ZEOTAP DATA SCIENCE ASSESSMENT

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

1. Perform EDA on the provided dataset.
2. Derive at least **5 business insights** from the EDA.
   - Write these insights in **short point-wise sentences** (maximum 100 words per insight).

Deliverables:

- A Jupyter Notebook/Python script containing your EDA code.
- A PDF report with business insights (maximum 500 words).

After performing and applying the extensive EDA Steps on the given csv files, I have come across some of the following Business Insights which might be crucial for the Business:

BUSINESS INSIGHT 1: CUSTOMER DISTRIBUTION:

The customer base is predominantly concentrated in South America, with a large number of customers compared to other regions.

This uneven distribution suggests a high level of engagement in specific regions, while other regions may be underserved.

To optimize the market presence, the company can focus on increasing its footprint in underrepresented regions such as North America or Asia by leveraging targeted

advertising, localized marketing campaigns, and region-specific promotions.

Expanding to new geographic areas could also drive sales and balance regional growth across the customer base

BUSINESS INSIGHT 2: PRODUCT PRICING

The product price distribution is heavily right-skewed, with most products priced lower, but a notable portion of items priced at higher ranges.

The median product price is around $300, indicating that the company might be missing out on potential revenue from higher-priced products.

Introducing more premium-priced options could create additional revenue streams, especially in high-demand categories like electronics and luxury goods.

We can use a more dynamic approach towards pricing to ensure the normality.

BUSINESS INSIGHT 3 : PRODUCT CATEGORIES:

Books and Electronics stand out as the top-performing product categories, consistently driving sales.

The high demand for these categories presents a strategic opportunity for the company to capitalize on their popularity by expanding the product offerings in these segments.

For example, bundling related products like e-readers with books or accessories with electronics could increase the average transaction value.

Additionally, seasonal promotions or loyalty programs targeting customers who frequently purchase from these categories could further increase repeat purchases.

BUSINESS INSIGHT 4 : TRANSACTION VALUES

The dataset reveals a broad range of transaction values, from small purchases to large, high-value transactions.

A few high-value customers contribute disproportionately to the overall revenue, suggesting that focusing on these customers could yield substantial returns.

By identifying these high-value customers, the company can design loyalty programs, VIP memberships, and personalized offers to retain them.

Targeted marketing campaigns could focus on these segments to nurture them further, offering incentives like discounts, exclusive products, or early access to sales.

Additionally, understanding the purchasing patterns of these high-value customers can help replicate their behavior in other customer segments.

BUSINESS INSIGHT 5 : CUSTOMER SEGMENTATION

There are 4 clusters of customers that I could segment after performing k-means clustering and here's what I found through that:

Cluster 0 : Low Spend, Low Transaction Frequency

Cluster 1 : High-Spending, Loyal Customers

Cluster 2 : Moderate Spend, High-Transaction Frequency

Cluster 3 : Moderate Spend, Low Transaction Frequency

By identifying and understanding these distinct customer segments, the company can design marketing strategies that cater specifically to each group.

This targeted approach can improve customer retention, increase sales, and ultimately enhance overall profitability.

After performing EDA to the best of my knowledge, these were the insights that I could find.

THANK YOU,

Sushrith Reddy.