

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

Expected Submission Submit a well documented notebook with these four sections:

- **Section 01: Exploratory Data Analysis**
 1. Are there any null values or outliers? How will you wrangle/handle them?
 2. Are there any variables that warrant transformations?
 3. Are there any useful variables that you can engineer with the given data?
 4. Do you notice any patterns or anomalies in the data? Can you plot them?
- **Section 02: Statistical Analysis** Please run statistical tests in the form of regressions to answer these questions & propose data-driven action recommendations to your CMO. Make sure to interpret your results with non-statistical jargon so your CMO can understand your findings.
 1. What factors are significantly related to the number of store purchases?
 2. Does US fare significantly better than the Rest of the World in terms of total purchases?
 3. Your supervisor insists that people who buy gold are more conservative. Therefore, people who spent an above average amount on gold in the last 2 years would have more in store purchases. Justify or refute this statement using an appropriate statistical test
 4. Fish has Omega 3 fatty acids which are good for the brain. Accordingly, do "Married PhD candidates" have a significant relation with amount spent on fish? What other factors are significantly related to amount spent on fish? (Hint: use your knowledge of interaction variables/effects)
 5. Is there a significant relationship between geographical regional and success of a campaign?
- **Section 03: Data Visualization** Please plot and visualize the answers to the below questions.
 1. Which marketing campaign is most successful?
 2. What does the average customer look like for this company?
 3. Which products are performing best?
 4. Which channels are underperforming?
- **Section 04: CMO Recommendations** Bring together everything from Sections 01 to 03 and provide data-driven recommendations/suggestions to your CMO.

Evaluation This is not a formal competition, so results won't be measured using a strict metric. Rather, what one would like to see is a well-defined process of exploratory and statistical analysis with insightful conclusions.

1. **Data Exploration** - Was the data wrangled properly? How well was the data analyzed? Are there any useful visualizations? Does the reader learn any new techniques through this submission? A great entry will be informative and thought provoking.

2. Statistical Analysis - Were the right statistical tests used? How well was the statistical output interpreted? A great entry will interpret results without the use of any statistical jargon.
3. Business Recommendation - Were the recommendations tied to your analysis in Sections 1-3? Are they data-driven and focused on marketing concepts such as targets, channels, or products?
4. Documentation - Are your code, and notebook well documented so a reader can understand what you did? Are your sources clearly cited? A high quality analysis should be concise and clear at each step so the rationale is easy to follow and the process is reproducible.

1. **ID**=Customer's unique identifier
2. **Year_Birth**=Customer's birth year
3. **Education**=Customer's education level
4. **Marital_Status**=Customer's marital status
5. **Income**=Customer's yearly household income
6. **Kidhome**=Number of children in customer's household
7. **Teenhome**=Number of teenagers in customer's household
8. **Dt_Customer**=Date of customer's enrollment with the company
9. **Recency**=Number of days since customer's last purchase
10. **MntWines**=Amount spent on wine in the last 2 years
11. **MntFruits**=Amount spent on fruits in the last 2 years
12. **MntMeatProducts**=Amount spent on meat in the last 2 years
13. **MntFishProducts**=Amount spent on fish in the last 2 years
14. **MntSweetProducts**=Amount spent on sweets in the last 2 years
15. **MntGoldProds**=Amount spent on gold in the last 2 years
16. **NumDealsPurchases**=Number of purchases made with a discount
17. **NumWebPurchases**=Number of purchases made through the company's web site
18. **NumCatalogPurchases**=Number of purchases made using a catalogue
19. **NumStorePurchases**=Number of purchases made directly in stores
20. **NumWebVisitsMonth**=Number of visits to company's web site in the last month
21. **AcceptedCmp3**=1 if customer accepted the offer in the 3rd campaign, 0 otherwise
22. **AcceptedCmp4**=1 if customer accepted the offer in the 4th campaign, 0 otherwise
23. **AcceptedCmp5**=1 if customer accepted the offer in the 5th campaign, 0 otherwise
24. **AcceptedCmp1**=1 if customer accepted the offer in the 1st campaign, 0 otherwise
25. **AcceptedCmp2**=1 if customer accepted the offer in the 2nd campaign, 0 otherwise

- 26. **Response**=1 if customer accepted the offer in the last campaign, 0 otherwise
- 27. **Complain**=1 if customer complained in the last 2 years, 0 otherwise
- 28. **Country**=Customer's location