**Senior Data Engineer Technical Task**

Thanks for your application to the above role.

The next step of the process is for you to complete the exercise below and send us your code.

To complete this exercise you will need at a minimum:

* An IDE or a text editor

The exercise:

1. XYZ Bank has multiple customers. Each customer may hold multiple accounts. Each account may have multiple account holders (customers). How would you model this?
2. Using your model, write a SQL query to return the total account balance for each customer where number of accounts is > 1
3. Using your model, write a SQL query to return all the customers with no accounts
4. ABC Ltd sells clothing via internet and in store. They want to determine the most profitable product categories across various regions. You need to take into account cost of product manufacture and any discounts applied to items sold. How would you model this (in a data warehouse)?
5. A data stage process was previously taking 10 minutes to run. Now it’s taking an hour to complete. Detail the possible causes and how you would determine which was responsible
6. Using the coding language of your choice (not SQL): import name.basics.tsv.gz from IMDb and find the number of who are still alive and have the first of their 3 primary professions as producer

Ideally, your code should compile and run without errors. If you are unable to produce workable code, we will accept a pseudo-code version.

If you have any questions, please contact Sarah Ryan [sarah.ryan@hl.co.uk](mailto:sarah.ryan@hl.co.uk)

To return the exercise email to [sarah.ryan@hl.co.uk](mailto:sarah.ryan@hl.co.uk). Alternatively, the code section can be committed to your public github repository (<https://github.com/>) and URL emailed.

### 