

Ten Years of Introducing Youth to the Power of Data Science



2014 - 2024

Board of Directors

Judy Cameron, PhD
Executive Director, The DataJam
University of Pittsburgh
Dietrich School of Arts and Sciences

Raja Sooriamurthi, PhD
Carnegie Mellon University

Beth Schwanke
PITT CYBER

Cheryl Begandy
Secretary, The DataJam
PSC

Devashish Saxena
Imperial Dade

Brian Macdonald
President, The DataJam
ORACLE

Beth Bauer
Vice President, The DataJam
PosiROI

Catherine Cramer
ADVANCED SCIENCE RESEARCH CENTER
THE GRADUATE CENTER
CITY UNIVERSITY OF NEW YORK
SDSC
SAN DIEGO SUPERCOMPUTER CENTER

The DataJam, a 501c3, was formed by a collaboration with industry software and data scientists and faculty from the University of Pittsburgh, Carnegie Mellon University and the joint Pitt-CMU Supercomputer Center to work together to inspire youth to enter the field of big data science.

Partners and Sponsors

aiu3
Allegheny Intermediate Unit

NORTHEAST BIG DATA INNOVATION HUB

WEST BIG DATA INNOVATION HUB

PITT CYBER

PPG

Carnegie Mellon University

PSC
PITTSBURGH SUPERCOMPUTING CENTER

SDSC
SAN DIEGO SUPERCOMPUTER CENTER

UNIVERSITY OF PITTSBURGH

WOODS HOLE INSTITUTE

PosiROI



The Origins of The DataJam

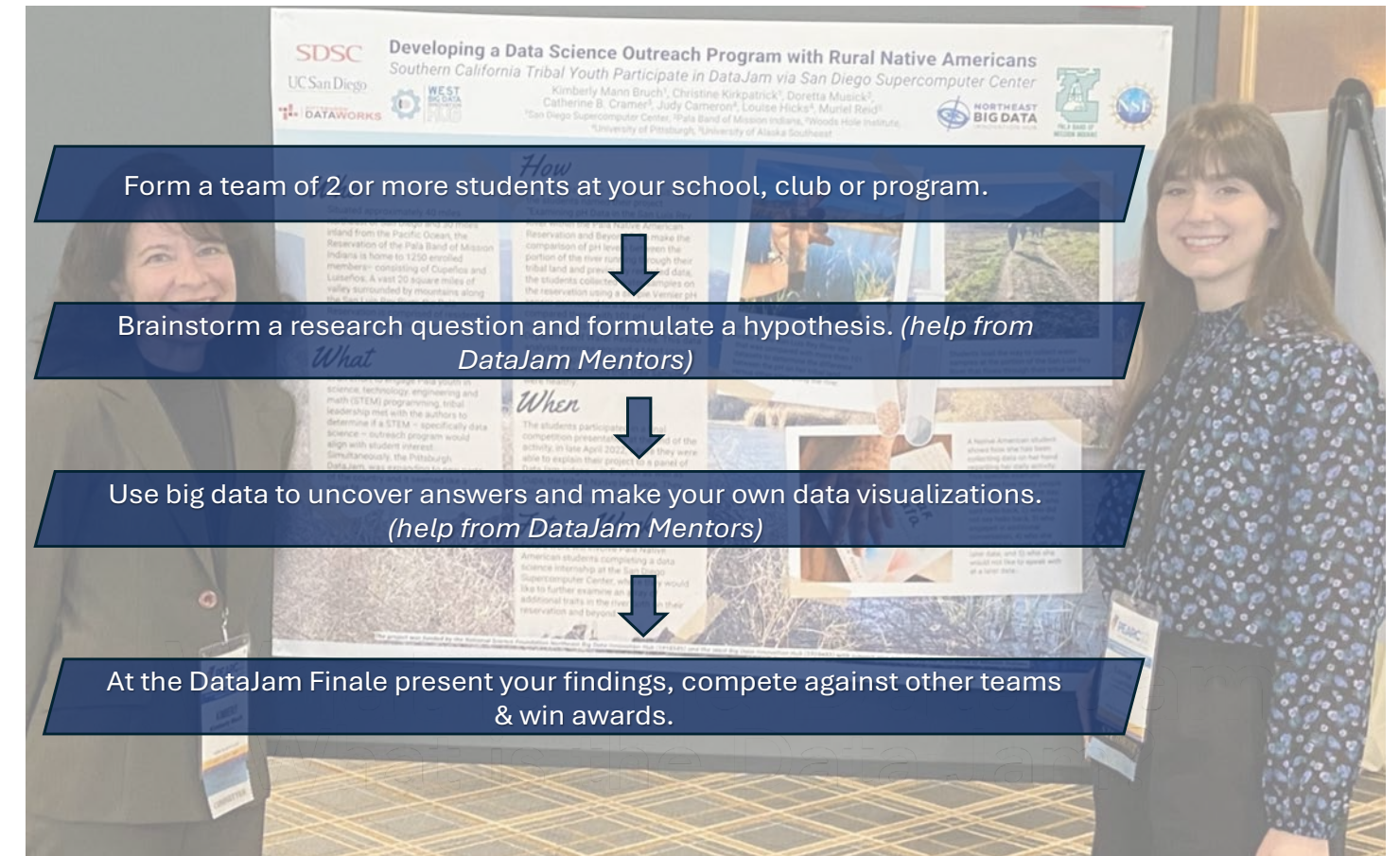
The DataJam is a mentored data science learning activity and competition for youth (high school, community college and middle school students) designed to introduce, encourage and engage young people in data science focusing on the topic area they are most interested in.

The DataJam is a nonprofit organization that was formed in the 2013-14 academic year in Pittsburgh, PA by software scientists from IBM & Oracle and faculty from the University of Pittsburgh, Carnegie Mellon University and the joint Pitt-CMU Supercomputer Center to work together to develop and run The DataJam to inspire high school youth to enter the field of big data science. The central goal was to familiarize youth with the field of big data science and make them more aware of the role of big data in their everyday lives. As virtually all companies and businesses rely heavily on big data and are helping shape the future using big data, ***the underlying concept was to build a pipeline of young people interested in and trained to participate in the big data science of the future.***

For high school students The DataJam is run annually throughout the academic year. For community college students The DataJam is run in the Fall (Sept - Dec), and in the Spring (Jan - May). Middle School DataJam Days are half day workshops designed to introduce students to the power of analyzing data to find answers to questions they are interested in.

All DataJam teams receive individualized mentoring from university students who receive specialized training in how to teach basic statistics, work with diverse communities and teach presentation skills.

How Does The DataJam Work?



Impact on Individuals

Vaishnavi Akella - student from the winning 2024 High School DataJam Team at San Ramon High School in Daville, CA

"Analzing data shows you new insights and reveals patterns in the world that are exciting to learn about."

Bob Moreland - teacher from Central Dauphin High School, Harrisburg, PA

"The DataJam projects not only motivate students, but they get them opportunities in the future. We had a student whose participation in The DataJam helped get him an internship with a very prominent insurance investment company."

Yuting Duan - student from the winning 2024 College DataJam Team at Skyline College in San Bruno, CA

"It's very good to work with a team because you can encourage each other and you can learn from each other."

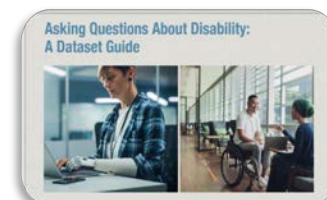
What are the top human talents needed by businesses?

Numerous surveys indicate that learning technology and data analytic skills are a key for future business success. They additionally show that innovators need the softer skills of collaboration, problem-solving, critical thinking and an understanding of the digital world. ***The DataJam teaches all of these skills.***

The DataJam teaches innovation, data analytic skills, collaboration and teamwork as well as presentation skills to tell a story with data.

Skills

Formulate a **research question**, find **data**, **analyze** data, make data **visualizations**, tell a **data story** and **present** findings to a panel of data science judges.

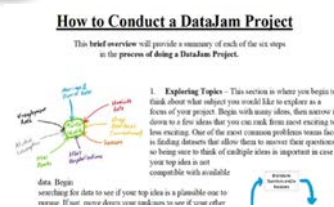


Excitement

A competition focusing on the **process** and **uses of big data** to answer research questions.

Resources

Free **resources** to all DataJam Teams **reduces barriers** to Data Science career entry.



The DataJam is the winner of the
2022 Carnegie Science Award
for “Best Interdisciplinary
Approach to STEM Education”

CARNEGIE **SCIENCE** CENTER



Receiving the award



Dr. Cameron & DataJam Mentors

Impact on Communities

High school teachers and college faculty value the **data science and presentation skills** that students participating in The DataJam learn. They are also very appreciative that **all resources are freely available** and that their **teams receive individualized mentoring**.

Bob Moreland, Teacher from Central Dauphin High School, Harrisburg, PA
“Just get involved, just get started with The DataJam. I guarantee not only will you have fun. Your students will be dragging you back into it the next year. It happens to me every year. As I’m getting the school year started, my students are saying “When will The DataJam be starting?” They may not even be students who were in it the year before. They’ve just heard about it from other students and they want to get involved.”

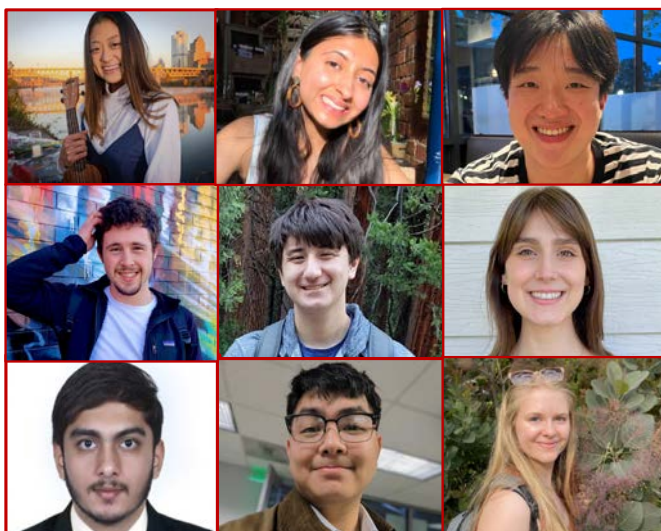
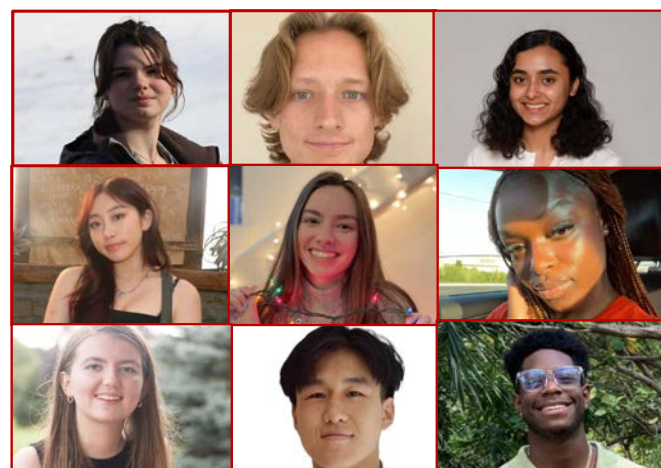
The Mentors

Training

The DataJam Mentors are university students interested in the power of data and data analysis, who are **trained to teach data science and communication skills, working with diverse** communities.

Culture

DataJam Mentors volunteer their time, creating a lifetime journey of **community** and **inspiration** based on **trust, caring** and **giving back** to their communities.



DataJam Mentoring

One of the unique features of The DataJam is that it is a mentored data science learning activity in which each DataJam team is paired with a university student who has been trained to teach data science and presentation skills to both high school and community college students. Mentors are also given the opportunity to assist with Middle School DataJam Days.

The DataJam has developed two college level courses to train DataJam mentors. The first is a course offered in the Fall semester each year at the University of Pittsburgh Frederick Honors College, which also provides a Blended Synchronous Learning Experience for students from other universities (currently Caldwell University in New Jersey, San Francisco State University and University of California San Diego), who join the class each week by Zoom. This course trains mentors to teach data science skills at various levels (from basic skills to beginners to advanced skills for students with significant statistics background). Mentors also learn how to prepare and present posters and slide presentations so they can assist DataJam teams with their presentations. A key feature of this course is training in how to work with diverse communities as DataJam teams come from a wide variety of communities (urban, suburban, rural, immigrant, homeless, Native American).

The same DataJam mentor training is also available for free online (<https://datajam-mentor-training.teachable.com/p/datajam-mentor-course>) in

Impact on Future Data Scientists

Sarah Sirakos, DataJam Mentor, University of Pittsburgh

"My love for statistics started when I took my first statistics class in high school. I hope to help The DataJam students share that same excitement and passion."

Carlos De Leon, DataJam Mentor, San Francisco State University

"I'm passionate about data and trying to find answers to questions we all have. I look forward to helping students find answers to their questions."

Alexia Cunningham, DataJam Mentor, Caldwell University

"My passion for data and its transformative potential drove me to seek out this experience, and it's been nothing short of fulfilling."

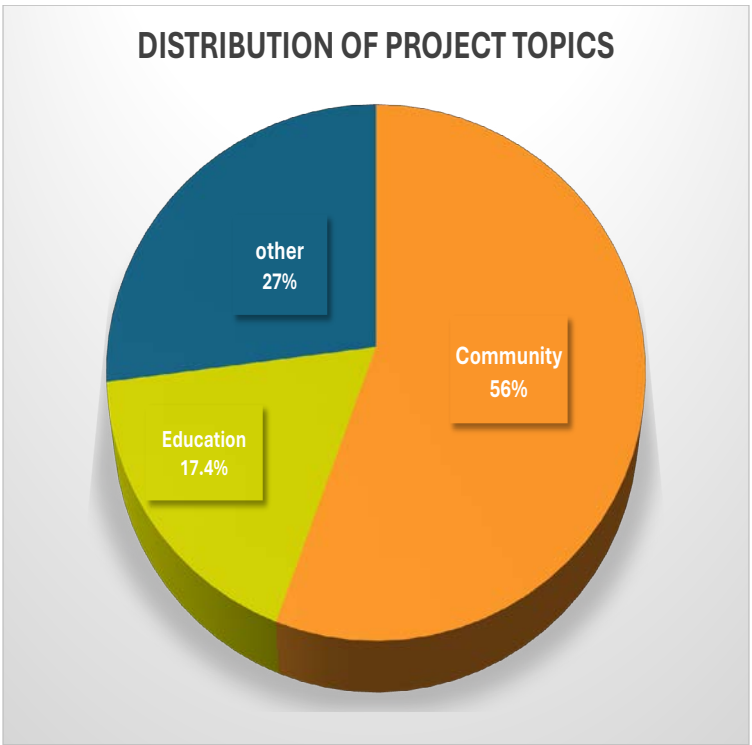
The DataJam Expansion

From the 2013-2014 academic year through the 2018-2019 academic year The DataJam grew every year, expanding within Pennsylvania. All mentoring, DataJam workshops and The DataJam finale were held in person. However, in the 2019-2020 academic year the Covid-19 epidemic started and many high schools closed or transitioned to classes being held online. The DataJam was fortunate to receive pilot funding from the National Science Foundation Northeast Big Data Innovation Hub, with further assistance from the West Big Data Innovation Hub, to transition to online mentoring and an online finale. With the onset of Covid, teachers from all over the country contacted us and asked if their students could participate in The DataJam. Having an online platform allowed The DataJam to rapidly expand to California and New Jersey.

We realized that it would be important to likewise expand mentor training across the country, to maintain the sense of community within The DataJam.

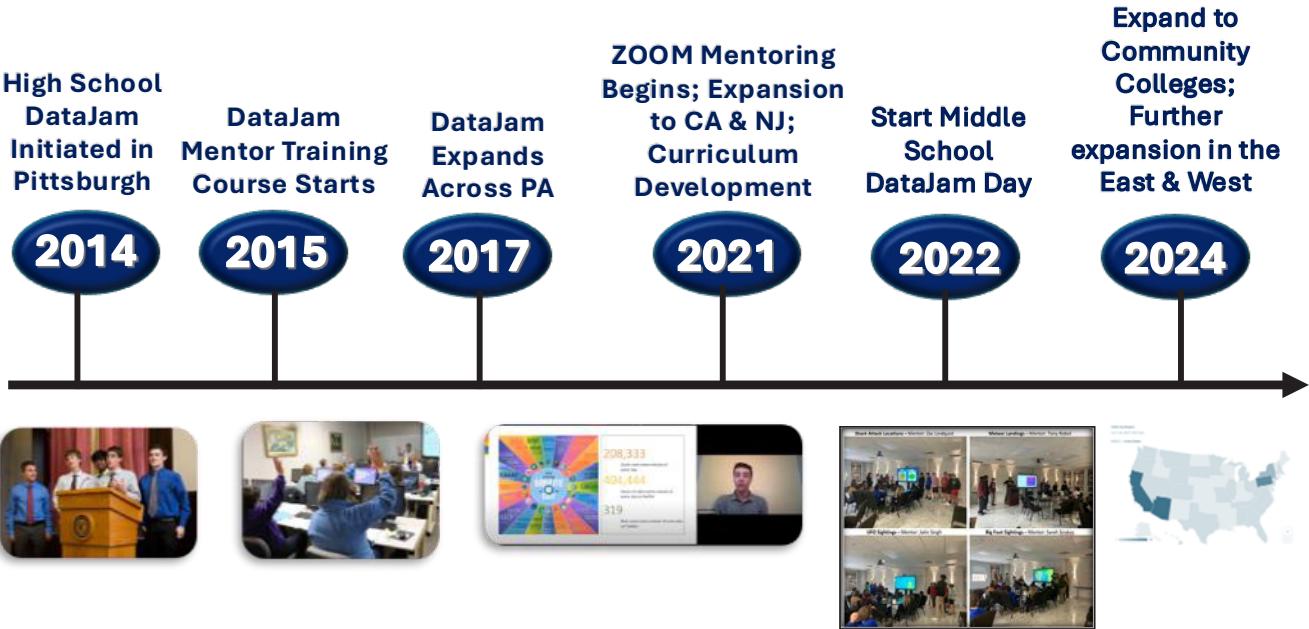
It is important to recognize that *The DataJam teams choose their own topics to study*, and *almost 75% of topics chosen have a civic focus --- they ask questions about their own community or their school.*

We are thus committed to strengthening the learning communities as The DataJam expands.



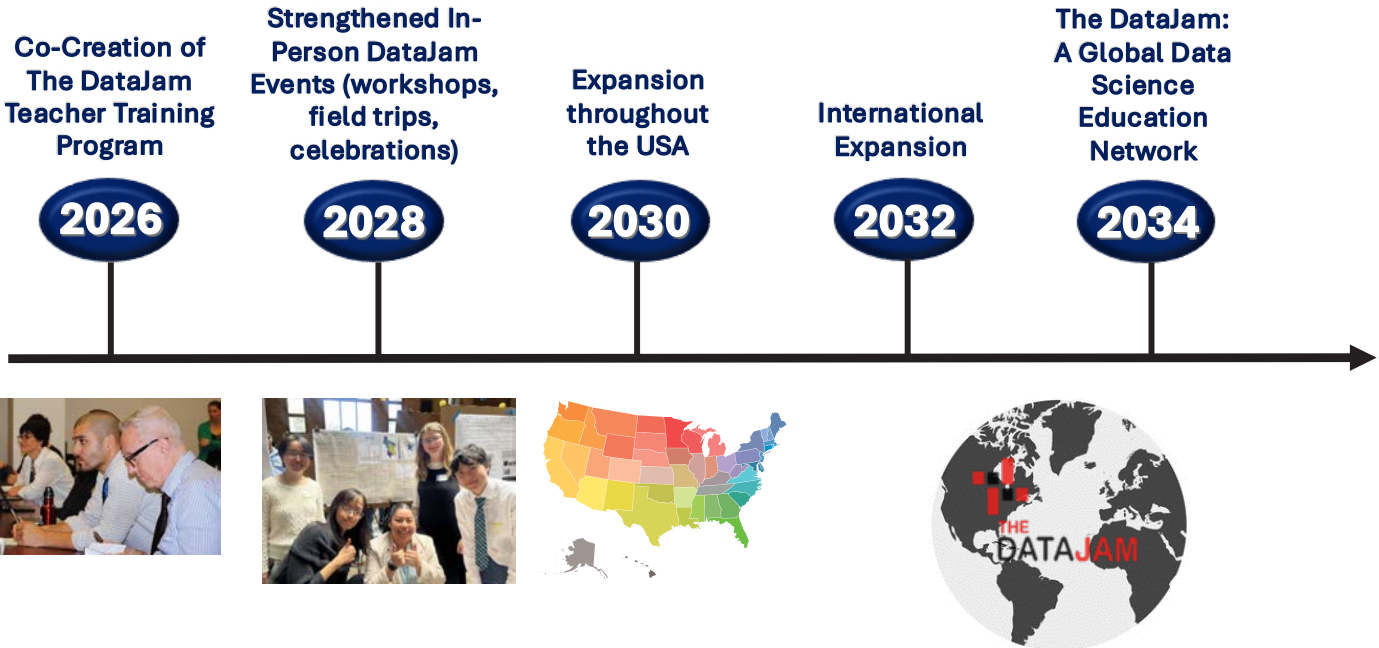
A SUCCESSFUL HISTORY ON A MISSION...

HOW WE GREW



...AND A VISION FOR THE FUTURE

ROADMAP ...CREATING A DATA FLUENT SOCIETY



Strengthening The DataJam Community

Creating Awareness of The DataJam

The DataJam Website (thedatajam.org)

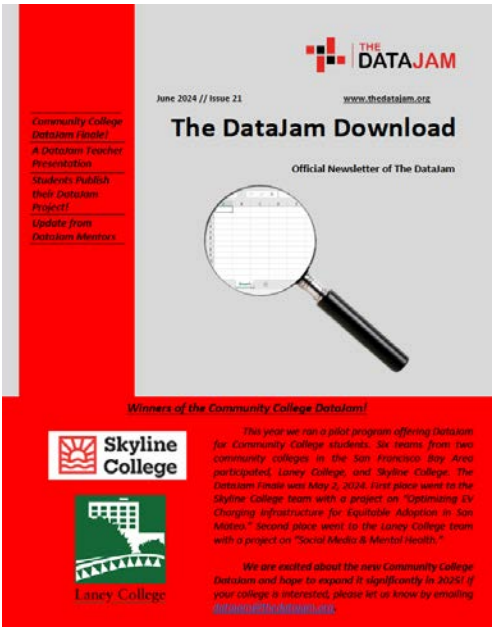
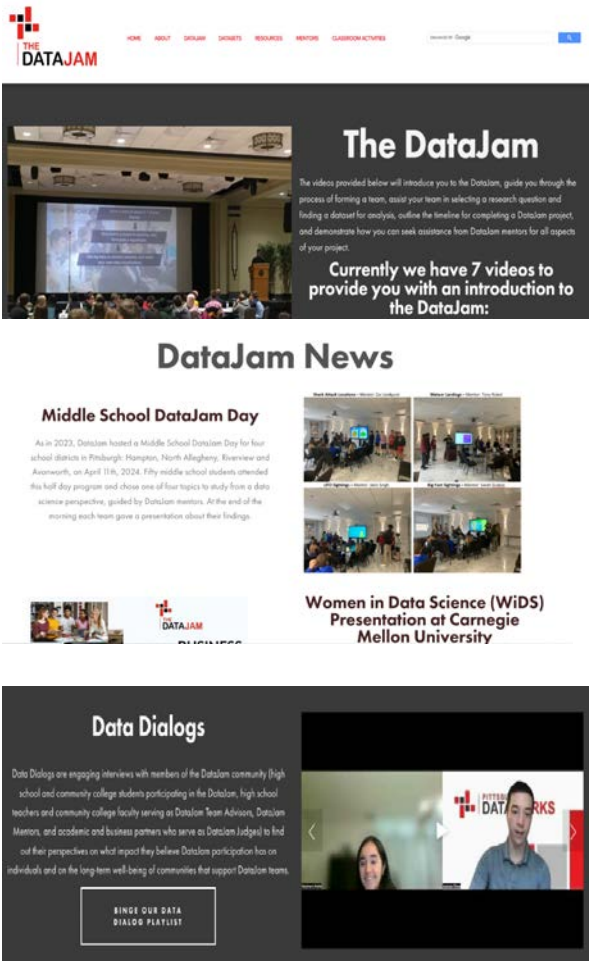
The DataJam website is updated every few days. It has The DataJam timeline, guide-book, informational videos, hundreds of resources for teams regarding designing and undertaking their projects, mentor’s bios, and posters from all past projects. There is also a News section with regular articles about what is happening in the various DataJam communities. The newest section of the website is “Data Dialogs”. These are interviews with members of The DataJam teams, teachers who serve as The DataJam advisors, The DataJam mentors, and professionals who serve on The DataJam Board of Directors and as The DataJam judges.

The DataJam Social Media

The DataJam posts regularly on its Instagram and Linked In accounts.

The Data Download

The DataJam has published a monthly newsletter for the past 21 months. Every month there are “Meet the Mentor” and “Meet the Data Science Professional” sections to share with The DataJam communities stories of the people involved with The DataJam and allow students to visualize pathways forward from where they are to using data science in their future. News highlights of what the various DataJam communities are doing keep everyone involved in The DataJam up to date.



Creating The DataJam Communities

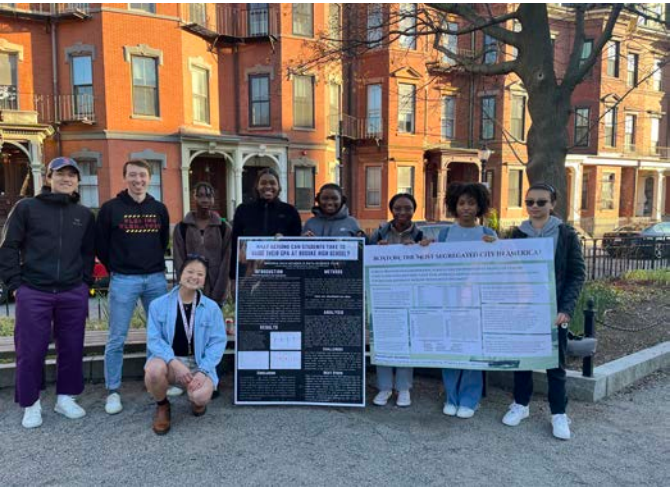
The DataJam in the San Francisco Bay Area

In 2023-2024 a strong DataJam community was developed in the Bay Area. The DataJam mentors were trained at SFSU. High School DataJam teams were active at San Lorenzo and San Ramon high schools. Community College DataJam teams were active at Laney and Skyline colleges. And in fact one of the SFSU mentors, Carlos De Leon, had transferred to SFSU from Skyline was delighted to help mentor the Skyline team. The Skyline team that won the 2024 College DataJam presented in person at the National Workshop on Data Science Education in June 2024 at UC Berkeley alongside Judy Cameron (director, The DataJam), Luella Fu (faculty at SFSU) and Kyla Oh (faculty at Laney).



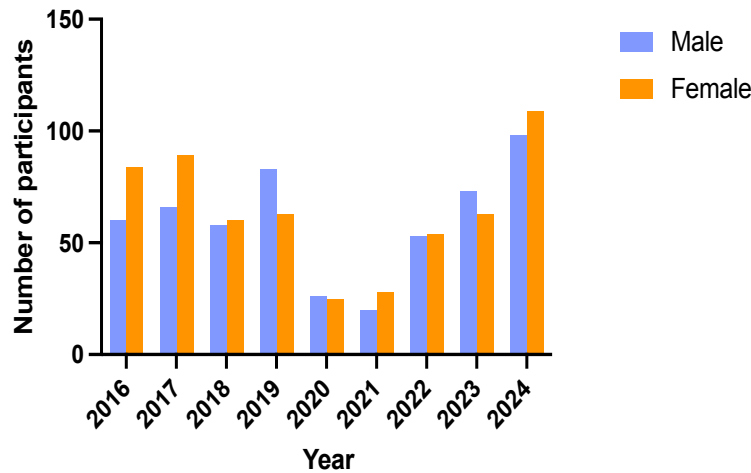
The DataJam in Boston

In 2023-2024 Yunge Xiao, who had been a DataJam mentor in Pittsburgh the previous year, moved to Boston as a graduate student at Harvard and accepted a DataJam Fellowship to expand The DataJam into Boston. Yunge recruited mentors from Harvard and Northeastern University and they took the online DataJam mentor training. She then recruited The DataJam teams from two high schools, Brooke High School and Boston Latin Academy (BLA). The BLA team received a “Best Presentation” and “Best New Team” awards at The DataJam finale. The Friday after the The DataJam finale all of those involved in The DataJam Boston gathered to celebrate in person. The Boston DataJam teams are planning to take DataJam field trips to visit companies using big data in the Fall of 2024.



Impact on Access to Digital Skills

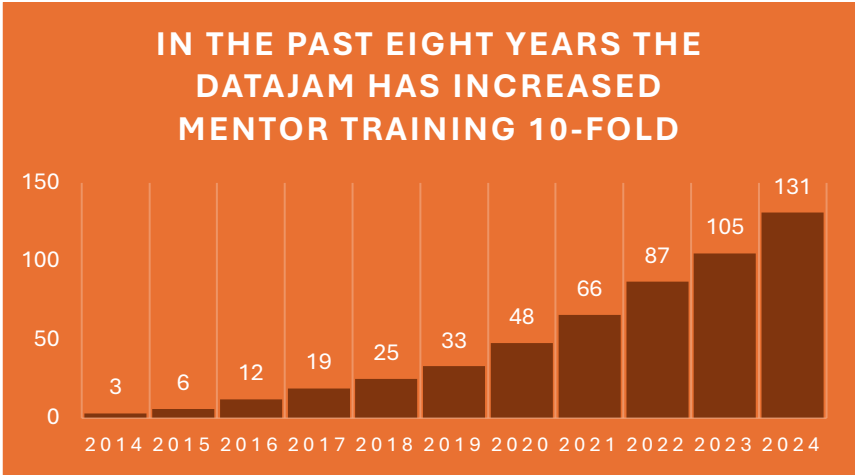
Over the past 10 years 1,112 students in the United States have participated in The DataJam, with similar numbers of females and males each year. We have been very pleased that overall 52% of The DataJam participants have been female.



In the 2024 annual student survey 98% of the students felt that participation in The DataJam made them more aware of the importance of data in all fields. 76% would recommend participating in The DataJam to others. Very exciting are students who have participated in The DataJam in high school and gone on to become mentors in college.

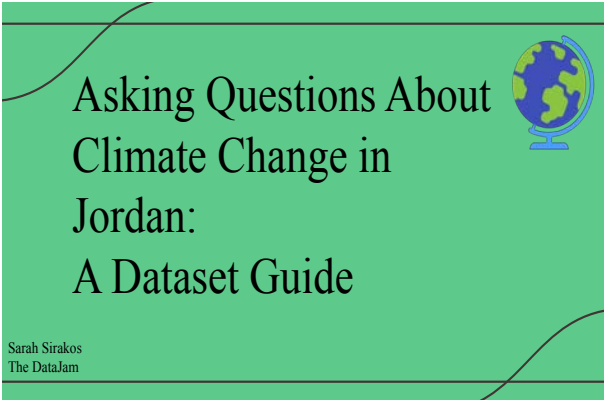
Most schools participating in The DataJam have done so repeatedly, with one school participating all 10 years. The average number of years schools have participated is 2.79 ± 4.24 .

Over the past 10 years 131 college students have been trained to serve as DataJam Mentors. The rate of training mentors has been escalating in the last two years as we have added mentor training in universities across the country and an online mentor training course.



Cummulative number of DataJam mentors over past 10 years

In June 2024 The DataJam was offered for the first time internationally, with 18 teams in Jordan participating. The DataJam Jordan was managed by the Global Center on Climate Change, Water, Energy, Food, and Health and implemented in Jordan by Climate Action Network Jordan (CAN), in cooperation with the the Jordan Ministry of Adaptation Fund, the United Nations Human Settlements Programme in Jordan and UNICEF Jordan. Resources, such as the “Asking Questions About Climate Change in Jordan: A Dataset Guide” were made by DataJam mentors in the USA (through the Fredrick Honors College at the University of Pittsburgh). The online DataJam mentor training course was translated to Arabic for use training mentors in Jordan. Teams first presented their projects to judges from the partner organization in Jordan and then first, second and third place awards were given after judging by the DataJam Board of Directors. True collaboration was key to the success of this initiative.



The DataJam Awards Ceremony in Jordan

Over the past 10 years The DataJam has proven to be a data science learning activity and competition that is widely enjoyed and valued by youth from 12-22 and highly appreciated by teachers. The DataJam has succeeded in introducing, encouraging and engaging young people in Data Science. How do we know this? Students, teachers, mentors and judges have eagerly communicated their enthusiasm for The DataJam with us!

The DataJam Vision for the Next 10 Years

Mission

Grow communities of learners who use data, analysis, and critical thinking skills to better understand and impact the world around us.

Vision

Build a data fluent society where data ethically and responsibly informs shared understanding.

The Next Ten Years...

Our goal is to build on the successes of the first ten years of The DataJam by *further creating awareness of The DataJam, continuing to strengthen the community focus of The DataJam, and facilitating access to digital skills for all youth.*

- We envision development of **networks within The DataJam communities** consisting of students participating in The DataJam, the high school teachers and community college faculty advising The DataJam teams, The DataJam mentors, The DataJam judges, and the businesses supporting The DataJam. *These networks will play a key role in conveying to youth the utility of data analytical skills and give them the abilities, motivation and connections to become data fluent and contribute meaningfully to the data driven world of the future.*

- We envision working with high school teachers and community college faculty to **co-create a training program for advisors of The DataJam** that will allow them to best facilitate participation in The DataJam by the wide diversity of students we hope to engage with The DataJam. *Just as we have adapted to meet the needs of student and mentor stakeholders, we recognize that teachers are an important stakeholder community essential for the success of The DataJam.*

- We envision **expansion of The DataJam to all communities of students** in urban and rural communities, immigrant communities, underserved and underresourced communities, as well as to international communities. *We recognize that the world is becoming more and more of a global community in which providing opportunities for each person compounds the opportunities for all.*

Vision for the future:

Community
focus with
national &
international
accessibility

