# The DataJam Download

**Official Newsletter of the Pittsburgh DataWorks**

## Academic Data Science Alliance (ADSA)



*Judy Cameron and Catherine Cramer from the Pittsburgh DataWorks Advisory Board attended the 2023 ADSA meeting in San Antonio, Texas and presented a poster about the DataJam, and participated in the "Learning and Doing Data for Good Workshop". The workshop was coordinated by DataWorks partner "The NSF West Big Data Innovation Hub" and included examples of how project-based learning of data science is pivotal in providing high school and college students with important experiences in using data science for community good. Presentations were about the DataJam, the Data Science Discovery Program at UC Berkeley, and Data Science for Social Good at the University of Washington. DataJam is proud to be a partner in the nationwide movement to introduce the next generation to the power of data science!*
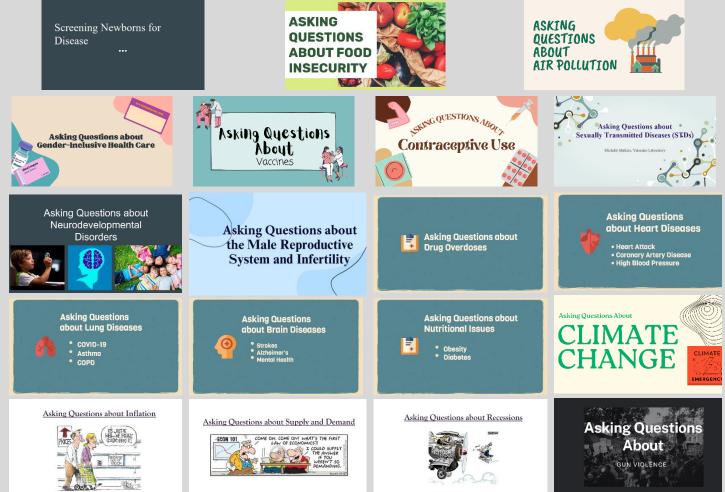
## *Unleash the Power of Data with Our Dataset Guides*

Are you ready to dive headfirst into work on your DataJam project? Before you start, it is important to recognize that the cornerstone of any successful data analysis is finding the right dataset. With the overwhelming abundance of data resources out there you may think this would be easy. But it is NOT! Finding the perfect data set to answer your research question can be a daunting task. But fear not, Pittsburgh DataWorks and our many collaborators are working hard to make data sets more available to you!

**Introducing Our Dataset Guides**

We understand the importance of finding the right dataset for your research or analysis. That's why we've curated a collection of DataSet Guides available on the DataSets page of the website. We urge you to look at them as they will help you access, understand, and harness the power of data on a variety of topics. Our guides offer more than just links to freely available datasets; they provide you with the guidance you need to formulate meaningful research questions related to specific topics. Whether you're exploring the realms of biomedical research, delving into climate change, navigating the intricacies of economics and finance, or pondering political science questions, our Dataset Guides have got you covered.

**Click any of the pictures below to see what information our guides can provide you!**

**Unlocking International, National and Regional Datasets**

In addition to our Dataset Guides, we've curated a wealth of datasets, both international, national and regional, to boost your research endeavors. Our curated International Datasets, maintained by international organizations, governmental bodies, and research institutions, offer a panoramic view of our interconnected world, allowing you to delve into diverse topics like global trade, climate change, health disparities, and much

more. You'll be equipped to analyze trends, foster international cooperation, and address the pressing challenges that affect us all.

Our National DataSets offer a wealth of data covering diverse topics pertinent to the United States. Explore demographic data, economic indicators, public health statistics, and more. Notable national databases include Data.gov, US Census Bureau, USA.gov, US Bureau of Labor Statistics, Centers for Disease Control and Prevention, National Institutes of Health & Human Services, and Google Trends.

Interested in diving into the characteristics and challenges of specific regions or localities? Regional DataSets provide insights into demographics, employment, health, education, and more. Regions we cover include California, Massachusetts, New Jersey, and Pennsylvania. With these datasets at your fingertips, you'll be well-equipped to inform policy decisions, track trends, and address the unique challenges of different regions. We will be adding new regions regularly. If you would like a specific region added, please email us, and let us know at datajam@pghdataworks.org.

## *Meet the Data Science Professional*

Hello! I'm Christopher, a current PhD candidate for the Human Genetics program at the University of Pittsburgh's School of Public Health. I started life in Illinois, as a First Generation American. From a young age, I have always enjoyed helping people, which I did so with volunteer service as member of the Boy Scouts, eventually earning my Eagle Scout Rank along the way. During that time, I also began to develop an interest in the sciences, rather specifically in the science of disease genetics. As a kid, I stumbled upon a TV show called "Mystery Diagnosis", where the show focused on stories of people and their journey to find the cause of their conditions. And although I would never be able to guess the illnesses, the process of putting each symptom together got me hooked. It was very similar to solving a puzzle, where I had to put the pieces together to obtain the whole picture. The diseases that caught my attention most often were genetic diseases, where all the various symptoms were tied to one small change. I had decided I wanted to mix my desire to help people with my interest in genetics early on in my education, though the path I took to get to where I am was rather roundabout.

I started by seeking a Bachelor of Arts degree with a Major in Chemistry, as I had a more passionate interest in chemistry than biology, despite my fascination with Genetics, a subject deeply rooted in biology. After earning that degree, I began to switch tracks, earning my first Masters of Science degree in Indianapolis, where I studied mainly medical biology related subjects. Afterwards, I decided to take a gap year in my education, which coincidently happened to line up with the start of the COVID Pandemic. During that time, I worked in a Grocery Store while I applied to Master programs specifically in Human Genetics, which is how I ended up at Pitt in the first place. At Pitt, I began to really learn about the scientific field I had been interested in for most of my life, I eventually found myself in Dr. Alexander Yatsenko's laboratory, where I would be able to help solve the "puzzle" that was patient's infertility, and hopefully improve the field's knowledge of the genetic traits that cause infertility.

I urge you to use DataJam, as it can be a great tool to discover an interest you may never have realized. While there may be stumbles on your journey to answer whatever question you may ask, you may be the only one curious enough to try and put the pieces together and solve the "puzzle" that is your question.

## _Meet the New Mentors_

This year for the first time DataJam is training mentors on the West Coast. Below you will meet our new mentors being trained at San Francisco State University. This is particularly exciting because we will now have more and more DataJam teams in California who will be participating in DataJam 2024.

### Dona Inayyah

Hi everyone! I'm Inayyah, a senior at SF State where I'm majoring in Applied Mathematics with an emphasis on Data Science. My passions lie at the intersection of math and social justice, and my goal is to become a data analyst so that I can contribute to solving pressing global issues. A little about me - I'm from Malaysia, and I have a deep love for travel and exploring new places. I'm a huge fan of sappy rom-coms and anything Harry Potter related, and I love spending quality time with people I care about. I am incredibly excited to work with my mentees, and I hope it is as enriching for them as I know it will be for me :)

### Kartik Patel

Hello, I'm Kartik Patel, a sophomore majoring in Applied Mathematics at SF State. My goal is to become a data scientist. In addition to my academic pursuits, I enjoy spending my time reading books and playing volleyball. I'm currently an active member of our volleyball team. My passion lies in working with data, especially when it involves big data and machine learning modules. I'm thrilled to embark on my first data science project and eagerly look forward to collaborating with your team.

### Carlos De Leon

Hi! My name is Carlos, I'm a Computer Science major here at San Francisco State University while also taking Data Science and Machine Learning for biotech. I'm passionate about data and trying to find answers to questions we all have which is why my end goal is becoming a data scientist. Other than school I also enjoy playing musical instruments such as the guitar and piano and reading mystery books. I look forward to being able to hear some interesting topics from the upcoming DataJam and help you find the answers to those questions.

### Akemi Smart

Hello everyone, my name is Akemi! I am a second year at San Francisco State where I am majoring in Statistics and minoring in Philosophy. My first exposure to statistics was in high school where I fell in love with its broad application and real-world usage, so I am so excited to maybe foster that love to other students. I've also found that statistics pairs very nicely with philosophy, which is one of my passions. A little bit about me, I am from San Diego, California and love the beach and being in the sun. My friends and family are the most important things in my life, and I love spending time with them as much as I can. Overall, I am so happy to have this experience and I'm looking forward to meeting you all!

## *DataJam Timeline for 2023-2024*

On the DataJam page of the website the new [2024 DataJam Timeline](#) has been posted. Click [here](#) to see the Timeline.

- **Proposals due Fri., Dec. 8, 2023**
- **Posters will be due Fri., March 29, 2024**
- **2024 DataJam Finale will be Thur., April 25, 2024**

**FRIDAY, DECEMBER 8, 2023**

**DataJam Proposals Due**

Teams should email their DataJam proposal to DataJam@pghdataworks.org. Use the template provided in the DataJam Guidebook, located at the bottom of this page.

Make sure the names of the team members (there must be at least 2 team members, but often are 5-8 members), and non-school email addresses are provided (note many schools block emails to student school emails from outside email addresses). Once team members have been identified, you and the students will be invited to join an Edtera Group for DataJam 2024. This is an online platform that every team can access on an internet browser where they can communicate with their team members about their project and also communicate with DataJam mentors to get help with their project.

The instructions for writing the DataJam Proposal are on page 5 of the [DataJam 2024 Guide Book](#), and a template for the one page DataJam Proposal is on page 6. The guidebook can be downloaded from the DataJam page of the Pittsburgh DataWorks website.

**PITTSBURGH DATAWORKS**

DataJam Guidebook - 2024

# We are looking forward to DataJam 2024!
# We Hope You Are Too!

Email us at [datajam@pghdataworks.org](mailto:datajam@pghdataworks.org) when you are ready to start working with a DataJam Mentor!