Business Case

A marketing manager is planning a new campaign designed to increase customer retention. Their plan is to deliver the campaign to a targeted audience defined as customers who are likely to abandon the brand within six months. You have been asked to develop a model which can identify these customers.

Data

The dataset provided is a snapshot of consumer purchase history taken in mid-June. Customers who made no additional purchases during the six months following the snapshot period were assigned a `churn` value of 1, to indicate that they had abandoned the brand.

No data dictionary was given with this file. There are no additional measurements that the manager can reliably provide, so your model must be built using only the information you've been provided here.

Instructions

Choose the language you are most proficient in and demonstrate how you would do the following:

- Explore the data.
- Determine appropriate modeling techniques.
- Build the model(s).
- Assess and explain the performance of your chosen solution.
- Predict the churn value (0 or 1) of the observations in `test.csv`

Expectations

- You should return two files:
    1. A rendered notebook (Jupyter, RMarkdown, or similar), showing all code and outputs.
    2. A version of the `test.csv` file with the predicted `churn` values included as a new column.
- Your results should be reproducible, in case we have questions and would like to run the code ourselves.
- Code should be annotated, where necessary, to explain why you are engaging in certain steps.
- Interpretations of plots, tables, and model outputs should be included.
- Your final product should be something you would be comfortable presenting and explaining to a director or similarly senior stakeholder.
- Please do not spend more than 4 hours on this task.
- This should be a fun exercise!