**Havas Media Group**

**Data Operations Manager**

**2nd Stage Technical Test**

The main aim of the task is for the candidate to prove data manipulation and automation ability using appropriate languages and tools.

**Scenario:** An important client, SAVAH, have a report they need to automate regarding their online media costs.

On a daily basis the following data is made available:

**Ad server media data:** The data is available on a third party FTP server that has new data added to it daily in the form of CSV files. There will be a new file for each day.

**Media cost data:** an Excel file on a network drive. The data is manually updated daily (data for weekends is added on Mondays) by another team.

There’s also a lookup table mapping campaigns to products.

The datasets will need to be joined together and output as a single dataset (CSV file, a table in MySQL, an Excel sheet or similar) using any scripting language (ideally R, Python, SAS, Bash scripts, PHP) and/or tool of candidate’s choice, talking us through how the process would be automated. Where possible, please be prepared to go into detail regarding how your code works.

The solution does not have to be developed end to end; parts of it may simply be described, as long as ability is demonstrated.

Optional bonus points for creating a (simple) visualisation on top of the output dataset.

**Presentation:**

There will be 30 minutes during the interview to take us through the solution. Candidates will have access to a PC to talk us through their thinking. Any files used should be emailed in advance.