Data Science Education Research Call for Dataset Proposals

Research Overview:

This project is sponsored by Arkansas's NSF EPSCoR DART Seed Grant program. Our objective is to collect a series of datasets from different disciplines related to the research of faculty at Arkansas Tech University (ATU). We will use those datasets in development of data science course projects, capstone projects/senior designs/practicum course projects. These projects will be used in data science classes offered in different educational institutions in Arkansas. Consequently, any dataset selected will be openly available in Arkansas. Finally, an incentive to participate we will compensate you if your dataset is selected.

Dataset Proposals:

If you have dataset(s) to contribute, please read and follow the instructions below.

1. You will be compensated if your dataset is selected.
2. Any dataset related to your research could be submitted. The dataset size could be big or small. The type could be numerical or categorical.
3. The datasets selected will be made public to DART and shared with all the data science courses offered in educational institutions in Arkansas. However, datasets that are not selected will not be made available. **Please do not contribute sensitive data.**
4. Please include the following in your submission:
    a. the objective of the research for which the data were collected,
    b. the source of the data,
    c. the number of variables in the data and variable descriptions,
    d. the size of the dataset,
    e. any background information available in text, image, and videos. -- *We want the students in data science courses to know the stories behind the data.*
    f. if data collection has been completed. If not, please estimate when it's going to be completed,
    g. any publications and analyses that have been completed that have involved the dataset,
    h. *if the dataset is fully cleaned -- We welcome datasets which are somewhat messy. We want data science students to practice working on the data wrangling and cleaning.*
    i. Any other information you feel is pertinent.

Stipend Information

Estimated Number of datasets to collect: 10-20 (pending availability of funds)

Anticipated All Datasets Funding Total Amount: $14,000

Dataset Proposal Deadline:

Review of the dataset proposals has started in December 2021. Submitting proposals early is strongly recommended.

Dataset proposals should be submitted via email as a single PDF email attachment by **March 31, 2022, no later than 5:00 PM** Central Daylight Time to any of the research team members (contact information listed below).

Research team members:

Dr. Jacob Grosskopf, Department of Physical Sciences, Email: jgrosskopf@atu.edu

Dr. Weijia Jia, Department of Mathematics, Email: wjia@atu.edu

Dr. Christopher Kellner, Department of Biological Sciences, Email: ckellner@atu.edu

Dr. Wan Wei, Department of Accounting and Economics, Email: wwei@atu.edu

Dr. Matthew Wilson, Department of Agriculture, Email: mwilson58@atu.edu

Dr. Xinli Xiao, Department of Mathematics, Email: xxiao@atu.edu

Please do not hesitate to contact us if you have any questions. We'd like to hear from you!