**TASK INSTRUCTIONS**

1. **DATA CLEANING / FORMATTING**

The example data file is similar to what we sometimes receive from clients when a survey has been programmed by a third party. There are a few logic errors and the structure is not quite how we need things to be for analysis.

The final data file should be in SPSS (.sav) format with all variable names, variable labels, codeframes (value labels) and formats in place so that the file is ready to import to Q.

The following are the actions required to clean the data file and get it in the correct format/structure *(These are in no particular order, so please read them all to understand what needs to be done and plan your approach accordingly)*:

* Identify any cases (rows) with logic errors or missing data. These will need to be removed from the final data set
* Multi-response questions have been provided with all selections contained in one column and delimited with a semi-colon. These columns will need to be transformed to a series of binary columns (“one-hot encoding”)
* All values in the data file are text labels, however these need to be converted to their numeric codes (matching the QNA)
* The data was exported with question texts as column names, but we need the column names to match the question numbers (as per the questionnaire):
  + Multiple response questions should have the question number and the option code in the column names separated by an underscore (e.g., “Q1” columns will need the names “Q1\_1”, “Q1\_2”, “Q1\_97”)
  + Any columns containing text entered for "Other (please specify)" options will need the question number, code and "Oth" in the name, separated by an underscore (e.g., “Q1\_97\_Oth”)
* A new constructed variable called “Wave” will also need to be created which uses the “CompletedDate” string to categorise responses into weekly waves (Mon-Sun). For example, code 1 should be week commencing 4th August, code 2 should be week commencing 11th August, etc.

In SPSS, the following formatting should be applied:

* variable names should match those in the QNA exactly, with multi-response and "Other" variable names following the naming convention above
* variable labels should contain the question text and multi-response variables will also require the option text, delimited with a hyphen (“ - ”)
* value labels (codeframes) should match the QNA exactly
* variable types and formats will need to be correctly set (e.g., 'Numeric' or 'String'). All numeric values are whole integers so 'Decimals' should be set to 0. String variables will need to be set to an appropriate width so that no text is truncated

You can use any method/tools you like to complete the task (e.g., Excel, Python, R). Once finished, please include a brief explanation of how you approached the task / what steps you took. Feel free to share any scripts you used!

1. **DATA VISUALISATION**

Extract the data from your SAV file as a CSV and import to a dashboard platform of your choice (e.g., Power BI or Tableau). Then, perform the necessary steps to format the data appropriately so that you can create the following visualisations and filters:

FILTERS (on all pages):

* S1 (Gender) - multi-response dropdown box
* S2 (Age Bracket) - multi response dropdown box
* D1 (Work Status) - multi response dropdown box
* D2 (Income) - multi response dropdown box
* D3 (Household Structure) - multi response dropdown box
* Wave - multi response dropdown box

MEASURES:

* PAGE 1 – “Brand Awareness & Usage”
  + Q1 (Awareness) – bar chart (% per brand)
  + Q2 (Main Provider) – bar chart (% per brand)
* PAGE 2 – “Brand Perceptions”
  + Q3 (Perception) – donut chart (% NET: Favourable (codes 4-5))
  + Q4a (Likelihood to Recommend) – stacked column chart (% of NETs, 0-2, 3-4, 5, 6-7, 8-10)
* PAGE 3 – “Brand Values over time”
  + Q5 (Brand Values over time) – line chart (x-axis: Wave / y-axis: % NET: Good (codes 4-5))

Please add titles to pages and charts and use the following style guide:

Font: Arial Narrow

Colour palette:

A screenshot of a screen

AI-generated content may be incorrect.

Feel free to add data labels to charts or anything else you think helps bring your visualisations to life!