# Analyzing Marketing Campaign Effectiveness and Customer Behavior at Maven Marketing

## Background

The dataset for this project comes from Maven Marketing, which includes marketing campaign data for 2,240 customers. This data encompasses various aspects, including customer profiles (age, education, income), product preferences, campaign successes and failures, and channel performance (web, store, catalog). The goal of this analysis is to provide actionable insights into customer behavior, campaign performance, product success, and channel efficiency to help improve future marketing strategies and optimize overall marketing efforts. By analyzing historical data, we aim to identify trends, understand the factors influencing customer actions, and evaluate the effectiveness of different marketing channels.

## Problem Statement

The goal of this analysis is to examine the marketing campaign data of 2,240 customers from Maven Marketing to uncover key insights regarding customer behavior, campaign performance, and channel effectiveness. We aim to identify factors influencing web purchases, evaluate the success of different campaigns, understand customer profiles, and pinpoint top-performing products and underperforming channels. These insights will help optimize marketing strategies and improve business outcomes.

## Data Overview

The dataset consists of marketing campaign data for 2,240 customers of Maven Marketing, with the following key attributes:

- **Customer Demographics**: Includes columns like age, education, marital status, income, number of children, and country.
- **Customer Behavior**: Tracks purchase behaviors, including number of web, catalog, store purchases, and web visits per month.
- **Campaign Performance**: Columns related to customer response to specific marketing campaigns (e.g., AcceptedCmp1 through AcceptedCmp5), campaign success (Response), and complaints (Complain).
- **Sales Data**: Contains monetary values for products purchased, such as MntWines, MntFruits, MntMeatProducts, etc.
- **Date & Time**: Customer registration date (Dt_Customer) and recency of interaction (Recency).

## Methodology

1. **Data Understanding**:
    - **Exploration**: Initially, the dataset was explored to understand its structure and contents. Key attributes were identified, including customer demographics, product purchases, and campaign responses.
    - **Data Profiling**: A descriptive analysis was performed to identify basic statistics (e.g., mean, median, mode, etc.) and determine data distribution, identifying columns that could contain missing or incorrect data.

2. **Data Cleaning**:
    - **Handling Missing Data**: Null values were handled, particularly in the 'Income' column, by imputing the mean value for missing entries.
    - **Outlier Detection**: Outliers in numeric fields such as income and purchase amounts were identified and addressed using statistical methods such as Z-scores and IQR.
    - **Data Transformation**: Some columns were converted to appropriate data types (e.g., datetime format for Dt_Customer, categorical encoding for Education, Marital_Status, etc.).
3. **Data Manipulation**:
    - **Feature Engineering**: New features were created based on existing columns, such as age groups or purchase categories, to enhance the analysis.
    - **Aggregating Data**: The data was aggregated at customer level to summarize overall purchasing behavior, campaign responses, and other metrics.
4. **Data Analysis**:
    - **Trend Analysis**: Statistical methods and visualizations were used to identify trends and relationships, such as factors that influence the number of web purchases and successful campaign responses.
    - **Segmentation**: Customer profiling was done to create segments based on demographics and purchase behavior, such as identifying the "average" customer based on key features.
    - **Campaign Performance Evaluation**: Marketing campaigns' success rates were analyzed to determine which campaigns yielded the best responses and drove the highest engagement.
5. **Data Visualization**:
    - **Visualizing Key Metrics**: Visualizations like bar charts, histograms, and heatmaps were used to present campaign success rates, product performance, and other key findings.
    - **Channel Performance Visualization**: Graphs were used to identify the performance of various marketing channels (e.g., web, catalog, store) and highlight areas that need improvement.

## Technical Processes

- Utilize Pandas Data Frame for data manipulation and analysis.
- Use Matplotlib and Seaborn for data visualization.
- Handle missing data using Pandas methods like fillna()
- Detect and manage outliers using statistical methods with Pandas and NumPy.

## Key Findings

1. **Null Values and Outliers**:

    - 24 null values were found in the Income column. These were handled by imputing the mean value to avoid skewing the analysis.
2. **Factors Influencing Web Purchases**:
    - **Income**: Customers with higher income levels are more likely to make web purchases, as shown by the correlation between Income and NumWebPurchases.
    - **Education**: Customers with higher education levels (particularly PhDs and Masters) show a higher number of web purchases, indicating a potential target for web-based campaigns.
    - **Age**: Age groups between 46 and 63 years have the highest number of web purchases, indicating that middle-aged customers are more likely to engage in online shopping.
3. **Most Successful Campaign**:
    - The marketing campaign with the highest success rate was **Response**, which recorded 334 successful responses. This indicates the effectiveness of the general campaign approach over specific ones like AcceptedCmp1 to AcceptedCmp5.
4. **Average Customer Profile**:
    - The average customer is typically **56 years old**, has a **Graduation** level of education, and falls into a moderate to high-income bracket.
5. **Top Performing Products**:
    - **MntWines** (Wine) had the highest sales, with **680,816 units sold**.
    - **MntMeatProducts** and **MntGoldProds** also performed well, indicating that food and luxury products are key drivers of sales.
6. **Underperforming Channels**:
    - **NumDealsPurchases** and **NumCatalogPurchases** were among the weakest-performing channels, with only 5,208 and 5,963 purchases, respectively.
    - **NumStorePurchases** and **NumWebVisitsMonth** were higher, but still below the expected threshold for optimal channel performance.

## Recommended Analysis

1. Are there any null values or outliers? How will you handle them?
2. What factors are significantly related to the number of web purchases?
3. Which marketing campaign was the most successful?
4. What does the average customer look like?
5. Which products are performing best?
6. Which channels are underperforming?

## Conclusion

The analysis of Maven Marketing's campaign data has provided valuable insights into customer behavior, campaign effectiveness, product performance, and channel efficiency. By identifying key factors influencing web purchases, profiling the average customer, and evaluating the success of various marketing campaigns, we have highlighted areas for

improvement and optimization. The data revealed that higher-income customers, those with advanced education, and middle-aged individuals are more likely to make online purchases, while certain marketing campaigns and product categories, such as wines and meat products, are more successful than others. Additionally, underperforming channels like catalog and in-store purchases suggest a shift toward enhancing digital and web-based strategies. These findings will help Maven Marketing make informed decisions, optimize future campaigns, and improve overall business outcomes.

**Project Owner**
**Name:** Vikant Kumar
**Date:** 17-07-2024