CDSS Data Discovery Partner Welcome

Spring 2026

Thank you for participating in Data Discovery! Because of you, students will be getting valuable hands-on experience to complement their coursework. Thank you for your partnership in mentoring these students to provide a valuable educational experience.

Contents and overview

Partner orientation

Please plan to join us for one of two virtual kick-off meetings to meet the Data Discovery team and learn more about the program

- **Date:** Tues Jan 20 | **Time:** 4:00 - 5:00 PM
  - Zoom registration link
- **Date:** Weds Jan 21 | **Time:** 1:00 - 2:00 PM
  - Zoom registration link

Important dates for Spring 2026

Data Discovery runs according to UC Berkeley's academic calendar.

Access student applications	**January 12**

You can begin reviewing applications and scheduling interviews.	12 pm
Student application closes No more student applications will be accepted by the program after this date. You may close your individual project application earlier if you've received enough applications.	**January 12** 12 pm
Wrap up sending offers in DDMatch Send offers to students by January 21 to ensure you're ready to form your team by January 23.	**January 21** 11:59 pm
Team selection deadline Help us start on time by confirming your team as soon as possible	**January 23** 11:59 pm
Project work can begin Schedule your first meeting with your students for the week of January 26. This will also be students' first day of required weekly class meetings, Mondays 3-5. Students will ask you to help them create a Mentoring Agreement as one of their first program requirements.	**January 26**
Feedback for students You'll be asked to provide feedback on student's commitment and engagement during the project so far. This feedback should be shared with the program, and a version should be shared directly with your students. Your feedback will also be factored into students' grades.	**March** (Dates TBD)
Project work ends by Team meetings and project work should end during the last week of classes. Students will begin working on their poster presentation in mid-April.	**May 1**
Data Discovery Symposium Students will be presenting posters about their work with you. We warmly encourage you to attend!	**May 7 (tentative)** 1:30 pm
Student holidays Students will not be attending classes during several holidays, and we ask that you not schedule meetings or deadlines during these days.	**January 19** **February 16** **March 23-27**

Forming your student team

- Please select 3-5 students to work on your project. You can extend up to 7 offers at a time to find your team. After 5 students have accepted, remaining offers will be rescinded.
- We recommend scheduling 15-minute video interviews with applicants after an initial review of submissions.
- Keep in mind that students have applied to multiple projects and may be considering other offers. If a student hasn't responded to an offer within 72 hours, you may rescind the offer so that you can invite another student.
- We recommend selecting a more senior student to serve as a student leader.
- Consider also including a student with less experience.
    - All students require specific training; less advanced students who have more time remaining at Berkeley can return to your project in the future.
    - Data Discovery projects are a unique and transformative learning opportunity. Help us offer it to as many students as possible and invest in their potential.

Tips for student interviews

- MIT's Undergraduate Research Opportunities Program's interview tips page has great advice relevant for Data Discovery interviews.
- The interview isn't just about assessing applicants' skills. Also:
    - Share important details and confirm students' understanding.
    - Introduce yourself and answer students' questions about you and your work.
    - Give students a chance to directly address any concerns you might have.
- Clearly communicate details and expectations for the interview ahead of time.
    - Send a calendar invite with relevant links.
- Interview preparation
    - Consider sending interview questions ahead of time to put students at ease and to get thoughtful responses.
    - Provide websites or short documents for students to review.
- Example interview questions
    - You wrote about *[experience or project]* in your application. Can you elaborate on that?
        - How did it prepare you for this project?

        - What did you learn from that experience?
    - What are you looking to get out of this project and Data Discovery?
        - Can you say more about why you applied to this project in particular?
    - The weekly commitment for this project is 6 hours per week. How do you see that fitting in with your other commitments this semester?
    - Can you tell me about a time when you practiced or demonstrated *[some relevant skill (e.g., collaboration)]*?

NDA and confidentiality

- Some projects may require students to sign a Non-Disclosure Agreement (see a UC Berkeley-aligned format for your reference).
- NDAs should clearly define what information is confidential and how it can be used.
- Avoid overly broad restrictions that could interfere with students' academic work.
- We **do not allow financial penalties** for students. These can create unnecessary anxiety and detract from the educational focus of the program.
- The goal of your NDA is to protect your project while maintaining an open, supportive environment.

Computational Resources

- Data Discovery projects often require substantial computing capacity for data processing, analysis, and modeling.
- We are committed to providing both students and partners with access to high-quality resources to ensure project success.
    - All students can request access to Savio, Berkeley's supercomputing infrastructure, to run compute-intensive Python and R jobs.
    - In addition, students and partners can request access to the CDSS DataHub, part of the NSF NAIRR (National Artificial Intelligence Research Resource), to run large-scale Python and R workloads, particularly those involving machine learning, AI model training, or high-volume data processing.
    - Both Savio and CDSS DataHub environments come with a wide range of pre-installed open-source and licensed software, including Python data science

libraries (NumPy, pandas, scikit-learn, TensorFlow, PyTorch) and R statistical packages (tidyverse, caret, ggplot2).
    - After project teams are formed, email cdss-datadiscovery@berkeley.edu to request access to Savio and the CDSS DataHub.
    - The Data Discovery team will provide onboarding instructions for students and partners.

Contact information

General inquiries can be directed to **cdss-datadiscovery@berkeley.edu**. This is a shared inbox, and you may get a response from:

- **George Obaido, PhD.** Director, Data Discovery
- **Emily Remirez, PhD.** Co-Director, Data Discovery

Communications may also come from CDSS personnel who advise Data Discovery, including

- **Prof. Narges Norouzi,** Associate Dean of Students
- **Robbie Powers,** Assistant Dean of Students
- **Prof. Deb Nolan,** Faculty Advisor

References

https://urop.mit.edu/mentors/resources/interviewing/
https://www.augusta.edu/curs/documents/interview-tips.pdf