

# Community Data Intern Recruitment Assignment

The purpose of this assignment is two-fold:

- To give you a better sense of the role of the **Community Data Intern**
- To give us a sense of whether you would be a good fit for the role

The assignment is designed to be intentionally diverse and fast-paced, much like the actual role.

As you work through this assignment, note that:

- The structure of the assignment is intentionally broad to give you the freedom to think about the assignment on your own terms. Make sure to include any assumptions you are making as you complete this assignment.
- We understand you don't have enough context on our operations and our business. Working with uncertainty and partial information is a part of this assignment. Be sure to highlight the assumptions you're making and leave comments so we can collaborate with you.
- We understand that in a perfect world, you'd have more time to research and complete this task. Understanding constraints, prioritising, and demonstrating a "done, not perfect" mindset is part of what we look out for. So, in places where you don't have enough time to complete the task to your standards, please share your thought process and your next steps if you have more time.

What you share here is only for assignment purposes. We do not intend to use your submission beyond this recruitment process.

Aside from the submission deadline, the parameters of this assignment are purposefully broad and open to your interpretation and understanding.

We'd like you to return your assignment submission in **one PDF document** on the date stated on the email. Please let us know **at least 24 hours in advance** if you won't be able to submit on time.

We do not take the evaluation for this role lightly. We appreciate your time, effort and understanding.

We look forward to receiving your assignment!

All the best!

## Task 1: Data Cleaning

### Context:

As part of our community efforts, it's crucial to maintain accurate and up-to-date data on graduates and learners. Recently, we discovered over 300 duplicates in our graduate data, which impacted our coaching strategy and peer-to-peer initiatives.

### Assignment:

You are provided with a sample dataset (**provided in Excel format**) that includes graduate data. This dataset has several duplicates and inconsistencies. Your task is to:

- Clean the data by identifying and removing duplicates.
- Explain your approach to data cleaning and the tools you used.
- Provide a summary of the cleaned data, highlighting the number of duplicates removed and any other anomalies found.

### Deliverables:

- A clean version of the dataset in Excel.
  - A brief report (300–500 words) detailing your process, tools used, and findings.
- 

## Task 2: Community Dashboard Update

### Context:

Our community dashboard tracks key metrics such as community size, engagement rates, and graduate placement. The data needs to be manually updated to ensure accuracy, especially when new graduates are added or removed and when graduates re-enroll in courses.

### Assignment:

You are provided with a mock-up of the current dashboard and a new dataset containing information on recent graduates and learners. Your task is to:

- Update the dashboard with the new data, ensuring that no duplicates or errors are introduced.
- Identify any potential issues that could arise from manual updates and suggest solutions to automate this process in the future.

### Deliverables:

- An updated dashboard (mock-up in Excel or a visualisation tool of your choice).
- A brief report (300–500 words) discussing the challenges of manual updates

and your proposed automation solutions.

### Task 3: Data Insights for Community Engagement

**Context:**

The community team needs insights to plan targeted outreach and engagement strategies. Accurate data is essential for identifying key groups and tailoring our communication efforts across different platforms.

**Assignment:**

Using the cleaned data from Task 1, analyse and provide insights that could be used for community engagement. Your analysis should help in identifying specific groups (e.g., graduates by city, course re-enrollees) and suggest how these insights could inform outreach plans.

**Deliverables:**

- A data insights report (500–700 words) that includes at least three actionable recommendations based on your analysis.

**GOOD LUCK!**