



## Data+ 2022: Request for Proposals (Faculty-sponsored Projects)

### Overview

Data+ is a ten-week summer research experience that welcomes Duke undergraduates interested in exploring new data-driven approaches to interdisciplinary challenges.

Students join small project teams, collaborating alongside other teams in a communal environment. Each undergraduate participant receives a \$5,000 stipend. They learn how to marshal, analyze, and visualize data, while gaining broad exposure to the modern world of data science. There are typically around 25-30 such teams. We hope to be in-person this summer. If so, all teams will have dedicated workspace in Gross Hall, provided by the Rhodes Information Initiative at Duke (iiD), the Social Science Research Institute (SSRI), and the Energy Initiative. The last two summers we have successfully run Data+ as a remote/hybrid experience, and we are prepared to do so if this is required.

Data+ 2022 will run from May 31<sup>st</sup> until August 5<sup>th</sup>. As the communal atmosphere is essential for student success, please **note that Data+ projects only run during these ten weeks, and that all student participants are required to contribute full-time efforts (no employment, no other classes)**.

This document is a call-for-proposals for faculty-sponsored Data+ projects in the Summer 2022 edition of Data+. We are especially interested in proposals that involve a partner from outside the academy, or a faculty member from a different discipline. We also encourage proposals that involve previously untested ideas or un-analyzed datasets, and we hope that the Data+ team can make a contribution with important proof-of-principle work that may lead to more substantial faculty work and/or connections in the future. We also welcome proposals that will lead to the undergraduates creating tools that might be used in the classroom or that might facilitate community engagement with data and data-driven questions. Finally, we are including a special segment within Data+ 2022 called Climate+, for projects connecting data science to issues around climate change. Please see the description at the end of this document to understand this additional opportunity.

## Deadline and Contact

The deadline for completing this application is **November 1st, 2021, end of day**. If you would like help in developing your proposal, please contact Paul Bendich ([bendich@math.duke.edu](mailto:bendich@math.duke.edu)) and/or Greg Herschlag ([gjh@math.duke.edu](mailto:gjh@math.duke.edu))

Please submit your proposal via email to Ariel Dawn ([ariel.dawn@duke.edu](mailto:ariel.dawn@duke.edu)) by **November 1st, 2022, end of day** for consideration.

## How to Apply

To apply, please prepare a document (three pages maximum) that responds to the following prompts, ideally in this order.

*Name of Project:* Name of Project: Please use a short name that succinctly describes the nature of the project and is not overly technical. If your project is selected for Data+, then this title will be used for the project web page and project listings, although we may reach out to suggest shorter and punchier names if required.

*Summary:* Please write a project summary, including the basic ideas behind the proposal.

*Faculty Leads:* Data+ is especially interested in projects that connect faculty from distinct disciplines, as well as projects that enable faculty to branch out in new directions. Please describe the intended faculty leads, and the expected benefits from their participation.

*Mentoring:* Day-to-day faculty involvement in Data+ is not expected. Instead, each Data+ project has a mentor, usually a graduate student or post-doc, who is on hand to give the student team more focused guidance. The time commitment tends to be 5-7 hours per week, and funding is generally available to cover this person's time. If you have a mentor in mind, please indicate who this is and why they are well-suited. If you do not, please describe the skills you would like this person to have (we are generally able to find faculty-mentor matches)

*Goals:* describe the intended goals and products of the project, in the following manner:

1) Describe entirely reachable goals that you fully expect the students to achieve: these could be answers to a question, explorations of a hypothesis, and things of that nature.

2) Describe a tangible product the students will create in the course of their research, which ideally will be of use both to further researchers at the university and to the students as something they can show off to future employers or graduate schools. This could be, for example, a good piece of well-commented software, or a visualization device, or a detailed curation of previously raw data.

3) Describe a more outrageous goal that you would be quite (pleasantly!) surprised to see the students achieve, along with a plan for them to build a potential roadmap towards that goal. For example, this goal might only be reachable if you had data that you currently do not have, and the students might build a speculative roadmap towards acquiring that data

*Data:* Most Data+ projects involve analysis of datasets. Some of these are publically available, and some are not. **As it is essential that students actually be able to analyze the needed data for the project, we are very interested in plans to ensure that this will happen.** Please address this in the following manner:

1) For each dataset that will be analyzed by the student team, please give a high-level description of the dataset (what's in it, how was it collected, and for which purpose, how large is it, etc. . . .)

2) For each dataset, indicate whether you anticipate IRB approval will be needed for student access, and if not, why not. If IRB approval will be needed, indicate whether a protocol already exists, and your plan for incorporating the student involvement. If it does not already exist, please describe your plan (including a timeline) for obtaining one.

3) For each dataset, indicate whether it is owned and/or is being provided by an outside party. If so, please describe the intended path towards ensuring that students will be granted the ability to access the dataset (we are often able to assist in crafting Data Use Agreements with outside parties, for example)

*Outside Partners:* Some of the best Data+ projects have a partner from outside of the university, or at least from outside the traditionally academic parts of the university. This might be someone who is invested in the data or the questions, and to whom the students will in essence “deliver analysis and insight” Ideally, this partner will be able to come two or three times during the summer to hear updates from your students and provide feedback.

For each such partner, please describe their expected interest in the project, how much they would interact with the team, and please also identify a point-of-contact for this partner.

*Funding:* Please indicate if you or some other entity, including an outside partner, would be able to contribute funds towards the student stipends on your team.

### **Special Call for Climate+ Projects:**

This year we are seeking to host 4-5 Data+ projects that are thematically focused on climate change, which we are calling “Climate+”. Climate+ is aligned with Duke University’s commitment to advancing interdisciplinary understanding of the causes and societal impacts of climate change as well as potential solutions for long-term sustainability including climate change mitigation and adaptation strategies. Climate+ takes a wide-angle view of relevant topics, that

extends well-beyond climate science stretching across disciplinary boundaries including (but not limited to) such areas as:

- Climate and inequality (jobs, justice, and economy)
- Climate and health (lungs, heat, and pollution)
- Climate and oceans (fisheries, hurricanes, and coasts)
- Climate and biodiversity (forests, species, and ecosystems)
- Climate and energy (systems, resources, and policy)

All projects must meet all Data+ requirements to be considered.

**Not sure if your project fits?** If your project is thematically close, but you are not quite sure if it's an exact fit, don't hesitate to submit it as a potential Climate+ project. Projects that are submitted as prospective Climate+ projects but are not deemed to be thematically aligned during the review process will receive consideration as a prospective Data+ project.

**Benefits to proposers:** Compelling projects focused in this thematic area will receive preferential consideration.

**Benefits to students:** Climate+ students will be part of a subcohort of thematically related projects that will engage regularly, in addition to the wider Data+ content, with climate, environment, and energy researchers and practitioners.

**How to be considered for Climate+:** Those projects proposing in this thematic area should note this on the top of their proposal. Data+ projects that fit topically as Climate+ may be considered for either program.