THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

CAESER Center for Applied Economic Social and Environmental Research

STC stcenter.net

Open for Innovation KNIME

INTERNATIONAL COMPETITION

ON DATA SCIENCE AND ARTIFICIAL INTELLIGENCE

Rules and Guidelines

ENQUIRY: caeser@ust.hk

International Competition on Data Science and Artificial Intelligence
Rules and Guidelines

Background and Objectives

In an era where data drives human progress and AI enhances every facet of work and life, the next generation's ability to harness this technology will become a defining skill. The *International Competition for Data Science and AI* responds to this imperative, creating a global stage for exceptional young minds to hone their skills through real-world, data-driven challenges.

Hosted by The Center for Applied Economic, Social and Environmental Research (CAESER) at the Hong Kong University of Science and Technology (HKUST), U.S. NSF Center for Spatiotemporal Thinking, Computing, and Applications (STC) at George Mason University, and KNIME AG, this premier challenge will empower high school and university students worldwide to solve global, data-driven challenges. This entails analyzing real datasets and enhancing AI by redesigning workflows to make models faster, more innovative, and more efficient.

The Competition Schedule

- Registration closes: December 31, 2025
- Written materials submission: January 20, 2026 (10 days before finals)
- Presentation slides submission for Live Final: January 28, 2026
- Live Final & Awards Ceremony: January 31, 2026, in the Hong Kong University of Science and Technology (Hybrid mode)

Briefing Session

- To facilitate preparation, an **online** briefing session will be held, introducing participants and participating teams to the general expectations for the competition and the principles of final submission.
- All registered participants are eligible to participate.
- This key session will take place on Zoom on January 5, 2026 (Monday); the specific time and Zoom link will be announced later to participants via email.

Competition Eligibility and Discipline

1. The competition is divided into two separate divisions: the University Division and the High School Division.

- High School Division: Current high school students (no restrictions on major or region) can participate as individuals or in teams of up to **three** persons.
- University Division: Current undergraduate students can participate individually or in teams of up to **two** persons.

Graduate students are not eligible to participate in this competition. Each team or individual participant may designate one faculty advisor, who is responsible for pre-competition guidance and academic support. However, the advisor is strictly prohibited from providing any advice or participating in discussions during the official competition period.

2. During the competition, participants are permitted to use publicly available resources from the internet, as well as computers and software. However, each team must independently complete its solution without external assistance.

3. Rules on the Use of AI Tools (including but not limited to large language models, generative AI, code assistants, and intelligent mathematical software): Participating individuals and teams are allowed to use AI tools during the competition, provided that such use adheres to the principle of transparency and that the core modeling and analysis of the submission must be independently completed.

- Any content generated with the assistance of AI tools must be clearly cited in the main text.
- All AI tools used must be listed in the bibliography or references section.
- Submissions utilizing AI tools must include a detailed "AI Tool Usage Statement" as part of their submitted materials.

Submissions that do not comply with these requirements will be considered in violation of the competition rules and will be disqualified from award eligibility.

Competition Content

This competition offers two distinct submission categories. A team or individual participant must select either category 1 or category 2 for their final submission.

Category 1: Solve Real-World Problems Using KNIME Workflows and AI Models

For this category, the Organizing Committee will provide the official datasets and their descriptions within three days following the briefing session. These datasets are sourced from diverse fields, including science and engineering technology, humanities, and social sciences (such as economics and management). Since high school and university students have different levels of knowledge and experience, different datasets will be assigned to the University Division and the High School Division. , datasets will be randomly allocated to individual or team participants.

Participants are required to write a comprehensive report based on their assigned datasets. This report must include: the model's hypothesis and establishment, computational implementation (using KNIME), results validation and interpretation, and potential implications and impact.

The final submission files must include:

- **Annotated Workflow Files:** Fully documented KNIME workflows with clear annotations detailing each step.
- **Data Analysis Report:** Containing the data cleaning process (how raw data was prepared for analysis), a clear problem statement (the real-world issue addressed), and the methodology (the AI models and workflows used and the rationale for their selection).
- **Impact Presentation:** Detailing key findings (patterns or insights emerged from the data), real-world relevance (how results reflect actual economic, societal, or environmental challenges), and future implications (potential applications of the solution).

Category 2: Optimize the Efficiency of KNIME Workflows and AI Models

This category challenges participants to enhance existing KNIME workflows and to make AI models faster, more efficient, or more accurate. Participants may compare different KNIME workflows applied to the same research question to evaluate relative performance in terms of speed, accuracy, and other practical advantages. Alternatively, they may refine a single workflow by introducing modifications or adding new functionality (for example, new components, Python-based nodes, or scripts) to increase its robustness and operational efficiency.

Participants are required to write a comprehensive documentation with details of the implemented improvements and their expected impact.

The final submission must include:

- **Annotated Workflow Files:** Fully documented KNIME workflows with clear annotations for both the original and the optimized versions.
- **Optimization Documentation:** Featuring a performance comparison (metrics showing before/after improvements, e.g., speed boost, resource savings) and a technical explanation (how these enhancements were achieved).
- **Demonstration Evidence:** Providing visual proof (screenshots/videos showing both workflows in action) and test results (data proving efficiency gains under identical conditions).

General Rules & Live Final

All written materials must be submitted in English and by the official competition deadline.

The competition also includes a Live Final session. Each participating individual/team must deliver a 10-minute presentation of their results. The final presentation slides must be submitted to the Organizing Committee at least two days before the Live Final. This event can be attended online or on-site at the HKUST campus.

Note: Participants are not required to possess advanced specialized knowledge before participation. A foundational understanding of statistics and econometrics is sufficient. The competition format is designed with significant flexibility to encourage creativity and innovation.

Organizational Structure

- This competition has led to the establishment of the International Competition in Data Science and Artificial Intelligence Global Organizing Committee. This committee is responsible for formulating the competition rules, launching registration, organizing the judging committee and award evaluation, printing certificates, and hosting the award ceremony.
- The competition will be divided into four main regions: Asia-Pacific, North America, Europe, and Other Regions. All participants not from the Asia-Pacific region, North America, or Europe will be automatically assigned to the "Other Regions" category. Each region will have its own Regional Organizing Committee, responsible for promotion, registration, and overseeing competition discipline within its region.

Award Evaluation Method

1. Awards will be comprehensively evaluated by an expert judging committee based on the following aspects of submissions:

- Reasonableness of hypotheses
- Creativity in modeling
- Correctness and foresight in data interpretation
- Clarity of result visualization
- Overall presentation quality

2. Awards are divided into two categories:

Global Prizes

- Champion: 1 team
- 2nd Place: 2 teams
- 3rd Place: 3 teams
- Excellence Awards: 10 teams

Regional Prizes

Gold, Silver, and Bronze awards will be presented separately in each competition region (The number of awards in each region will be scaled based on the number of participants)

Note: Participants from all regions are eligible for both Global and Regional prizes. However, teams that win a Global Prize will not be additionally granted a Regional Prize.

All awards will be evaluated and determined by a Global Expert Judging Committee appointed by the Organizing Committee.

Any participating team found violating the competition rules will be disqualified from award eligibility. The Global Organizing Committee (or the respective Regional Committee) will take appropriate action based on the specific circumstances.

Judging Committee

The Judging Committee will be composed of world-renowned scholars and leading professors in the field of data science, including but not limited to experts from the organizing universities and institutions.

Registration

The competition is free of a registration fee.
Interested participants should register for the competition at the following link by 23:59, 31 December 2025 (Wednesday):
https://ust.az1.qualtrics.com/jfe/form/SV_ah1YHUpTnRewOWO
Important note: *Please be advised that for team entries, each team member must complete the registration process individually and pay the separate fee.*

Enquiries

Organizing Committee, International Competition on Data Science and Artificial

Intelligence
Address: Room 2338, Lift 3, Academic Building, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong
Email: caeser@ust.hk

- Competition affairs: Mr. Nicky TANG
- Global outreach & communications: Ms. Gisele ZHOU
- judging and partnerships: Ms. Caroline Susan Korterud

(Please include the relevant contact name in enquiries for quicker response)
Web: https://caeser.hkust.edu.hk/en/node/381

Copyright

- All materials distributed and produced in the competition, including but not limited to the competition dataset as well as the photos and video recordings of the competition, shall remain the rightful property of the International Competition on Data Science and Artificial Intelligence.
- The above materials must not be reproduced or disseminated without the written permission of the Committee.
- Events on live finals will be photographed and/or video recorded. Consent will be sought from the participating teams at the time of registration.