

The DataJam Download

Official Newsletter of The DataJam



***Watch DataJam
Instagram for
News from the
DataJam Finales***

***High School
DataJam Finale!***

***Community College
DataJam Finale!***

***Middle School
DataJam Day***

***New DataSet
Guides***

***Meet the Data
Science
Professionals***

Watch DataJam Instagram Every Day in the Month of May



THE_DATAJAM

Keep an eye on DataJam Instagram for updates throughout May! The early part of May we will be featuring the DataJam projects done by Community College students in preparation for the Community College DataJam Finale --- May 2nd, 5:30-6:30 PDT, 8:30-9:30 EDT, Zoom Link: <https://pitt.zoom.us/j/99862628516>

From May 6th to the end of the month we will feature short videoclips made by each high school DataJam team. These presentations were part of the High School DataJam Finale on April 25th. Check out DataJam Instagram each day to see when your team is being featured.

High School DataJam Finale – April 25th, 2024

The largest DataJam Finale in our ten-year history was held on Thursday, April 25th on Zoom, with 29 teams who fully completed their 2024 DataJam projects competing. New this year, we highlighted the work of each of the teams on the DataJam Instagram for two weeks leading up to the Finale! Also new this year we featured 30 second videoclips produced by the teams, themselves, giving the highlights from their projects, and a “fan favorite” award that viewers on Zoom were able to vote for during the Finale.

Watch the Finale on YouTube by clicking on the link in the [NEWS section](#) on the Homepage of the DataJam website, thedatajam.org. Posters from all winning DataJam teams are also featured in this section:

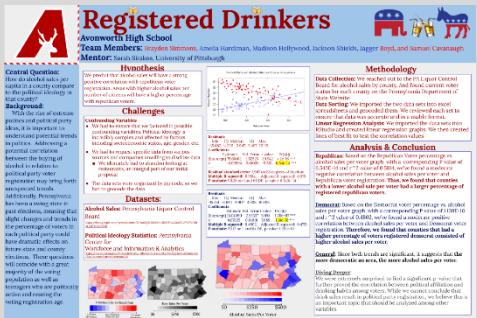
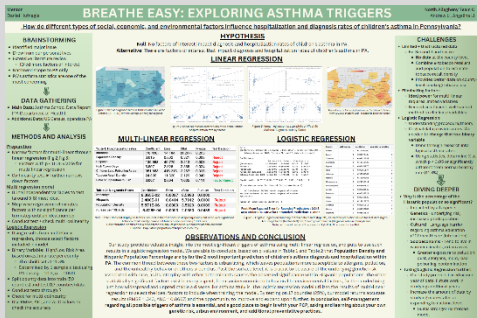
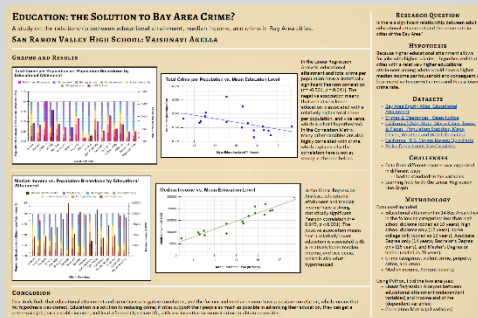
First Place – San Ramon Valley High School, Danville, CA

Second Place – North Allegheny High School Team C, Pittsburgh, PA

Third Place – Avonworth High School Team 1, Pittsburgh, PA

Best Presentation (a tie) – Boston Latin Academy, Boston, MA & North Allegheny High School Team C, Pittsburgh, PA

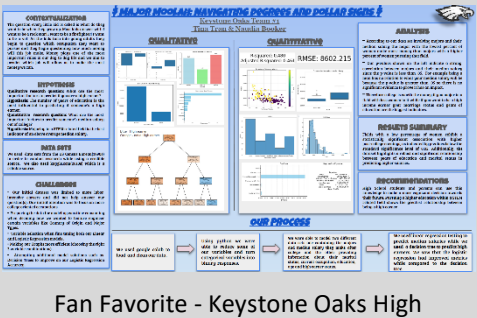
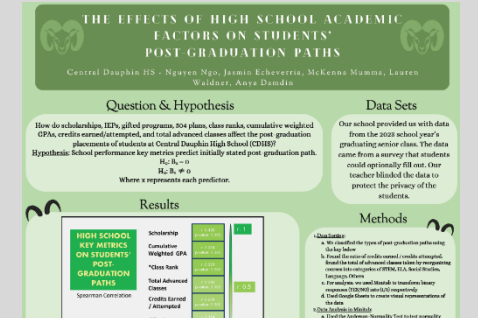
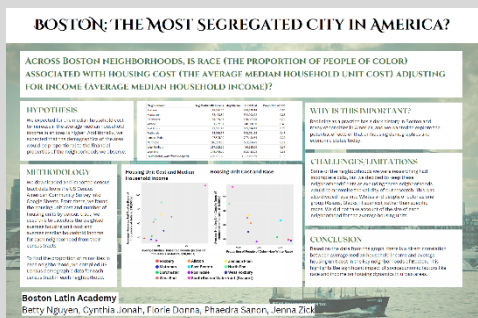
Fan Favorite (a tie) – Central Dauphin High School Team Bob, Harrisburg, PA & Keystone Oaks High School Team 1, Pittsburgh, PA



1st Place - San Ramon Valley High School, Danville, CA

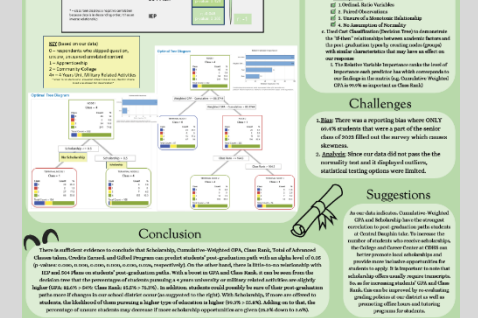
2nd Place & Best Presentation - North Allegheny High School Team C, Pittsburgh, PA

3rd Place - Avonworth High School Team 1, Pittsburgh, PA



Best Presentation - Boston Latin Academy, Boston, MA

Fan Favorite - Keystone Oaks High School Team 1, Pittsburgh, PA



Fan Favorite - Central Dauphin High School Team Bob, Harrisburg, PA

Community College DataJam Finale!

This year for the first time we are running a pilot program offering DataJam for Community College students. Six teams from two community colleges in the San Francisco Bay Area have participated in this pilot study, Laney College, and Skyline College. This pilot has been run in collaboration with the Data Science Discovery Program at the University of California at Berkeley (<https://cdss.berkeley.edu/discovery>), who has provided DataJam Mentors for the college teams.

The Community College DataJam Finale will be May 2nd, and highlights will be posted on the DataJam Instagram.

Addressing Diagnostic Errors in Medical Care: A Focus on Maternal Health Disparities

Dorajha Holland, Sofia Ponder, Essa Taleb
The DataJam, Laney College

1. Introduction
Background: Diagnostic errors in medical care are a significant and under-researched area of concern. This study aims to explore the prevalence and impact of diagnostic errors in maternal health care, with a specific focus on disparities among Black and Hispanic women. The study is structured into several sections: Introduction, Methodology, Results, Research Limitations, and Conclusion & Future Work.

2. Methodology
This study employed a cross-sectional design, analyzing data from a large-scale survey of pregnant women. The data was collected through a series of structured interviews and questionnaires. The study population was divided into two groups: Black and Hispanic women, and White women. The study was conducted in a large urban hospital system.

3. Results
The study found that the prevalence of diagnostic errors in maternal health care is significantly higher among Black and Hispanic women compared to White women. The most common types of diagnostic errors were related to missed or delayed diagnoses of conditions such as preeclampsia, gestational diabetes, and fetal growth restriction. The study also found that the impact of diagnostic errors on maternal and fetal health outcomes was more severe among Black and Hispanic women.

4. Research Limitations
The study has several limitations, including its cross-sectional design, which does not allow for the establishment of causal relationships. Additionally, the study was limited to a single hospital system, which may not be representative of all maternal health care settings.

5. Conclusion & Future Work
The study highlights the need for improved diagnostic accuracy in maternal health care, particularly for Black and Hispanic women. Future research should focus on identifying the underlying causes of these disparities and developing targeted interventions to reduce diagnostic errors and improve maternal and fetal health outcomes.

Laney Community College Team 1

Social Media & Mental Health

Team 2
Laney College Team 2

1. Introduction
The purpose of this study is to explore the relationship between social media usage and mental health. The study is structured into several sections: Introduction, Methodology, Results, Research Limitations, and Conclusion & Future Work.

2. Methodology
This study employed a cross-sectional design, analyzing data from a large-scale survey of young adults. The data was collected through a series of structured interviews and questionnaires. The study population was divided into two groups: high social media usage and low social media usage. The study was conducted in a large urban hospital system.

3. Results
The study found that high social media usage is associated with increased mental health issues, including anxiety, depression, and stress. The study also found that the impact of social media usage on mental health is more severe among young adults who are also experiencing other life stressors.

4. Research Limitations
The study has several limitations, including its cross-sectional design, which does not allow for the establishment of causal relationships. Additionally, the study was limited to a single hospital system, which may not be representative of all young adults.

5. Conclusion & Future Work
The study highlights the need for improved mental health support for young adults, particularly those who are also experiencing other life stressors. Future research should focus on identifying the underlying causes of these associations and developing targeted interventions to reduce social media usage and improve mental health outcomes.

Laney Community College Team 2

"Fleet of the Future" in perspective. EV's vs BART

Paris Middleton Berry, Shilong Pan, Sanchez, Shurong J. Terasaki
Instructors: Julia Ch. Poon, Monica Kim, Sara Yu, Yuhao
Laney College, UC Berkeley

1. Introduction
The purpose of this study is to compare the environmental impact of electric vehicles (EVs) and BART (Bay Area Rapid Transit). The study is structured into several sections: Introduction, Methodology, Results, Research Limitations, and Conclusion & Future Work.

2. Methodology
This study employed a life cycle assessment (LCA) approach, analyzing the environmental impact of EVs and BART across their entire life cycle. The data was collected through a series of structured interviews and questionnaires. The study population was divided into two groups: EVs and BART. The study was conducted in a large urban hospital system.

3. Results
The study found that EVs have a lower environmental impact than BART across most life cycle stages, including manufacturing, operation, and disposal. The study also found that the impact of EVs on the environment is more severe among those who are also driving other vehicles.

4. Research Limitations
The study has several limitations, including its reliance on secondary data, which may not be representative of all EVs and BART. Additionally, the study was limited to a single hospital system, which may not be representative of all life cycle stages.

5. Conclusion & Future Work
The study highlights the need for improved environmental support for EVs, particularly those who are also driving other vehicles. Future research should focus on identifying the underlying causes of these associations and developing targeted interventions to reduce the environmental impact of EVs and BART.

Laney Community College Team 3

ELECTRIC VEHICLES: Optimizing EV Charging Infrastructure for Equitable Adoption in San Mateo

Angela Garcia, Toronto Johnson, Yanyang Guo, Edan Huang, Pin Chao Zhu
Skyline College

1. Introduction
The purpose of this study is to explore the relationship between electric vehicle (EV) adoption and charging infrastructure in San Mateo. The study is structured into several sections: Introduction, Methodology, Results, Research Limitations, and Conclusion & Future Work.

2. Methodology
This study employed a cross-sectional design, analyzing data from a large-scale survey of EV owners in San Mateo. The data was collected through a series of structured interviews and questionnaires. The study population was divided into two groups: high EV adoption and low EV adoption. The study was conducted in a large urban hospital system.

3. Results
The study found that high EV adoption is associated with increased charging infrastructure, including public charging stations and home charging stations. The study also found that the impact of charging infrastructure on EV adoption is more severe among those who are also driving other vehicles.

4. Research Limitations
The study has several limitations, including its cross-sectional design, which does not allow for the establishment of causal relationships. Additionally, the study was limited to a single hospital system, which may not be representative of all EV owners.

5. Conclusion & Future Work
The study highlights the need for improved charging infrastructure in San Mateo, particularly for those who are also driving other vehicles. Future research should focus on identifying the underlying causes of these associations and developing targeted interventions to reduce the environmental impact of EVs and BART.

Skyline Community College Friday Team

Factors Impacting Muni Ridership

Camille Catolcs, Zaw San
Skyline College, DATAJAM

1. Introduction
The purpose of this study is to explore the relationship between various factors and Muni ridership. The study is structured into several sections: Introduction, Methodology, Results, Research Limitations, and Conclusion & Future Work.

2. Methodology
This study employed a cross-sectional design, analyzing data from a large-scale survey of Muni riders. The data was collected through a series of structured interviews and questionnaires. The study population was divided into two groups: high Muni ridership and low Muni ridership. The study was conducted in a large urban hospital system.

3. Results
The study found that high Muni ridership is associated with increased factors such as income, education, and age. The study also found that the impact of these factors on Muni ridership is more severe among those who are also driving other vehicles.

4. Research Limitations
The study has several limitations, including its cross-sectional design, which does not allow for the establishment of causal relationships. Additionally, the study was limited to a single hospital system, which may not be representative of all Muni riders.

5. Conclusion & Future Work
The study highlights the need for improved Muni ridership support, particularly for those who are also driving other vehicles. Future research should focus on identifying the underlying causes of these associations and developing targeted interventions to reduce the environmental impact of EVs and BART.

Skyline Community College Tuesday Team

Socioeconomic Dynamics in San Francisco

SOAR
Tiber Kavalec, Hannah Tsai, Clara Balderas, Ashley Gutierrez, Carro
Skyline College

1. Introduction
The purpose of this study is to explore the relationship between socioeconomic dynamics and public health in San Francisco. The study is structured into several sections: Introduction, Methodology, Results, Research Limitations, and Conclusion & Future Work.

2. Methodology
This study employed a cross-sectional design, analyzing data from a large-scale survey of San Francisco residents. The data was collected through a series of structured interviews and questionnaires. The study population was divided into two groups: high socioeconomic status and low socioeconomic status. The study was conducted in a large urban hospital system.

3. Results
The study found that high socioeconomic status is associated with increased public health outcomes, including lower rates of chronic diseases and higher rates of health care utilization. The study also found that the impact of socioeconomic status on public health is more severe among those who are also driving other vehicles.

4. Research Limitations
The study has several limitations, including its cross-sectional design, which does not allow for the establishment of causal relationships. Additionally, the study was limited to a single hospital system, which may not be representative of all San Francisco residents.

5. Conclusion & Future Work
The study highlights the need for improved public health support, particularly for those who are also driving other vehicles. Future research should focus on identifying the underlying causes of these associations and developing targeted interventions to reduce the environmental impact of EVs and BART.

Skyline Community College Wednesday Team

Middle School DataJam Day

As in 2023, DataJam hosted a Middle School DataJam Day for four school districts in Pittsburgh: Hampton, North Allegheny, Riverview and Avonworth, on April 11th, 2024. Fifty middle school students attended this half day program and chose one of four topics to study from a data science perspective, guided by DataJam mentors. At the end of the morning each team gave a presentation about their findings.

Shark Attack Locations – Mentor: Zac Lindquist



Meteor Landings – Mentor: Tony Robol



UFO Sightings – Mentor: Jatin Singh

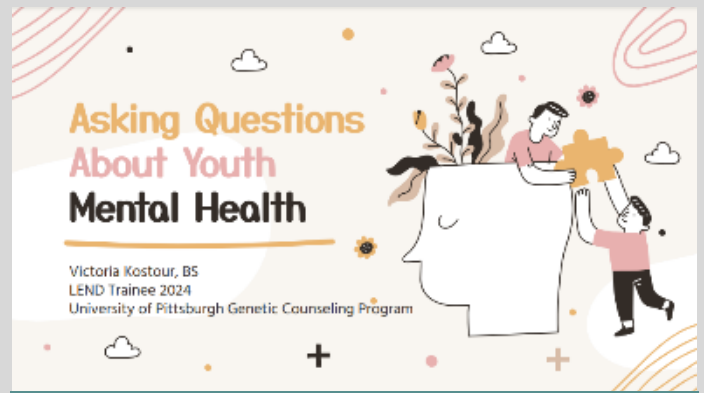


Big Foot Sightings – Mentor: Sarah Sirakos



New DataSet Guides

We have posted five new [DataSet Guides](#) on the website this month, that were produced by LEND fellows at the University of Pittsburgh participating in the LEND (Leadership Education in Neurodevelopmental Disabilities and Related Disorders) program (lend.pitt.edu). The goal of the LEND program is to educate the next generation of leaders in the field of disabilities and to forge partnerships to improve the lives of children with special needs and their families. For 2024, DataJam worked with five young professionals in the LEND program (featured this month in the “Meet the Data Science Professional” section of this newsletter), who produced DataSet Guides in their area of expertise. We are very excited at DataJam to expand the resources available for answering research questions pertaining to neurodevelopmental disorders using data analysis.



Meet the Data Science Professionals

Hi everyone! My name is **Marissa Lahr (MS CF-SLP)** and I am a newly graduated speech-language pathology masters student from The University of Pittsburgh. My love of speech-language pathology began when I spent a summer working as a camp counselor at Victory Junction — a camp for children with several medical diagnoses and disabilities. While working there, I saw my first ever augmentative and alternative communication (AAC) device and realized the importance of ensuring that every individual has access to communication. This drove me to the extremely versatile field of speech-language pathology.

While completing my undergraduate degree at Case Western Reserve University, I became interested in immersing myself in research. I joined two research labs and completed an independent honors research project across my senior year. I presented this project at the Intersections: SOURCE Symposium, and I am proud to say that my project won the first-place prize! I loved how I had the opportunity to discover an area that hadn't been explored,



dive into the data, obtain significant results, and share those findings with my wider community. From that moment on, I knew I wanted to remain involved in research throughout my studies and career. While in graduate school, I joined the Bilingual Phonology Lab (mi familia), became a Graduate Student Researcher (GSR), and completed a master's thesis under an amazing mentor. Once again, I was able to present this project to the greater SLP community during a technical session at the ASHA conference in Boston (2023), which was the best experience ever!

I sought out the leadership education in neurodevelopmental disabilities (LEND) program because of how it centers serving individuals with disabilities, a passion of mine since at least 2018. From the moment we got our community partner placements, I knew I had been placed into the best group. Five of us were given the incredible opportunity to create a dataset guide for The DataJam. My hope is that my dataset guide gets you interested in researching a field you may or may not have heard of before. My guide highlights speech-language pathology but lays the groundwork so you can investigate any profession that you want to further explore. I hope that my guide sparks some interest in a field you hadn't previously considered and supports your diverse research projects, which you will go on to share with others in and beyond your personal communities. And always remember that good research is SHARED research!



Hi there! My name is **JP Marrero-Rivera**, and I am a Graduate Student Researcher at the University of Pittsburgh. I am a rising 4th-year PhD student studying Exercise Physiology, and I have an on-going project funded by a Diversity Supplement from the NIH/NHLBI ([3R01HL164662-01A1S1](#)) that investigates the effects of maternal and early childhood physical activity on child development during the first 24 months of life. My work and expertise are rooted in the cross-section of clinical exercise physiology and applied developmental psychology, with a newly-acquired specialty in neurodevelopmental disabilities and disorders. This work requires extensive knowledge on evidence-based practices and “following” the data to provide best recommendations for optimizing pediatric health and development.

My interest in research and data science dates to November 2nd, 2018. On this day, I presented some of my undergraduate research at a regional conference, where I met an admissions counselor from the University of Pittsburgh. They asked me if I was considering a PhD. At that time, however, I was still focused on attending physical therapy school. Before I could tell them that I was not thinking about a PhD, one of my mentors chimed in and said “yes, he is!” with a large grin on his face. I went home that evening thinking, “does he see something in me that I am not aware of?” 3 days later, I was in class learning about muscle physiology when it hit me... I wanted to pursue a PhD to become an academic and educate the next generation of exercise scientists while also conducting my own, independently-funded research – just like my mentors!

With a bit of hard work, a guiding hand from some wonderful mentors, and some good luck along the way, I ended up at the University of Pittsburgh 3 years and some change later! Ultimately, my doctoral training and the connections I made led me to participating in [leadership education in neurodevelopmental disabilities and related disorders](#). Through this training, I had the opportunity to work with the DataJam to create a dataset guide. The one I created focuses on the CDC's Youth Risk Behavior Surveillance System dataset!

As you consider a career in STEM and possibly in data science, I leave you with a quote from Albert Einstein: “Education is what remains after one has forgotten what one has learned in school.” Never stop learning, and never be afraid to ask questions!

Hi! My name is **Julianna Braun**, and I am a second year Doctor of Physical Therapy Student at the University of Pittsburgh planning to work in a pediatrics setting upon my graduation in December. I have always been passionate about learning, helping others, and giving back to the community - which makes the DataJam such a great organization in my eyes! I have worked with many community organizations in the past, but most of my undergraduate career was spent volunteering with Pitt Dance Marathon, the largest student-run philanthropy at the University of Pittsburgh, which raises money for UPMC Children's Hospital of Pittsburgh. I have continued to work with this organization occasionally since I graduated with my bachelor's degree from Pitt in 2022.



I became interested in the health sciences when I was in high school, and eventually decided to pursue physical therapy because I wanted to help people do the things they love again and help them come up with strategies to minimize the risk of any future injury or pain. Physical therapists also can work with fitting patients for assistive devices such as wheelchairs and power chairs, which is great because it helps people become independent again and access their environments.

I began learning about data science through a project with the Leadership Education in Neurodevelopmental Disabilities (LEND) program through Pitt. The program partnered with the DataJam to create dataset guides on topics related to disability. I created the dataset guide titled "Asking Questions About Special Education." I chose this topic because I have always been interested in school PT and the special education process. I thought it would be an interesting topic for students, as special education is likely something they see every day! There are many different questions that could be asked about trends and data related to this area.

Because I am in a field that isn't typically thought of as a "data-science" field, I think the DataJam is so awesome for raising awareness about data in a variety of topics and fields! It is important for students to learn how to recognize trends and analyze the data in whatever field they choose to study. I hope that as the organization continues to expand, students have the opportunity to research any topic that interests them!



Hello there! My name is **Saad Akhtar**, and I am a recent graduate of the University of Pittsburgh's Masters in Public Health program, focusing on human genetics. My love for public health and genetics really started in high school, where I had a strong interest in STEM from my coursework and a desire to use my abilities to help others. Volunteering with different organizations around my community in a suburb near Boston was something I initially did to get professional experience for my college application, but as time went on, I realized I had a strong sense of satisfaction from helping people in need.

Going into my undergraduate career at UMass Amherst, I explored that urge to help others through working with different organizations around campus such as working with a Boy Scout troop of boys with disabilities or making cards for sick children at the local hospitals. At the same time, I was pursuing a degree in Biology as that love for STEM really evolved into a love of learning about the complex systems that make us who we are. Genetics as an area of science has always been interesting to me because it influences so

much about us as who we are, and yet it's an always growing field that requires you to be a lifelong learner and to adapt to new information as it is researched and analyzed.

After graduating from UMass, my career took me in a few directions as I tried to figure out how I could best combine my love of working with people and patients and my love of genetics. I worked as a laboratory technician at a DNA sequencing company, an EMT and a site coordinator of a small doctor's office, taking something from each of the positions I had. Throughout these different roles I had a strong feeling that I could be doing more with the skills I had and helping people on a larger scale than I had been so far. All those feelings eventually led me here to Pittsburgh to study Public Health and Human Genetics, a rare opportunity as only a few schools across the country offer this type of degree. In my studies I have learned about the complex public health issues facing communities across the country and the different scales to which these issues affect us, ranging from a personal level down to our DNA all the way to a national and global level.

As part of my studies here at Pitt, I got the chance this year to be a part of the LEND training program, and as part of that program I had the pleasure of working with DataJam to create my dataset guide regarding study the topic of disability as a topic. In my LEND training this year I learned so much about people with disabilities and the issues that they face and how we can help break down barriers facing them. My hope is that through this guide, you too can learn about these issues and can feel inspired to be a part of this change in your future careers as well.

It's nice to meet you! My name is **Victoria Kostour**, and I am a recent [MS in Genetic Counseling](#) graduate from the University of Pittsburgh's School of Public Health. I am originally from Connecticut and completed my undergraduate degree from the University of Connecticut (go Huskies!) in both Molecular and Cell Biology and Psychological Sciences. I have always been drawn to biology but started developing my interest in genetics during my high school genetics class. I was fascinated by the way that complex molecular interactions could dictate so much of our daily lives.

During my undergraduate experience, I sought out research experience and ended up in a lab ([NP-OILD Lab](#)) investigating organic elements' and histone modifications' impact metabolic pathways. Although this work was fascinating, I really wanted to have more of a direct impact on people's health. I learned of genetic counseling through my genetics courses and felt it was the best career to leverage my interest in molecular biology and my interpersonal skills. While learning more about the field from shadowing, webinars, and internships, I heavily connected with the genetic counselors' mission to not only educate, but also help patients absorb emotions and adapt to diagnoses. If this sounds like an interesting profession to you, you can learn more at nsgc.org!

As I navigated the Pitt Genetic Counseling Program, I quickly recognized that working with individuals with disabilities and their families was a crucial part of our profession. Many health professionals lack the confidence and education to properly assist these families and I wanted to leverage my position as a provider to help combat these health inequities. Recognizing my gaps in knowledge, I pursued the [Leadership Education in Neurodevelopmental and Related Disabilities \(LEND\)](#) fellowship to better understand the lived experience and public policy related to disability. As a part of this program, I was able to work intimately with Dr. Judy Cameron to develop a novel Dataset Guide on Youth Mental Health. With the increasing prevalence of anxiety and depression among adolescents, I hope this guide can help you learn more about the factors that influence the rates of these conditions. As I begin the job hunt to kickstart my genetic counseling career, I hope you can share in the exciting opportunities that await you in the STEM fields! Best of luck!



We are looking forward to DataJam 2025!

We Hope You Are Too!

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