OMD EMEA Technical Challenge

Challenge outline

Our client, XYZ Limited, have carried out an experiment, advertising their service for the last 10 weeks across social, blog, and user content driven websites. Their digital marketing manager now wants a recommendation on how to better allocate investment across different types of website.

Data

The file *data.csv* contains a sample of 10,000 rows of user activity from the client's CDP (Consolidated Data Platform), showing the number of impressions (defined in a marketing context as an 'ad view') per site per user. Also joined to this data, is whether or not the user clicked on the ad, shown as 1 for a successful click, and 0 otherwise. The Excel file *Website Descriptions*, prepared by a junior marketing executive, provides the link between sites and website types. The data has some discrepancies which you will have to decide the best way to handle.

Challenge specifics

You will need to build a model to predict clicks based on website types. Care should be taken to ensure your model is both accurate and interpretable. You are free to use any tool, package, coding language or system you see fit, but the work must be entirely your own.

Your final output will need to consist of the following:

- A prediction of clicks from a further 1,000 rows of user data in the file *prediction.csv*. Add a final column with your 1/0 predictions as per the main file.
- Your code, in any language, well commented and in a single executable file
- A short condensed report of your findings and recommendations. This should be suitable for XYZ Limited's client media planning team at OMD, who are familiar with media terminology and strategy, but have no statistical or data science knowledge much beyond correlations and simple regression. You should clearly highlight your recommendations and findings, while offering confidence that your methodology and results are robust. You may want to think carefully about how to balance the two.

Assumptions

You may assume the cost of every impression is the same, as is the value of every click. You should state any other assumptions you are making in your recommendation.

Assessment

In the report we are primarily looking for evidence of critical thinking and problem solving, your ability to convert the technical into the everyday, and your judgement of what you should show given the requirement to keep it brief. Your code will be assessed for errors, methodology, and commenting in roughly that order. Finally, the prediction will be tested against the known data and assessed for accuracy, precision and recall.