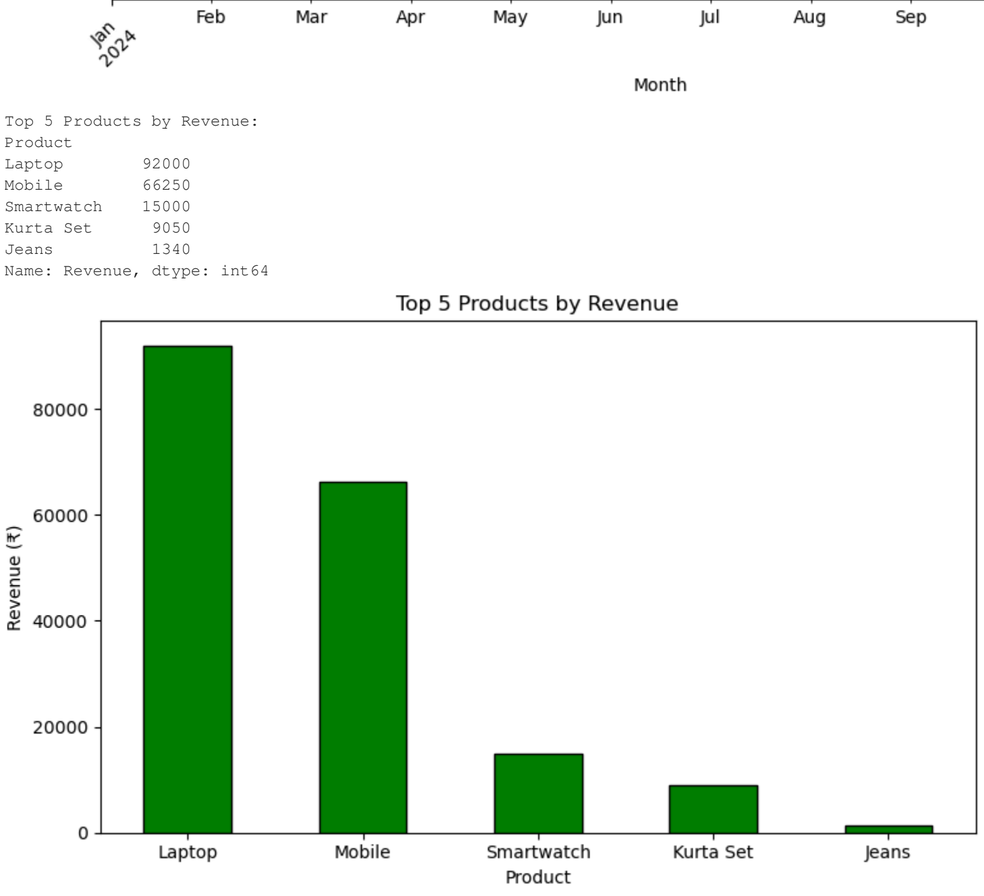
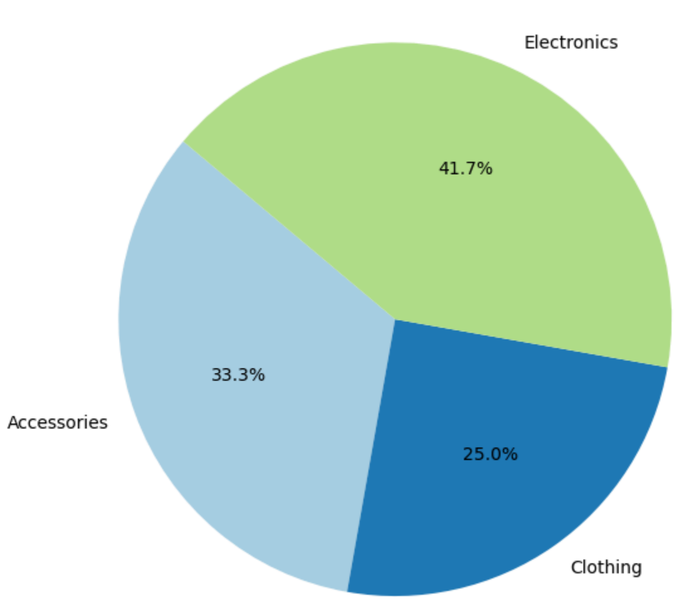
PROJECT TITLE : Sales Performance Analysis and Visualization for an E-commerce Platform

In [3]:
```python
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

dataset = {
    "Order_ID": [1001, 1002, 1003, 1004, 1005, 1006, 1007, 1008, 1009, 1010, 1011, 1012],
    "Product": ["Laptop", "Mobile", "Headphones", "Laptop", "Mobile", "Smartwatch", "Mobile", "Headphones", "Kurta Set", "Kurta Set", "Kurta Set", "Jeans"],
    "Category": ["Electronics", "Electronics", "Accessories", "Electronics", "Electronics", "Accessories", "Electronics", "Accessories", "Clothing", "Accessories", "Clothing", "Clothing"],
    "Customer_ID": [201, 202, 203, 201, 204, 205, 202, 203, 201, 206, 207, 207],
    "Order_Date": ["2024-01-15", "2024-01-20", "2024-02-12", "2024-02-26", "2024-03-05", "2024-03-21", "2024-04-19", "2024-04-23", "2024-05-17", "2024-06-20", "2024-08-02", "2024-12-16"],
    "Quantity": [1, 2, 1, 1, 1, 3, 1, 2, 3, 1, 2],
    "Price_per_Unit": [50000, 23500, 450, 42000, 17600, 15000, 550, 800, 1500, 1250, 2300, 670],
}

df = pd.DataFrame(dataset)
df["Order_Date"] = pd.to_datetime(df["Order_Date"])
df["Revenue"] = df["Quantity"] * df["Price_per_Unit"]

total_revenue = df["Revenue"].sum()
average_order_value = df["Revenue"].mean()
print(f"Total Revenue: ₹{total_revenue:,.2f}")
print(f"Average Order Value: ₹{average_order_value:,.2f}")

df["Month"] = df["Order_Date"].dt.to_period("M")
monthly_revenue = df.groupby("Month")["Revenue"].sum()

plt.figure(figsize=(10, 5))
monthly_revenue.plot(kind="line", marker="o", color="blue", lw=2)
plt.title("Monthly Sales Trends")
plt.xlabel("Month")
plt.ylabel("Revenue (₹)")
plt.grid(alpha=0.3)
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()

product_revenue = df.groupby("Product")["Revenue"].sum().sort_values(ascending=False)
top_5_products = product_revenue.head(5)
print("\nTop 5 Products by Revenue:")
print(top_5_products)

plt.figure(figsize=(8, 5))
top_5_products.plot(kind="bar", color="green", edgecolor="black")
plt.title("Top 5 Products by Revenue")
plt.xlabel("Product")
plt.ylabel("Revenue (₹)")
plt.xticks(rotation=0)
plt.tight_layout()
plt.show()

category_count = df.groupby("Category")["Product"].count()
plt.figure(figsize=(8, 6))
category_count.plot.pie(autopct="%1.1f%%", startangle=140, colors=plt.cm.Paired.colors)
plt.title("Product Category Sales Distribution (Count)")
plt.ylabel("")
plt.tight_layout()
plt.show()

customer_revenue = df.groupby("Customer_ID")["Revenue"].sum()
average_customer_spending = customer_revenue.mean()
print("\nCustomer Behavior Analysis:")
print(f"Average Spending Per Customer: ₹{average_customer_spending:,.2f}")

plt.figure(figsize=(8, 5))
customer_revenue.sort_values(ascending=False).plot(kind="bar", color="pink", edgecolor="black")
plt.title("Customer Spending")
plt.xlabel("Customer ID")
plt.ylabel("Revenue (₹)")
plt.xticks(rotation=0)
plt.tight_layout()
plt.show()
```

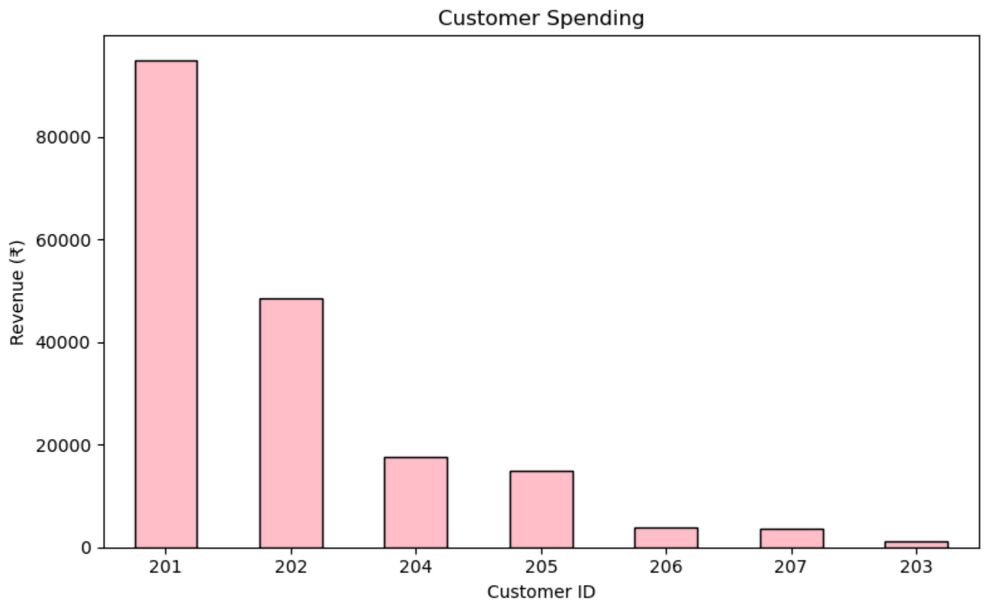
```
Total Revenue: ₹184,890.00
Average Order Value: ₹15,407.50
```


```
Top 5 Products by Revenue:
Product
Laptop        92000
Mobile        66250
Smartwatch    15000
Kurta Set      9050
Jeans          1340
Name: Revenue, dtype: int64
```






```
Customer Behavior Analysis:
Average Spending Per Customer: ₹26,412.86
```


Customer Spending
Revenue (₹)
0
20000
40000
60000
80000
201
202
204
205
206
207
203
Customer ID

In [ ]: