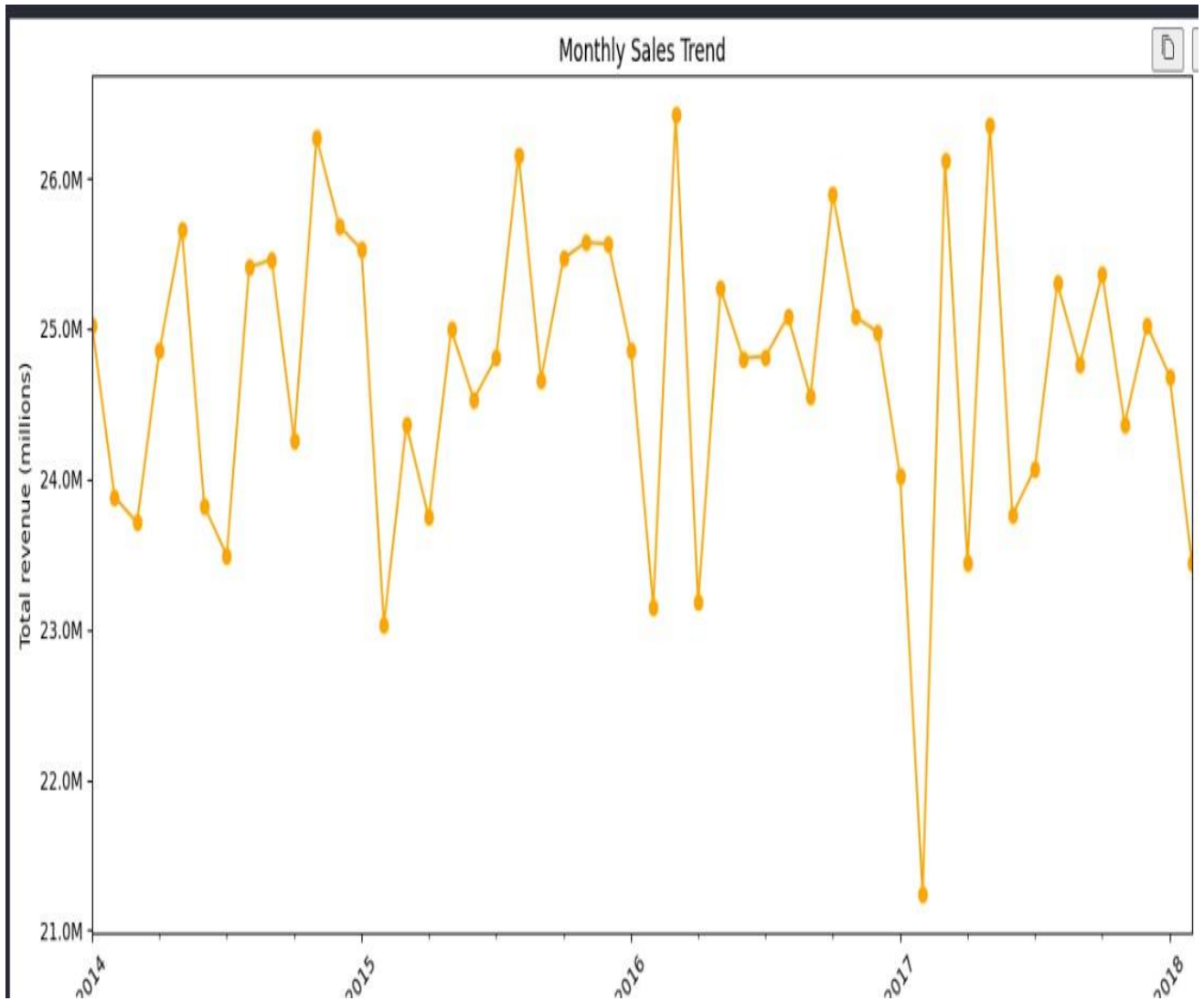
- Consistent sales cycle: $24M to $26M.

- Seasonal peaks: Late spring/early summer (May-June).

- Annual low: January.

- Notable outlier: Sharp revenue drop in early 2017.

2. Top 10 Products by Revenue

Top 10 Products by Revenue (in Millions)
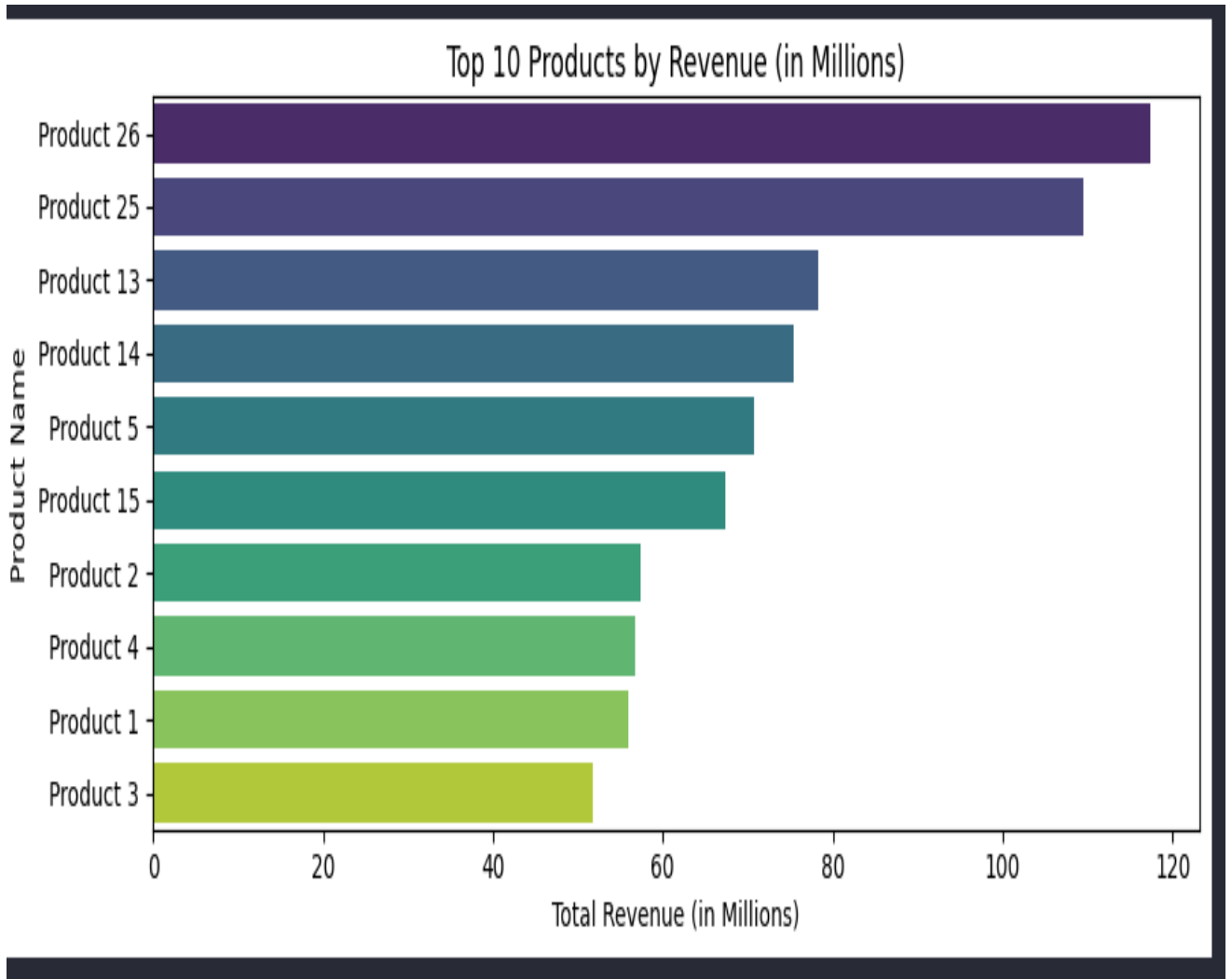
- Revenue leaders: Products 26 & 25 dominate.

- 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

3. Sales by Channel

Sales Distribution by Channel

Wholesale

54.1%

Export

14.6%

31.3%
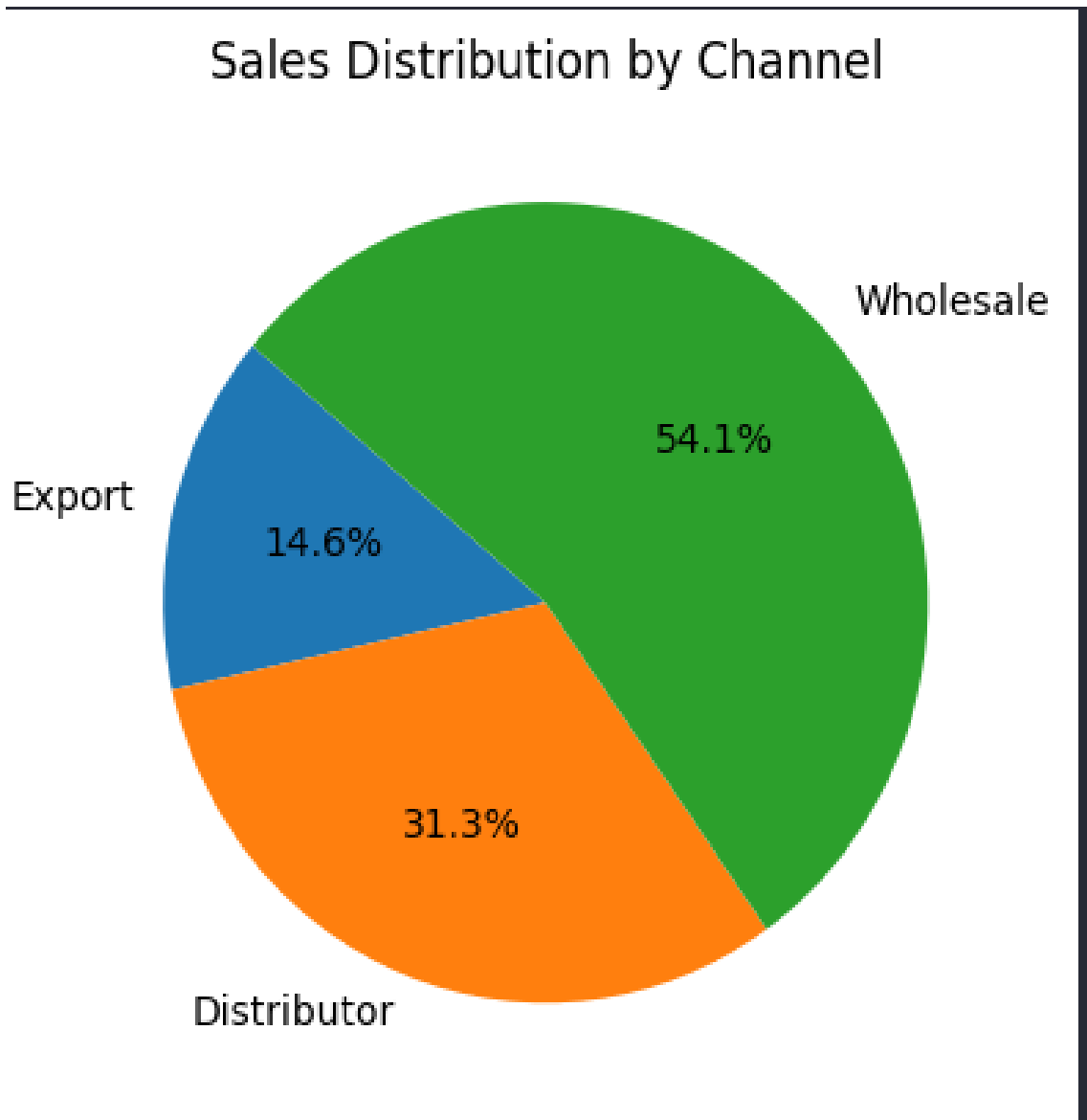
Distributor

- 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.

4. Total Sales by State (Choropleth Map)
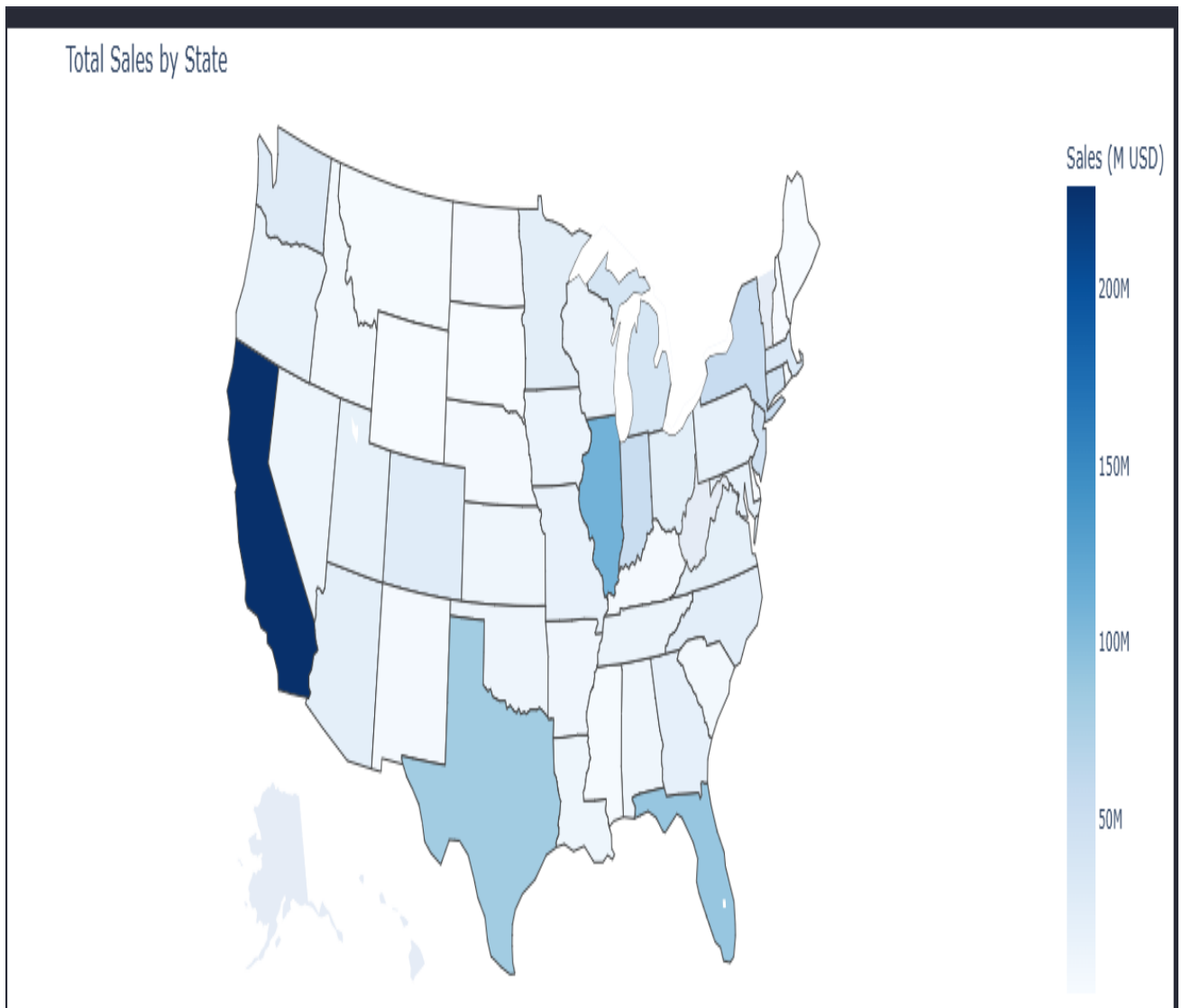
- California leads: Highest total sales.

- Texas, Florida, Illinois: Significant sales.

- Varying sales: Other states show moderate to low sales.

- Visual pattern: Higher sales in larger and some coastal states.

5. **Total Sales by US Region**

Total Sales by US Region
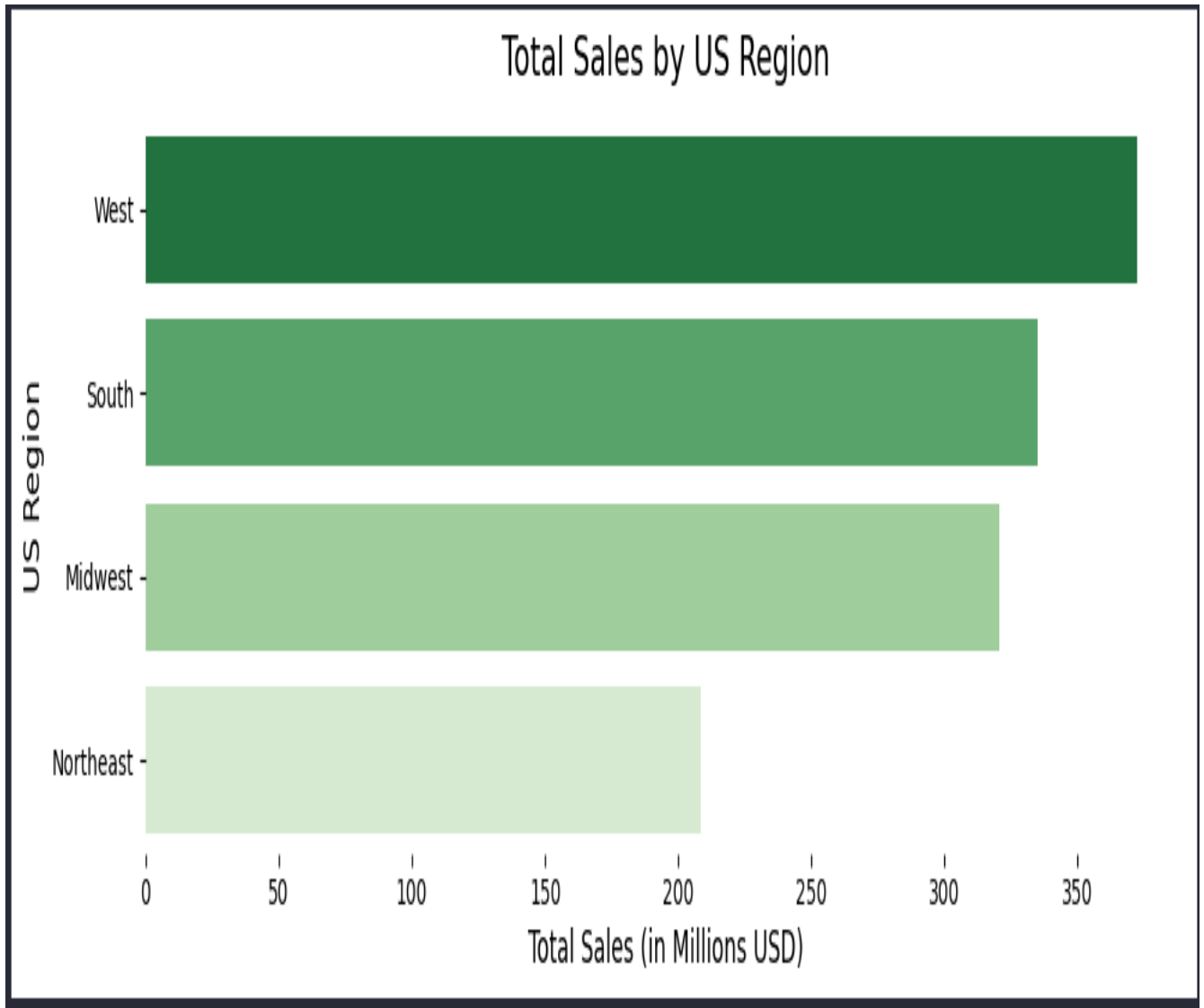
- West: Highest sales, strong market influence.

- South: Major sales contributor, key market area.

- Midwest: Steady sales performance, moderate market size.

- Northeast: Lowest sales, suggests need for deeper market understanding.

6. Average Order Value (AOV) Distribution
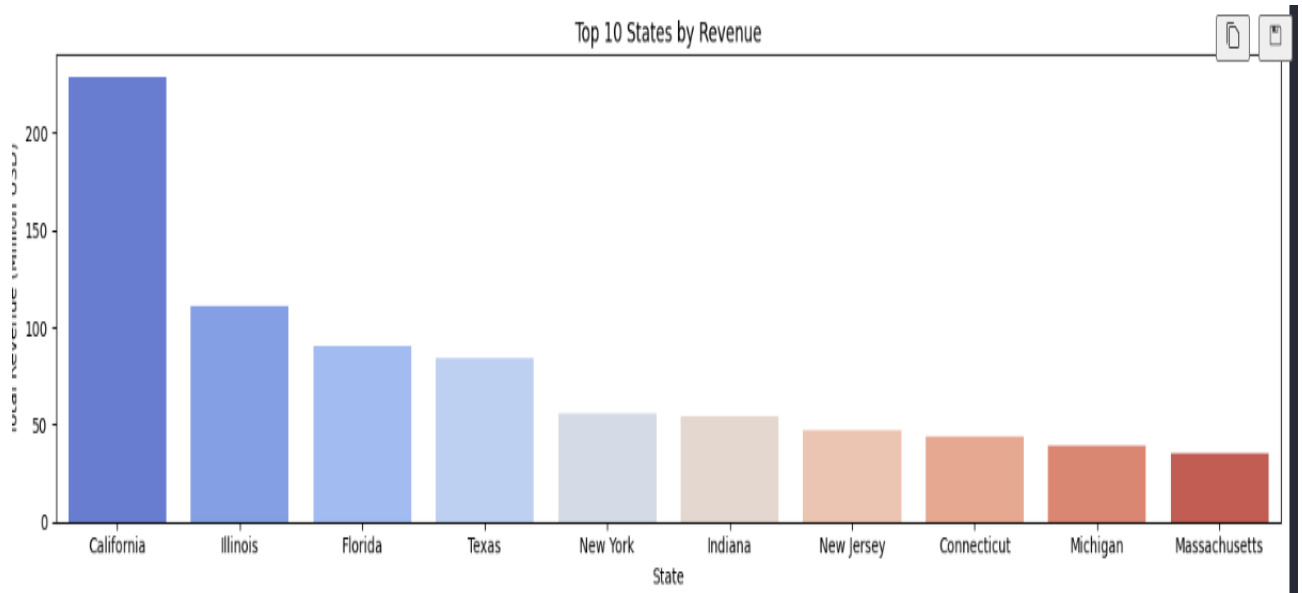
Average Order Value Distribution
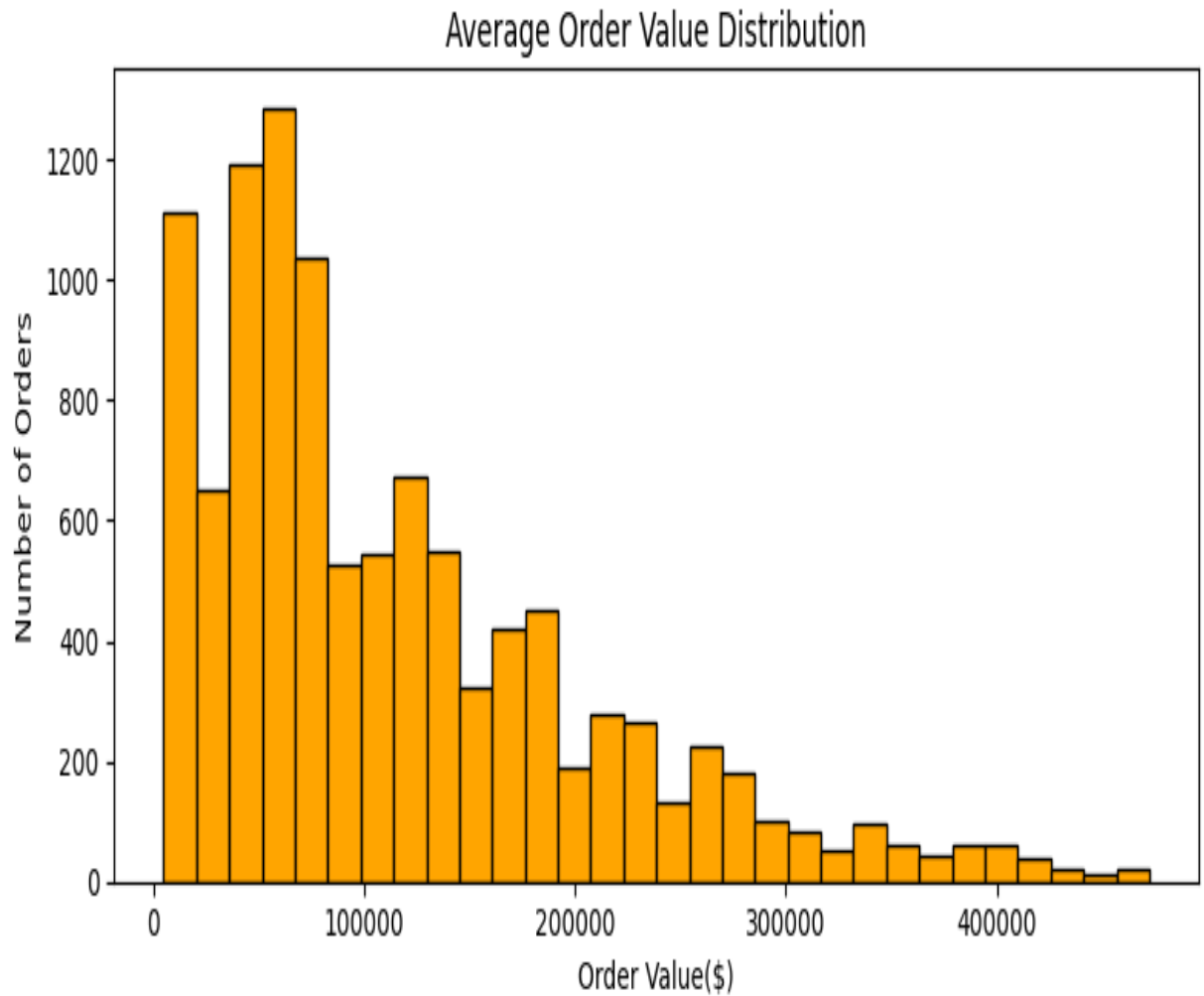
- Low average order values are frequent.

- Distribution is right-skewed (long tail of high-value orders).

- Multiple order value clusters exist.

- Higher order values are less common
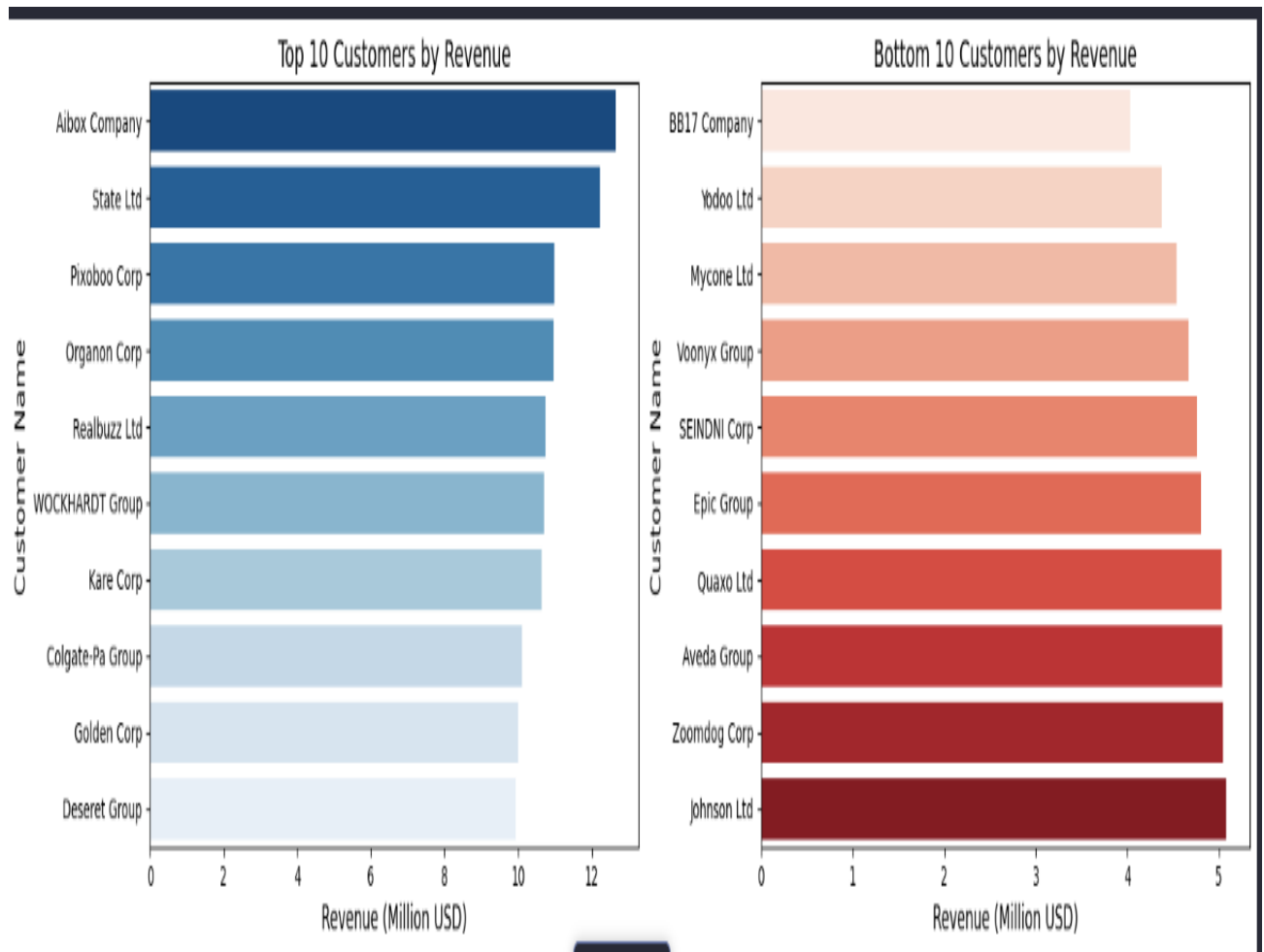
7. Top State Performance: Revenue vs. Orders

Top 10 States by Revenue

Top 10 States by Number of Orders
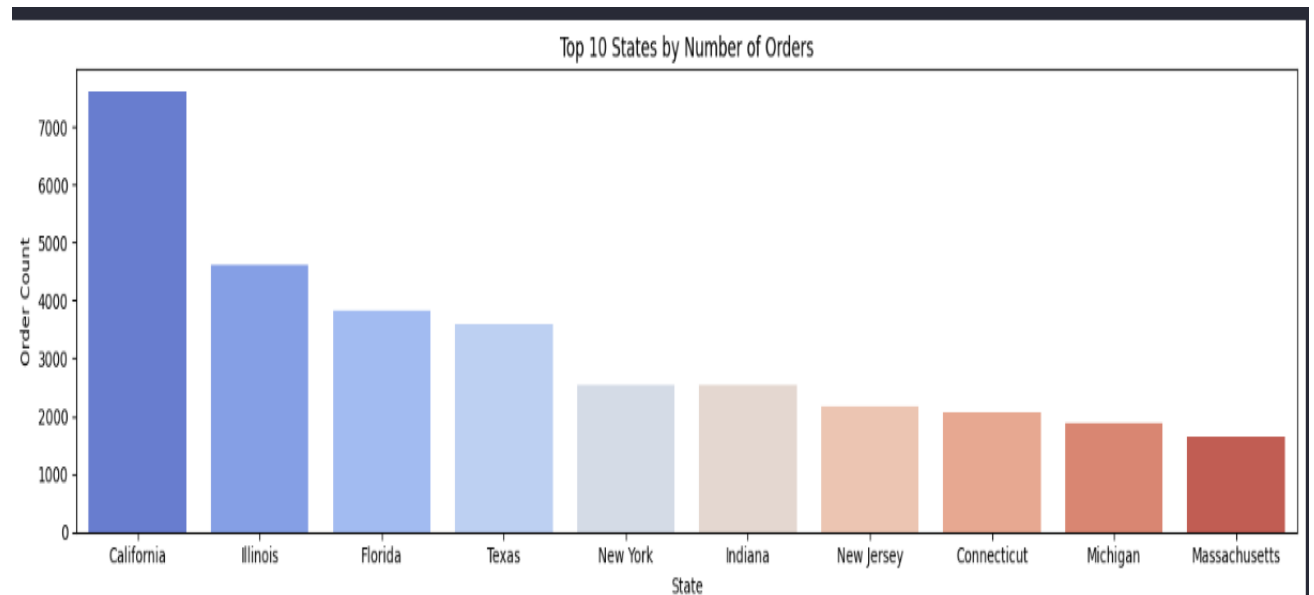
- California tops revenue & orders.

- IL, FL, TX: High in both.

- Revenue & orders linked.

- Other top states: Lower contribution.

8. Top and Bottom 10 Customers by Revenue

- 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

9. Customer Segmentation: Revenue vs. Profit Margin
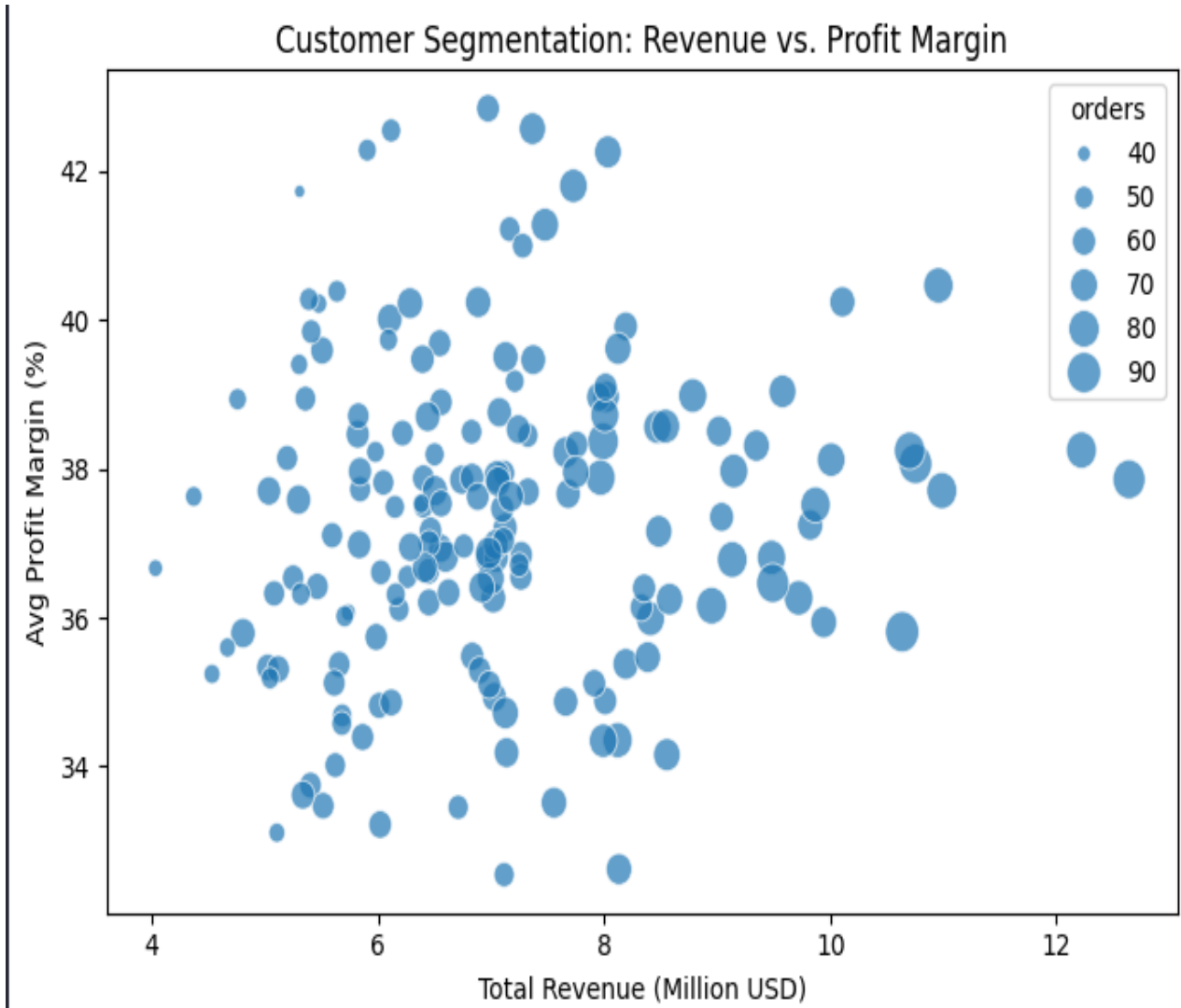
- Those Uniform 35–40 % margins confirm strong, consistent pricing and cost control.

- >$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.

10.   Correlation Heatmap of Numeric Features

Correlation Matrix
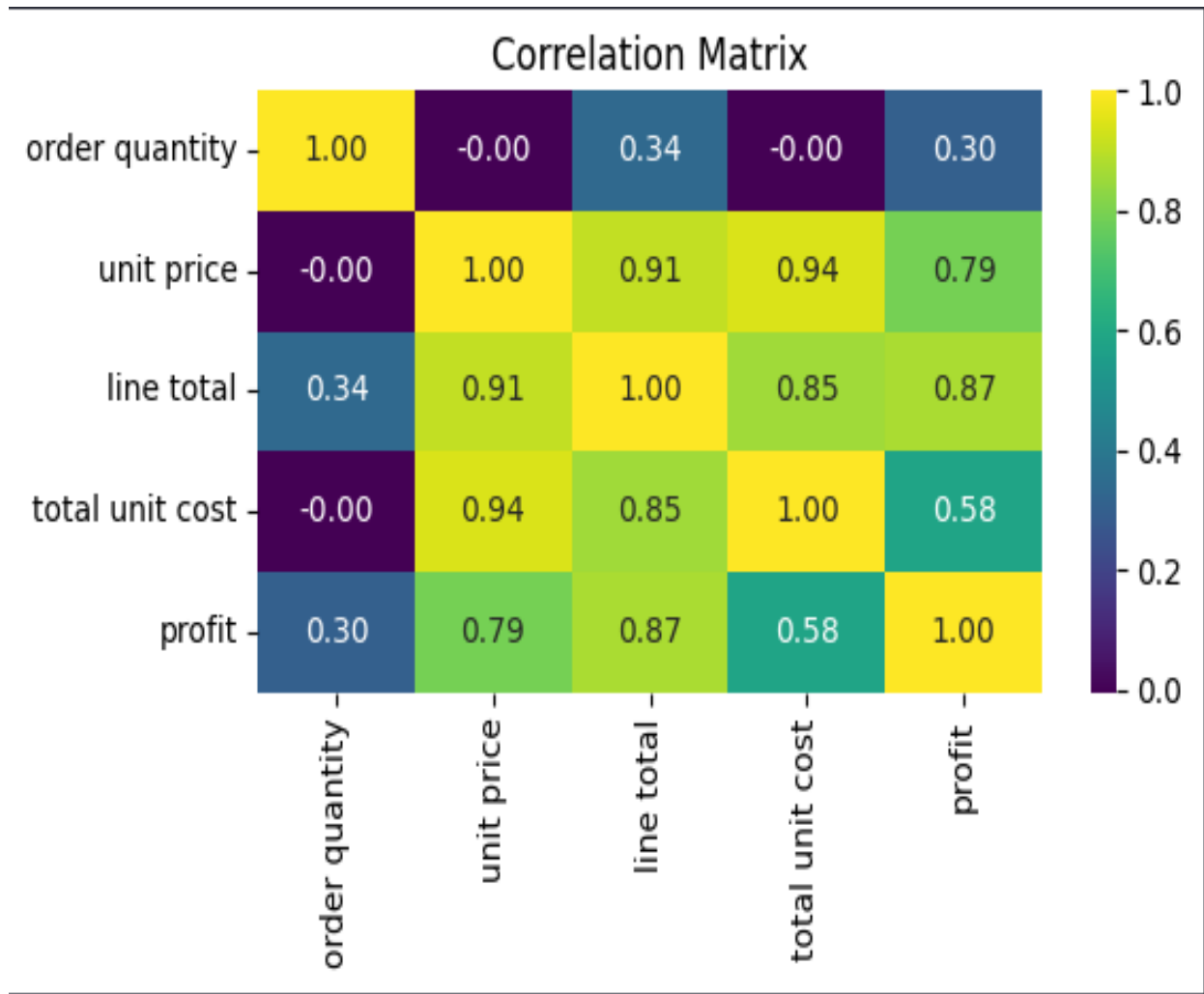
- Unit price is the primary driver, showing very strong correlations with cost (0.94), revenue (0.91) and profit (0.79).

- Revenue & profit maintain a high link (0.87), underscoring direct profitability gains.

- Quantity's impact is minimal (≤ 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.

Dashboard Preview

Page 1 - Performance Summary

Page 2 - Customer Segmentation

Executive Overview & Trends | **Products & Channel Performance** | **Geographic & Customer Insights**

Revenue Champions

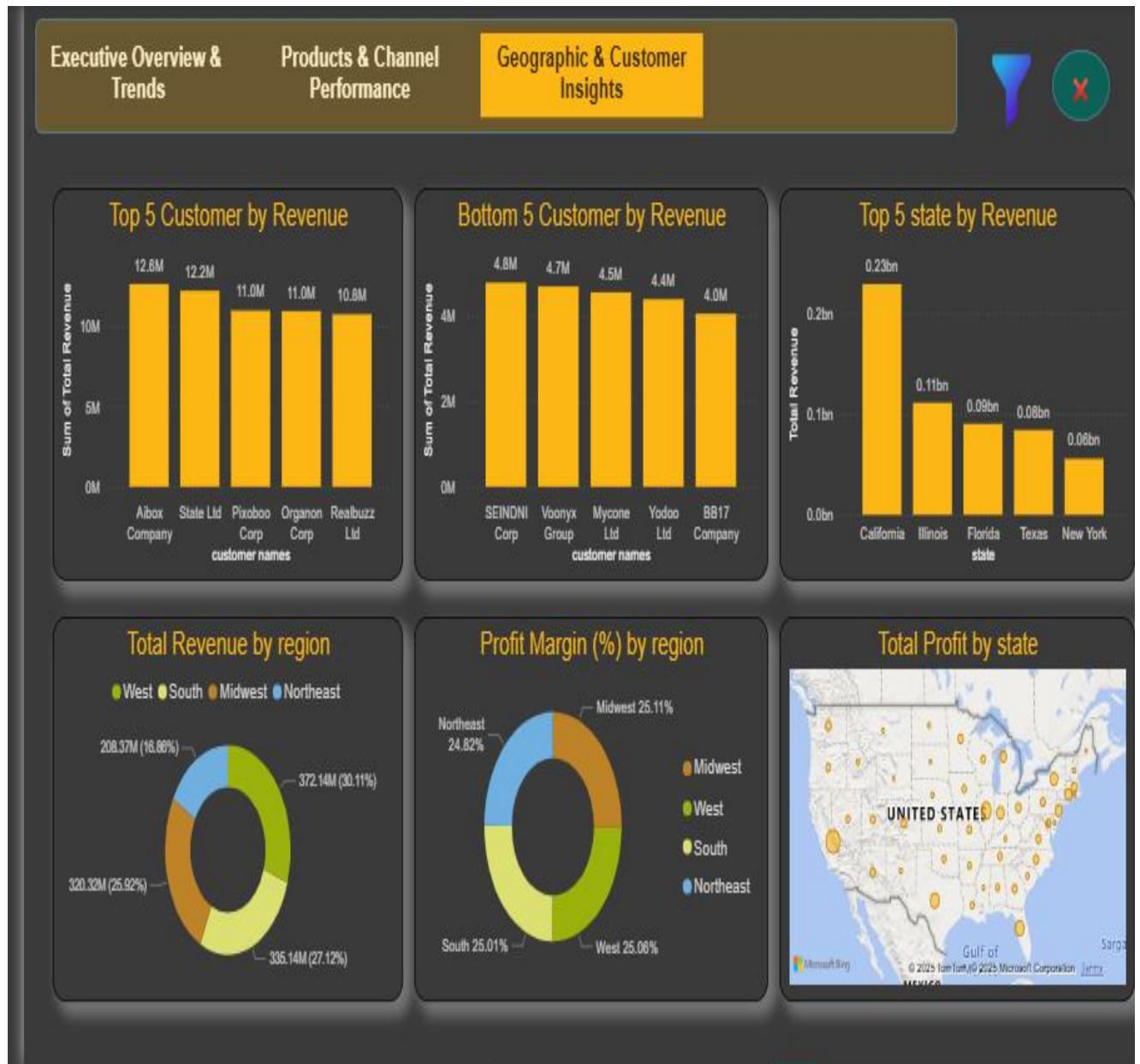| Product	Revenue
Product 26	117M
Product 25	109M
Product 13	78M
Product 14	75M
Product 5	71M
Product 15	67M
Product 2	57M
Product 4	57M
Product 1	56M
Product 3	52M

High Margin Heroes

Product	Value
Product 9	40
Product 30	39
Product 28	39
Product 7	39
Product 21	39
Product 18	38
Product 1	38
Product 16	38
Product 29	38
Product 25	38

Sum of Order Count by Total Revenue and Avg Profit Margin %

Total Revenue by channel
- 180.63M (14.61%)
- 387.14M (31.32%)
- 668.2M (54.06%)

channel
- Wholesale
- Distributor
- Export

Profit Margin (%) by channel
- 37.0... (32...)
- 38.0... (33...)
- 37.6530... (33.42%)

channel
- Export
- Distributor
- Wholesale

Total Profit by channel
- 68.85M (14.87%)
- 145.77M (31.57%)
- 247.35M (53.58%)

channel
- Wholesale
- Distributor
- Export

Page 3 - Revenue Scenarios

Conclusions

- Completed end-to-end EDA and interactive Power BI dashboard, surfacing seasonality, SKU, channel & regional insights.

- Insights inform sales policies and operational planning (store & warehouse prep aligned with annual trends).

- Stakeholders can self-serve real-time analysis and confidently onboard new datasets for additional use cases.

Thank You