Analytics Engineer - Technical Task

Welcome to the Multiverse Data Modelling Project!

This task is designed to assess your ability to interpret source data and propose a fit-for-purpose data model based on stakeholder needs. We're interested in how you approach ambiguity, design models, and think through implementation — not just the final output. We've provided:

- some sample dummy data in csvs
- a description of the source tables
- a set of sample business questions that your model should help answer
  (Note: you don't need to answer these questions — they're there to guide your design)

Please don't spend more than 1-2 hours on this task.

Please produce:

- An ERD of your proposed data model based on the provided source data
- A short description or diagram of your transformation steps (i.e. how you go from source tables to final model e.g. via dbt transformation layer)

These can be simple diagrams (e.g. on Miro, Figjam, any whiteboarding tool), markdown docs, slides, or notes — whatever is clearest for you.

In the technical follow-up interview, you will be expected to:

- Share any assumptions you made whilst completing the task
- Present your data model and explain your reasoning and tradeoffs
- Discuss how your model is designed to support stakeholder needs
- Walk us through your implementation plan (including cleaning; transformations; testing; optimisation; naming)

There are no right or wrong answers here - we believe that data modelling is as much art as it is science. We'll discuss your rationale and trade-offs in the interview, so don't worry about perfection.

Source Data Context

We have two systems for sending surveys to apprentices, Ash and Elm. We have provided some samples of data we from these two systems. Both systems have been used for the same types of surveys, but they capture the response data in slightly different ways.

We've also provided a list of apprentice users. This dataset has been manually maintained.

Sample Business Questions

- What is the average Net Promoter Score for each employer? How does this change month on month, or quarter on quarter?
- Are some apprenticeship programmes perceived as more difficult than others? Is this changing over time?
- How is the number of apprentices on-break distributed by age?

- Is the content of a particular apprenticeship programme suited to a particular age group?