**Marketing Optimization through Customer Personality Analysis**

(Alessio Baldini, Florian Derchain, Antonella Mele, Sofia-Zoi Sotiriou)

**Objective:**

The aim is to optimize marketing strategies by identifying customer segments based on preferences and personal characteristics. This approach seeks to categorize each client into ideal clusters aligned with company products, enabling more precise targeting and efficient budget allocation in marketing efforts.

**Approach and Dataset:**

Using a dataset of over 2,000 customer profiles, we analyze correlations between static characteristics (demographics, income, etc.) and purchasing behavior. This analysis involves four key segments:

**1. People**

◦ **ID (Integer):** Unique identifier for each customer, essential for individual-level tracking. ◦ Year_Birth (Integer): Customer's birth year, useful for calculating age.
◦ **Education (Categorical):** Education level (e.g., PhD, Master's, Bachelor), potentially affecting purchasing behavior.
◦ **Marital_Status (Categorical):** Marital status, indicating lifestyle attributes.
◦ **Income (Continuous):** Annual income, vital for segmenting by spending power.
◦ **Kidhome and Teenhome (Integer):** Number of children or teenagers in the household, which may influence family-related purchasing.

**2. Products**

◦ **MntWines** *(Expectancy = 303,9)***, MntFruits** *(Expectancy = 26,3)***, MntMeatProducts** *(Expectancy = 166,9)***, MntFishProducts** *(Expectancy = 37,5)***, MntSweetProducts** *(Expectancy = 27,1)***, MntGoldProds** *(Expectancy = 44)* **(Continuous):** Total spending in each product category, reflecting customer preferences and spending habits. Those variables are especially important to determine if

**3. Promotion**

◦ **AcceptedCmp1 to AcceptedCmp5 (Binary):** Indicators of response to specific campaigns (1 = accepted, 0 = not accepted), helping measure campaign success. ◦ Response (Binary): General campaign response, showing overall promotional effectiveness per customer.

**4. Place**

◦ **NumWebPurchases, NumCatalogPurchases, NumStorePurchases (Integer):** Purchase counts by channel, offering insights into preferred shopping methods.
◦ **NumWebVisitsMonth (Integer)**: Frequency of website visits in the past month, indicating digital engagement.

**Analysis on the amount of products :**

We see that the expectancies of said purchased products may vary but it can be separated between two ranges : 25-45 for fruits, fish products, sweet products and gold products, can be considered as "minor" part of the budget as far as expectancy is concerned in this

dataset, whereas the two remaining : wines and meat products are “major” part of the budget for most people.

However, within those we can see that the people registered in the dataset have different ways of consuming since the standard deviation for each variable is quite high compared to the expectancy.

For the “minor part of budget” variables :
-Fruits (Expectancy = 26,3 Std = 39,8)
-Fish Products (Expectancy = 37,5 Std = 54,6)
-Sweet Products (Expectancy = 27, Std = 41,3)
-Gold Products (Expectancy = 44, Std = 52,2)

And the “major” ones :
-Wines (Expectancy = 303,9, Std = 336,6)
-Meat Products (Expectancy = 166,9, Std = 225,7)

Thus, we can see the importance of clustering the dataset considering the large discrepancy that the standard deviation shows when taking two random individuals within the dataset as far as their consuming habits are concerned.

**Goal:**
Using a range of analytical techniques—including clustering, correlation analysis, and predictive modeling—we aim to create detailed customer segments with distinctive attributes tied to purchase behaviors, demographic factors, and marketing responses. This segmentation will allow the company to design a data-driven marketing strategy that optimally targets each customer segment with relevant offers, enhances personalization, and improves conversion rates.