# **Analysis Report**

**Data preparation**

1. Additional columns that be added from the data for analysis
   1. Clean the enroll\_date column / find argument to read date
   2. Add Age column using Year\_birth
   3. Classify the person as teenager, adult, senior citizen using Age column.
   4. Add total spend using sum of wines, meat … columns.
   5. Created a new column Enrollment\_Age in years using Dt\_Customer and Birth\_Year
   6. What is Z\_CostContact? And Z\_Revenue?

**To find**

1. Using the data and analysis find
   1. Biggest buyer
   2. Frequent buyer
   3. Low buyer
   4. Non buyer
2. Which customers (marital status) are churning?
3. Who to send campaign emails to?
4. Category-wise buying trend.
5. Relationship between income and alcohol and marital status and alcohol consumption?

GPT Guidance:  
  
Certainly, I'm here to help you with your analysis. Let's go through the columns you provided one by one, and I can provide guidance on how to approach each of them for your customer behavior analysis project.

1. \*\*ID:\*\* This seems to be a unique identifier for each customer. It will be important for tracking individual customer behavior and transactions.

2. \*\*Year\_Birth:\*\* This column likely represents the birth year of each customer. You can use this information to calculate the age of each customer, which might be useful in your analysis.

3. \*\*Education:\*\* This represents the educational level of the customers. You could categorize this data to understand the education level distribution among your customers.

4. \*\*Marital\_Status:\*\* Indicates the marital status of the customers. You can analyze how marital status might influence purchasing behavior.

5. \*\*Income:\*\* This column represents the income of the customers. Analyzing income levels can provide insights into the spending capacity of different customer segments.

6. \*\*Has\_Kid:\*\* Indicates whether the customer has kids. This binary variable can be useful for segmenting customers based on family status.

7. \*\*Has\_Teen:\*\* Indicates whether the customer has teenagers in the family. Similar to "Has\_Kid," this variable helps in understanding the family structure of the customers.

8. \*\*Dt\_Customer:\*\* This column likely represents the date when the customer was added to the database. You can analyze customer acquisition trends over time.

9. \*\*Recency:\*\* Represents how recently a customer made a purchase. This could be important for customer segmentation and targeting.

10. \*\*Wines, Fruits, MeatProducts, FishProducts, SweetProducts, GoldProds:\*\* These columns likely represent the quantity of different product categories purchased by customers. You can analyze the purchasing patterns for each product category.

11. \*\*NumDealsPurchases, NumWebPurchases, NumCatalogPurchases, NumStorePurchases:\*\* These columns represent the number of purchases made through different channels. Analyzing these columns can provide insights into customer preferences for shopping methods.

12. \*\*NumWebVisitsMonth:\*\* Indicates how many times the customer visited the website in a month. This can be important for understanding online engagement.

13. \*\*AcceptedCmp3, AcceptedCmp4, AcceptedCmp5, AcceptedCmp1, AcceptedCmp2:\*\* These columns likely represent whether a customer accepted different marketing campaigns. Analyzing these columns can help assess the effectiveness of marketing strategies.

14. \*\*Complain:\*\* Indicates whether a customer made a complaint. Monitoring this can provide insights into customer satisfaction and service quality.

15. \*\*Z\_CostContact, Z\_Revenue:\*\* These columns seem to have constant values and might not provide meaningful information for your analysis.

16. \*\*Response:\*\* Represents whether the customer responded to the marketing campaign. Analyzing this column can help evaluate campaign success rates.

For a successful analysis, you might want to:

- \*\*Clean Data:\*\* Ensure that data is clean, handle missing values if any, and resolve inconsistencies.

- \*\*Feature Engineering:\*\* Create new features if necessary. For example, you can calculate customer age from the birth year.

- \*\*Exploratory Data Analysis (EDA):\*\* Visualize data distributions, correlations, and trends. EDA will help you understand your data better.

- \*\*Customer Segmentation:\*\* Segment customers based on demographics, purchasing behavior, or response to campaigns.

- \*\*Predictive Modeling:\*\* If you have a specific goal, like predicting customer response to future campaigns, you might consider building predictive models.

Feel free to ask more specific questions about any of these steps or any particular analysis you're interested in performing!