

Teaching Data Science and Building Data Literacy – Nationally and Globally

With the Northeast Big Data Innovation Hub & National Student Data Corps

Florence D. Hudson, Executive Director, PI, Founder of NSDC

Emily Rothenberg, NSDC Program Manager

National Student Data Corps

Northeast Big Data Innovation Hub at Columbia University

Northeast Big Data Innovation Hub Team

Columbia University Students - Part Time



Florence D. Hudson

Executive Director

Northeast Big Data Innovation Hub



Sanket Bhandari
MS Data Science
'23



Sneha Dahiya
MS Business
Analytics '23



Anna Eggers
BA Film & Writing '24



Lauren Close
*Operations and
Communications
Manager*



Emily Rothenberg
*National Student
Data Corps
Program Manager*



Stephanie Guo
MS Business
Analytics '23



Hannah Yang
MS Data Science
'23

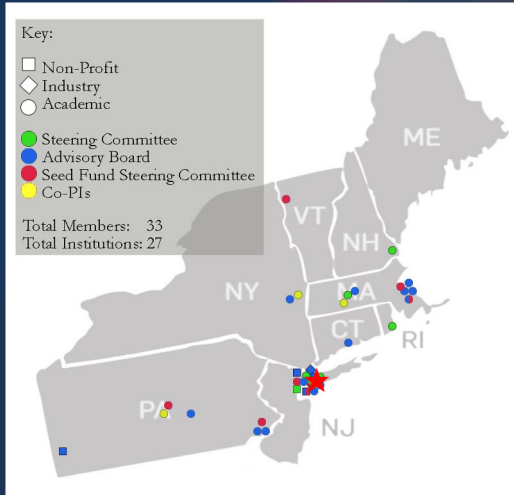


Varalika Mahajan
MS Data Science
'23



Lylybell Teran
MS Data Science
'23

★ PI Institution



- The Northeast Hub is a community convener, collaboration hub, and catalyst for data science innovation in the Northeast Region and around the world.
- Building a diverse, equitable, and inclusive community with accessible resources is central to our work.
- We are a community of over 8,300 individuals from 1,375 organizations, all 50 U.S. States and 61 countries. Join us!

Four Focus Areas guide our community engagement and activities.



Education +
Data Literacy



Health



Urban to Rural
Communities



Responsible Data Science:
Security + Privacy + Ethics

Funded by NSF grants
#1550284, 1748395, 1916585,
2028999, 2139391, and
NIH AIM-AHEAD.

<http://nebigdatahub.org/about/>

Northeast Hub Focus Areas Guide our Activities

The NSDC is the Hub's premier Education & Data Literacy Program

			
Education & Data Literacy	Health	Urban to Rural Communities	Responsible Data Science
National Student Data Corps (NSDC)	COVID Information Commons (CIC)	Smart Cities Data Exchange	Cybersecurity Risk Initiative
OpenDS4All - Call for Participation	CIC Student Paper Challenge	Driver Video Privacy Challenge	Responsible Data Science Comic Books
Data Science Resource Repository (DSRR)	CIC Working Groups	Transportation Projects	IEEE/UL P2933 TIPPSS Standard Working Group
NSDC Global Chapter Community	AIM-AHEAD Data Science Training Core Portal	Smart Cities and Buildings	IEEE Connected Healthcare Cybersecurity Workshops
Data Science Panel Video Series	Clinical IOT and TIPPSS at HIMSS	Urban to Rural Communities Working Group	Framework for Integrative Data Equity Systems (FIDES)
Data Science Mentorship Program	Blockchain & Data Science in Healthcare - Call for Papers	Climate and Nature	Data Sharing and Cyberinfrastructure WG
Education & Data Literacy Seed Fund Projects	Health Seed Fund Projects	Urban to Rural Seed Fund Projects	Responsible Data Science Seed Fund Projects

National Student Data Corps

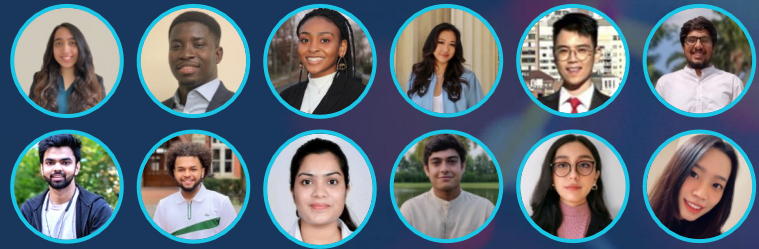
The National Student Data Corps (NSDC) is a community-developed initiative that provides resources and opportunities for students to learn data science, in a community of support, with a special focus on engaging underserved institutions, students, and communities.

There are three teams from the NSDC Founding Committee leading the effort.

Content and Pedagogy Team



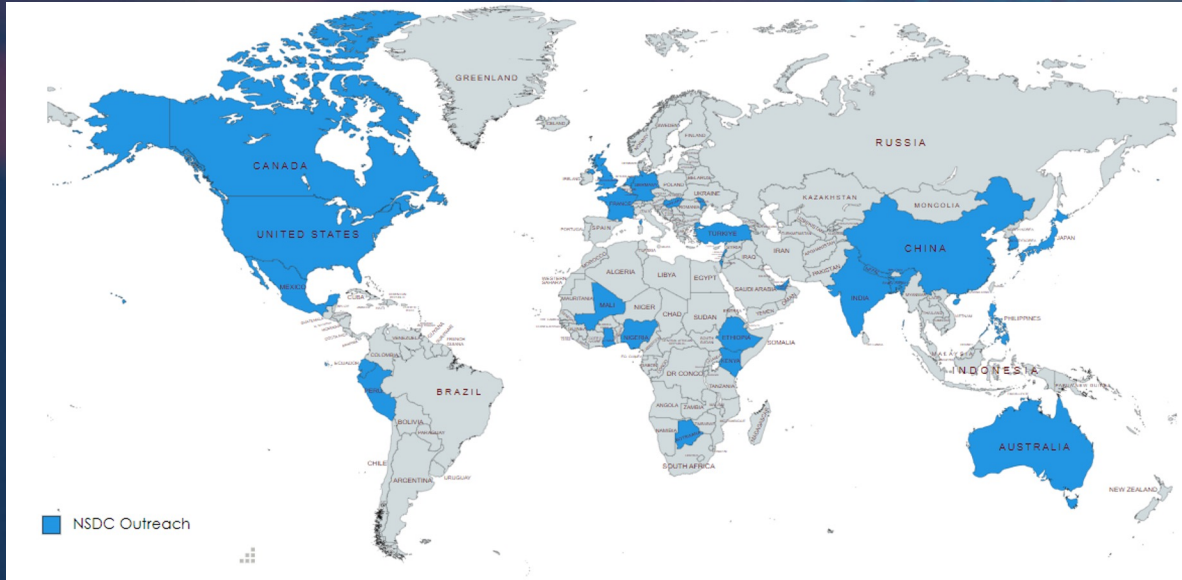
Peer Instructor and Mentoring Team



Outreach Team



NSDC Enables Data Science Literacy - Now National with an International Reach!



As of June 2023, the NSDC community is made up of 4,430 individuals from 613 institutions representing the U.S. and 28 other countries – Australia, Bangladesh, Botswana, Canada, China, Ecuador, Ethiopia, France, Germany, Ghana, Hungary, India, Israel, Japan, Kenya, Lebanon, Mali, Mexico, Moldova, Nepal, Netherlands, Nigeria, Peru, Philippines, South Korea, Turkey, United Arab Emirates, and the United Kingdom.

NSDC Provides over 200 Free Online Data Science Resources for Learners & Educators

National Student Data Corps – Learner Central

Jump right in! Learn data science at your own pace by leveraging this collection of free resources for learners of all backgrounds. Whether you're just getting started or want to learn advanced machine learning techniques, you'll be able to find something for your data science journey.

Data Science Ethics

Learn about the most important ethical issues facing data scientists today and how you can obtain and use data responsibly.

Prerequisites: None!

Getting Started

Get acquainted with the basics. Review the essential arithmetic and algebraic skills you'll need to begin working with data.

Prerequisites: None!

Introductory to Intermediate

Dive into data collection and analysis. Learn about statistical procedures, draw conclusions from data, and develop an intuition for common machine learning methods.

Prerequisites: Arithmetic, Pre-Algebra.

Intermediate to Advanced

Want to expand your data science knowledge? Learn about advanced statistical tools, state-of-the-art machine learning methods, and mechanics of optimization.







National Student Data Corps Video Library

Check out these videos on data science topics based on IBM's OpenDSAI curriculum and presented by Columbia University Master's students.

Introduction to Data Science

With [Yusen Wang](#)

Introduction to Data Science – Part 1 (Yusen Wang)

What Does Big Data Analytics Involve?

- Aggregate, store, and analyze online big data
- Organize, analyze, and visualize data
- Apply machine learning algorithms
- Develop predictive models
- Visualize, track, and report
- Test and evaluate

Introduction to Data Science – Part 2 (Yusen Wang)

Data Science is Interdisciplinary

- Statistics
- Mathematics
- Computer Science
- Engineering
- Business
- Law
- Medicine
- Agriculture
- Environmental Science
- Social Science
- Humanities

Data Acquisition and Wrangling

With [Renjin Zhang](#)

Data Acquisition and Wrangling – Part 1 (Renjin Zhang)

Importing Data from CSV

- Importing Data from CSV
- Cleaning Data
- Handling Missing Data
- Handling Duplicate Data
- Handling Outliers
- Handling Inconsistent Data
- Handling Incomplete Data
- Handling Inaccurate Data
- Handling Inconsistent Data
- Handling Incomplete Data
- Handling Inaccurate Data

Data Acquisition and Wrangling – Part 2 (Renjin Zhang)

Handling Missing Data

- Handling Missing Data
- Handling Duplicate Data
- Handling Outliers
- Handling Inconsistent Data
- Handling Incomplete Data
- Handling Inaccurate Data

Data Acquisition and Wrangling – Part 3 (Renjin Zhang)

Handling Duplicate Data

- Handling Duplicate Data
- Handling Outliers
- Handling Inconsistent Data
- Handling Incomplete Data
- Handling Inaccurate Data

Supervised Machine Learning

With [Tomislav Galjanic](#)

Introduction (Tomislav Galjanic)

Then Prediction

- Then Prediction
- Linear Regression
- Logistic Regression
- Decision Trees
- Random Forests
- Support Vector Machines
- Naive Bayes
- K-Nearest Neighbors
- Artificial Neural Networks

Introduction to Decision Trees (Tomislav Galjanic)

As Example

- As Example
- Linear Regression
- Logistic Regression
- Decision Trees
- Random Forests
- Support Vector Machines
- Naive Bayes
- K-Nearest Neighbors
- Artificial Neural Networks

Decision Trees (Tomislav Galjanic)

Decision Trees (Tomislav Galjanic)

- Decision Trees
- Random Forests
- Support Vector Machines
- Naive Bayes
- K-Nearest Neighbors
- Artificial Neural Networks









National Student Data Corps – Educator Central

Welcome Educators!

Data science is more important than ever in today's digital world and has applications in all disciplines. The resources below can help you whether you are building a new data science program at your school, revising a lab and looking for datasets, or just interested in how you can incorporate data science topics in your classroom.

Open Courses and Resources

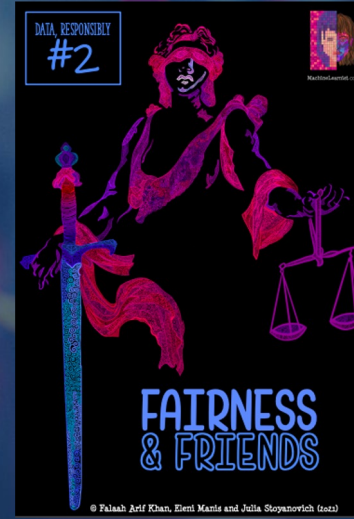
Build a data science curriculum that works for your students using these materials from different universities and organizations.

 <p>DataKidds</p> <p>To assist educators in supporting student data science learning, DataKidds created five data stories. Educators can freely access and modify for their own use and students' opportunities.</p>	 <p>Data 8: Foundations of Data Science</p> <p>This course from UC Berkeley covers statistics, data visualization, machine learning, and basic programming in Python.</p>	 <p>Data: Past, Present, Future</p> <p>This Columbia course integrates the teaching of algorithms and data manipulation with the political, scientific, and ethical controversies from which those techniques emerged.</p>	 <p>Data Science Ethics</p> <p>Visit NSDC – Learner Central for resources on data science ethics issues including privacy, informed consent, and bias.</p>
 <p>Institute for Data-Driven Dynamical Design (ID²)</p> <p>ID² aims to transform how scientists and engineers harness data when designing materials and structures.</p>	 <p>International Data Science in Schools Project</p> <p>This international collaborative project provides a pre-calculus course for high school students and a course for educators on teaching Data Science.</p>	 <p>Introduction to Data Science Course</p> <p>Created and taught by NSDC Founding Committee member Amy Arndt, this introductory course was adapted from the Data 8 course. Assignments and corresponding code are available on GitHub.</p>	 <p>Investing in America's Data Science and Analytics Talent</p> <p>This joint report from the Business Higher Education Forum and PwC looks at eight actions for change to put the supply of skills in balance with the demand.</p>

Data, Responsibly Comic Book Series

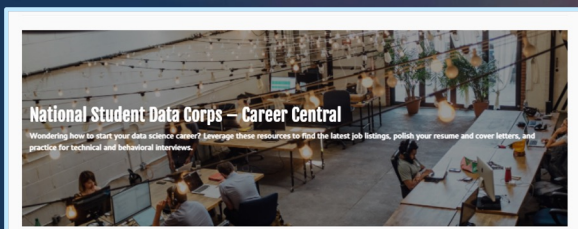
Teaching data science ethics and fairness. In English, Spanish, French, Portuguese, and Greek.

Included in the NSDC Educator and Learner Central, and Data Science Resource Repository (DSRR) (<https://nebigdatahub.org/learner-central-data-science-ethics/>)



Falaah Arif Khan and Julia Stoyanovich. “Mirror, Mirror”, “What is AI?”, “Fairness and Friends”, “Who lives, Who dies, Who decides?”. Data, Responsibly Comics, Volume 1 and 2 (2020). <https://dataresponsibly.github.io/comics/>

NSDC Free Online Resources for Career & Professional Development



Job Search Websites

<p>Data Umbrella</p> <p>Data Umbrella is a community for underrepresented persons in data science, machine learning, and artificial intelligence.</p>	<p>Analytic Talent</p> <p>Find answers to salary questions, build a resume that stands out, and set up personalized job alerts.</p>	<p>Diversify Tech</p> <p>Explore jobs from companies committed to fostering a diverse and inclusive culture.</p>	<p>Jobtensior</p> <p>Explore jobs in tech, science, and engineering tailored for you by Jobtensior's AI algorithms.</p>
--------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------

Data Science Career Resources

<p>DATA SCIENCE DREAM JOB</p>	<p>SQL TUTORIAL</p>	<p>Geeks for Geeks</p>
--------------------------------------	----------------------------	-------------------------------

Career Opportunities for Women and Minorities in STEM

<p>Ain't B Org</p> <p>Connect with, and guide women in computing in advancing their careers and improving the technology sector.</p>	<p>ASPIRE</p> <p>NSF INCLUDES Aspire Alliance</p> <p>Aspire works to increase the diversity of STEM Faculty across the nation.</p>	<p>BC</p> <p>Best Colleges Scholarships for Women</p> <p>BestColleges.com has assembled a list of scholarships available to women in STEM.</p>	<p>Correlation One</p> <p>Correlation One</p> <p>Data science and analytics training, assessments, technology, booting programs, and resources to increase diversity initiatives, Data Science for All.</p>	<p>Girls in Tech</p> <p>Become a part of numerous inspiring programs with a diverse community designed to help women enter into and excel in Tech.</p>
<p>HACE</p> <p>Hispanic Alliance for Career Enhancement (HACE)</p> <p>HACE represents more than 300 colleges and universities committed to Hispanic higher education success in the U.S., Puerto Rico, Latin America, and Spain.</p>	<p>HACU</p> <p>Hispanic Association of Colleges & Universities (HACU)</p> <p>HACU represents more than 300 colleges and universities committed to Hispanic higher education success in the U.S., Puerto Rico, Latin America, and Spain.</p>	<p>Hispanic Heritage Foundation</p> <p>The Hispanic Heritage Foundation identifies, inspires, prepares, and connects Latino Leaders in the community, classroom, and workforce.</p>	<p>HISPANIC SCHOLARSHIP FUND</p> <p>HSP empowers students and parents with the knowledge and resources to successfully complete a higher education.</p>	<p>March 2022 Women in STEM Campaign</p> <p>For Women's History Month, the NSDC launched a campaign to share a message and opportunities for women in data science.</p>
<p>NSBE</p> <p>National Society of Black Engineers (NSBE)</p> <p>NSBE supports and provides the aspirations of collegiate and pre-collegiate students and technical professionals in engineering and technology.</p>	<p>NYU Ability Project</p> <p>An interdisciplinary research space dedicated to the collaboration between disability and technology.</p>	<p>R Ladies</p> <p>R Ladies works towards appropriate representation by encouraging, inspiring, and empowering people of gender currently underrepresented in the R Community.</p>	<p>Texas Hispanic Serving Institutions Consortium</p> <p>THSIC holds conferences annually in the spring in Texas for the State's HSIs.</p>	<p>Society for the Advancement of Chicano and Hispanics and Native Americans in Science</p> <p>SACNAAS builds with culture and identity as the means to advance their diversity in STEM.</p>

Get REAL with the Northeast Big Data Innovation Hub

Do you want to get involved with the NEBDHub or [National Student Data Corps \(NSDC\)](#) but aren't sure how to get started? Browse our *Research Experience and Leadership (REAL)* directory for opportunities, programs, and resource offerings for individuals of all professional and educational backgrounds to find the perfect opportunity for you!

Use the search and filter functions to find opportunities for current students and teachers, professors, and data science professionals. You can also toggle opportunities by focus area and discover opportunities for those looking to learn, teach, or practice data science, grow their careers, or find a mentor.

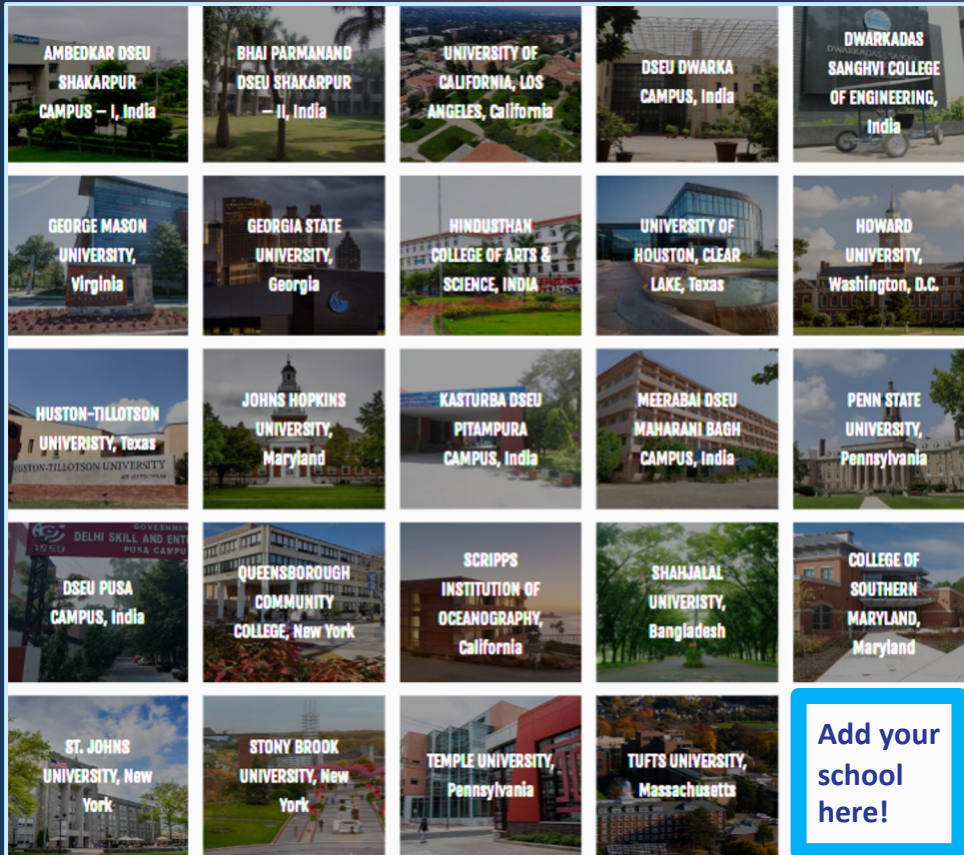
Reach out to our team at nebdh_info@columbia.edu to learn more about how you can jumpstart your data science journey!

⊗ Customize cards Filter IT Sort

<p>Browse Data Science Resources</p> <p>DESCRIPTION</p> <p>Browse the Data Science Resource Repository (DSRR) library of open-source resources and best practices.</p> <p>OPPORTUNITY TYPE</p> <p>Learn Data Science Grow Your Data</p> <p>PERFECT FOR</p> <p>Community College Students Data</p> <p>PROGRAM</p> <p>NEBDHub</p> <p>WEBSITE</p> <p>https://nebigdatahub.org/dsrr/</p>	<p>Build a Data Science Career</p> <p>DESCRIPTION</p> <p>Find career development resources and job search tools targeted towards data science professionals and recent graduates.</p> <p>OPPORTUNITY TYPE</p> <p>Grow Your Data Science Career</p> <p>PERFECT FOR</p> <p>Recent Graduates Data Science Pro</p> <p>PROGRAM</p> <p>National Student Data Corps (NSDC)</p> <p>WEBSITE</p> <p>https://nebigdatahub.org/nsdc/nsdc-...</p>	<p>Find or Become a Mentor</p> <p>DESCRIPTION</p> <p>The National Student Data Corps Data Science Mentorship Program facilitates mentorship opportunities for data science enthusiasts of all ages...</p> <p>OPPORTUNITY TYPE</p> <p>Find or Become a Mentor Develop</p> <p>PERFECT FOR</p> <p>High School Students Community</p> <p>PROGRAM</p> <p>National Student Data Corps (NSDC)</p> <p>WEBSITE</p> <p>https://nebigdatahub.org/nsdc/nsdc-...</p>	<p>Join a COVID Working Group</p> <p>DESCRIPTION</p> <p>Join a community of students, young professionals, and new researchers interested in the future of COVID-19 and ongoing pandemic research.</p> <p>OPPORTUNITY TYPE</p> <p>Join a Data Science Community De</p> <p>PERFECT FOR</p> <p>Undergraduate Students Commu</p> <p>PROGRAM</p> <p>COVID Information Commons (CIC)</p> <p>WEBSITE</p> <p>https://covidifcommons.datascienc...</p>
<p>Join OpenDS4All Leadership</p> <p>DESCRIPTION</p> <p>The NEBDHub and the Linux Foundation AI & Data invite applications from graduate students and data science professionals to join...</p> <p>OPPORTUNITY TYPE</p> <p>Develop Leadership Skills</p> <p>PERFECT FOR</p> <p>Graduate Students Data Science Pro</p> <p>PROGRAM</p> <p>NEBDHub</p> <p>WEBSITE</p> <p>https://nebigdatahub.org/opends4all/</p>	<p>Learn & Teach Data Science</p> <p>DESCRIPTION</p> <p>Strengthen your data science skills by teaching it to others! Reach out to our team if you'd like to create materials for the Data Science Resource ...</p> <p>OPPORTUNITY TYPE</p> <p>Learn Data Science Teach Data Sci</p> <p>PERFECT FOR</p> <p>Adult Learners Graduate Students</p> <p>PROGRAM</p> <p>National Student Data Corps (NSDC)</p> <p>WEBSITE</p> <p>https://nebigdatahub.org/nsdc/</p>	<p>Make Data Science Content</p> <p>DESCRIPTION</p> <p>Help make COVID-19 research accessible to everyone! Email our team to join an international group of translators and transcribers who belong...</p> <p>OPPORTUNITY TYPE</p> <p>Teach Data Science Learn Data Sci</p> <p>PERFECT FOR</p> <p>Professors / Faculty Recent Graduat</p> <p>PROGRAM</p> <p>National Student Data Corps (NSDC)</p> <p>WEBSITE</p> <p>nebdh_info@columbia.edu</p>	<p>Make Research Accessible</p> <p>DESCRIPTION</p> <p>Help make COVID-19 research accessible to everyone! Email our team to join an international group of translators and transcribers who belong...</p> <p>OPPORTUNITY TYPE</p> <p>Develop Leadership Skills Support</p> <p>PERFECT FOR</p> <p>Graduate Students Recent Graduat</p> <p>PROGRAM</p> <p>COVID Information Commons (CIC)</p> <p>WEBSITE</p> <p>nebdh_info@columbia.edu</p>

<https://nebigdatahub.org/nsdc/nsdc-career-central/>
<https://nebigdatahub.org/nsdc/nsdc-volunteer-central/>

Start or Join an NSDC Chapter! In US and Globally!



Add your school here!



- Chapter members benefit from:
- Data science resources
 - Mentorship
 - Research, volunteer and career opportunities

NSDC Data Science Career Panels - 14 panels to date, 40 expert panelists, 3800 viewers!

Representing a diverse range of careers and paths to data



National Student Data Corps
Data + Earth Sciences Career Panel
APRIL 21ST, 2023
3:00 - 4:30 PM (ET)
[BIT.LY/APR2023-PANEL](https://bit.ly/APR2023-PANEL)



Jason Holmberg
Wild Me



Dr. Laurie Nowatzke
USDA Midwest Climate Hub



Dr. Jacob Pasner
California State Water Resources Control Board




NATIONAL HISPANIC HERITAGE MONTH
Data Science Career Panel



Ashley Atkins
West Big Data Innovation Hub



Fred Uquillas
Princeton University; NeurON



Gabriela Cavazos
Laredo College



Raul Cosio
IBM; Hispanic Engineer National Achievement Awardee



Isabella Graham Martinez
Columbia University; NEBDHub

This event is designed to highlight Hispanic-identifying data enthusiasts, Hispanic-serving Institutions (HSIs), and the opportunities HSIs provide to their communities.



Haleigh Stewart
Undergraduate Senior at Columbia University



Benjy Sango
Undergraduate Senior at Columbia University



Edwood Brice



Julia Stoyanovich



Yoni Malchi



Shane Chang



Brandon Sepulvado



Asya Spears



Amen Ra Mashariki



Daniel Balouek-Thomert



Zoran Obradovic



Marija Stanojevic

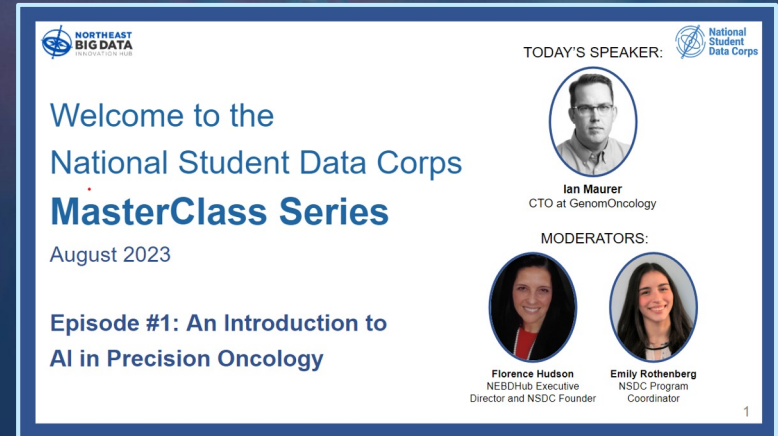
Watch past panels on demand here: bit.ly/nsdc-panels

NEW NSDC MasterClass Series


The new NSDC MasterClass Series is designed to showcase how data science is being leveraged in various domains.


Highlighted domains may include:

- Health
- Finance
- Athletics
- Entertainment, Music & Media
- Technology
- Education
- And more!





NORTHEAST BIG DATA
INNOVATION HUB

TODAY'S SPEAKER: 


Ian Maurer
CTO at GenomOncology

MODERATORS:

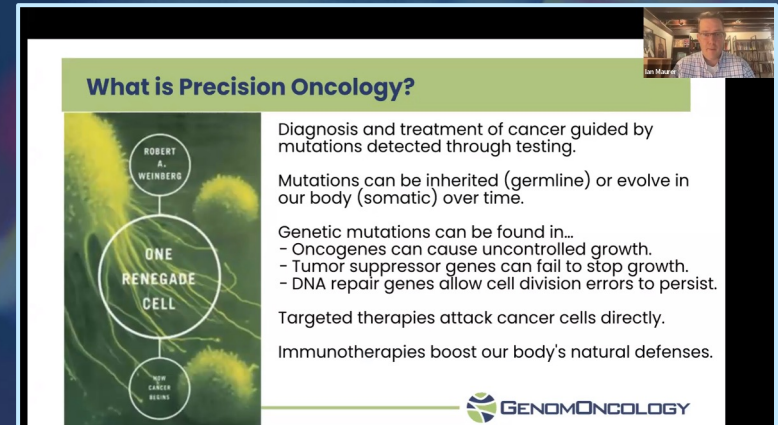

Florence Hudson
NEBDHub Executive Director and NSDC Founder


Emily Rothenberg
NSDC Program Coordinator

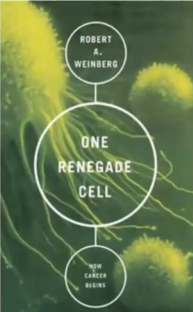
1

Welcome to the
National Student Data Corps
MasterClass Series
August 2023

**Episode #1: An Introduction to
AI in Precision Oncology**



What is Precision Oncology?



Diagnosis and treatment of cancer guided by mutations detected through testing.


Mutations can be inherited (germline) or evolve in our body (somatic) over time.


Genetic mutations can be found in...

- Oncogenes can cause uncontrolled growth.
- Tumor suppressor genes can fail to stop growth.
- DNA repair genes allow cell division errors to persist.

Targeted therapies attack cancer cells directly.

Immunotherapies boost our body's natural defenses.

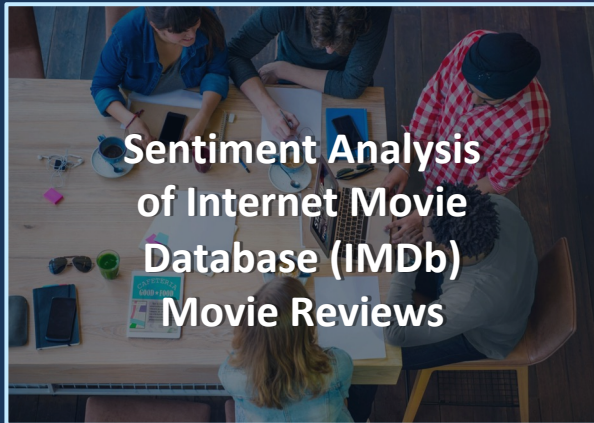




NSDC Data Science Projects (DSP)

The NSDC DSP provide **practical data science project experience** for learners of all skill-levels and education-levels, and highlight data science as a team effort!

Through a guided **Google Colab Notebook**, and supportive weekly mentorship calls, cross-institutional teams of students learn data science techniques and gain leadership and professional development skills.



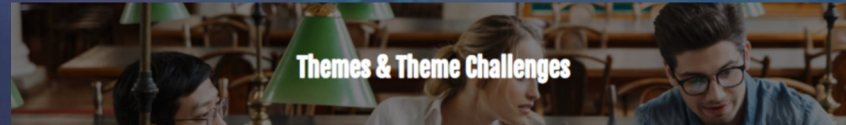
colab
New DSPs planned!

Join an upcoming NSDC Data Science Symposium!

The NSDC Data Science Symposium (DSS) invites students to learn about data science research and grow their research community.

Students from higher ed institutions are invited to create a data science research project and showcase their findings in a virtual poster board presentation.

The NSDC Data Science Symposium is held annually, and we welcome you to join the upcoming challenges!



Challenge	Topic	Icon
Challenge 1 (Health)	COVID Research	–
Leverage the COVID Information Commons (covidinfocommons.net) to identify the use of data science tools and techniques in COVID research.		
Challenge 2 (Health)	Single-Cell Transcriptomics Research	–
Leverage the Human Cell Atlas Data Portal (https://data.humancellatlas.org) to identify the use of data science tools and techniques in single-cell transcriptomics research.		
Challenge 3 (Health)	Immunology Research	+
Challenge 4 (Health)	Cancer Genomics Research	+
Challenge 5 (Health)	Pathogen Genome Research	+
Challenge 6 (Urban to Rural Communities)	General Bikeshare Feed Specification Research	+
Challenge 7 (Responsible Data Science/Urban to Rural Communities)	Motor Vehicle Crash Research	+
Challenge 8 (Responsible Data Science)	Cybersecurity Best Practices Research	+
Challenge 9 (Responsible Data Science)	Social Media Bias Research	–

Based on a topical collection of recent social media posts, use data analytics and innovative thinking to explore the data and identify, track and highlight any form of bias as seen in social media data.

Your submission should seek to answer the following questions:

1. Does social media data embody gender bias or other forms of bias? Identify and track such trends and demonstrate it with your data analysis and visualizations.
2. Using the text of the social media posts, identify examples of bias related to behavioral, opinion and emotion trends by implementing methods for textual and sentiment analysis on your chosen dataset.

Data:

You can use original social media datasets for this challenge. Data must be relevant to exploring bias, gender bias, bias in education and current trends and events associated with bias. Submissions can be based on any relevant social media dataset such as tweets or other social media posts. A link to tutorials and basic code (R / Python) for downloading tweets by topic (keyword) will be made available. Participating students and teams can build upon this or create something entirely on their own within the framework of this challenge.

NSDC Mentorship Program

Find or become an NSDC Mentee or Mentor to support a data science journey! The National Student Data Corps Data Science Mentorship Program facilitates mentorship opportunities through the NSDC Slack Channel for over 610 data science students, professionals, and enthusiasts across the U.S. and beyond!

NationalStudentDataCorps ▾



@channel, I hope you're having a fun weekend and wish you great holidays ahead.

I'm Dashansh, a Project Coordinator at NEBDHub and in charge of the Mentorship Program. If you're in need of guidance or mentorship, I would be happy to connect you to our amazing volunteers. Don't be shy, just make a request and I'll try my best to set you up with someone whose expertise aligns with your needs.

That being said, have a great weekend. (edited)



4



Interested in becoming an NSDC Mentor?
[bit.ly/become-mentors](https://nebigdatahub.org/nsdc/nsdc-mentorship-program/)

Interested in becoming an NSDC Mentee?
[bit.ly/become-mentees](https://nebigdatahub.org/nsdc/nsdc-mentorship-program/)

<https://nebigdatahub.org/nsdc/nsdc-mentorship-program/>

Data Science Resource Repository

The DSRR is an open access platform with 800+ resources including data science best practices and content to broadly increase knowledge capacity, with a specific focus on underserved communities.

Topical areas include:

- Accessibility & Data Science
- AI
- Apache
- Azure
- Big Data
- Bioinformatics
- Cities
- Climate Change & Environment
- Coding
- COVID
- CSS
- Cybersecurity & Privacy
- Data Analytics
- Data Modeling
- Data Science Basics
- Data Science Careers
- Data Science for Good
- Data Visualization
- Database Management
- Deep Learning
- Diversity
- Ethics & Bias
- Java
- Kubernetes
- Linear Algebra
- Machine Learning
- Math Fundamentals
- Natural Language Processing
- Pandas
- Power BI
- Probability
- Public Health
- Public Policy
- Python
- Quantitative Reasoning
- Quantum Computing
- Research Tools
- Rural Populations
- Splunk
- SQL
- Statistics
- Tableau
- TypeScript
- Water

Volunteer opportunity to add more content, including Spanish content.

<https://nebigdatahub.org/data-science-resource-repository/>

Data Science Resource Repository

Resources are grouped by subjects, or topical areas.

Users can browse, search, and filter resources by keyword or tag.

Data Science Resource Repository

All DSRR Resources Learn & Teach Data Science Research & Practice Data Science Network & Explore Data Science

Browse and Search All DSRR Resources

Use filters to find data science resources by keyword or tag.
To the left, resources are grouped by Audience Category for easy browsing.

Reset

Title contains Enter a value

Description contains Enter a value

Author has any of Select an option

Subject has any of Select an option

Format has any of Select an option

Language has any of Select an option

Data Science Resource Repository

	Title	Subject	Format
	AUDIENCE CATEGORY		
	Learn & Teach Count 336		
1	10 Minutes to Pandas	Coding Python Pandas	Article or Journal
2	40 Questions on Probability for Data Science	Statistics Math Fundamentals Probability	Data Science Challenge / Project
3	200+ Machine Learning Tutorials	Machine Learning Natural Language Processing (NLP)	Tutorial
4	A Beginner's Guide to Apache Spark	Coding Data Analysis Apache	Article or Journal
5	A Comprehensive Tutorial on Video Modeling	Data Modeling	Video
6	A Gentle Introduction to Data Science	Data Science Basics AI Python Machine Learning	Course Video
7	A Gentle Introduction to Graph Embeddings	Data Analysis Machine Learning	Article or Journal

Learning formats include articles, journals, tutorials, videos, textbooks, courses, challenges, and podcasts.

OpenDS4All - Call for Participation

The NEBDHub & NSDC, in collaboration with the Linux Foundation AI & Data, is delighted to **invite you to join the leadership and development team of the OpenDS4All GitHub repository.**

The OpenDS4All GitHub repository includes hundreds of resources from slide presentations on data science ethics and supervised machine learning, to Jupyter notebooks which provide hands-on learning opportunities.

If you would like to join this volunteer team, the three volunteer roles include:

- Technical Steering Committee Member
- Content Review Team
- Content Developers



<https://nebigdatahub.org/opends4all/>

Thank you!

Florence D. Hudson, Executive Director, PI, Founder of NSDC; fh2417@columbia.edu
Emily Rothenberg, NSDC Program Manager; er3101@columbia.edu
National Student Data Corps
Northeast Big Data Innovation Hub at Columbia University

- Sign up for the NEBDHub & NSDC mailing list: bit.ly/HubNewsSignup
- Join the NSDC Community: nebigdatahub.org/nsdc
- Join the NSDC Slack Channel: bit.ly/nsdc-slack-channel
- Keep up with the NEBDHub & NSDC on the NEBDHub YT Channel:
youtube.com/@NortheastBigDataInnovationHub
- Join a Hub or NSDC event: nebigdatahub.org/events-calendar
- Any questions, reach out to us at contact@nebigdatahub.org

All available through: <https://nebigdatahub.org/about>