Data Expeditions

Call for Proposals

The purpose of this call is to introduce more undergraduate students to exploratory data analysis early in their Duke experience, and to involve graduate students in thinking about the way classes can interact with data. Our hope is that expeditions will encourage students to be more adventurous in exploring the Duke curriculum and that students with deeper skills will be capable of deeper insights.

What are Expeditions?

The Information Initiative at Duke (iiD), in partnership with the Social Science Research Institute (SSRI), will support pairs of graduate students to prepare a data set for use in an undergraduate class and then assist the faculty instructor by supervising the data expedition within the class.

Graduate students who participate become Pathfinder Fellows and receive a (tax free) grant of $1500 for academic-related travel or computers. They add to our undergraduate curriculum through expeditions and we reciprocate by investing in their intellectual development

How are Expeditions Organized and Funded?

iiD provides resources, SSRI stores the datasets for later use, and representatives from affiliated departments provide direction. Departments interested in participating are encouraged to contact the iiD Director Robert Calderbank (robert.calderbank@duke.edu).

Application Process

Applications will be reviewed by a faculty committee, those received by May 25 will receive full consideration, and funding decisions will be made by June 25.

Graduate students are encouraged to contact Paul Bendich (bendich@math.duke.edu) for help in developing ideas.

We particularly encourage exploration of data sets that bring different intellectual communities together, and we also place a special emphasis on expeditions in the introductory undergraduate classroom.

Application Details

Email Kathy Peterson (kathy.peterson@duke.edu) a PDF, at most 2 pages, with the following information.

1. Sponsoring faculty member and target undergraduate class
2. Title of dataset
3. Description: A brief data description that includes (at least) the following information:
   - one-two sentence description of data file
   - source(s) where the data come from
   - why the data were collected in the first place
   - how the dataset was put together
   - dimensions of the dataset
4. Potential Classroom Exercises: List of potential questions that can be explored using this dataset, and description of pathways towards answers the students can take.
5. Techniques: List of computational techniques – this is an opportunity to ask for access to a virtual machine that comes pre-loaded with different software packages
6. Source(s): Properly formatted citation of data source(s)

Expeditions recommended for an award will be asked to provide a Markdown or HTML document that contains, in addition to the information listed above,

7. List of variables: A list of the variable names and brief description for each (with hyperlinks to Codebook below)
8. Codebook: Description of each variable and its values

We will provide templates in .md, .Rmd, and .html formats so that expeditions can be documented consistently

Robert Calderbank and Jerry Reiter (iiD)

Mine Cetinkaya-Rundel (Statistics)   **Emma Rasiel** (Economics)
Merlise Clyde (Statistics)   **Tom Nechyba** (Economics and SSRI)
Candice Odgers (Political Science)   **Victoria Szabo (**ISIS**)**
Anita Layton (Mathematics)   **Hans Van Miegroet** (Art, Art History and
Harold Layton (Mathematics)   Visual Studies)
Paul Bendich (Mathematics)
Richard Lucic (Computer Science)   **Scott deMarchi** (Political Science)
Lisa Huettel (Electrical Engineering)   **Larry Carin** (Electrical Engineering)