

Exploratory Data Analysis (EDA) Report for Task 1

Objective:

The objective of this task is to analyze the data from the given datasets to identify key trends and insights into customer behavior, product performance, and overall sales trends.

Data Sources:

1. **Customers.csv:** Contains customer details, such as CustomerID, CustomerName, Region, and SignupDate.
 2. **Products.csv:** Contains product details, including ProductID, ProductName, Price, and Category.
 3. **Transactions.csv:** Contains transaction data, such as TransactionID, CustomerID, ProductID, TransactionDate, Quantity, and TotalValue.
-

Steps Performed:

Data Loading and Preprocessing:

- Merged the Transactions, Customers, and Products datasets into a single unified DataFrame using CustomerID and ProductID as keys.
 - Converted the TransactionDate column to datetime format for better date-based analysis.
 - Extracted the TransactionMonth column from TransactionDate to analyze monthly sales trends.
 - Checked for missing values across all columns and confirmed that there were no missing values.
-

Analysis Performed:

a) Monthly Sales Trend:

- Grouped the data by TransactionMonth and calculated the total sales (TotalValue) for each month.
- **Observation:** Sales vary across months, with specific months showing higher sales peaks. For example, months like December and January exhibit strong sales trends, possibly due to seasonal factors like holidays or promotions.

b) Regional Sales Analysis:

- Grouped the data by Region and calculated the total sales for each region.
- **Observation:** Certain regions contribute significantly more to overall sales compared to others. This provides insights into where the business performs best geographically.

c) Sales by Category:

- Grouped the data by Category and calculated the total sales for each category.
- **Observation:** Some product categories consistently generate higher revenue than others. These top-performing categories could benefit from focused marketing efforts or increased stock availability.

d) Top 5 Products by Sales:

- Identified the top 5 products based on total revenue by grouping data by ProductName and calculating total sales.
- **Observation:** These products are likely the most popular or high-demand items. Ensuring availability and strategically promoting these products can further boost sales.

e) Top 5 Customers by Purchases:

- Identified the top 5 customers who have made the highest total purchases by grouping data by CustomerName and calculating total sales.
 - **Observation:** These customers represent the most valuable segment for the business. Focusing on customer retention strategies, such as loyalty programs or exclusive discounts, could help maintain their engagement and spending.
-

Visualizations:

1. **Monthly Sales Trend:** A line plot showing the variation in total sales across months.
2. **Regional Sales:** A bar plot visualizing total sales for each region.
3. **Category Sales:** A bar plot highlighting the revenue contribution of each product category.
4. **Top 5 Products:** A horizontal bar plot showcasing the top 5 products contributing the most to sales.
5. **Top 5 Customers:** A bar plot displaying the highest-purchasing customers and their contribution to total sales.

Key Insights:

1. **Monthly Sales Performance:**
 - Sales peak during specific months possibly driven by holiday sales and promotions.
 - There are low-sales months that may require targeted campaigns to boost revenue.
2. **Regional Performance:**
 - Regions such as {Top Region} contribute the highest revenue, indicating a strong market presence in these areas.
 - Regions with lower sales can be investigated to identify potential growth opportunities.
3. **Category Insights:**
 - The category {Top Category} is the highest revenue generator and could benefit from increased marketing focus.
 - Low-performing categories might require product redesign, better marketing, or discontinuation.
4. **Product Insights:**
 - The top 5 products account for a significant portion of the sales. Ensuring these products are always in stock can prevent revenue loss.
 - Cross-selling these products with complementary items could increase overall sales.
5. **Customer Insights:**
 - The top 5 customers account for significant total purchases, highlighting their importance to the business.
 - Implementing strategies like loyalty programs, personalized offers, or exclusive discounts can further enhance customer satisfaction and retention.