**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**

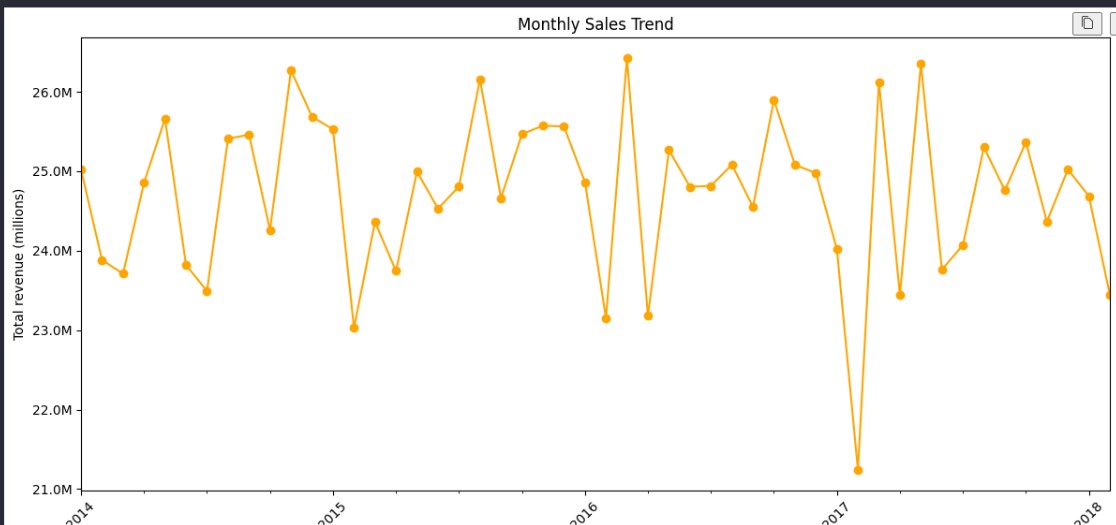
* Set header row for state – region table
* Merge Sales, Customers, Products, Regions, State–Region & Budgets tables
* Drop redundant columns
* Standardize column names to lowercase
* Select key columns that are used for that analysis
* Rename columns to more sensible names
* Create profit and profit\_margin\_pct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

**Final Dataset Structure – Ready for Analysis**



* 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**



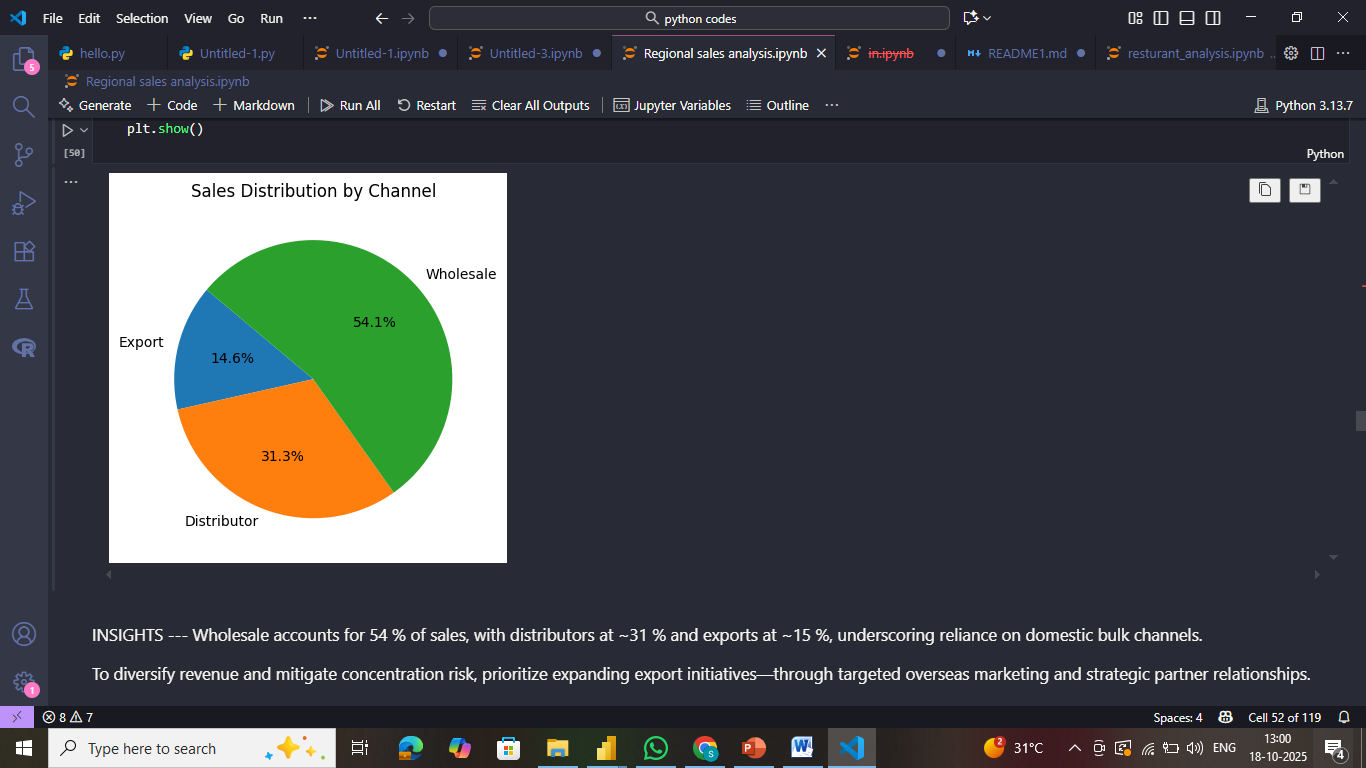
* Consistent sales cycle: $24M to $26M.
* Seasonal peaks: Late spring/early summer (May-June).
* Annual low: January.
* Notable outlier: Sharp revenue drop in early 2017.

1. **Top 10 Products by Revenue**



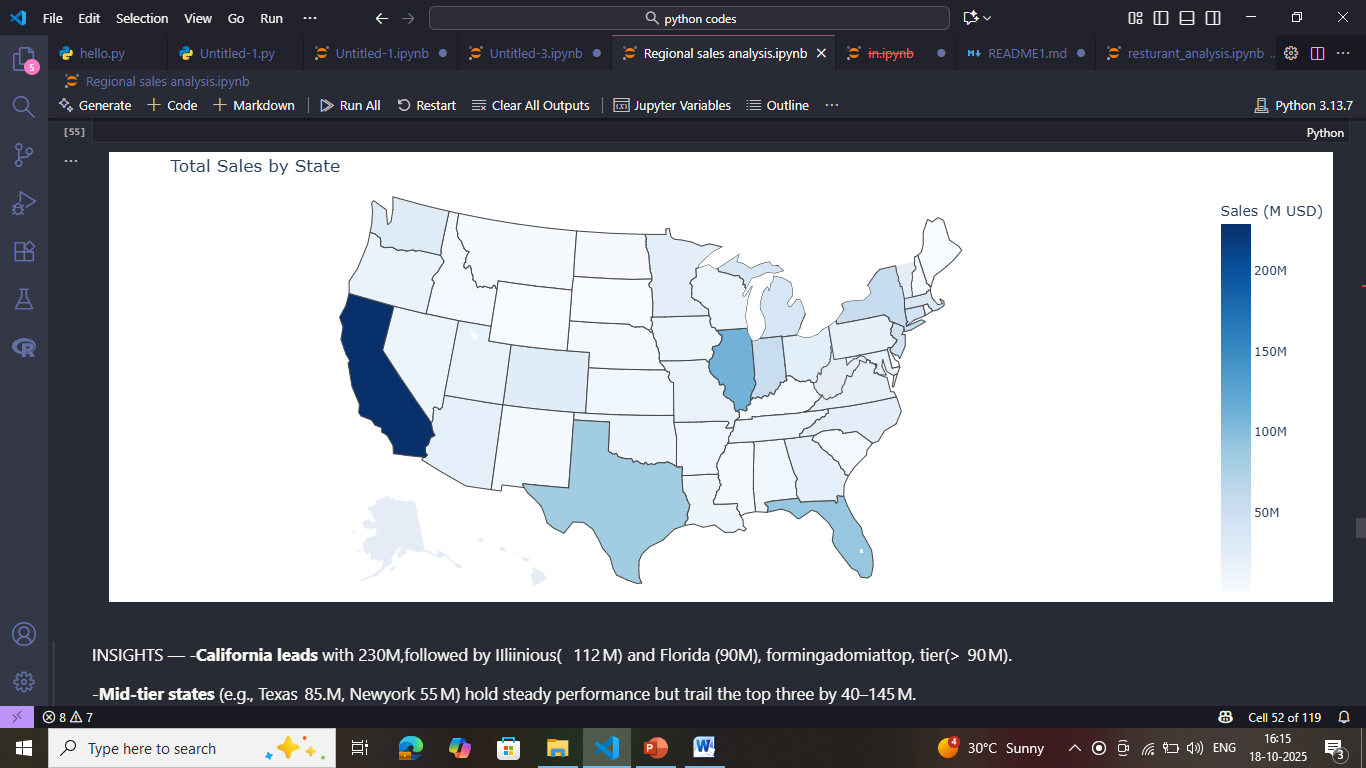
* Revenue leaders: Products 26 & 25 dominate.
* Mid-range: Products 5, 13, 14, 15 show similar revenue.
* Bottom cluster: Products 1, 2, 3, 4 have the lowest revenue.
* Strategy: Grow mid-tier, improve lower performers

1. **Sales by Channel**



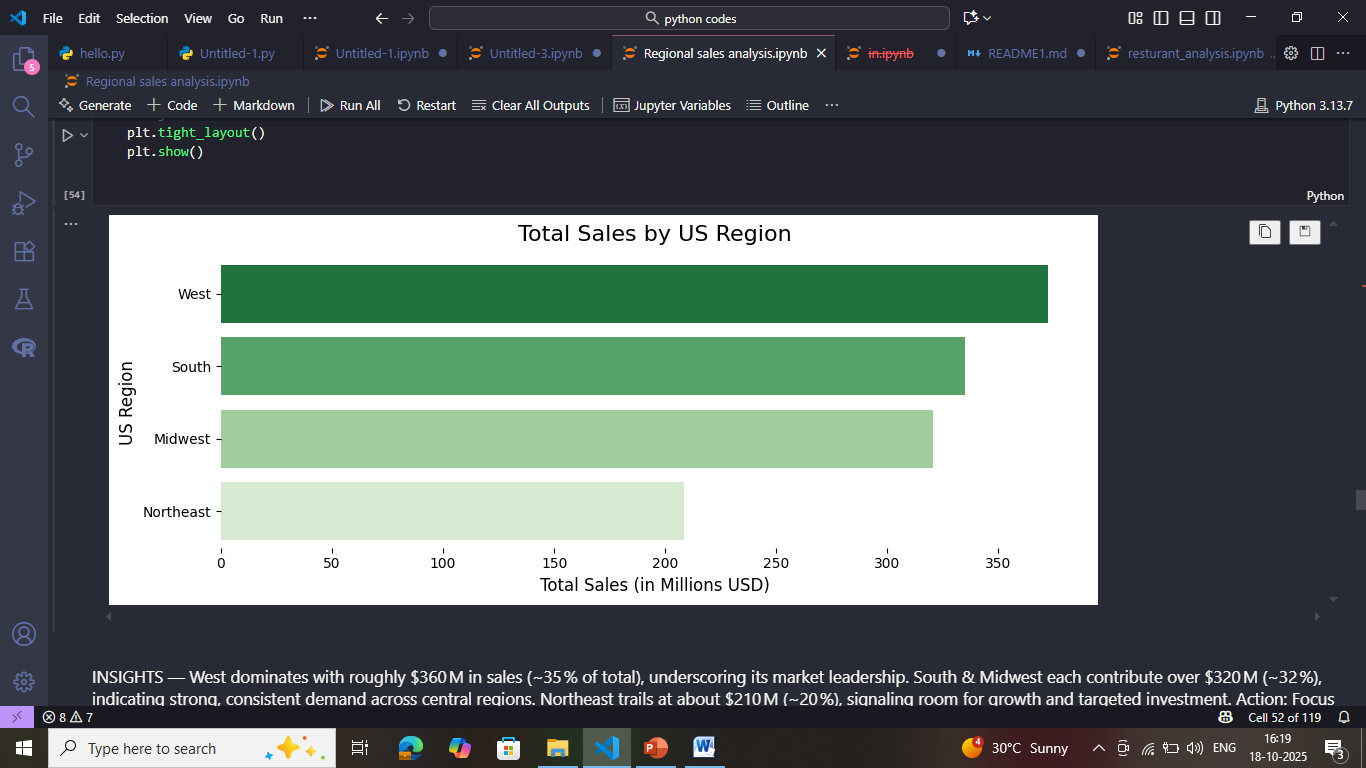
* Wholesale dominates: Generates the majority of total sales at 54.1%.
* Distributor is significant: Contributes a substantial 31.3% to total sales.
* Export is a smaller portion: Accounts for 14.6% of the total sales.

1. **Total Sales by State (Choropleth Map)**



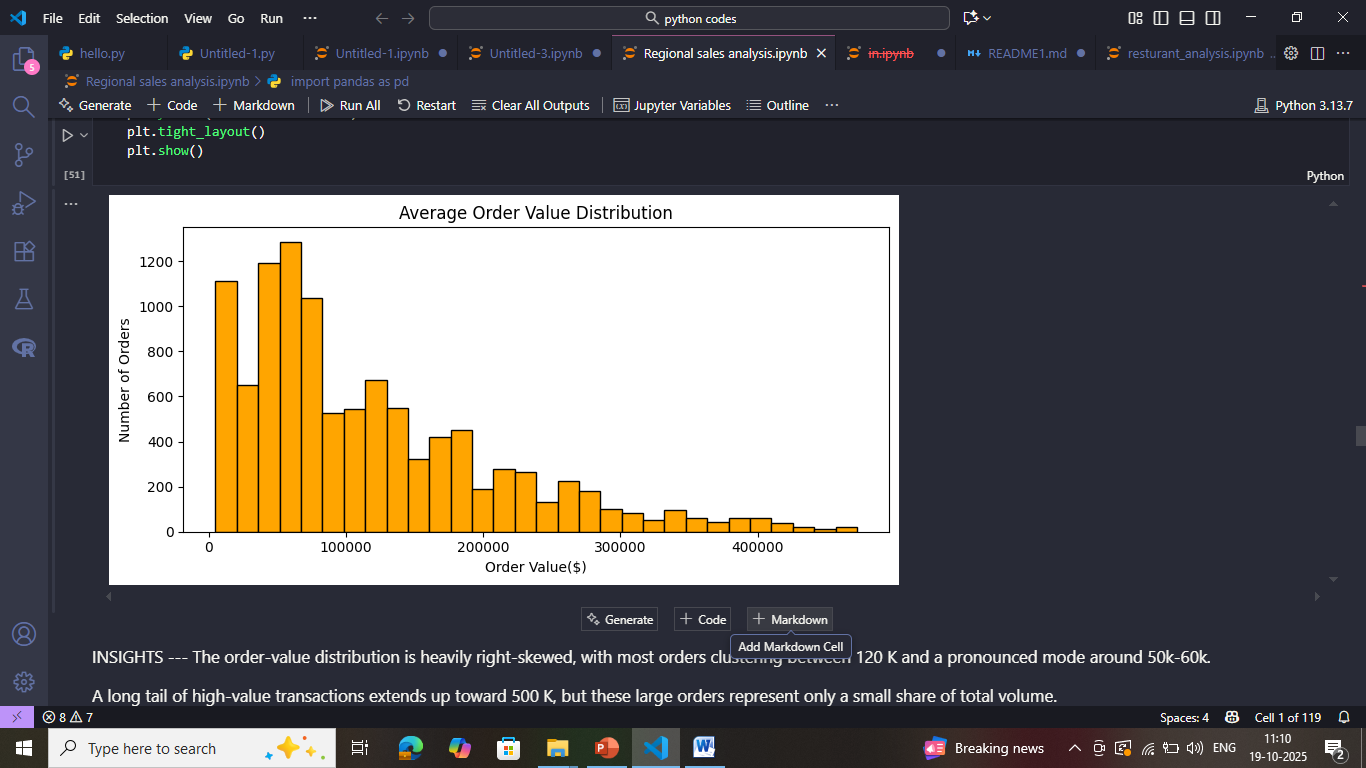
* California leads: Highest total sales.
* Texas, Florida, Illinois: Significant sales.
* Varying sales: Other states show moderate to low sales.
* Visual pattern: Higher sales in larger and some coastal states.

1. **Total Sales by US Region**



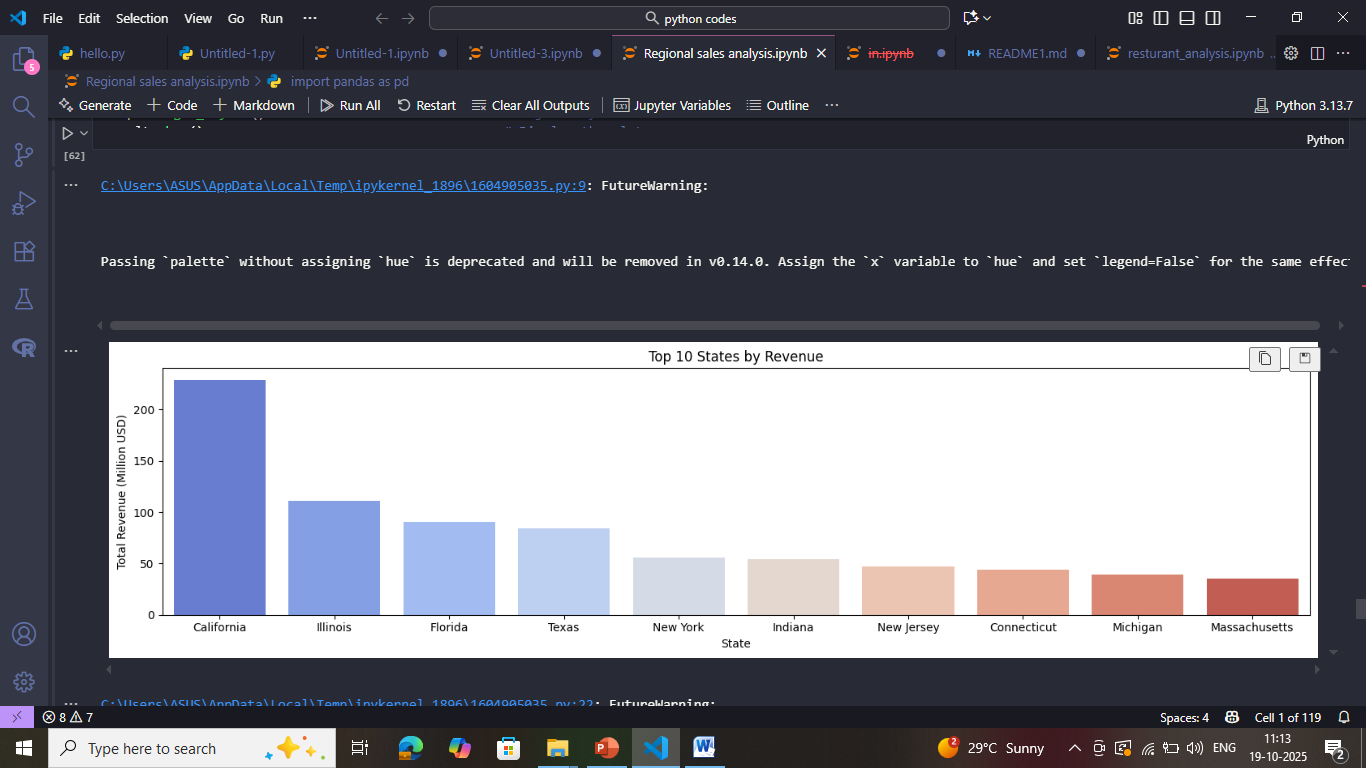
* West: Highest sales, strong market influence.
* South: Major sales contributor, key market area.
* Midwest: Steady sales performance, moderate market size.
* Northeast: Lowest sales, suggests need for deeper market understanding.

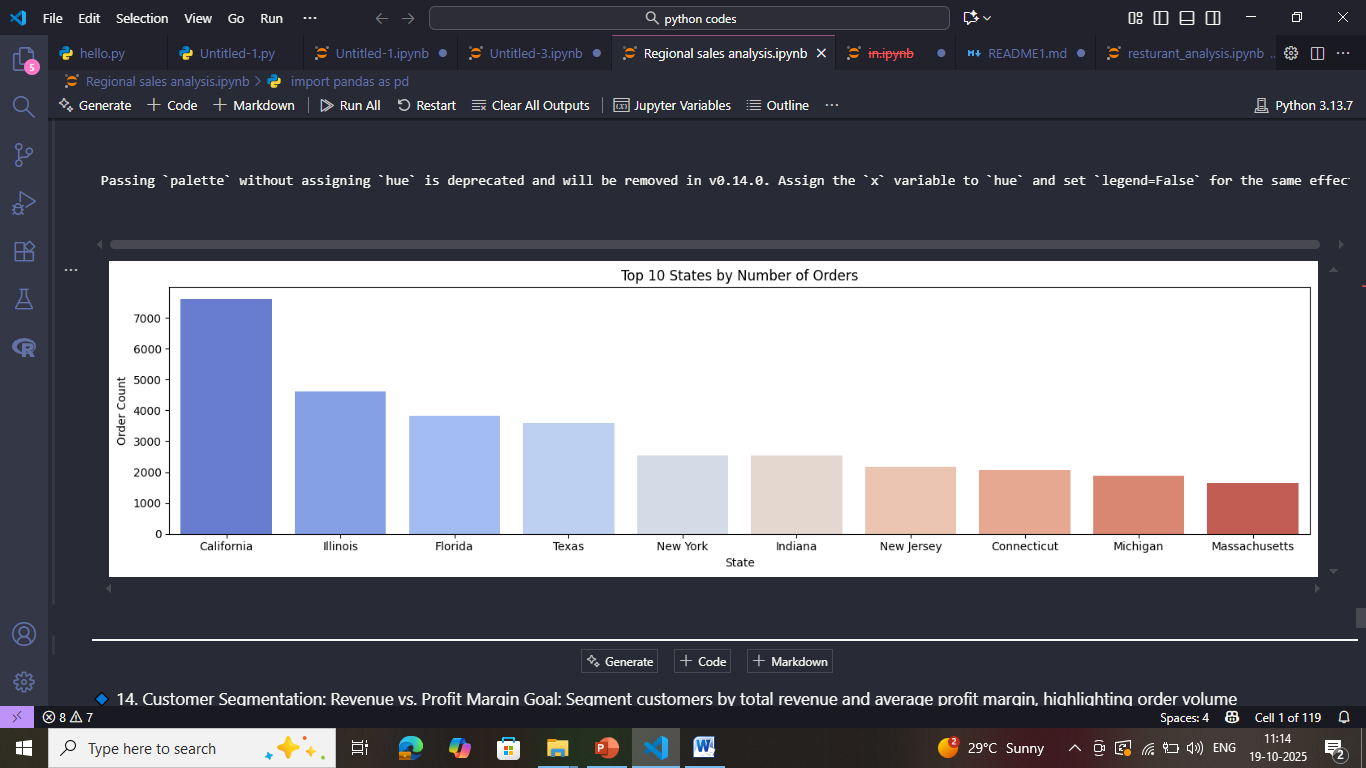
1. **Average Order Value (AOV) Distribution**



* 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

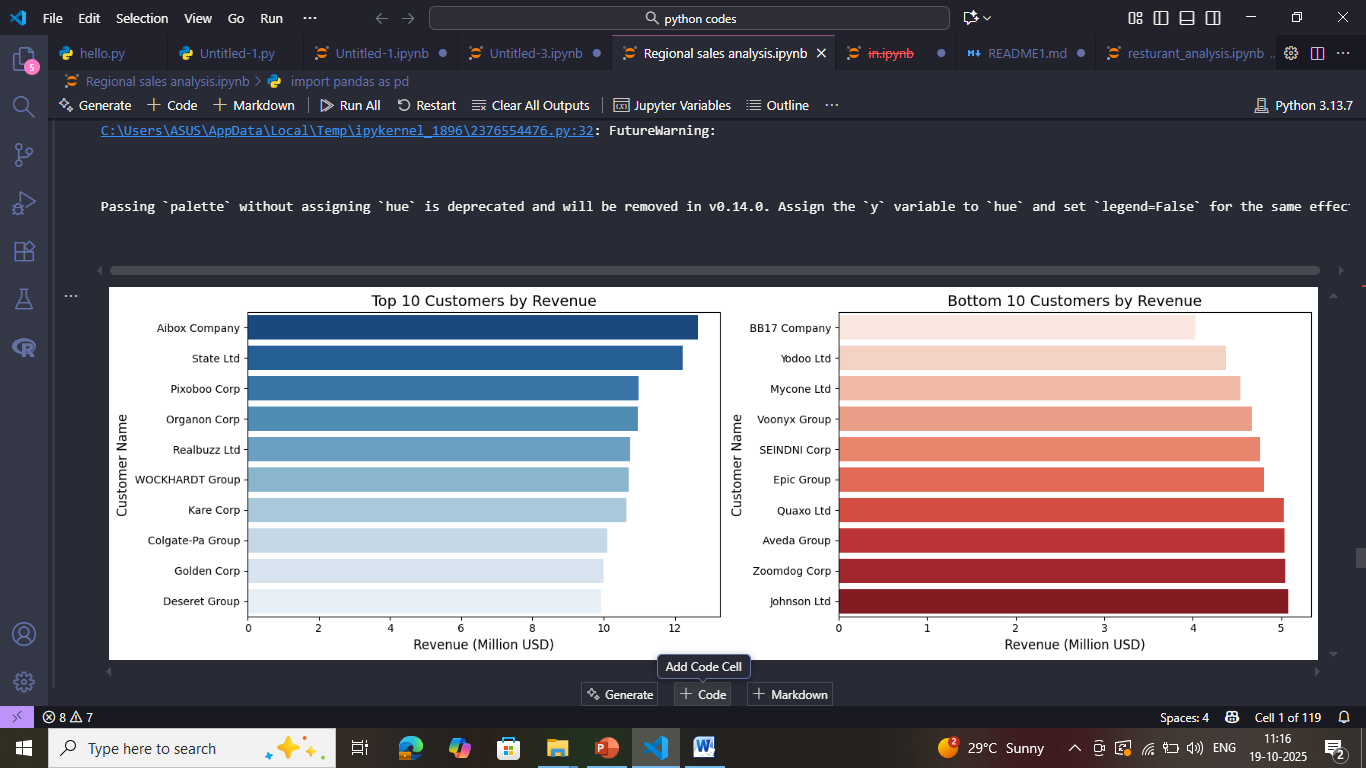
1. **Top State Performance: Revenue vs. Orders**





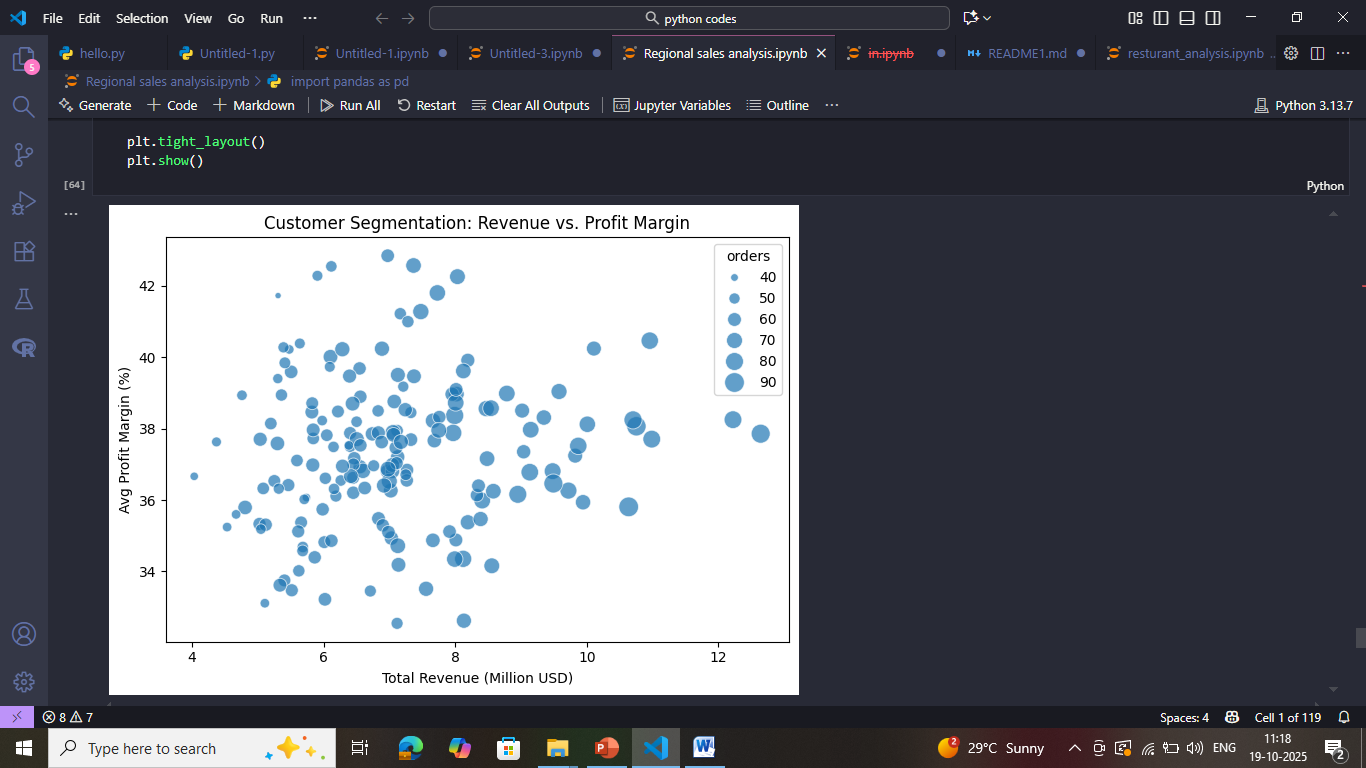
* California tops revenue & orders.
* IL, FL, TX: High in both.
* Revenue & orders linked.
* Other top states: Lower contribution.

1. **Top and Bottom 10 Customers by Revenue**



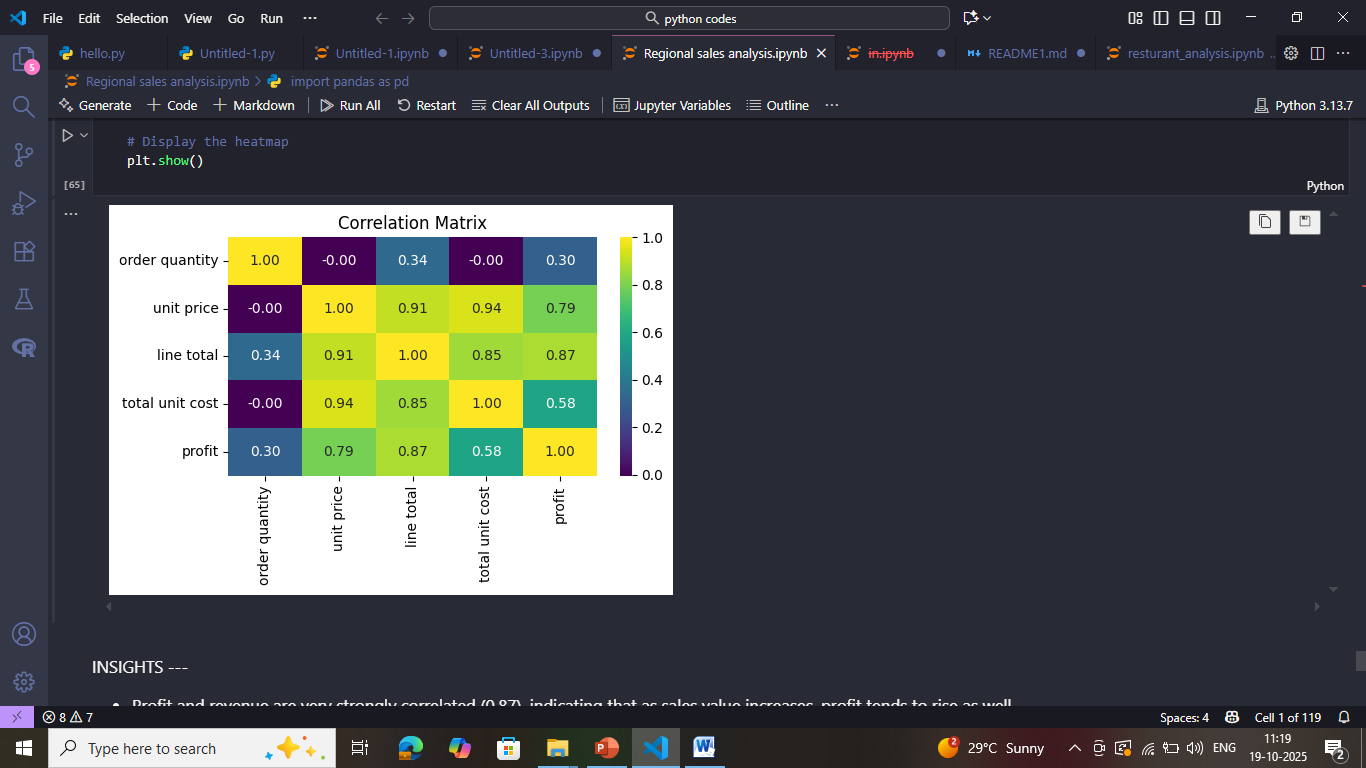
* Aibox Company leads significantly as the top revenue generator.
* Bottom 10 customers generate substantially less revenue (around $4-5M).
* Revenue concentration: Top customers drive a disproportionate share.
* Large gap: Exists between the revenue of top and bottom tier customers

1. **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.

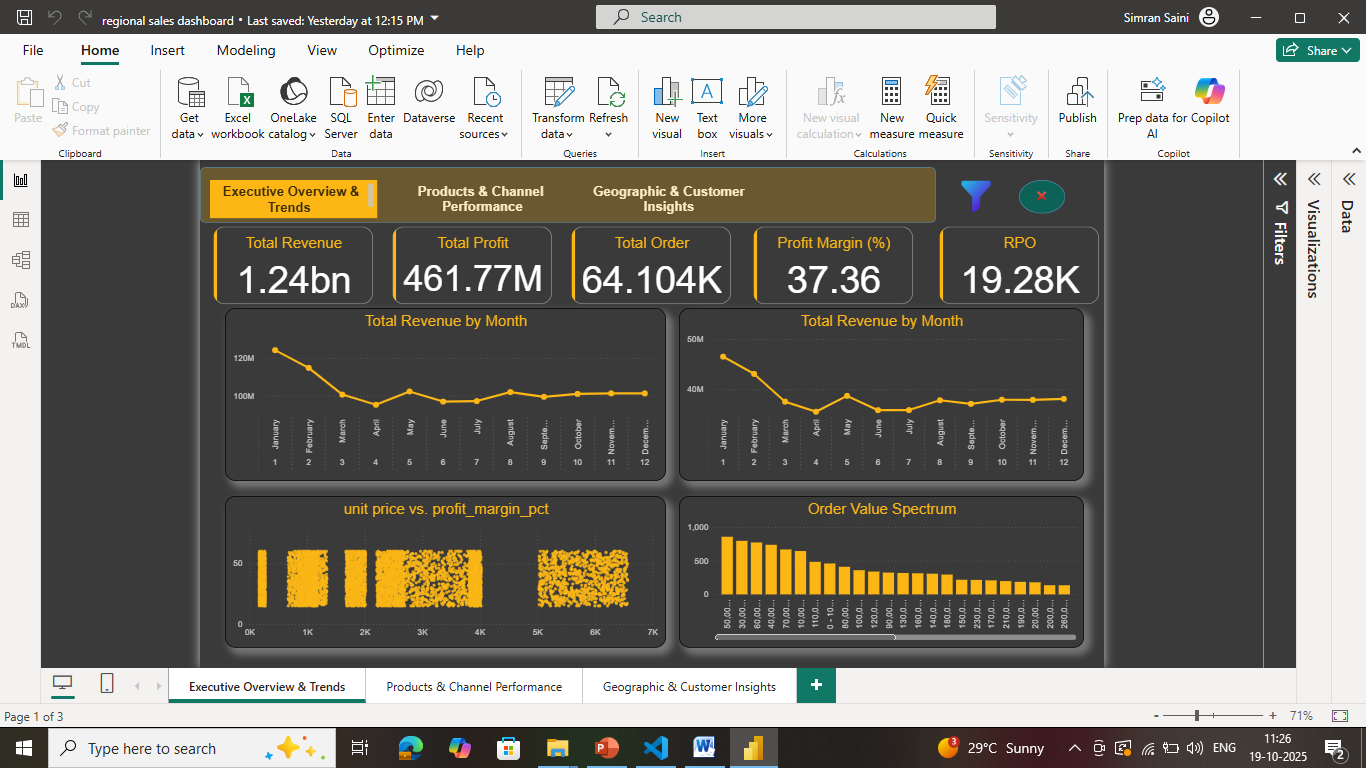
1. **Correlation Heatmap of Numeric Features**



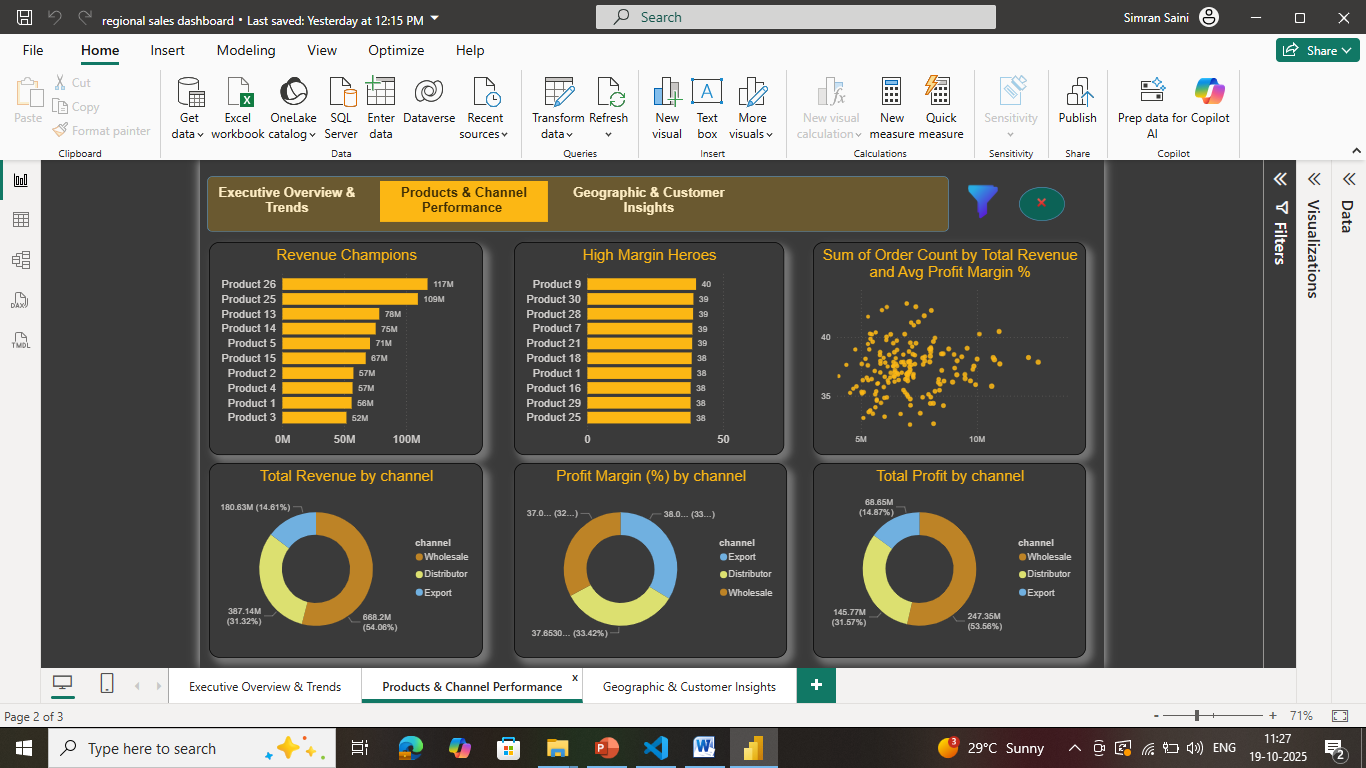
* Unit price is the primary driver, showing very strong correlations with cost (0.94), revenue (0.91) and profit (0.79).
* Revenue & profit maintain a high link (0.87), underscoring direct profitability gains.
* Quantity’s impact is minimal (≤ 0.34 vs. financials), indicating volume plays a secondary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

**Dashboard Preview**

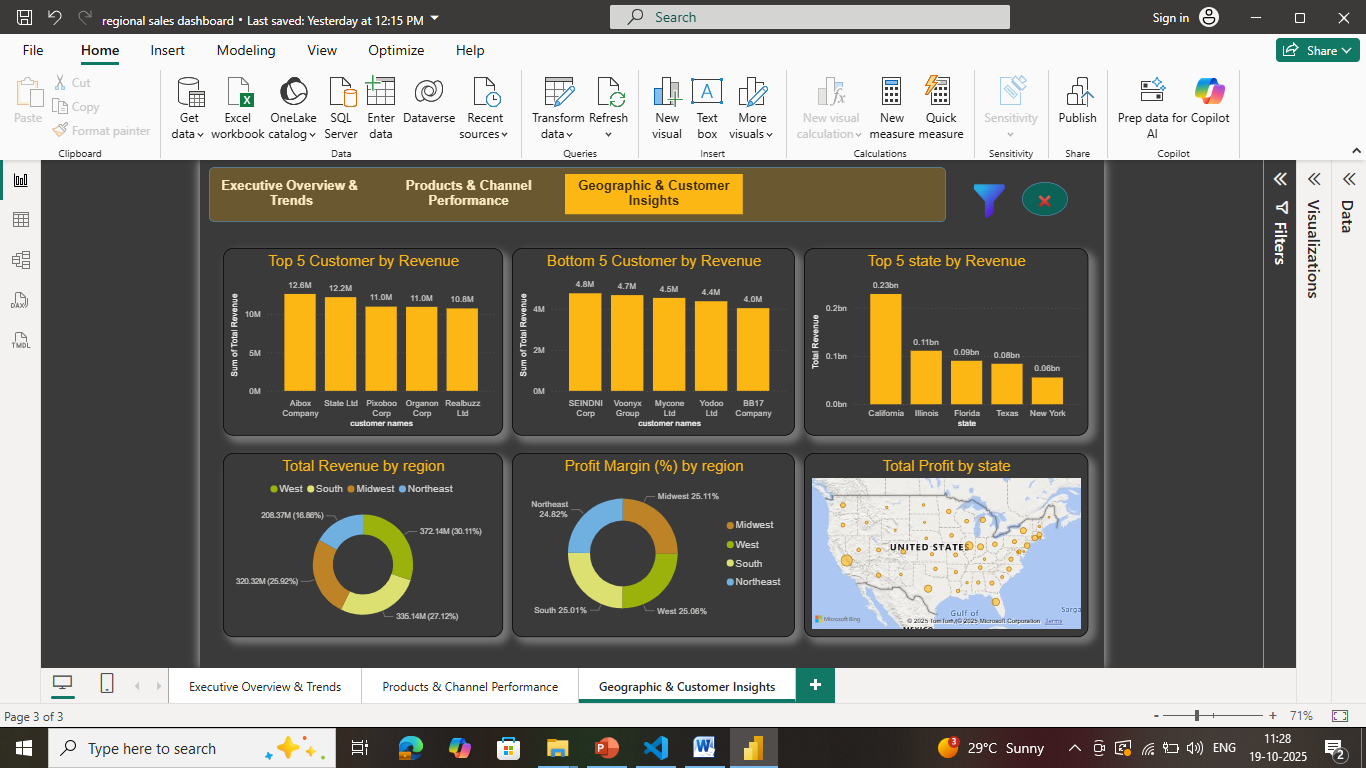
**Page 1 - Performance Summary**



**Page 2 - Customer Segmentation**



**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**