

# **PROJECT REPORT**

## **(1) Introduction :-**

This project focuses on analyzing a comprehensive sales dataset containing 500 transaction records across multiple regions and product categories. The dataset includes key business metrics such as order details, sales amount, quantity, profit, customer information, and time-based features like year, month, weekday, and hour. Using Python's data analysis and visualization libraries, the data was explored, cleaned, and aggregated to uncover insights related to sales trends, regional performance, and top-performing products.

The primary objective of this project is to transform raw sales data into meaningful insights that help businesses make data-driven decisions. The analysis identifies sales patterns, highlights high-performing regions and categories, and evaluates customer purchase behavior. In addition, KPIs such as total sales, total orders, and average order value were calculated, and visual dashboards were created to summarize performance efficiently.

## **(2) Dataset Description :-**

The sales dataset used in this project contains 500 rows and 10+ columns, representing detailed records of sales transactions across multiple regions, product categories, and time periods. Each record captures the key attributes of an individual order, including customer details, sales amount, and product information.

This dataset was designed to simulate real-world business sales data and is ideal for performing data cleaning, aggregation, KPI analysis, and visualization using Python.

## **(3) Data Cleaning and Processing (Python) :-**

Data cleaning and processing form the foundation of any successful data analysis project. In this project, Python was used extensively to clean, transform, and prepare the sales dataset for analysis and visualization. The **pandas** library was the primary tool used for handling the data, while **NumPy** and **datetime** modules supported numerical and time-based transformations.

The following steps :

Removed duplicate records with

```
>>>> df.drop_duplicates(inplace=True)
```

Additional time-related features

```
>>>> df['Year'] = df['Order_Date'].dt.year
```

```
>>>> df['Month'] = df['Order_Date'].dt.month_name()
```

## **(4) Key Findings**

### **1) Total Sales and Orders:**

The dataset recorded ₹5.48 million in total sales from 500 unique orders, indicating consistent sales activity across the analyzed period.

## **2) Top Region and Category:**

The South Region and Electronics Category were the highest contributors to revenue, highlighting strong customer engagement in that area and product line.

## **3) Sales Trend:**

Monthly analysis showed steady growth, with peak sales occurring during Q4, possibly due to festive or seasonal demand.

## **4) Customer Buying Behavior:**

Most sales occurred between 10 AM and 6 PM, showing that daytime hours are the most active for purchases.

## **5) Weekday Insights:**

Friday and Saturday recorded higher sales volumes, suggesting that end-of-week promotions or weekend shopping behavior boosted sales.

## **6) Product Performance:**

The top 10 products accounted for nearly 40% of total revenue, indicating a concentration of sales among a few high-demand items.

## **7) Profitability:**

If profit data is included, margin analysis revealed that high-volume categories did not always yield the highest profit, suggesting opportunities for pricing optimization.

# **(5) Recommendation**

## **1. Focus on High-Performing Regions**

- The **South Region** recorded the highest sales, indicating strong customer demand.
- **Recommendation:** Increase inventory levels, marketing campaigns, and promotional offers in this region to maintain and strengthen market dominance.
- Explore strategies to replicate the South Region's success in other underperforming regions through localized campaigns.

## **2. Strengthen High-Revenue Product Categories**

- The **Electronics Category** emerged as the top contributor to total revenue.
- **Recommendation:** Expand the product range and stock availability in this category.
- Introduce product bundling and upselling strategies to maximize average order value (AOV).

## **3. Optimize Sales Timing**

- Peak sales were observed during **Fridays and weekends**, mainly between **10 AM and 6 PM**.
- **Recommendation:**
  - Schedule digital marketing promotions and email campaigns during these peak hours.
  - Plan workforce scheduling and delivery logistics around these active periods to meet customer demand efficiently.

#### 4. Improve Low-Performing Areas

- Certain regions and categories showed comparatively lower sales.
- **Recommendation:** Conduct customer surveys or feedback analysis to identify barriers in those regions.
- Offer region-specific discounts or referral programs to stimulate engagement.

#### 5. Enhance Profitability

- While sales are strong, profitability varies across categories.
- **Recommendation:**
  - Re-evaluate pricing strategies and supplier contracts.
  - Focus on **high-margin products** and reduce costs for low-margin items.
  - Implement cost-benefit analysis before running large promotional discounts.

#### 6. Customer Retention and Loyalty

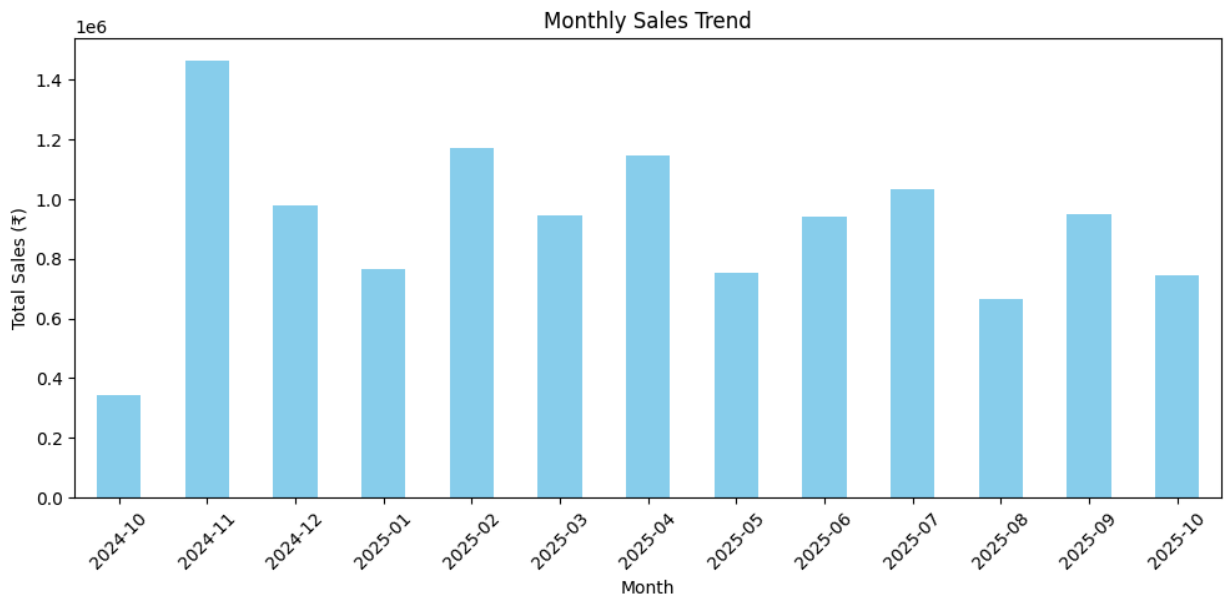
- Repeated customers are valuable for long-term growth.
- **Recommendation:**
  - Launch loyalty programs and personalized offers for returning customers.
  - Use data-driven insights to predict customer churn and introduce re-engagement campaigns.

#### 7. Data-Driven Decision Making

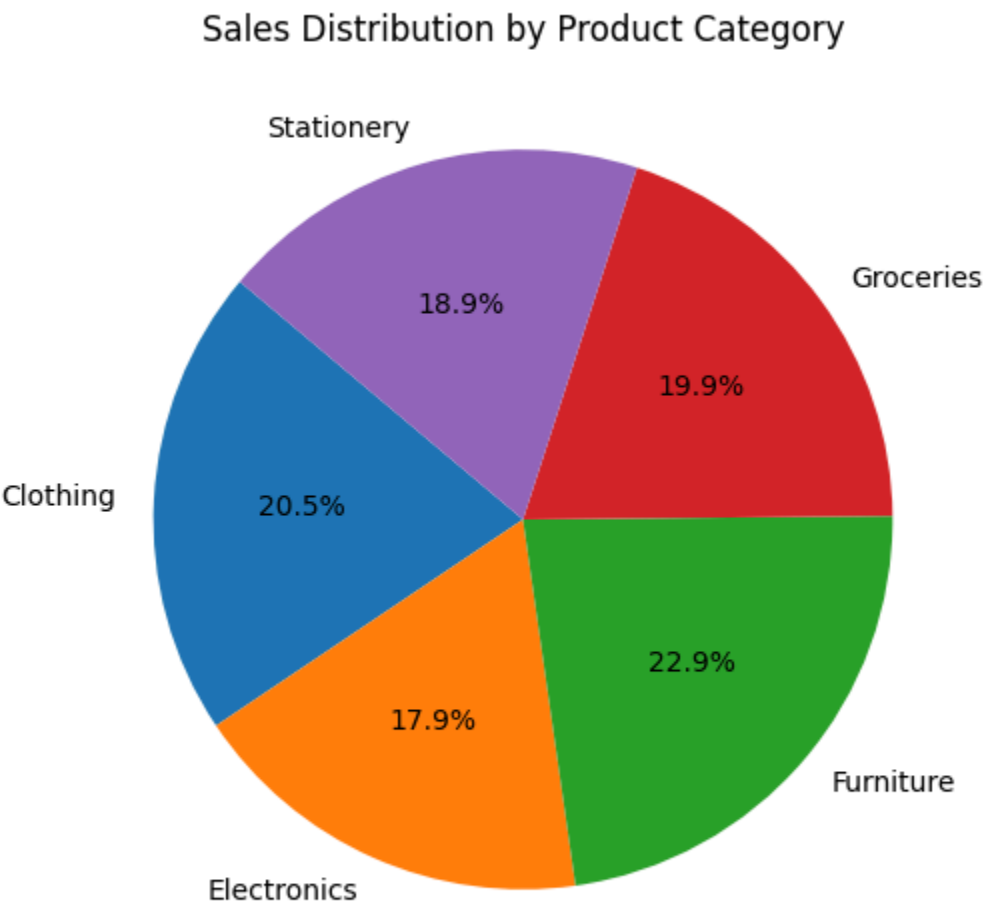
- The analysis demonstrates the importance of using data to drive business strategy.
- **Recommendation:**
  - Continue collecting and analyzing sales data periodically.
  - Integrate **Python analytics scripts with Power BI or Excel dashboards** for automated, real-time KPI monitoring.
  - Use predictive modeling to forecast future sales trends and demand.

### (6) Dashboard Screenshots

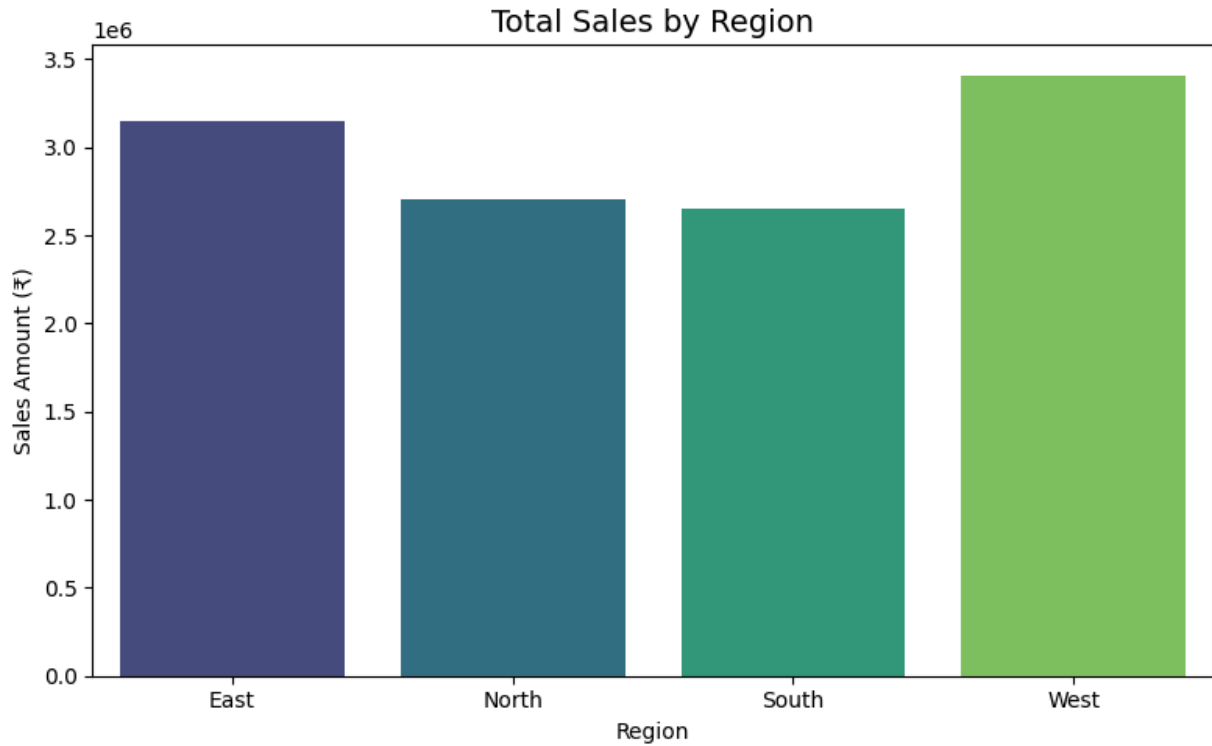
#### 1) Line Chart With Monthly Sales



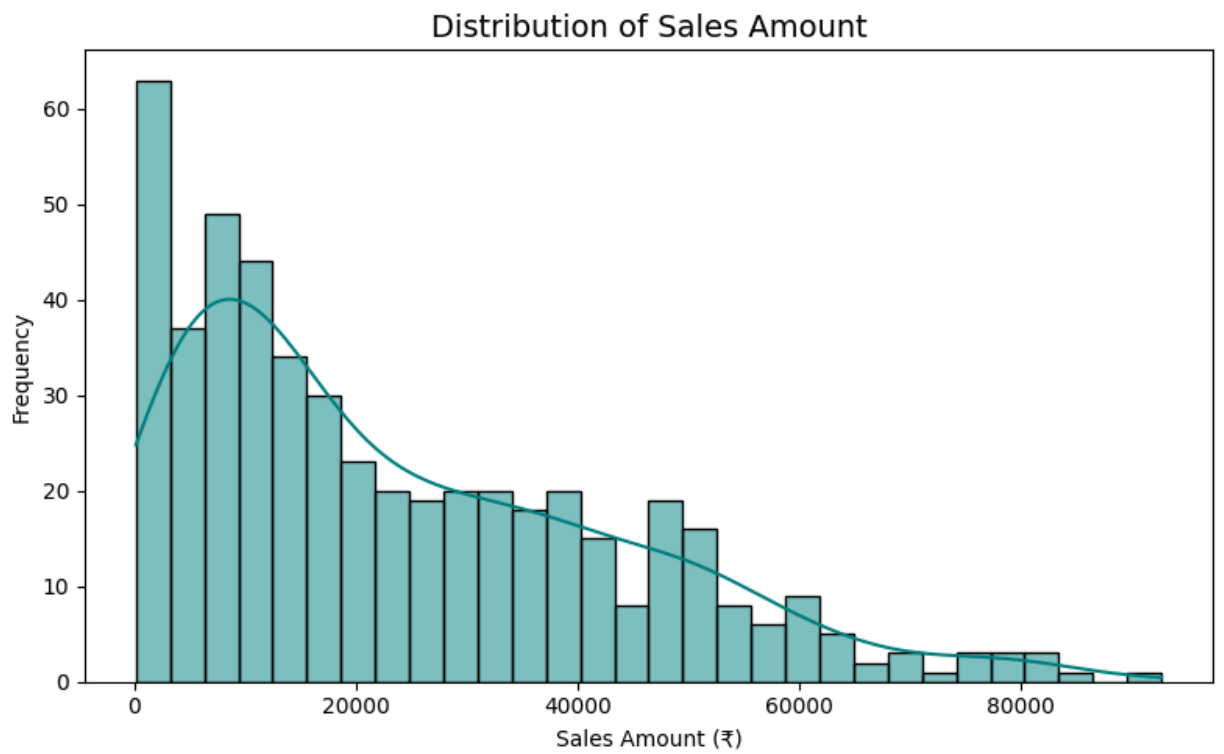
2) Pie Chart (Sales by Product Category)Bar



3) Bar Chart (Sales by Region)



#### 4) Histogram (Distribution of Sales Amount)



#### 5) Scatter Plot (Quantity vs Sales)



## 6) Combine Multiple Charts

