

Marketing Analytics EDA

<https://drive.google.com/file/d/1XYj3hcJc9aTLPDFJ5cWsbufGmYEOEO/view?usp=sharing>

This is a CSV file of 2240 observations (customers) with 28 variables related to marketing data. More specifically, the variables provide insights about:

- Customer profiles
- Products purchased
- Campaign success (or failure)
- Channel performance



Data Dictionary

- ID: Customer's unique identifier
- Year_Birth: Customer's birth year
- Education: Customer's education level
- Marital_Status: Customer's marital status
- Income: Customer's yearly household income
- Kidhome: Number of children in customer's household
- Teenhome: Number of teenagers in customer's household
- Dt_Customer: Date of customer's enrollment with the company
- Recency: Number of days since customer's last purchase
- MntWines: Amount spent on wine in the last 2 years
- MntFruits: Amount spent on fruits in the last 2 years
- MntMeatProducts: Amount spent on meat in the last 2 years
- MntFishProducts: Amount spent on fish in the last 2 years
- MntSweetProducts: Amount spent on sweets in the last 2 years
- MntGoldProds: Amount spent on gold in the last 2 years
- NumDealsPurchases: Number of purchases made with a discount
- NumWebPurchases: Number of purchases made through the company's web site
- NumCatalogPurchases: Number of purchases made using a catalogue
- NumStorePurchases: Number of purchases made directly in stores
- NumWebVisitsMonth: Number of visits to company's web site in the last month
- AcceptedCmp3: 1 if customer accepted the offer in the 3rd campaign, 0 otherwise
- AcceptedCmp4: 1 if customer accepted the offer in the 4th campaign, 0 otherwise
- AcceptedCmp5: 1 if customer accepted the offer in the 5th campaign, 0 otherwise

- AcceptedCmp1: 1 if customer accepted the offer in the 1st campaign, 0 otherwise
- AcceptedCmp2: 1 if customer accepted the offer in the 2nd campaign, 0 otherwise
- Response: 1 if customer accepted the offer in the last campaign, 0 otherwise
- Complain: 1 if customer complained in the last 2 years, 0 otherwise
- Country: Customer's location

Problem Statement

You're a marketing analyst and you've been told by the Senior Marketing Manager that recent marketing campaigns have not been as effective as they were expected to be. You need to analyze the data set in order to understand this problem and propose data-driven solutions.

Note: You have to explain all the questions using graphs and charts.

Download the dataset from the above link and perform the following operations:

1. Import necessary libraries and load the dataset and display random 5 samples. Check the info of the data and write your findings.
2. Check the following using an appropriate method and write your findings
 - a. Check how spread out or varied your data set is.
 - b. Check where the middle 50% of your data lies.
 - c. Check boundaries for the lower, middle and upper quarters of data.
3. Check for any missing values in the dataset and handle them using an appropriate method.
4. Check for any presence of special characters in any variables. If present, clean/replace and change the datatype of the variable if required.
5. The Marketing Manager wants to know the 'Age' of the customers. Extract the age feature from the given dataset and display the statistical summary of the age?
(Statistical Summary: Show the dispersion of the data in the form of graphs.
Perform outlier treatment
Categorize data in Age Groups)
6. The Marketing manager wants to understand the total amount spent on various products so that we can find what percentage of the amount is spent on which product.
 - a. Find out the total amount spent by a customer.
 - b. Display the Percentage of the amount spent on Wines and other products.
7. Being an Analyst understands the total number of purchases made through different channels which can help find the percentage these channels contribute.
 - a. Find out the total purchases done by a customer through different channels.
 - b. Display the percentage of the store and other channels' contribution to the total purchases.
8. The marketing manager wants to understand the performance of different marketing campaigns. Find out which marketing campaign is most successful?
Use suitable graphs for visualization. (Hint: - use features like AcceptedCmp for campaign information)
9. The marketing manager wants to understand which products are performing the best and which are performing the least in terms of revenue. Being an analyst, analyse the data and plot a suitable graph to display a report on revenue generated by different products.
10. The team wants to understand if there's any pattern between the age of customers and the last campaign acceptance rate. Plot a suitable graph to visualize the distribution of the age with respect to customers who accepted the last campaign.

11. The Chief Marketing specialist wants to visually see which Country has the most number of customers who accepted the last campaign. What is your approach?
12. Analyse visually and write your inferences about the relationship between the following:-
 - a. Total amount spent Vs Dependents. (Dependents=['Kidhome']+['Teenhome'])
 - b. Total Purchases Vs Dependents.
13. Convert the Year_Birth column into categorical column. Create categories according to the analysis done so far.
14. Perform Correlation Analysis and write your key inferences.
15. Understand the Education background of the customers who complained in the last 2 years. State the Education background of the customers who have registered the most number of complaints.

