



EUROPEAN COMMISSION
EUROSTAT

European Big Data Hackathon 2025

- event description

The European Big Data Hackathon 2025 – event description

What is a hackathon?

Data hackathons, known as ‘data dives’, are intense events where teams of data scientists, programmers, graphic/interface designers and project managers seek to creatively tackle data problems and prototype data analytics products. Some data hackathons serve mainly educational or social purposes, although in many cases, the goal is to create usable data tools. Data hackathons tend to have a specific focus, e.g. on data sources, methodologies, technologies and applications. Hackathon organisers, however, frequently leave the type of final data product open, to expand a range of possible data solutions.

What is the purpose of the European Big Data Hackathon?

The European Big Data Hackathon (‘the Hackathon’) has three main objectives:

- to solve statistical problems, by leveraging algorithms, exploring novel, big and open data sets as potential sources for official statistics, and engaging with practitioners, developers and data scientists across Europe, to generate ideas and proposals;
- to devise innovative products and tools, including for data visualisation, to stimulate the use of open data and public use files and to engage with new audiences and users;
- to promote partnerships with the research community and private sector, by raising awareness about big data developments in official statistics in the EU and by initiating, extending and reinforcing collaboration with the private sector and universities.

Teams

Each National Statistical Institute (NSI) of the European Statistical System (ESS) is invited to nominate, or select, one team to represent them in the 2025 European Big Data Hackathon. The team should count three members, chosen among the staff of the NSI, and/or experts from the national data science community and/or from collaborating associations, organisations or partners. In case of particular interest, an NSI can also propose a second team whose participation in the Hackathon is subject to confirmation once all interested countries have nominated their teams¹.

Additionally, master’s programmes carrying the label ‘European Master in Official Statistics’ (EMOS) can nominate up to four teams. These teams can be composed of enrolled students and staff teaching in EMOS-labelled master’s programmes. In case there is a larger number of teams proposed, the EMOS Board is responsible for selecting four teams that will participate in the Hackathon. NSIs are also welcome to include in their teams students and teaching staff from EMOS-labelled programmes.

¹ The second team will be invited to the Hackathon provided that fewer than the maximum number of 35 teams have been registered by the defined deadline.

Up to 35 teams will compete in the European Big Data Hackathon, 31 nominated by the NSIs, one per each country of the ESS (the 27 EU Member States and Iceland, Liechtenstein, Norway and Switzerland), and four representing EMOS-labelled master's programmes. The final list of competing teams and their members will be announced as soon as possible after the registration date.

Event outline and preparation

Teams will compete for the best application of novel, big and open data addressing a pressing European policy or statistical need. Information on the policy domain(s) and a catalogue of datasets will be provided in advance, while the specific data challenge will be announced at the beginning of the Hackathon. This stepwise approach to disclosing the Hackathon topic permits early preparation of team members by deepening their knowledge about the highlighted policy needs and preparing own datasets, which teams can bring to the Hackathon under certain conditions (see section 'Data sources').

The Hackathon will take place in Brussels, as an in-person four-day event and back-to-back to the New Techniques and Technologies for Statistics (NTTS) conference².

During the Hackathon, the teams will put together a data product's prototype supporting the policy analysts in their work on identified policy needs.

At the beginning of the Hackathon, the data challenge will be announced and the teams will then work for about two-and-half days developing their prototype. In the last day of the work on their projects, the teams will have an opportunity to discuss them bilaterally with organisers and evaluators in 'pre-evaluation' meetings. At the end of the hacking time, the teams will have 7-10 minutes to present their data product/application to everyone. After the presentations, a panel of evaluators will rank proposed data products based on a set of criteria. The Hackathon will finish with the announcement of the winning team and five runners-up, selected among the NSI hacking teams, and of an EMOS winning team.

The winner and the first two runners-up, as well as the EMOS winning team, will present their prototypes at a plenary session of the NTTS conference on 11 March 2025.

Data sources

For the 2025 Hackathon edition, Eurostat proposes to explore Earth Observation data made available by the Directorate General for Defence Industry and Space (DG DEFIS) and the European Space Agency (ESA) using the [Copernicus Data Space Ecosystem \(CDSE\)](#).

Use of Earth Observation data, integration of spatial information with statistical data and exploiting geospatial capabilities for European statistics are an important work strand of the ESS Innovation Agenda.

Eurostat will provide a catalogue of datasets consisting of a description of the datasets and instructions on how they can be accessed. Teams are allowed to use additional data sources, which they may prepare in advance. These additional data sources should have an international applicability, i.e., they either cover EU Member States widely or they are

² [NTTS 2025 | Eurostat CROS \(europa.eu\)](#)

national data sources but are available in several Member States. They should also allow the particular data product to be possibly released as a policy tool at European level.

Data product

The teams are free to propose any type of data products. The conditions are that the data products:

- help in answering the policy needs
- address the statistical challenge
- include a front-end (i.e. visualisation) which allows the policy analysts and /or statisticians explore the answers provided by the data product.

IT infrastructure

The European Commission will ensure an internet connection and power plugs in the room where the teams will hack, as well as access to a cloud data infrastructure. However, participants will need to bring their own laptops.

Evaluation

The evaluation will be done in two stages. In the first stage, up to 8 ‘*pre-evaluation*’ teams of 2-3 persons, composed of the members of the final evaluation panel and organisers, will make a pre-assessment of the proposals of the participating teams during the third day of the hackathon (Sunday). The participating teams will be assigned randomly to the pre-evaluation teams. The pre-assessment will last for 30-40 minutes. It will consist of a questionnaire to be completed by the teams and an interview, where the teams will present the current status of their data product and their final target. Based on the questionnaire and the interview, the pre-evaluation teams will make a preliminary assessment of the applications of the participating teams. The assessment results will feed into the final evaluation as qualitative input (no points will be assigned based on the pre-evaluation).

In the second stage of the evaluation (‘*pitching*’), the teams will present their data product in a 7-10-minute pitch. Based on the pre-evaluation results and the pitching, the evaluation panel will rank the data products of the teams.

Evaluation criteria

The panel of evaluators will assess and rank the proposals of the teams based on the following criteria:

- *Relevance*: the extent to which the data product provides an effective answer to the policy needs;
- *Methodological soundness*: the extent to which the data product provides an unbiased answer to the policy needs;
- *Communication*: the ability of the product to provide the information in a clear and user-friendly way, including through data visualisation;
- *Innovative approach*: the extent to which the data product introduces or uses new ideas or methods;

- *Replicability*: the extent to which the data product can be implemented in several EU Member States.

Panel of evaluators

An independent panel of evaluators will be responsible for the evaluation of the data products proposed by the competing teams.

The panel will consist of 10 - 16 persons having no links with the NSIs or the universities running EMOS-labelled programmes which nominated the teams, and will include individuals from both the statistical and the policy domain(s).

Funding

The European Commission will reimburse the travel and accommodation cost for one team per NSI and for up to four teams nominated by EMOS-labelled programmes. The cost of participation of a second (additional/‘reserve’) country team nominated by an NSI will not be reimbursed by the European Commission and the NSI will need to secure its funding.

The cost reimbursed will include the return trip to Brussels and accommodation for up to 5 nights in Brussels. Meals will be provided during the event at the venue.

Prize

Prizes will be awarded to the members of the winning NSI team and five runner-ups as well as to the winning EMOS team. For the 2025 Hackathon, Eurostat is planning to introduce additional ‘audience’ prizes, awarded based on the votes of the teams’ members.

Venue

The Hackathon will take place back-to-back to the 2025 NTTS conference in Brussels, during four days and a half, in a venue situated close to the Charlemagne building (where the NTTS will take place).

Indicative calendar

Registration of teams and on-boarding on the data platform	Autumn 2024
Announcement of the final list of competing teams	January 2025
Release of a data catalogue	November 2024
Hackathon	6-10 March 2025
NTTS 2025	11-13 March 2025