Thinking Machines Support Engineering Take Home Exam

I. Exam Overview

As Support Engineers, it is our primary job to maintain, troubleshoot, and optimize data platforms in production by delivering high quality and reliable solutions efficiently to our customers.

This exam will test your abilities on:
- Familiarizing with different components of a data platform
- Designing robust solutions with focus on maintainability, availability, and performance
- Communicating your solution to client engineering teams and business stakeholders

In this exam, you will design, present, and demo a solution for a sample project.

II. Project Overview

We live and breathe data. So naturally, we try to make our internal processes as data-driven as possible. Time-tracking allows us to determine our teams' bandwidth and throughput, so we checkin religiously. We do this using a handful of tools that we can use without interrupting our work.

For this exam, you are given an anonymized dump of our checkins for the past year. Your goal is to:
- Clean the data
- Load it to your choice of database
- Create a web service that returns the checkin data associated to a given user

🌟**Deliverable 1.** Your first deliverable is the solution code repository, and its documentation.
- Bonus points if your solution code uses Python, contains tests, and has a publicly accessible deployment of your service.
- We value communication, reproducibility, pragmatism, and code hygiene.
- The documentation should contain:
    - What is in the environment and how you created the environment.
    - Draw a data architecture showing different components of your ETL process.
- For the purposes of this exam, a written documentation is enough — but you will get bonus points if you use an infrastructure-as-code tool (such as Terraform) to create the environment.

🌟**Deliverable 2.** Your next deliverable is a presentation for two different audiences:
1. Business: A short, high-level overview of your solution for business stakeholders (around 2-3 slides)
2. Technical: A detailed presentation with a live solution demo to the client engg team (max 10 slides)

🌟**Deliverable 3.** These are additional design questions, which will test your understanding of the problem.
1. If the data is to be ingested periodically, what changes will you make to your current approach?

2. How will you verify the correctness of the ingested data?
3. As a Support Engineer, what type of information do you need from the project team in order to have a successful handover?
4. Outline a plan for maintaining and adding new features to the solution that you designed.
5. In the near future, what possible changes can you foresee being asked by the client? What other factors can affect the current setup?
6. What would you change in the current implementation to accommodate possible changes and factors that can affect the current setup?

III. Deliverables

In summary, these are the deliverables for this exam:
1. **Solution Code Repository (50%)**
   - Documentation
   - Solution Code (in repo)
2. **Presentation Deck (20%)**
   - Business deck (2-3 slides), Slides/PPT/PDF format
   - Technical deck (max 10 slides), Slides/PPT/PDF format
   - You will present these slides to us in person if you are accepted to the interview stage of the application
3. **Design Questions (30%)**
   - Answers to questions

You are given **72 hours** to submit your deliverables. The time allocation is tough by design, so budget your time wisely, deliver as much as you can. Don't worry, we'll still review your work even if not everything is there.

For all documentation or presentation files, please upload them in the shared Drive folder, under *Exam*. For all code files, either send us a link to your **private repository**, or upload them in the Drive folder as well. If you're using Github's private repos, please share to the following:
- https://github.com/florobarotjr/
- https://github.com/santosnarom
- https://github.com/vincentpaul12
- https://github.com/t-eevee