

Blue Health Intelligence

Data analytics internship: project brief

Supervisor: Serghei Mihailov, CTO

About us

Blue Health Intelligence runs preventive health checks in Amsterdam. In about an hour, we collect hundreds of data points per person, including blood markers, ECG, body composition, vital signs and questionnaire responses. Doctors review the results and turn them into a personal health report.

We build our own software for booking, clinic operations and reporting. Our data helps us understand both how the service works and the health measurements collected during each check.

The project

Build reliable analytics that the team can use regularly, while making it easier to investigate new questions as they arise.

We regularly need to pull data, investigate an issue or set up a report. Doing this well requires understanding our systems, interpreting the data correctly and checking the results. Our analytics infrastructure is in place, but much of that understanding still needs to be captured in models, reports and documentation.

The core of the internship is practical data analysis: answering real requests and building tested data models and dashboards that colleagues can refer to routinely. Alongside this, the intern will help evaluate our AI analytics assistant as a more flexible way to explore the data.

The initial two-month project will produce:

- Tested, documented data models for selected business and operational priorities.
- Dashboards and recurring reports that the team can use and maintain.
- Ad hoc analyses addressing questions raised during the internship.
- A repeatable assessment of the AI assistant's accuracy and limitations.

Our analytics setup

The technical foundation consists of:

- **PostgreSQL:** a separate analytics database updated nightly from production.
- **dbt:** SQL-based transformations, automated tests and documentation.
- **Metabase:** dashboards, reports and self-service exploration.
- **A data catalog:** documentation and lineage to help understand where data comes from.

- **An AI analytics assistant:** a way to ask questions in plain English.
- **GitHub:** version control and code review.

The intern will build on this platform rather than set up the infrastructure. Stable models and dashboards will support regular reporting; ad hoc analysis and the AI assistant will support questions that do not already have a prepared report.

What the intern will work on

Data analysis

Work with colleagues to understand what they need to learn or decide, then investigate the relevant data.

Starting areas include bookings, conversion, rebooking, cancellations, clinic capacity and patient flow. The exact priorities will be agreed with the supervisor and the people using the results.

For each analysis, the intern will identify suitable sources, examine data quality, account for edge cases and check the findings. Results should explain the answer, the method and any important limitations.

Data models, dashboards and reporting

Turn recurring needs into tested dbt models and useful Metabase dashboards or reports.

This work includes:

- Modelling the data needed for the agreed analyses.
- Checking calculations and interpretations with the relevant colleagues.
- Adding automated tests and documenting assumptions.
- Building dashboards for information the team needs to monitor regularly.
- Reviewing the outputs with users and improving them based on feedback.

The aim is to give the team dependable reference points and reduce repeated investigation. Not every request needs a dashboard: some are best answered through a one-off analysis or report.

AI-assisted exploration

Evaluate whether the AI assistant can answer a representative selection of real business questions correctly.

The intern will assemble checked reference answers, define scoring criteria and record a baseline. They will investigate failures, improve the supporting models or documentation where appropriate, and rerun the evaluation.

This work should establish where the assistant is useful and where its answers still require checking. It complements the models and dashboards rather than replacing them.

Proposed milestones

The supervisor and intern will agree on a manageable scope during onboarding, based on business priorities, data availability and complexity.

Weeks 1 to 2

Learn the platform and relevant data sources; complete a first checked analysis; deliver a first tested model and dashboard for a recurring need.

Weeks 3 to 4

Build the initial set of models and reports; review dashboards with their users; establish an initial evaluation of the AI assistant.

Weeks 5 to 8

Complete and refine the agreed models and dashboards; deliver priority ad hoc analyses; document the assistant's performance and limitations; prepare the handover.

Ad hoc requests are part of the role. They will be prioritised with the supervisor so that the core modelling and reporting work remains achievable.

What success looks like

By the end of the initial project:

- Colleagues have useful dashboards and reports they refer to regularly.
- Selected recurring data needs are supported by tested, reproducible models.
- Priority ad hoc requests have received checked answers with clear explanations.
- Important knowledge about the data and its limitations is documented.
- The team has evidence of where the AI assistant is reliable and where human checking remains necessary.
- Another team member can maintain the work and extend it.

Optional extension: research and health analytics

If the internship extends to months 3 and 4, there may be an opportunity to explore research using our health data, alongside continued business analytics.

The intern would help assess how we could study that data within the terms of our research consent and applicable privacy requirements. This could include evaluating privacy-preserving methods such as differential privacy and the use of a trusted research environment.

The first output would be a proposed approach, including its technical requirements, analytical limitations and questions requiring specialist review. A first research analysis would follow only if the necessary approvals and data arrangements are in place. The intern would not be responsible for determining legal compliance.

Supervision and working practices

The intern will report to Serghei Mihailov, CTO, and use the existing onboarding guide and platform documentation.

Code changes go through review before they are accepted. The intern will work with the colleagues requesting analyses to clarify requirements and check interpretations. Questions involving sensitive data, additional access or changes to the privacy boundary will be handled by the CTO.

The role involves independent investigation, with an expectation that the intern raises blockers, uncertainty and unexpected results early.

Candidate profile

We are looking for:

- Strong SQL and data-modelling skills.
- Python and a working knowledge of statistics.
- A careful approach to checking results and explaining limitations.
- Clear written English.
- The ability to investigate unfamiliar data and ask precise questions.
- Responsible handling of sensitive information.

Experience with dbt, Metabase or another BI tool, LLM evaluation, health data or privacy-preserving analytics would be useful but is not required.

The internship offers practical experience in analysis, dashboard development, tested data modelling and communicating findings, with additional exposure to AI evaluation and potentially health research.

Data access and privacy

The intern will work in a restricted analytics environment with directly identifying fields removed. They will not receive access to raw production data.

Removing identifying fields does not automatically make data anonymous. Access arrangements, including any access from outside the EU, must be reviewed before the internship begins.