EDA ANALYSIS POINTS

The dataset given were customer.csv,product.csv and transaction.csv. Neither of them had any null values or duplicates.

So proceeding onto EDA on products.csv

Insights were:

- Checked for total products in each category which gave books and electronics with ,most 26 each. This was done because while analysing the revenue if any of the category leads it must be analysed how many products that category had.
- Then plotted a histogram to check the price range which most of popular product falls in terms of buying: (" 160-180 rs " ranged as the most popular bought range) which eventually gives the idea that expensive products remain as least bought considering a general community.
- Books category remained as the most expensive category with price nearly 500rs followed by clothing.
- Price distribution was analysed and we got similar histograms this suggests that the pricing top and bottom 100 products might have similar distributions, which could be unusual.
- BOOKS category is more in number along with electronics as they rule the store as they move very fast compared to other categories

  Now on customers.csv

- South America is having most customers: By analysing the plot we can clearly see South America leads the no of customers followed by Europe.
- With customer count of 59 south America region had most no of customers.
- 2024 AUG AND 2024 generated more no of signups in the store history this was evaluated using the plot for each month.

 Finally we merged the first two tables with third for final analysis as there value and total revenue columns here in transactions.csv

- South America generated more revenue reaching 2lakhs as we see they already had more no of customers.
- Sales peaked on 2024 july and 2024 september when we plotted the cost and month.
- Now to check specifically the customers c0156 was the top customer
- Finally to finalise the category which ruled the sales throughout we plotted the revenue by category : we got that books category lead approx. 1.9lakh sales it did.

FINAL INSIGHTS

- BOOKS is the main category in the store with highest revenue generated because as we already saw the expensive products are in books category.
- South America is the region which contributes more sales in the store with ,most no of customers.
- Considering whole data C0156 customer along with C0109 were the top consumers in the store who has most frequent buys.
- 2024 JULY and 2024 September remained highest grossing month ever in store history.
- 2024 aug and nov made most no of signups ever.