

US Regional Sales Data Analysis

About Dataset

This dataset provides comprehensive insights into US regional sales data across different sales channels, including In-Store, Online, Distributor, and Wholesale. With a total of 17,992 rows and 15 columns, this dataset encompasses a wide range of information, from order and product details to sales performance metrics. It offers a comprehensive overview of sales transactions and customer interactions, enabling deep analysis of sales patterns, trends, and potential opportunities.

Columns In Dataset

- **OrderNumber:** A unique identifier for each order.
- **Sales Channel:** The channel through which the sale was made (In-Store, Online, Distributor, Wholesale).
- **WarehouseCode:** Code representing the warehouse involved in the order.
- **ProcuredDate:** Date when the products were procured.
- **OrderDate:** Date when the order was placed.
- **ShipDate:** Date when the order was shipped.
- **DeliveryDate:** Date when the order was delivered.
- **SalesTeamID:** Identifier for the sales team involved.
- **CustomerID:** Identifier for the customer.
- **StoreID:** Identifier for the store.
- **ProductID:** Identifier for the product.
- **Order Quantity:** Quantity of products ordered.
- **Discount Applied:** Applied discount for the order.
- **Unit Cost:** Cost of a single unit of the product.
- **Unit Price:** Price at which the product was sold.

Initial Plan for Data Exploration

1 Understand the Dataset

Verify the structure and data types of each column. Ensure appropriate data types. Check for missing values and inconsistencies across the dataset

3 Validate Data Quality

Look for missing or invalid data entries in key columns (e.g., OrderDate, Unit Cost, CustomerID). Identify any duplicate orders or records.

2 Explore Column Properties

Ensure OrderNumber uniqueness, analyze sales channels, and evaluate warehouses. Validate dates and calculate durations. Analyze ID patterns, explore order quantities, and detect outliers.



DATA CLEANING AND FEATURE ENGINEERING



1 Remove Duplicates

2 Handle Outliers

3 Create Derived Features

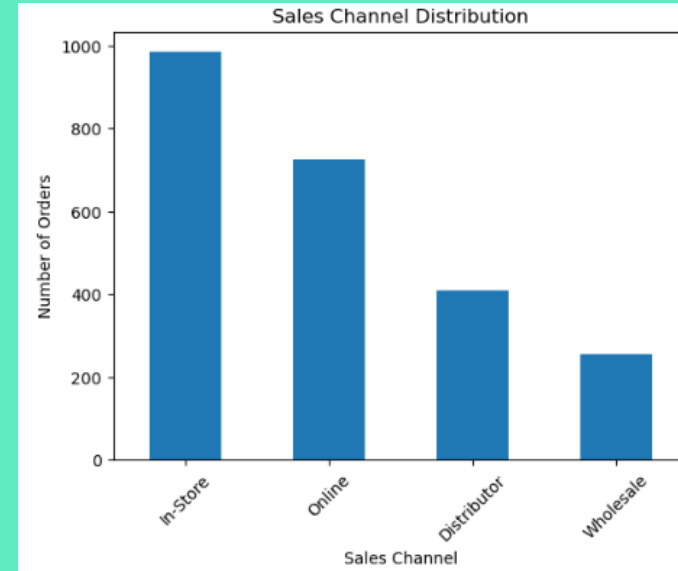


Key Findings and Insights

1 Sales Channel Distribution

Finding: Most orders were placed in-store (984 orders), followed by online (724 orders), distributor (410 orders), and wholesale (256 orders).

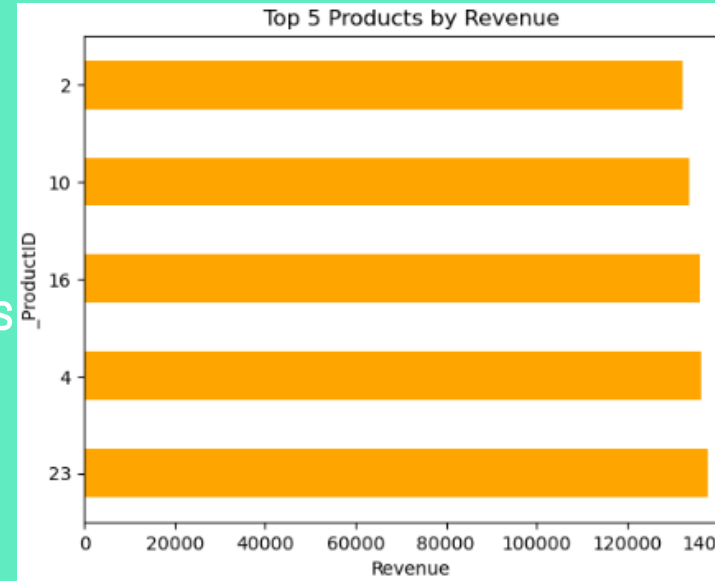
Insight: In-store sales lead the distribution, suggesting strong physical presence and customer engagement. However, online sales are significantly lower, indicating potential for growth through digital channels. Focus on improving the online shopping experience and digital marketing strategies to increase sales.



2 Product and Customer Analysis

Finding: A small number of products contribute significantly to total sales, with the top 10 products generating the highest revenue: Product 23: 137,671.6, Product 4: 136,472.3, Product 16: 136,157.4, Product 10: 133,624.8, Product 2: 132,251.3,

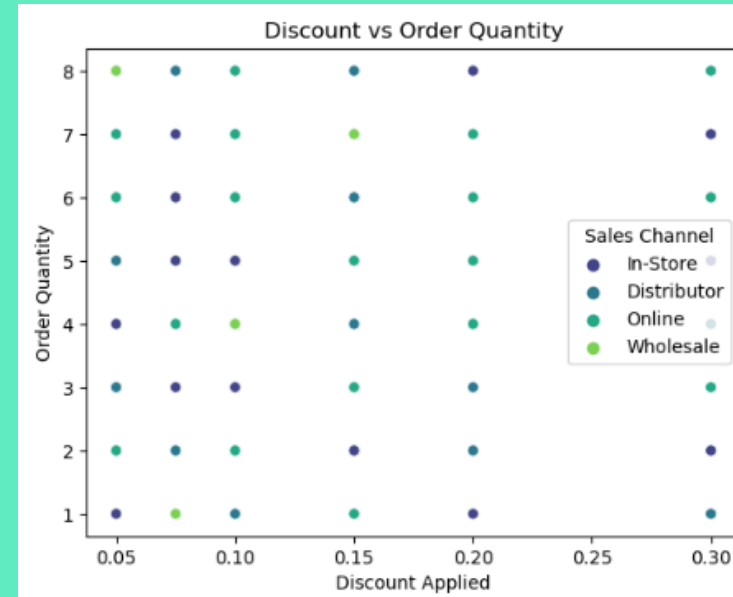
Insight: A small set of products generates the bulk of revenue. Strategies should focus on promoting these top-performing products, optimizing inventory, and enhancing customer targeting for products with high sales potential.



3 Discounts and Sales Correlation

Finding: There isn't a significant correlation between discount and order quantity.

Insight: Since there is no strong relationship between discounts and order quantities, focus should shift to improving other factors such as product appeal, customer engagement, and inventory management rather than relying heavily on discounts to drive sales.



4

Warehouse Utilization

Finding: The distribution of orders across warehouses is as follows: WARE-NMK1003: 749 orders, WARE-PUJ1005: 399 orders, WARE-UHY1004: 382 orders, WARE-XYS1001: 369 orders, WARE-MKL1006: 246 orders, WARE-NBV1002: 229 orders.

Insight: Warehouse WARE-NMK1003 processes the most orders, while other warehouses have significantly lower activity. Consider redistributing inventory and optimizing operations at underused warehouses to balance the workload and improve operational efficiency.

