Sales Order Dataset

Profit & Sales Analysis Dataset

About Dataset

This dataset contains retail sales transaction data, including customer details, product information, geographic locations, and financial metrics. It is designed primarily for exploratory data analysis (EDA) to uncover sales trends, profitability patterns, and customer behavior across regions and product categories..

**Information related to dataset :- **

**Features :- **

Row ID: A unique identifier for each row in the dataset.

Order ID: A unique identifier assigned to each customer order.

Order Date: The date on which the customer placed the order. (YYYY-MM-DD)

Ship Date: The date on which the order was shipped. (YYYY-MM-DD)

Ship Mode: The shipping method used to deliver the order (e.g., Standard, Second Class, First Class).

Customer ID: A unique identifier for each customer.

Customer Name: The full name of the customer who placed the order.

Segment: The market segment the customer belongs to (such as Consumer, Corporate, or Home Office).

Country: The country where the order was placed or delivered (United States overall).

City: The city associated with the order’s delivery address. where around 531 distinct cities are there.

State: The state or province of the delivery location. (Around 49 States)

Postal Code: The postal or ZIP code of the delivery address.

Region: The broader geographic region where the order was shipped. (East,West,North & South)

Product ID: A unique identifier assigned to each product.

Category: The high-level classification of the product (e.g., Furniture, Office Supplies, Technology).

Sub-Category: A more detailed classification of the product within a category.

Product Name: The name or description of the product sold.

Sales: The total revenue generated from the order line item.

Quantity: The number of units of the product sold in the order.

Discount: The discount applied to the product (usually expressed as a fraction or percentage).

Profit: The net profit earned from the sale after accounting for discounts and costs.

Use Cases :-

1. Exploratory Data Analysis
2. Time Series Analysis
3. Basic Level Spatial Analysis
4. Customer Segment Behavior Analysis
5. Practicing SQL or Power BI .

6. Not Intended for any Machine Learning Application has there is no clearly defined target features or creating the ones will not be an smart move.

EDA Questions

1. Which regions generate the highest sales and profit?
2. How do discounts affect profitability?
3. Which product categories are most profitable?
4. What shipping mode is used most frequently?
5. Which customer segments contribute the most revenue?
6. Are higher sales always associated with higher profit?

KPI Queries

1. Total Sales (Overall revenue generation)
2. Total Profit (Actual earnings after cost)
3. Profit Margin (Profitability efficiency)
4. Avg Discount (Pricing impact on margins)
5. Loss Orders % (Risk & revenue leakage)
6. Sales per Order (Customer value)
7. Profit per Order (Order quality)