

The DataJam Download

Official Newsletter of the Pittsburgh DataWorks



***Carnegie Science
Center Award***

***It's a Great Time to
Start Working with
a DataJam Mentor***

Meet the Mentors

***Meet the
DataScience
Professionals***

***Evaluation of
DataJam 2022***

***DataJam Timeline
for 2022-2023***

***Pittsburgh DataWorks is pleased to have won the 2022 Carnegie Science Award
for
Best Interdisciplinary Approach to STEM Education***



Dr. Judy Cameron, Director of Pittsburgh DataWorks (second from right), with mentors Taylor Mathis (right), Jackson Filosa and Callie Labaucher (left).

It's a great time to start working with a DataJam Mentor

Teams interested in participating in DataJam 2023 are starting to work on their DataJam proposals. **See the DataJam Project Proposal Template on page 5 of the DataJam 2023 Guide.** You can download a copy of the Guide by going to the DataWorks website (pghdataworks.org), to the DataJam tab and scrolling to the bottom of the page. **Proposals are due Friday, December 2, 2022.**

Now would be a really great time to email Pittsburgh Dataworks at pittsburghdataworks@pghdataworks.org and ask to be paired with a DataJam mentor. Just tell us when your team wants to meet and we will work to pair you with a mentor to help your team narrow in on a research question you are interested in studying, look for datasets you will be able to analyze to address your question, write your hypothesis and think about what analyses you will do to answer your question. DataJam mentors can help with all these things and are eager to help your team!

You can get even more help from DataJam mentors if your team joins the DataJam 2023 Slack Workspace. To join the workspace, have your parents fill out and sign the permission slip that is on page 6 of the DataJam 2023 Guidebook. Make sure you put a non-school email address on the permission slip because most schools block use of websites like Slack so Slack will not be able to send you and invitation to a school email address. Email your signed permission slip to pittsburghdataworks@pghdataworks.org, and we will send you an invitation to join the DataJam 2023 Slack workspace. This is a place where you and your teammates can work on your project together, even when you are not at school. Any time you have a question, you can just send the question to your mentor and get answers quickly.



DataJam Project Proposal Template

School Name (& Team Number if there is more than one team/school):

Project Title:

Team Member Names:

Problem:

IN THE FORM OF A QUESTION STATE WHAT PROBLEM YOUR DATAJAM PROJECT WILL ADDRESS?

Why is it Important?:

IN A PARAGRAPH EXPLAIN WHY THIS IS AN IMPORTANT QUESTION AND WHO WOULD BE INTERESTED IN THE ANSWER.

Hypothesis:

A STATEMENT SHOULD BE MADE ABOUT WHAT THE TEAM THINKS THE ANSWER TO THEIR QUESTION WILL BE.

Data Required:

PROVIDE A LIST OF DATASETS YOU HAVE FOUND THAT YOU WILL ANALYZE TO ANSWER YOUR QUESTION. PLEASE INCLUDE THE FULL LINKS SO WE CAN LOOK AT THE DATASETS WHILE WE ARE LOOKING AT YOUR PROPOSAL.

Analysis Plan:

IN ONE OR TWO PARAGRAPHS DESCRIBE HOW YOU PLAN TO ANALYZE YOUR DATA, WHAT ANALYTICAL STRATEGIES YOU WILL USE AND WHAT DATA VISUALIZATIONS YOU THINK YOU WILL USE? (No work needs to have been done at the time the proposal is submitted)

Meet the Mentors

Taylor Mathis



Hi everyone! My name is Taylor Mathis, and I am a senior here at the University of Pittsburgh majoring in Statistics with a minor in Economics and a certificate in Global Health. Outside of school, I am on the club rowing team here at Pitt and am a research associate at the Institute for Clinical Research Education. I have been a DataJam all 3 years of my college experience, and I am so excited to continue working with all of the teams this year!

I began my mentoring journey in a course taught by Dr. Cameron, “DataJam: Using Big Data for Community Good.” The class not only refreshed my analytical skills but also taught me how multifaceted data is. Learning about the ethical issues when collecting data and discrimination that can arise from the way data is collected really intrigued me. It allows me to bring that into my mentoring style, making sure teams ask, “Where did my data come from?” and, “What implications may arise from what we uncover?”

My favorite part about mentoring is experiencing how creative all the teams are. The brainstorming portion of the DataJam is a huge section. Having so many different students, with so many different passions can make it difficult to settle on a topic. But the way that teams work together to mesh their interests in such innovative ways is really cool to see. Creativity not only plays a part in choosing a topic but in poster design at the end. It allows students to showcase their creativity and analytical skills in a tangible form. There is really something for every student whether they are into the arts or analytics or both! They are able to build up their collaboration, data analytics, and problem-solving skills in ways that aren’t necessarily available through the traditional school curriculum.

Through DataJam, I have been able to work on these skills as well. I have worked with other mentors over the years to present on Big Data as well as DataJam’s initiatives to improve diversity, equity, and inclusion (DEI). Being exposed to DEI initiatives so early in my college career has led me to my current research position. I am currently working with the Institute for Clinical Research on a workforce diversification study. I have been able to work on my quantitative as well as my qualitative data skills to evaluate interventions for underrepresented early-career scientists and scientist-physicians. Collaborating with epidemiologists throughout the Summer and the school year has made me more confident in my decision to apply to public health school. I’m planning to pursue an MPH in Epidemiology in the Fall, and my experiences with research and DataJam have prepared me well.

Tony Robol



Hello everyone! My name is Tony Robol, and I am a junior at the University of Pittsburgh majoring in Industrial Engineering and minoring in Economics. I am very passionate about data analytics and its applications in the modern world, and I believe it is a field everyone should be exposed to – I got exposed to this field myself through the DataJam and I never looked back!

My origins with the DataJam began when I was a senior in high school, before I even got to the University of Pittsburgh. During the hardships of COVID-19, I had the privilege of participating in a program through the Pittsburgh DataWorks in which we analyzed available COVID-19 data for trends and insights. I then came to discover the beauty of data – there were many ways you could approach using and analyzing it, and this fascinated me! As such, I was inspired to enter the field of statistics at Pitt and become a DataJam mentor to teach others about the versatility and applicability of data.

Throughout my career as a DataJam mentor, I have had the privilege of mentoring various teams, including my alma mater, Peters Township High School. Projects I worked on included determining the factors influencing charge-off rates in banking and collecting pH data from a river in California. As you can see, DataJam projects can be on anything you want them to be – provided there is enough data out there to answer your question! This allows you to pursue your passion as a student in the DataJam and, in my opinion, is a huge part of what makes the DataJam fun. In addition to mentoring other high school teams, I also spent a summer creating a written manual to guide high school students along with their projects. The manual contains background information on statistical concepts for students that have not yet taken a statistics course, as well as a detailed step-by-step guide as to how to conduct a DataJam project. I am very proud of the finished product and hope that many DataJam teams continue to use it in the future!

This past summer, I completed an internship at FedEx Ground, and am currently still working there part-time, with arrangements to return next summer. Data analysis is a huge part of my work, as I help to analyze the prices of contracts that are given to independent service providers to deliver our packages. As the world continues to turn towards data-driven decision making, it is becoming more and more important that job candidates have exposure to data analysis techniques/skills. Especially in a field such as data science, I believe it is never too early to have exposure to a career field and in turn insight into what you might want to spend the rest of your life doing. The DataJam is a fantastic experience for gaining such insight (along with the many other benefits), and I highly, highly recommend it! I look forward to working with all of you and am excited for the future of the DataJam!

Meet the DataScience Professionals

This month we are starting a new section of the newsletter where we will introduce you to Data Science professionals who work with Pittsburgh DataWorks and play important roles in running the DataJam. These professionals play important roles in recruiting schools in new areas to participate in the DataJam, coordinate training of DataJam mentors at new universities, connect the DataJam to businesses that help support DataWorks, often serve as judges for DataJam posters and presentations, and offer field trip opportunities to DataJam teams.

Beth Bauer



Hello DataJammers! My name is Beth Bauer, and I am a member of the Pittsburgh DataWorks Advisory Board. As Founder and CEO of PosiROI, a strategy and analytics consulting firm, my mission is to unleash the power of people using data and analytics to evolve our collective future purposefully and positively. I got hooked on statistics while a Psychology and Professional Writing major at Carnegie Mellon University. My Stats101 professor taught statistics like Indiana Jones taught archaeology—every day was an adventure! It was so exciting that I became a Statistics minor and later earned a Masters in Statistics from Rutgers University. I'm amazed that in 30+ years I've never lost that feeling of adventure. Working in data and analytics is a continuous learning and innovation experience. I'm grateful to bring this same passion to the Pittsburgh DataWorks Advisory Board. I hope to share data's possibilities with anyone who'll listen—so connect with me!

We've historically had few words to describe working with data. So, before Data Science was a common term, I used to call myself a Data Detective. Using scientific methods and figuratively taking a magnifying glass to big data, my teams and I revealed the sources of opportunities and measured the effectiveness of solutions for 26 Fortune100 companies. Working in data and analytics at organizations like the Robert Wood Johnson Foundation, IQVIA and Merck, and now PosiROI, has been a journey of shared discovery and trust. By also honing soft skills, like communicating complex ideas, problem-solving and relationship building, data professions provide the possibility to better understand our world. Relationships, subject expertise, and diverse lenses provide the context to connect insight dots. When we organize, analyze and share data insights, we can begin to hypothesize, test, learn and evolve toward ever more positive outcomes. A common data language is just gaining traction now and the DataJam is helping to spread it.

Cloud computing and open-source programming are helping democratize data and its hidden stories. I believe putting responsible data usage into many hands will positively change our world in ways we cannot yet imagine. And that is what excites me

most about the DataJam and its mission. We all have a responsibility to know where our data comes from, how reliable it is, and how we can best use it to exceed both our professional and personal goals. I am truly honored to join Pittsburgh DataWorks and help spearhead its New Jersey DataJam expansion. Data we both create and use is everywhere. You don't need to be a data scientist to learn to use data wisely.

Kimberly Bruch



Hello everyone! My name is Kimberly Bruch, and I am a DataJam Team Advisor and the Senior Writer at the San Diego Supercomputer Center.

As a science writer that often references “big data” for the articles that I write in conjunction with my position at the San Diego Supercomputer Center (SDSC) at UC San Diego, I have found my experience with DataJam extremely beneficial. I would like to share with you a bit about how the DataJam has enlightened my career path, as the program provides opportunities to both students and professionals.

When I was in graduate school studying telecommunications at San Diego State University's School of Communication, I truly enjoyed my statistics courses and used SPSS for my thesis project as well as many evaluation projects at SDSC. As the years went by, I changed my career course from using high-tech distance learning tools for virtual explorations to pure science writing, which does not require evaluations or the use of “hands-on” statistics. So, when I had the opportunity to work with Dr. Judy Cameron on expanding DataJam to rural California Native American sites, I was elated to find that I could be involved with statistics again! As a member of DataJam's Advisory Board, I truly enjoy working on the expansion of DataJam in southern California with Native Americans and I especially love teaching the students about the use of data science on projects that are important to them.

Evaluation of DataJam 2022

During the DataJam 2022 academic year, Pittsburgh DataWorks was very fortunate to receive a pilot grant from the NSF Northeast Big Data Innovation Hub to expand the DataJam outside of the Pittsburgh region. We successfully expanded participation in the DataJam to other regions of Pennsylvania (outside of Harrisburg), to New Jersey, Massachusetts, and California.

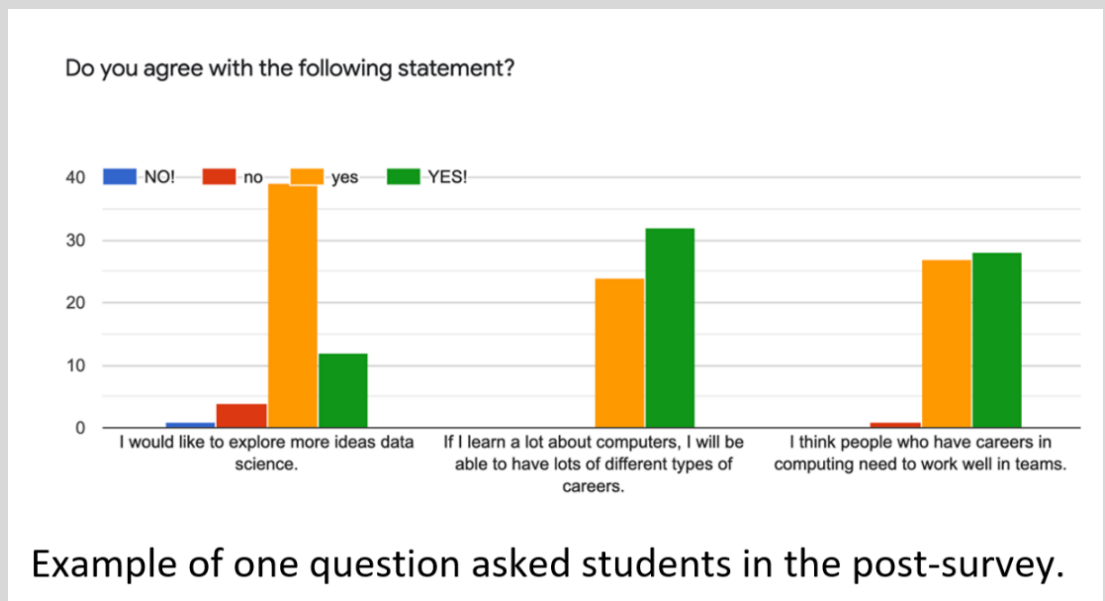
How was the evaluation conducted?

The evaluation was done using 3 instruments: (1) A pre-survey of 24 high school youth who participated in DataJam teams in New Jersey, Massachusetts, and Southern California; (2) A post survey of these same students and an additional 36 students who participated in the DataJam from schools in Pennsylvania, where the DataJam has been running since 2014; and (3) a focus group conducted with participating teachers. Data was gathered using Google forms and cleaned and tabulated using Excel. No identifiable information was gathered from the students through these instruments. Focus group data was gathered through transcripts of Zoom meetings with the participating teachers and using narrative analysis to extract general sentiment toward the program and the value that teachers believed their students derived from the program.

Lessons learned through formal evaluation of the DataJam pilot project

The formal evaluation of the pilot project to expand the DataJam outside of Pennsylvania, to more diverse populations of youth, showed a lot of promise for national expansion of the DataJam as an effective educational program to engage underserved communities in learning about data science. The pilot project was very effective in engaging females and youth from underserved populations.

The percentage of students knowing what they wanted to do in their career increased after participation in DataJam with 64% being interested in math/science careers. There was greater recognition of what data science was and students were able to give much more detailed answers to what data science was used for. Students clearly developed a good understanding that data science is used in virtually all fields of study. Although the pre-survey showed that the majority of students wanted to explore data science there was a marked increase in students interested in data science after the DataJam. Overall, we believe that the results of the pilot study suggest real promise for the DataJam to expand to a wide diversity of communities and a national platform for engaging youth interest in data science.



See the DataJam Timeline for 2022-2023

On the DataJam page of the website the new 2023 DataJam Timeline has been posted. Click [here](#) to see the Timeline.

- **Proposals due Fri., Dec. 2, 2022**
- **Feedback will be received by Fri., Dec. 16, 2022**
- **Posters will be due Fri., March 31, 2023**
- **2023 DataJam Finale will be Thur., April 27, 2023**

–

MID-AUGUST THROUGH SEPTEMBER

Schedule a DataJam Presentation at Your School

DataJam mentors are available to give a 15-minute presentation at your school to inform students about what the DataJam is, how to form a team, and what is involved in participating in the DataJam. Presentations can be scheduled by requesting a presentation at Pittsburghdatajam@pghdataworks.org. Presentations will be given by zoom. Alternatively, watch the 15-min DataJam Mentor Overview to the DataJam video found at the top of this page.

Register to be a 2023 DataJam Team

It is easy to register your team. Send an email to PittsburghDataJam@pghdataworks.org to let us know you have a team that would like to compete in the 2023 DataJam. Include: (1) the name of your school or group, (2) the name, email address and phone number for the team advisor (a

We are looking forward to DataJam 2023! **We Hope You Are Too!**

Email us at datajam@pghdataworks.org when you are ready to start working with a DataJam Mentor!