

## Data Integrity

- The datasets (Customers, Products, and Transactions) are clean and well-maintained, with no missing or duplicated values. This ensures the accuracy and reliability of subsequent analyses.

## Customer Distribution by Region

- The distribution of customers across regions provides insights into market penetration:
  - **South America:** 59 customers
  - **Europe:** 50 customers
  - **North America:** 46 customers
  - **Asia:** 45 customers
- This data shows relatively even distribution, with South America having the highest number of customers. Companies might focus on expanding their reach in Asia and North America to balance regional customer bases.

## Product Categories Distribution

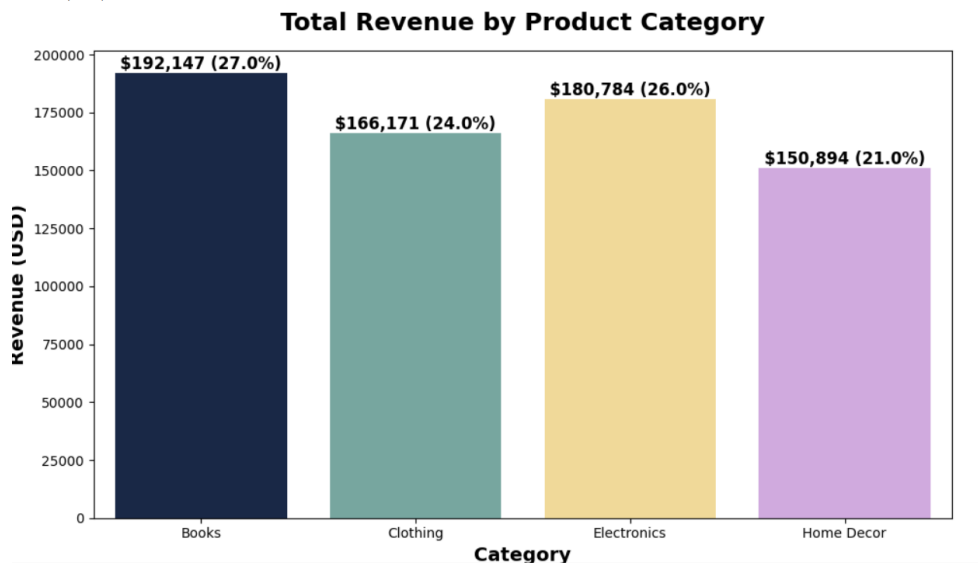
- The dataset includes four major product categories with nearly equal representation:
  - **Books:** 26 products
  - **Electronics:** 26 products
  - **Clothing:** 25 products
  - **Home Decor:** 23 products

## Transactions Per Month

- The monthly distribution of transactions reveals trends over time:
  - **Highest Transactions:** January 2024 (107 transactions) and July to September 2024 (96 transactions each).
  - **Lowest Transactions:** December 2023 (4 transactions).
- The sharp increase from December 2023 to January 2024 indicates seasonal demand, possibly influenced by post-holiday sales. The steady performance from July to September suggests a sustained period of high activity, which could be due to marketing campaigns, product launches, or other factors.

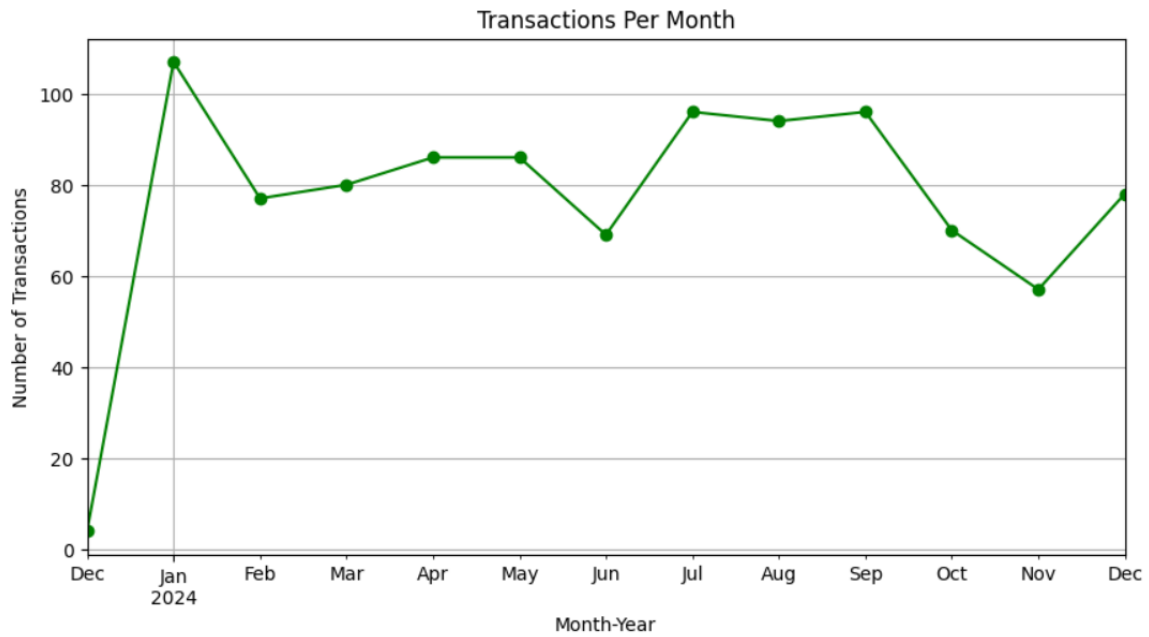
## Correlation Matrix

- The relationships between key numerical variables provide meaningful insights:
  - **Quantity and Total Value:** A strong positive correlation (0.61) suggests that higher quantities sold result in higher total revenue.
  - **Price and Total Value:** A strong positive correlation (0.72) highlights the importance of pricing strategies in revenue generation.
  - **Quantity and Price:** A negligible correlation (-0.009) implies that product price does not significantly affect the quantity sold, indicating potential pricing flexibility.



## Total Revenue by Category

- The total revenue generated by each product category highlights top-performing segments:
  - **Books:** \$192,147.47
  - **Electronics:** \$180,783.50
  - **Clothing:** \$166,170.66
  - **Home Decor:** \$150,893.93
- Books lead the revenue contribution, followed by Electronics. The relatively high revenue from Books could be due to consistent demand or higher margins in this category.
- **Books Dominate:** The Books category generates the highest revenue, accounting for 27% of the total.
- Electronics follows closely with 26% revenue share.
- Clothing and Home Decor contribute 24% and 21%, respectively, indicating lower revenue compared to Books and Electronics.
- There's a noticeable revenue gap between the top two categories (Books & Electronics) and the bottom two (Clothing & Home Decor).
- Categories with lower revenue share (Clothing & Home Decor) present opportunities for growth.



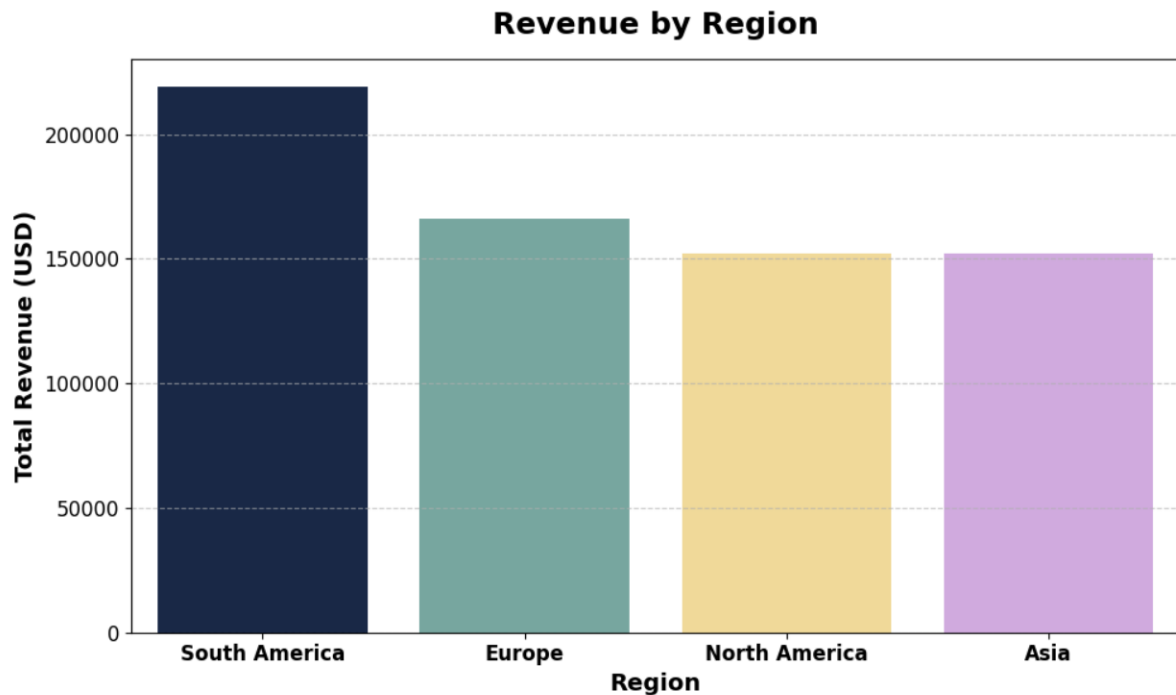
- The year begins with a high transaction volume in January, reaching 100 transactions.
- The number of transactions then decreases in February and remains stable in March and April.
- A slight increase occurs in May, followed by a more significant rise in June and July, reaching a peak of 95 transactions in July.
- From July onwards, the number of transactions starts to decline, reaching a low of 60 transactions in October.
- The number of transactions increases again in November and remains steady in December.

```
sns.barplot(x=customers_by_region.index, y=customers_by_region.values, palette='viridis')
```



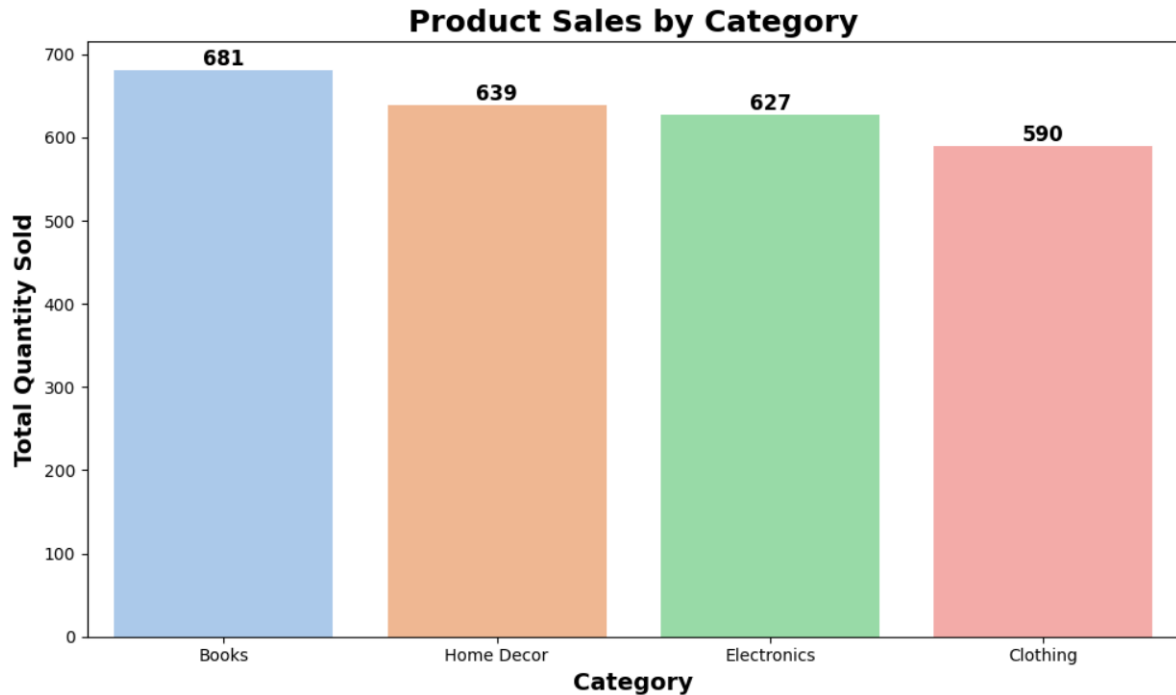
- South America has the highest number of customers, represented by the tallest bar in the chart.
- Europe has the second-highest number of customers, with a bar noticeably shorter than South America's.
- Both North America and Asia have similar customer counts, with their bars being almost equal in height.
- The number of customers ranges from the highest in South America to the lowest in Asia.
- The use of different colors for each region makes it easy to visually distinguish between the regions and their respective customer counts.
- The bar chart effectively represents the categorical data (regions) and their corresponding numerical values (number of customers).
- This chart could be used to inform marketing strategies, resource allocation, and business decisions by highlighting the regions with the highest customer concentration.

```
sns.barplot(
```



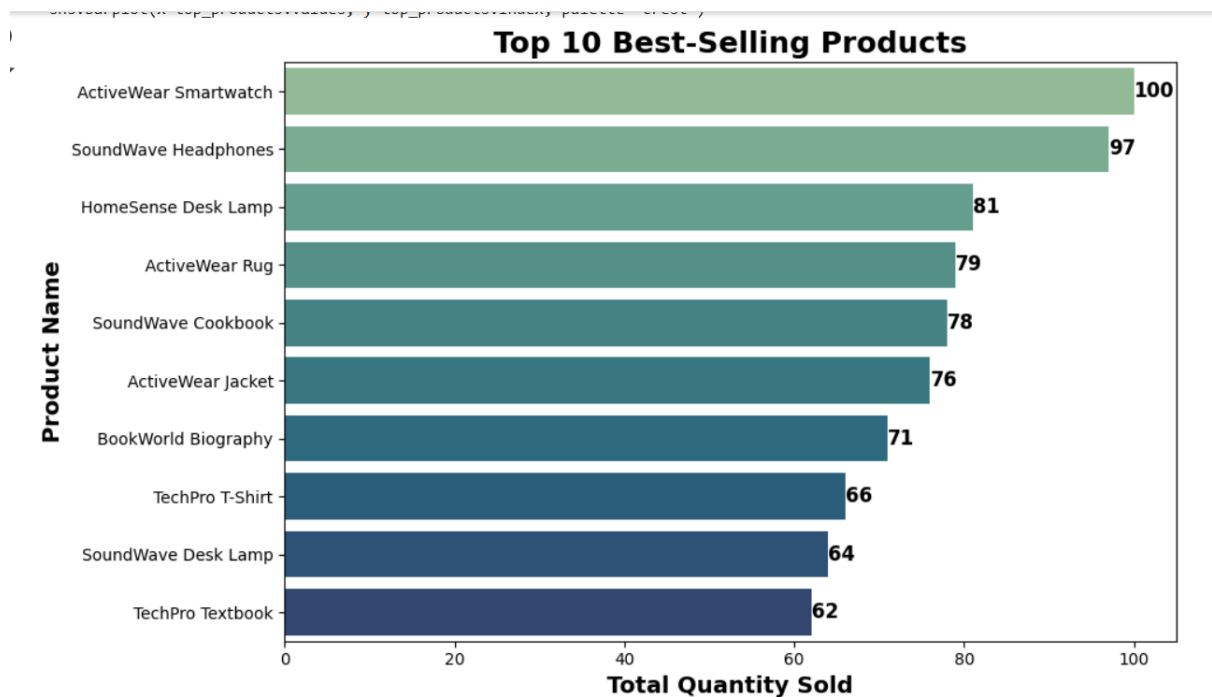
- **Revenue Disparity:** The chart clearly shows a significant difference in revenue generated across the four regions. South America leads with the highest revenue, followed by Europe, North America, and Asia.
- **South America Dominance:** South America stands out as the top revenue-generating region, significantly surpassing the other regions.
- **European Market:** Europe holds the second position in terms of revenue, suggesting a considerable market presence in that region.
- **North America and Asia:** These regions demonstrate comparable revenue levels, both contributing less than South America and Europe.

```
sns.barplot(x=sales_by_category.index, y=sales_by_category.values, palette="pastel")
```



### Most Popular Products (Top 5 by Quantity Sold)

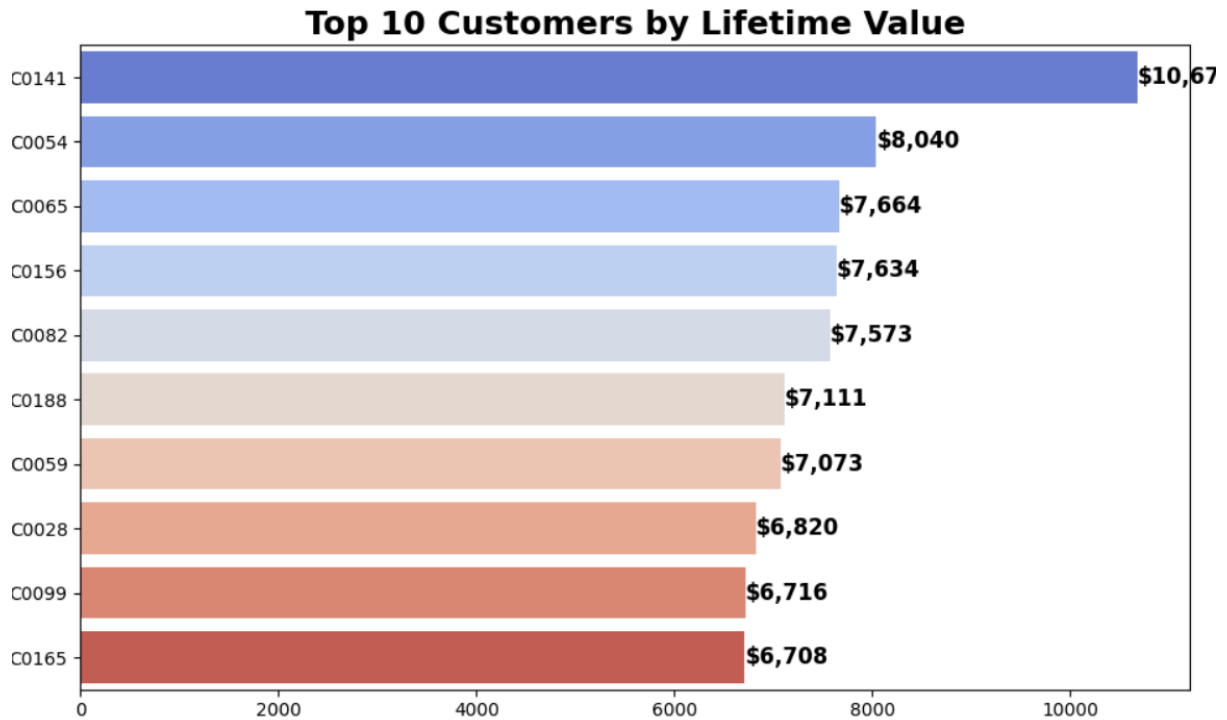
- The products with the highest sales volume are:
  - **ActiveWear Smartwatch:** 100 units
  - **SoundWave Headphones:** 97 units
  - **HomeSense Desk Lamp:** 81 units
  - **ActiveWear Rug:** 79 units
  - **SoundWave Cookbook:** 78 units
- The dominance of tech-oriented products like smartwatches and headphones suggests a strong demand for electronics. Expanding the Electronics category with similar products could capture additional market share.
- Books are the clear leader in terms of sales quantity, with 681 units sold.
- **Clothing Sales:** Clothing appears to have the lowest sales quantity among the four categories, with 590 units sold.



- **ActiveWear Smartwatch is the Top Seller:** It stands out with the highest quantity sold at 100 units.
- **SoundWave Products are Popular:** Two SoundWave products (Headphones and Desk Lamp) are among the top sellers, indicating strong brand recognition or customer appeal for this product line.
- **ActiveWear Brand Success:** Besides the smartwatch, the ActiveWear Rug and Jacket also rank high, suggesting a strong market for this brand's activewear products.
- **TechPro and BookWorld Representation:** Both brands have two products in the top 10, demonstrating their consistent performance across different product categories.



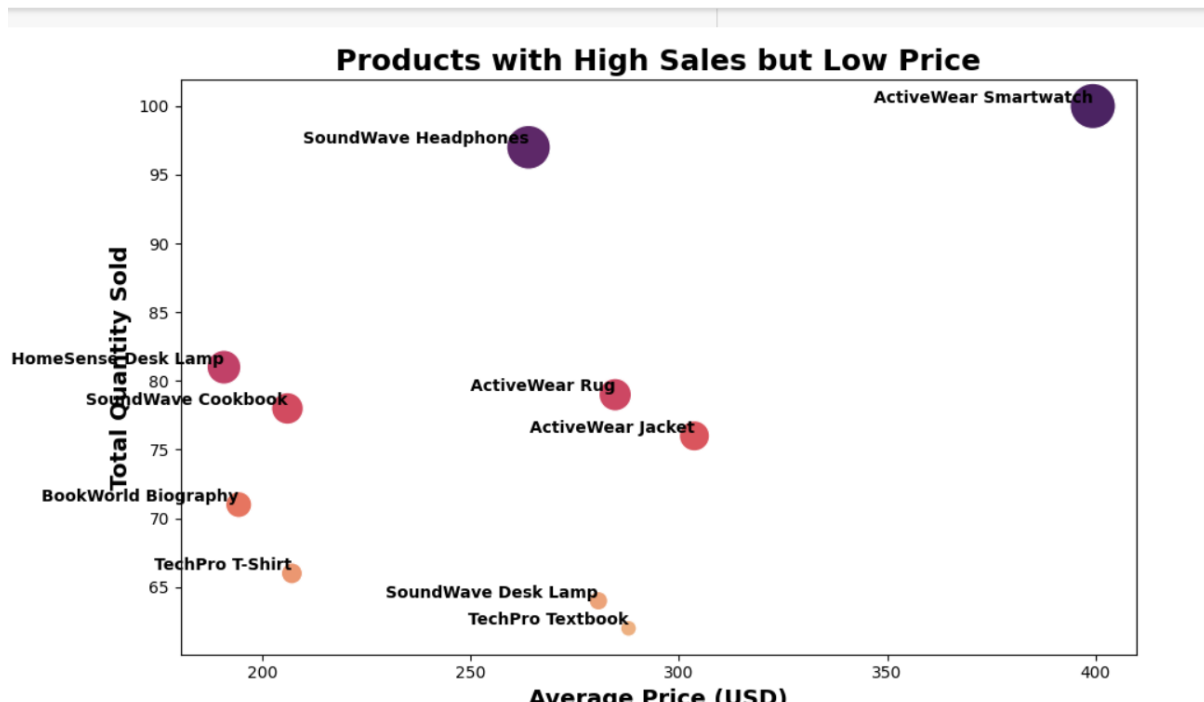
```
..barplot(x=top_customers.values, y=top_customers.index, palette= coolwarm )
```



- **High-Value Customers:** Customer C0141 stands out as the most valuable customer with a lifetime value exceeding \$10,000.
- **Value Range:** The lifetime value of the top 10 customers spans a range from approximately \$6,700 to over \$10,000, indicating a significant difference in their value to the business.



- The data shows an overall upward trend in revenue from 2023-12 to 2024-12, indicating growth in the business.
- While there is an upward trend, the revenue also exhibits fluctuations month-over-month. For example, there's a significant drop in revenue from 2024-01 to 2024-02, followed by a spike in 2024-03.
- The highest revenue month is 2024-07, while the lowest is 2024-10.
- The data may suggest seasonal patterns, with peaks and troughs occurring at certain times of the year. However, more data points are needed to confirm this.
- The chart shows a notable increase in revenue in the last two months (2024-11 and 2024-12), indicating positive momentum in the business.



- **High Sales, Low Price:** Products like SoundWave Headphones and ActiveWear Smartwatch are positioned in the top-left corner, indicating high sales and low prices.
- **Moderate Sales, Moderate Price:** Products like HomeSense Desk Lamp, SoundWave Cookbook, and ActiveWear Rug fall in the middle, with moderate sales and prices.
- **Lower Sales, Higher Price:** Products like TechPro T-Shirt, SoundWave Desk Lamp, TechPro Textbook are positioned towards the right, indicating lower sales and higher prices.
- The size of the dots might represent the actual sales volume, with larger dots indicating higher sales quantities.