

SISU SOLUTIONS CODING ASSESSMENT – DATA ENGINEERING

Please find below finalised set of questions, approach and how to deliver the solutions for the coding test:

Using the attached dataset:

- Augment the data with a hash key
- Filter out the questionable data
- Write the data in a JSON file one record at a time
- Watch the data emitted in the previous task
- When the number of events reach 1000, output the events to a JSON file
- The output filenames should have a batch number e.g. the second 1000 records will go into a file called 2.json
- List post codes based on fastest response . Hint (Refer columns Request date and implementation Date)
- Top Agents based on postcode and amount

To submit a code please send a link to your git repo with the solution, code and a diagram to describe the flow.

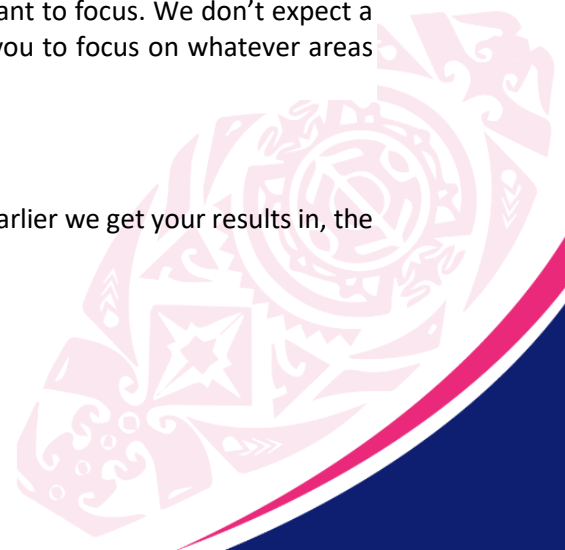
RECOMMENDATIONS

- Preferred language is Python but feel free to use what you like
- Approach this as you are writing code for your development work therefore testing the code is vital
- Ensure you are building a pipeline with more than one step.
- Feel free to use frameworks and libraries, but keep in mind that we are looking for something that demonstrates that you can write ETL code, not just wire up a framework.
- Include a README file with clear build instructions that we can follow.
- Include in your README any other details you would like to share, such as trade-offs you chose to make, what areas of the problem you chose to focus on and the reasons for your design decisions.

Beyond meeting the minimum requirements, it's up to you where you want to focus. We don't expect a fully finished, production-quality data pipeline; rather, we're happy for you to focus on whatever areas you feel best showcase your skills.

Due date

Must complete by **Monday 20th of September @ 8:30am**, however, the earlier we get your results in, the faster we move along the recruitment process.



How to submit:

To submit a code please send a link to your git repo with the solution, code and a diagram to describe the flow. If you do not have git account please Go to <https://github.com/join>

Please send result link to your SISU representative, so that we may progress with next steps.

We value your time and thank you for your interest and look forward to discussing this opportunity further with you shortly.



+61 1 300 00 SISU

sisusolutions.com.au



527 Spencer St, West Melbourne VIC 3003