Task 1: Data Cleaning and Preprocessing (Excel-Based Report)

Objective: The main objective of this task was to clean and prepare the **Customer Personality Analysis** dataset for analysis using Microsoft Excel. The dataset contained raw marketing data including customer demographics, spending patterns, and responses to campaigns. The goal was to ensure the data was accurate, consistent, and ready for analysis by removing errors, handling missing values, and formatting inconsistencies.

Dataset Used: Customer Personality Analysis (marketing_campaign.csv)
Total Records: 2,240 rows   **Total Columns:** 29

Tool Used: Microsoft Excel

Data Cleaning Summary:

The dataset was cleaned thoroughly in Microsoft Excel to remove inconsistencies and ensure accuracy. The following steps were performed:

1. 1. Removed all duplicate records using Excel's 'Remove Duplicates' feature to maintain data uniqueness.
2. 2. Identified missing values using filters; numeric fields such as 'Income' were filled with the average value, while categorical fields were filled using the most frequent (mode) value.
3. 3. Standardized text entries by correcting inconsistent capitalization (e.g., Education and Marital Status categories).
4. 4. Converted date formats in the 'Dt_Customer' column into a consistent format (dd-mm-yyyy) for uniformity.
5. 5. Renamed columns to more readable names using underscores and lowercase letters (e.g., 'Year Birth' → 'year_birth').
6. 6. Verified data consistency using conditional formatting, logical checks, and data filters to ensure accuracy across all records.

Results and Insights:

After cleaning, the dataset became well-structured and free from errors. All 24 missing income values were handled, date formats standardized, and column headers unified. This improved dataset now provides an accurate foundation for further analysis such as customer segmentation, targeted marketing, and behavior pattern recognition. It is now ready for visualization and model building tasks.

Conclusion:

The **Customer Personality Analysis** dataset has been successfully cleaned and organized using Microsoft Excel. Through systematic handling of missing values, duplicates, and inconsistent formats, the data is now complete, reliable, and ready for in-depth marketing insights and data analysis.