



# TM Internship Intern Analyst (Tech Track) Take Home Case Study 2023

## General Instructions

We're sharing this Case Study with you as a means to get to know your skills, interest and fit for a Business Intern Role at Thinking Machines Data Science. Aligned with our [Core Values](#), and to make the implicit, explicit, your output should be yours alone.

## Case Study mechanics

- **Duration:** You have **168 hours (7 days)** to complete the Case Study. Submit your output no later than the deadline indicated in the email.

Albeit the week-long lead time, this is designed to be **doable within 24 hours**.

- **Preparation.** Before starting, ensure you have access to your unique Case Study folder in the Google Drive shared to you by the TM Hiring Team. You will upload your output here. If for whatever reason you are not able to access the folder, please inform the Hiring Team immediately.
- **Expectations.**
  - We highly encourage you to act “in-character”: as if you are really working as an Intern at Thinking Machines. There is no need to assume a specific Tech Track role in order to accomplish these tasks.
  - A good submission has a core theme, and uses several different analytical angles to answer interesting questions on that theme. Even better submissions properly frame why their insights are important, and/or have well-thought out questions or recommendations based on their findings.

Make your presentation effective by applying a storytelling principle called the [Rule of Three](#).

- You are allowed to create assumptions to fill in information that is not given to you, as long as you document them and your reasoning for doing so.

- **Outputs**

- **Main Deliverables.** We expect you to understand and extract the objectives and tasks from the case study scenario -- no further instructions will be provided.
- **Attachments:** You are free to craft any attachments that you find appropriate to meet the objectives, accomplish the task, and present your best work. Upload these files in your Case Study folder and hyperlink these in your Main Deliverable where possible.
- Incomplete outputs will not be accepted.

- **Components.** The exam has three sections. We expect you to understand and extract the objectives and tasks from the prompts below -- no further instructions will be provided.

- You are free to do your own research to supplement your deliverable. Provide links to the sources you used.
- You may use technology to do your work (i.e. ChatGPT) following these ground rules:
  - **Be transparent about the use of the technology.** On top of your deliverables, we expect you to provide end-to-end documentation on your process of technology for your work (i.e. ChatGPT, or similar). Provide screenshots of your working prompt(s) and the generated response(s). These will be compared with your final output.
  - **Use technology as a tool, not a replacement for your creativity and critical thinking.** Your outputs must be yours and yours alone. Use technology (i.e. ChatGPT, or similar) as a prompt for your writing, but we expect you to make it your own.

- **How to submit and organize within the Folder:**

- Create 1 subfolder per scenario and upload your deliverables in the corresponding Scenario Folder. We are expecting that there will be at least 3 subfolders in the Google Drive folder provided to you.

- Within each folder, there should be at least 1 Google Document (Doc, Slides, PDF, etc.) which is your main output. Link the appropriate attachments accordingly.
- Submit your exam folder through the Google Form linked [here](#).
- **Next Steps:** Should you pass the exam, you will be invited to self-schedule a panel interview slot. Given limited slots, we encourage you to submit your exam output as early as possible!

**If selected**, you will be invited to a panel discussion to present the outputs and your thought process.

## Scenario 1

Thinking Machines is competing in an international research conference that aims to promote the good and ethical use of AI to drive social impact in our country. As part of your internship experience, your manager assigns you to join our research team.

To kick-off and find a research angle, the team assigns you to do an **exploratory data analysis** on the [dataset](#) provided by the conference organizer.

## Exam Deliverables

**Goal: Analyze the data, and uncover at least 3 main insights.** You are expected to surface interesting insights from your dataset, so please *refrain from simply describing its contents and characteristics*.

**Tip!** A good submission has a core theme, and uses several different analytical angles to answer interesting questions on that theme. Even better submissions properly frame why their insights are important, and/or have well-thought out recommendations based on their findings.

### 1. Data Analysis

This involves all code, dashboard and other tools that you used for cleaning and analysis.

You are also free to use any tool, language, or library you are comfortable with to do your analysis. We encourage you to go above and beyond our minimum deliverables as it will make it easier for your submission to stand out. Some ideas:

1. Expand your analysis by combining relevant datasets apart from the ones provided using SQL, appropriate machine learning techniques, etc.
2. Create an interactive dashboard or webapp to let the users play around with your data and insights

3. Exhibit your Python or programming skills by sharing your code repo to [hiring@thinkingmachin.es](mailto:hiring@thinkingmachin.es)
  - a. Either **upload them in our shared Drive folder**, or share your **private repo** to <https://github.com/tm-hiring>
  - b. Give our team permission to share the private repo to others so we can add in anyone else that needs to review your code in our team

## 2. Presentation deck

Uncover **at least 3 main insights** under a maximum of 6 slides. You are free to angle your insights however you like. This deck will be presented at the international research conference. As such, create a deck that is suited for the audience.

You are also allowed to create assumptions to fill in information that is not given to you, as long as you document them and your reasoning for doing so.

Make sure your insights are discussed inside the slides. Make it so that it can discuss your ideas even without you talking through it, as we will only be able to read it at this stage.

## 3. Technical Documentation

Document your **thought process and assumptions**. Maximum of 1 page. Walk us through your thought process by answering the following questions:

- What steps did you take to understand the dataset?
- How did you prepare the data for analysis? (Include any code snippets used, if any)
- What was your rationale behind the insights you chose to highlight?
- What was your design process when building the visualizations?

Feel free to be more creative and think of other ideas apart from these, as long as it will showcase your own data science skills. We're looking forward to being impressed by your work!