PROBLEM STATEMENT

Ace Supermarket is a mall that is currently in debt and has realized that sales must grow to boost revenue. To increase sales, they want to campaign targeted ads at top-buying customers. Their definition of top-buying customers is an individual with a higher annual income than the average and a higher spending score than three-fourths of the database’s total spending score.

Additionally, they added that getting the gender, age, annual income, and spending income of such customers will be instrumental to them in the final proposal we come up with. From the requirements specified, it is evident the segmentation to be performed on this challenge will be demographic segmentation.

The main dataset sent by the mall to make predictions from has 5 columns namely; CustomerID, Gender, Age, Annual Income, and Spending Score. The dataset can be found here; [Mall Dataset](https://github.com/Manlikebigb/Mall-Customer-Segmentation/blob/main/Main%20Mall%20Data.csv).

This analysis is done with Microsoft SQL Server Management Studio.

ANALYSIS BREAKDOWN AND INSIGHT

The SQL Query Solution file for this analysis can be found here; [Mall Customer Segmentation](https://github.com/Manlikebigb/Mall-Customer-Segmentation/blob/936cc0eeb6e3c66b7483a7afaef4fbb5e1e2a6ff/Mall%20Customer%20Segmentation.ssmssln)

I wrote an SQL query to extract the customers with an annual income greater than the average annual income of all customers in the data table.

I also wrote an SQL query to extract the customers with a spending score greater than three-fourths of the overall spending score of all customers.

1. I combined both queries as this would give us the Top-Buying Customers for the mall based on the requirements given by Ace Mall. I realised there were a total of 29 top-buying customers with 13 being male and 16 being females.
2. I broke this gender segmentation into age segments of 20s, 30s, and 40s using a CASE query. This revealed that most of the top-buying customers are women and men in their thirties totalling 22 of the 29 top-buying customers.
3. I segmented the customers into High, Medium, or Low-income customers. There were 10 Middle-Income earners, 1 High-Income earner, and 18 Low-Income earners; this aligns with the count of 29 top-buying customers and shows 11 customers are above the average income level.
4. A further breakdown into 4 unique customer segments was analyzed. The segments are:

* P1 (high annual\_income, high spending\_score)
* P2 (low annual\_income, high spending\_score)
* P3 (high annual\_income, low spending\_score)
* P4 (low annual\_income, low spending\_score)

where high means the customer is above average (medium or high earner or spender). We also ran a count showing 52 P1 shoppers, 50 P2 shoppers, 56 P3 shoppers, and 42 P4 shoppers.

1. The best customers to target are in P1 and a higher ad budget should be allocated to targeting these high-income, high-spending customers.
2. P2 is the next best to target while P3 can be further researched and analysed to find out why they earn so high but spend so little.
3. P4 should have the least budget and focus.