# Customer Personality Response Report

## What We Did

-----------

We built a campaign response classifier using the Customer Personality dataset.

The data was cleaned, categorical variables were one-hot encoded, and numeric features were imputed and scaled.

We trained a Random Forest classifier with an 80/20 train–test split to predict whether a customer would respond to a marketing campaign.

## How the Model Performed (Test Set)

----------------------------------

* Accuracy: 0.8884
* Precision (Responder): 0.8400
* Recall (Responder): 0.3134
* F1 Score (Responder): 0.4565
* ROC–AUC: 0.8824

Interpretation:

* The high AUC of 0.8824 indicates strong ability to separate responders from non-responders.
* Precision is high (0.84), meaning when the model predicts a responder, it is usually correct.
* Recall is lower (0.3134), meaning many true responders are missed.
* The model is effective for cost control (avoids wasteful outreach) but could be improved if maximizing campaign reach is the goal.

## Key Drivers of Response

-----------------------

From Random Forest importance analysis, the strongest drivers included:

* Income and overall spending power.
* Customer recency (more recent customers tend to respond more).
* Spending patterns in categories such as wine and gold products.
* Campaign history and complaint history (positive past interactions increase response likelihood).

## Mistakes the Model Makes

------------------------

* False Positives: Some customers were predicted as responders but did not respond (increasing campaign cost).
* False Negatives: Many true responders were missed (lost opportunities and revenue).

## Recommendations

---------------

1. Tune the classification threshold to increase recall if maximizing responders is more valuable than reducing costs.
2. Use class balancing techniques (e.g., class weights or SMOTE) to improve detection of responders.
3. Engineer RFM (Recency, Frequency, Monetary) features and engagement ratios to capture customer behavior better.
4. Try Gradient Boosting (XGBoost/LightGBM) with hyperparameter tuning for better performance.
5. Calibrate probabilities to align decision-making with business costs and benefits.

## Conclusion

----------

The Random Forest model achieved strong AUC (0.8824) and high precision,

providing a reliable tool for cost-effective targeting.

With adjustments to improve recall and additional feature engineering,

the model can help marketing teams both reduce wasted spend and increase campaign effectiveness.