Internship Interview Brief

Thank you for your application, we are excited to meet you.

This guide aims to provide transparency to the interview process.

We understand that we won't be able to offer an internship to everyone.

That said, we view this interview as a chance for mutual learning. We encourage you to make the most of your time with our team—ask questions, share your thoughts, and don't hesitate to lean on your interviewers for insights. Our goal is to bring out the best in you.

The Role

Depending on who is your interviewer, your main role will be different.

Interviewer 1: Keyword1, Keyword2
Interviewer 2: Keyword1, Keyword2

All of us have skills in each other's fields.

Joining a team does not mean you are excluded from other teams' work. You'll have access to our internal communication, documentation, and code repositories. However, we do hope whoever we extend offer to and accept it has a true interest. In extremely rare scenarios, you may be offered a team who did not interview you.

Interview Process

1. All applicants interviewed by <date>
2. Offer sent by <date>

Interview Outcomes

1. A clear understanding—for both sides—of expectations, motivations, and goals
2. Engaging, thoughtful discussions on real-world challenges and technical topics relevant to our work
3. A shared sense of having learned something valuable from the experience
4. The possibility of future collaboration, whether or not it takes place through this internship

Interview Structure

1. Introductions and discussing past work (10 min)

2. Behavioural questions (10 min)

3. Technical Problem Solving (30 min)

    1. Code review (10-15 min)

    2. Data Analysis (15-20 min)

    3. ETL Pipeline and AWS (15-20 min)

4. Your Questions (5 min)

We prefer your camera to be on so we can meet virtually first, also as an anti-cheating measure.

During the call, we will be taking notes. We encourage you to take notes too.

Your Setup

1. Our codebase is mainly Python, with some C++, Matlab, Javascript. We prefer you work in Python

2. If you do use an IDE (most of us use VSCode), please setup launch.json and prepare to show how you debug through screensharing

3. The Data Analysis exercise is easier done in Jupyter than python script (you can debug with Jupyter extension when opening ipynb in VSCode)

4. Login to your LLM accounts already so you can paste questions

5. Prepare a notepad to take notes, do back-of-the-envelope estimations (more to help yourself than show us)

Behavioural Questions

This segment consists of about 10 prepared questions. We are looking for concise answers that show real examples of lifelong learning, curiosity, grit, teamwork, empathy, leadership, creativity. It's also a chance for you to reflect on whether you would enjoy working with your interviewer.

We believe with great mentorship, knowledge is not a limiting factor. However, desirable personal attributes are a rarer find and more likely the dealbreaker.

Technical Problem Solving

All technical problems come from actual work, so no esoteric data structures and algorithms. We strive to guide you to complete all exercises, so you leave with a good understanding of what we do.

As you have already seen, the timings in Interview Structure don't add up. We don't expect anyone to complete all 3. A single exercise with well-considered trade-offs beats unstructured wild guessing on 3 exercises.

It is better to say upfront if you don't have much experience with a certain technique/library method than to fake non-existent experience, to help the interviewers decide whether to move on to other questions or teach you right there. Curiosity and initiative to fill in any gaps on your own later is highly valued.

The following section provides details on the 3 Technical Exercises. The signals shown here are a subset of what the interviewers have in mind, so really outstanding applicants would satisfy these and more.

Exercise 1: Code Review (10-15 min)

Given sample code, comment on its performance characteristics using python/computer science concepts

What signals this exercise provides

- Familiarity with IDE, filesystem and basic I/O
- Sense of high performance, scalable code
- Ability to read and understand code
- Ability to communicate and explain code

Exercise 2: Data Analysis (15-20 min)

Given 2 dataframes, we want to join them and analyze correlations between signals across the dataframes at regular intervals of 0.1 seconds.

All necessary code is provided, but first, there is a bug to investigate.

What signals this exercise provides

- Familiarity with Joins and common problems
- How you debug and test
- Discussion of multiple strategies for fixing problems

Exercise 3: ETL Pipeline and AWS (15-20 min)

We want to download X-minute long binary log files across a timespan of 1 year from cloud, process them into parquet files and upload the results back to cloud. Each log file is about Y MB.

Ideally, we want the pipeline to be as fast as possible, using as little memory/disk as possible, so we want to use concurrency (optional) to download, process and upload files in parallel. However, a first iteration does not have to satisfy this requirement, we can always start single process single thread and improve it later.

What signals this exercise provides

- Estimation skill

- Ability to break down problem and define function signatures

- Experience with ETL, concurrent programming, IPC

- Ability to generate multiple architectures and discuss trade-offs

- Full stack thinking across data, software, cloud (compute/storage/networking)

Your Questions

Ask us anything, ideally something not available on public sources.

We can field questions offline if we run out of time during the 1 hour call.

LLM Usage

The company has a LLM subscription, so LLMs (chatgpt, deepseek, claude, gemini) are fine. We prefer you to try to tackle the problem yourself first, until your interviewer prompts you to move to LLM if you get stuck for too long. We are looking for the kind of questions you ask, and how you interpret the responses.

Ending on a high note

Your interviewers may write you feedback on our interpretation of the interaction. Similarly, your interviewers may request you to write your reflections on the experience. This serves a few purposes:

1. You get practice on reflecting on an experience, which could serve you in future endeavours

2. In case we cannot decide, this does sway decisions

3. Technical writing is a crucial skill, since the international offices often communicate asynchronously through writing

Resources

1. Technical writing: https://refactoringenglish.com/

2. Python fundamentals: https://www.amazon.sg/Powerful-Python-Patterns-Strategies-Modern/dp/1098175700

3. Workplace fundamentals: https://www.amazon.sg/dp/1718501838

4. Building reliable pipelines: https://sre.google/workbook/data-processing/