

Data Science Intern Assignment | Zeotap

Task 1: Exploratory Data Analysis (EDA) and Business Insights

Github- https://github.com/PriyabrataBehera-24/eCommerce_Transactions

Github code - [https://github.com/PriyabrataBehera-24/eCommerce_Transactions/blob/main/Priyabrata Behera EDA.ipynb](https://github.com/PriyabrataBehera-24/eCommerce_Transactions/blob/main/Priyabrata%20Behera%20EDA.ipynb)

EDA Data Analysis

1. Data Loading and Preprocessing

- **Libraries Used:** pandas, matplotlib, seaborn.
- **Datasets Loaded:** Customers, Products, and Transactions.
- **Date Conversion:** The 'SignupDate' and 'TransactionDate' columns were converted to datetime objects.
- **Data Cleaning:** No missing values were found across the datasets, ensuring data quality for analysis.

.

2. Descriptive Statistics

- The dataset includes key metrics such as quantity, total transaction value, and product prices.
- **Key Statistics:**
 1. **Quantity:** Mean: 2.54, Min: 1, Max: 4
 2. **Total Value:** Mean: \$689.99, Min: \$16.08, Max: \$1991.04.
 3. **Price:** Mean: \$272.55, Min: \$16.08, Max: \$497.76.

3. Data Visualizations

- Bar Chart: Total revenue by region, highlighting South America as the top performer.
- Bar Chart: Top 5 products by revenue, emphasizing Product P029's dominance.
- Pie Chart: Customer distribution by region, showing South America as the largest market.
- Line Chart: Monthly revenue trends, indicating seasonal sales peaks.

Business Insights

1. Total Revenue Generated: \$689995.56
2. Top 5 Products by Revenue:

ProductID

P029 19513.80

P079 17946.91

P048	17905.20
P020	15060.92
P062	14592.24

3. Regional Revenue Distribution:

Region

Asia	152074.97
Europe	166254.63
North America	152313.40
South America	219352.56

4. Customer Distribution by Region:

Region

South America	59
Europe	50
North America	46
Asia	45

5. Monthly Revenue Trend:

The monthly revenue shows seasonal peaks in specific months, indicating high sales periods.

1. Revenue Performance

The total revenue generated from transactions is \$689,995.56. South America is the leading region, contributing \$219,352.56 to overall revenue. This highlights South America as the most lucrative market, suggesting a focus on retention and expansion strategies.

2. Product Performance

The top 5 products (P029, P079, P048, P020, P062) collectively contributed significantly to total revenue, with Product P029 generating the highest revenue at \$19,513.80. Prioritizing marketing, inventory management, and bundling offers for these products could further boost sales.

3. Regional Insights

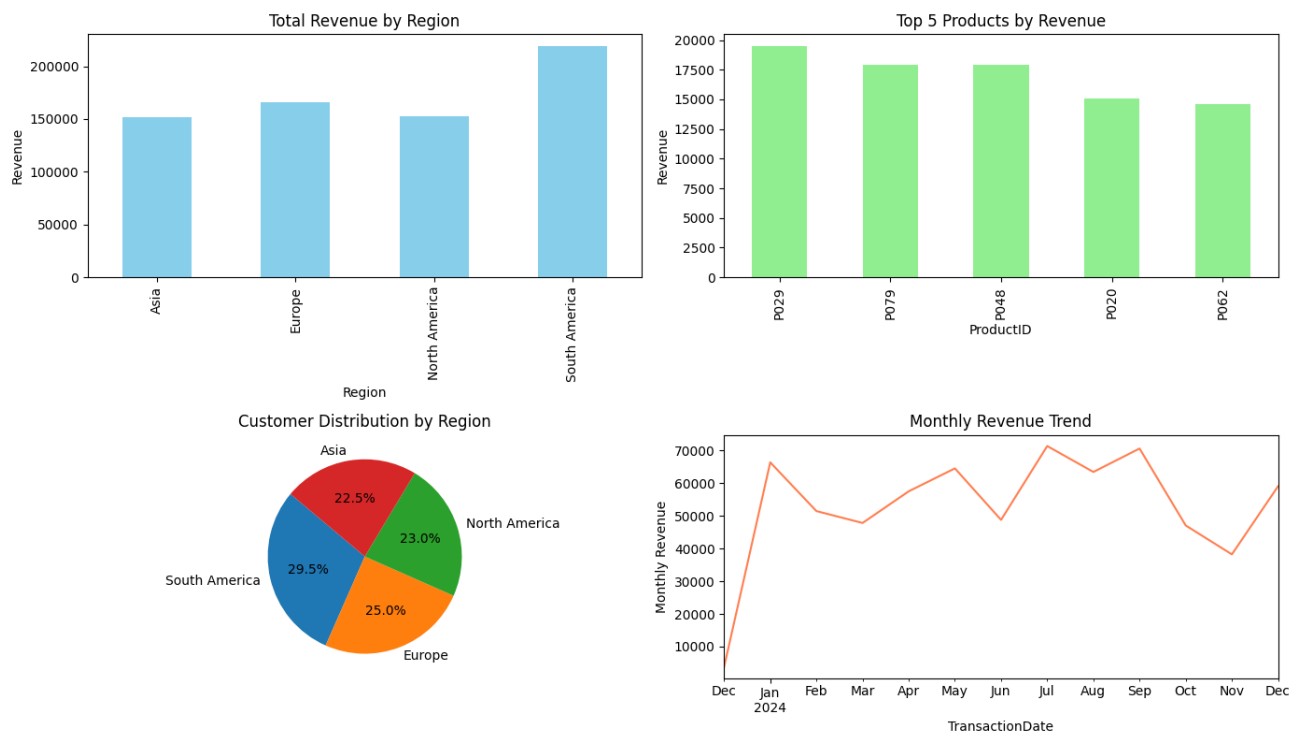
South America leads in customer concentration with 59 customers, followed by Europe (50), North America (46), and Asia (45). Customer-focused strategies in South America could maximize returns, while marketing efforts in underperforming regions like Asia and North America could enhance growth.

4. Seasonal Trends

Monthly revenue trends reveal seasonal sales peaks in specific months, indicating periods of high consumer demand. Aligning inventory management and marketing campaigns with these trends can optimize operational efficiency and maximize profits.

5. Customer Behavior

Customer distribution aligns with revenue contributions, with South America leading in both. Tailored retention strategies such as loyalty programs and personalized campaigns in South America can enhance customer satisfaction and lifetime value.



Recommendations

- Top Product Optimization:** Focus on promoting and optimizing inventory for the top-performing products to sustain and increase their revenue contribution.
- Regional Strategies:** Develop targeted marketing campaigns for South America while exploring growth opportunities in regions like Asia and North America.
- Seasonal Planning:** Prepare for high-demand periods by analyzing monthly trends and adjusting inventory, staffing, and marketing efforts accordingly.
- Customer Retention:** Leverage CRM tools to create personalized experiences for customers, especially in South America, to build loyalty and drive repeat purchases.
- Underperforming Markets:** Identify barriers in underperforming regions and develop strategies like localized promotions to address these challenges.

Conclusion

The analysis highlights the critical role of regional and product-specific strategies in driving business performance. South America's dominance in revenue and customer base positions it as the focus for retention and expansion. At the same time, leveraging insights into seasonal trends and customer behaviors across all regions can lead to comprehensive growth and enhanced profitability.