PROJECT DESCRIPTION

Project X is based on data from a clothing store chain that uses various marketing initiatives to promote sales. Classify which customers will respond to mail marketing campaigns based on data collected from past customers. The ultimate objective is to the increase the profitability of the store, and you are given a cost-benefit table below to help determine the optimal classification threshold. A false negative is a more costly error than a false positive for this business scenario, as the cost of sending a marketing mail is lower than the profit generated by a responded customer. The response variable is "customerIsResponse". The "customerID" field is unique, automatically generated, and has no relation to the other variables. Please use Google Colaboratory/Jupyter Notebook to perform your analysis, add comment blocks to explain as needed, and submit the .ipynb file (download if using Colab, code blocks should be run in sequence and results/visualisations should be displayed) for review.

Outcome	Classification	Actual Response	Cost/Benefit
True positive (profit)	Response	Response	+13.97
False positive (direct mail cost loss)	Response	Non-response	-2.00
False negative (profit opportunity loss)	Non-response	Response	-13.97
True negative (direct mail cost avoided)	Non-response	Non-response	+2.00