

## Assessment Task

Please use R (or another language of your choice) to answer the below questions concisely using statements, graphs and tables where applicable, making use of any packages of your choice. R script(s), R Markdown or Jupyter Notebook are all acceptable. You have 48 hours to complete this test so please do not feel rushed and it is “open book”, so feel free to use any online or offline resources during this time.

After completion please email back an archived folder with all your answers (data, scripts, graphs etc.) and your initials in the file name.

If anything is unclear or you are unsure how to answer a question, please note this down in the email and do not worry.

### Data:

Two files are provided, **Offers\_sent.csv** and **Offers\_accepted.csv**. They respectively contain data on offers sent to, and accepted by, customers for a specific appliance.

#### **Offers\_sent.csv variables are as follows:**

- A. OfferContactDate – Contact date, DD/MM/YYYY;
- B. CustomerID – Anonymized unique customer ID;
- C. ApplianceID – Anonymized unique appliance ID;
- D. Brand – Anonymized brand code;
- E. V1-V8 – Anonymized categories and counts.

#### **Offers\_accepted.csv variables are as follows:**

- A. OfferAcceptanceDate – Offer acceptance date, DD/MM/YYYY;
- B. CustomerID – Anonymized unique customer ID;
- C. ApplianceID – Anonymized unique appliance ID;

### Tasks:

- a) The *conversion rate* is defined as the number of offers accepted divided by the number of offers sent. Calculate the conversion rate for each brand and assess whether there is a statistically significant difference between them.
- b) Graphically present how the total acceptance of offers develops over time, i.e. how many offers have been accepted after a specified time from being sent (daily or weekly graph).
- c) Using any technique, build a simple predictive model to predict whether an offer will be accepted using columns V1-V8. Briefly explain your choice of model, and decide which variables are most important in the model. Furthermore, examine the model and report on an accuracy metric(s) of your choice. Please do not spend too much time optimising model parameters.
- d) V7 is the treatment category and currently, the business only applies treatment A. The business has decided to introduce new treatment categorical values to V7, factors B, C, and D. Design an experiment to assess the potential impact of the new factors on conversion rates and include which statistical methods you would use. Please state any assumptions made. (R is not required here, you can answer this question in comments)