

# Derived Business Insights from EDA

## 1 Profit for each product category

- Books are the most profitable category, generating a total profit of 192147.47, likely due to steady demand and lower operational costs.
- Electronics 180783.50 and Clothing 166170.66 follow closely, reflecting their consistent demand and higher price points.
- Home Decor contributes 150893.93 in profit, indicating strong growth opportunities, particularly with targeted marketing and seasonal promotions.

## 2 Regions Contributing the Most Revenue

- South America is the top-performing region with a total revenue of 219352.56, indicating strong customer engagement and potential for expansion.
- Europe ranks second with 166254.63, suggesting a well-established market with consistent sales opportunities.
- North America 152313.40 and Asia 152074.97 generate nearly equal revenues, showcasing a balanced market demand in these regions.

## 3 Seasonality of transactions over months

- The company's revenue peaks in the month of July.

- The company's revenue is generally higher in the second half of the year than in the first half of the year.
- Overall, the graph shows that the company's revenue is seasonal, with peaks in the summer and troughs in the winter.

## 4 High-value customers top 10 percent

- It first calculates the total spending for each customer by grouping the data by "CustomerID" and summing up the "TotalValue" column.
- Then, it determines the threshold value for the top 10 percent by calculating the 90th percentile of the customer spending distribution.
- In essence, this code snippet provides a way to identify and analyze the top-spending customers within a given dataset, which can be valuable for targeted marketing and customer relationship management strategies.

## 5 Relationship between product price and quantity sold

- The data points are clustered in distinct horizontal bands. This indicates that there are specific price points at which the quantity sold remains relatively constant. It's possible that these price points correspond to common pricing strategies or consumer expectations.
- There doesn't appear to be a strong or consistent trend between product price and quantity sold. The data points are scattered without a clear upward or downward slope. This suggests that price alone might not be the primary factor influencing the quantity of a product sold. Other factors, such as product features, marketing, or consumer preferences, may play a more significant role.
- The correlation between price and quantity sold is negative but very weak -0.009378. The relationship between these is so weak that it's practically negligible.

## 6 Plot daily transactions

- Wednesday generates the highest revenue for the company.
- Saturday has the lowest revenue generation.
- The graph shows a declining trend in revenue from Wednesday to Saturday.
- The graph highlights that the company experiences its highest revenue generation mid-week and that revenue gradually decreases towards the weekend.
- 

## 7 Plot peak transaction times

- The graph clearly shows that the highest revenue is generated around the 14th and 16th hours of the day (likely 02:00 PM and 04:00 PM). This indicates that these time periods contribute most significantly to the overall revenue.
- Revenue levels dip slightly after the peak but remain substantial throughout the afternoon. It then gradually declines towards the evening hours.
- Revenue generation significantly drops during the morning hours (from approximately 05:00 AM to 13:00 PM), suggesting minimal revenue contribution during this period.