

# Data4All High School Bridge Program

2022-2024 Evaluation Report

## The Impact in Students' Own Words

*Made possible with funding from the Hymen Milgrom Supporting Organization and the University of Chicago's Office of Civic Engagement and Data Science Institute*

Prepared by:

Evelyn Campbell, Ph.D.  
Data Science Institute  
[campbelle@uchicago.edu](mailto:campbelle@uchicago.edu)

Julia Koschinsky, Ph.D.  
Center for Spatial Data Science  
[jkoschinsky@uchicago.edu](mailto:jkoschinsky@uchicago.edu)

In collaboration with:



# Table of Contents

About this Report	<u>3</u>
<b><u>ABOUT DATA4ALL</u></b>	<b><u>5</u></b>
Data4All in Numbers	<u>6</u>
Why This Workshop	<u>7</u>
Data Science Reasoning Framework	<u>8</u>
Workshop Activities	<u>9</u>
Activities & Learning Objectives	<u>10</u>
<b><u>EVALUATION</u></b>	<b><u>11</u></b>
Participant Demographics	<u>13</u>
Learning Outcomes	<u>17</u>
<u>Evaluation of Staff &amp; Activities</u>	<u>23</u>
Student Highlights & Challenges	<u>24</u>
Small Groups, Mentors & Guest Speakers	<u>25</u>
Spark Activities & Lectures	<u>26</u>
Notebooks & Kepler Activities	<u>27</u>
Case Studies	<u>28</u>
<b><u>BACKGROUND</u></b>	<b><u>29</u></b>
About the Hosts	<u>30</u>
About the Surveys	<u>32</u>
Open Teaching Materials	<u>33</u>
Acknowledgements	<u>34</u>



# About this Report

The goal of the **Data4All High School Bridge workshops** (Data4All) has been to teach high school students how to use computation, statistics, and mapping to address real-world data problems. This focus on reasoning with data seeks to broaden students' understanding of what DS (DS) is, beyond a narrower technical focus on programming and statistics. It is also intended to highlight the relevance of DS to a broad variety of science, technology, engineering, and mathematics (STEM) and other fields, including college and career options students did not associate with DS before.

The target audience for the workshops are sophomores, juniors, and seniors from Chicago Public Schools who are underrepresented in DS in terms of race/ethnicity or gender and who are less exposed to computer science (CS) curricula at their schools. Data4All considered trauma-informed education frameworks and provided technological and financial support to students. For instance, Mentors, instructors, and curriculum developers received several hours of sensitivity training regarding trauma-informed education (co-led by the Office of Civic Engagement, Tyler Skluzacek, and Evelyn Campbell). Student participants are provided with technical support (e.g., loaned Chromebooks), as well as financial support in the form of a stipend to offset some of the costs for travel and potential lost earned income.

So far, **we have held six iterations of Data4All**. The first workshop was a 1-week virtual training during COVID in spring of 2021, followed by in-person workshops on eight Saturdays for 4 hours in the fall of 2022, spring of 2023, fall of 2023, spring of 2024, and fall of 2024. Data4All was hosted at the University of Chicago (UChicago)'s Data Science Institute (DSI), in collaboration with Argonne National Labs (ANL), the Center for Spatial Data Science (CSDS), and the Office of Civic Engagement (OCE). OCE and ANL aided in recruitment of students, ANL instructors taught the workshop, and DSI recruited Mentors and managed Data4All. The curriculum was co-developed by ANL, CSDS, and Tyler Skluzacek (Oak Ridge National Labs). Due to changes in the development of evaluation instruments, all visualizations include data that was analyzed from the spring of 2023 to the fall of 2024. Where possible, data for fall of 2022 will be presented and indicated.



# About this Report

This is an internal evaluation of the last four workshops held in spring 2023 (with 24 students), fall 2023 (30 students), spring 2024 (30 students), and fall 2024 (30 students). Because we revised the survey questions for 2023 to take the [logic model](#) into account that we developed after the 2022 workshop, all data excludes data from the 2022 workshop, unless otherwise indicated. The evaluation is based on student surveys that we disseminated before, in the middle, and at the end of the workshops in 2022, 2023, and 2024 (more details [here](#)).

This evaluation highlights Data4All's impact in the students' own words. We conducted this evaluation for several reasons: (1) To provide an overview of our efforts so far, (2) to summarize how the workshop works, and (3) to assess if its learning objectives were reached or not. The audience for this report include instructors and Mentors who will teach the workshop in the future (at UChicago and elsewhere), funders to whom we are accountable, and other stakeholders interested in teaching DS to high schoolers.

Three workshop aspects are evaluated in this report: 1) [Target audience](#), 2) short-term [learning objectives](#), and 3) [workshop activities and staff](#). The report starts with an [overview of Data4All's rationale and approach](#), followed by the above [evaluation](#) based on student surveys, and ends with a [background](#) section with further details about the hosts, surveys, and logic model.

All workshop materials will be [publicly released](#) after peer review in fall 2024. Data4All has been funded to continue through spring of 2026 at the University of Chicago's Data Science Institute.

This workshop was made possible by Milgrom's Computer Coding Fellowship Program Grant, which provided stipends to students and Mentors, and OCE's and DSI's operating support. We are also grateful for the invaluable contributions of everyone who worked on Data4All, [acknowledged at the end of the report](#).



# ABOUT DATA4ALL



# Data4All in Numbers

2022-2024

1,920 Minutes of Workshop Content

394 Program Applications

139 Pages of Instructor Guide

138 Student Participants

48 High Schools

21 Mentors

10 Jupyter Notebooks

6 Iterations

4 Curriculum Developers

3 Project Managers

3 Graduate Student Teaching Assistants

# Why Data4All?

## Data4All addresses 3 gaps in DS education:

1

While DS is one of the fastest growing employment sectors, DS courses are often not yet taught in high schools. For instance, of the students who participated in Data4All in 2022, almost two thirds had taken a computer science (CS) course and three quarters had some programming experience. Most students (86%) had been introduced to scientific concepts, but only a third of students had been exposed to data analysis or data visualization – and if data analysis is introduced, it is taught separately from scientific reasoning. Data4All seeks to close these gaps.

2

At the college level, many DS students get proficient at the technical skills in computation and statistics but still struggle to critically solve problems with data. To address this challenge early on, Data4All students learn foundational programming, statistical methods, and mapping within the context of a DS reasoning framework. For eight weeks, students are working through puzzles of two real-world case studies while analyzing data in a programming environment using the Python language in Jupyter and Google Colab notebooks. They learn to apply technical skills to assess the plausibility of evidence in relation to competing explanations and to communicate their findings to non-technical audiences.

3

There is a need for more data scientists in general, and for more data scientists from underrepresented groups in particular. **Between 2022 and 2024, almost all of the Data4All students (82%) were underrepresented ethnic/racial OR gender minorities.** The goal of the workshop is to introduce these students to DS-related college and career opportunities – not only in computer science or DS but also in other fields. This is done through the case studies, which apply DS to public health problems, through guest speakers who demonstrate the use of scientific reasoning and data analysis in other disciplines (e.g., biology, physics, social science, etc.) , and through talks about college and career preparation.

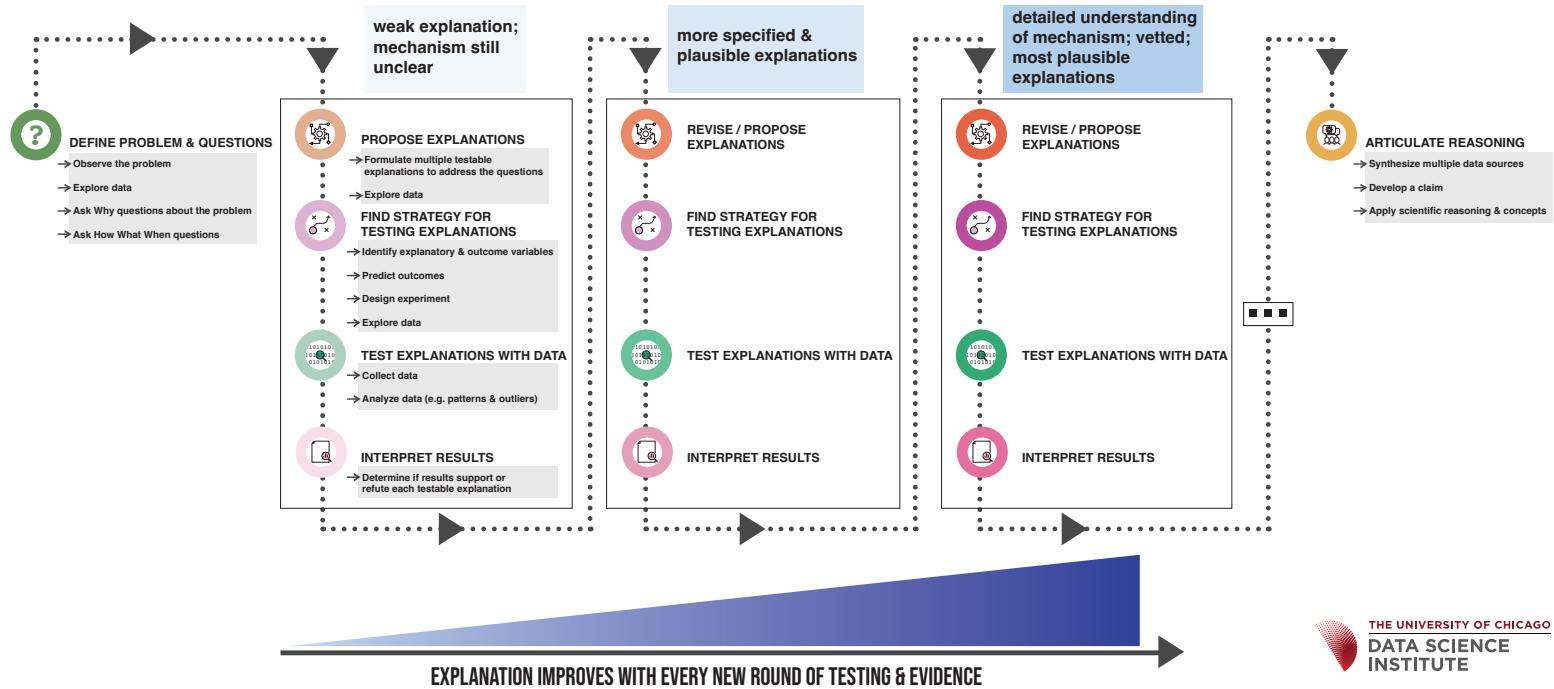


# Data4All's 'Special Sauce': The Data Science Reasoning Framework

The Data Science Reasoning Framework (DSRF) drives the design of the curriculum. It embodies the overall progression of using data to answer questions and solve problems through describing a phenomenon, identifying interesting patterns or outliers, proposing explanations for these patterns and testing these explanations against each other as one moves toward a claim that is supported by evidence and reasoning.

The DSRF is meant to provide structure to doing DS in a way that encourages the discovery of unexpected patterns and developing explanations for these patterns as well as addressing some of its pitfalls that plague DS such as confirmation bias and focus on expected patterns.

## THE DATA SCIENCE REASONING FRAMEWORK



# Workshop Activities

*Data4All offers a variety of activities for learning: They range from near-peer Mentors in small groups to experiential & social learning through games (spark activities) to traditional classroom formats (lectures), and guest speakers to demonstrate DS in other real-world applications and highlight related college and career opportunities. These activities are summarized here.*

## Small groups with near-peer mentors

Small groups with near-peer mentors give students the opportunity to ask questions and work through coding roadblocks with their Mentors and peers, so all students can get personal attention tailored to their level of understanding.

## Jupyter/Colab notebooks & Kepler

Jupyter notebooks are web-based programming platforms that allow students to learn and work with the reasoning framework while learning to analyze real data in Python with statistical methods. Kepler is a web-based spatial data visualization platform where students can learn to analyze data in map formats.

## Spark Activities

Spark activities give students an intuitive sense of the reasoning framework and statistical concepts through collaborative games, e.g. illustrating the structure of a difference-in-difference research design through two teams competing to hit can pyramids where one team's results are manipulated through a can magnet intervention.

## Lectures

Lectures are presentations an instructor to the whole class to introduce and explain key concepts used in a given class, such as the difference between correlation and causation or between demographic characteristics and explanatory mechanisms.

## Lunch Speakers

Lunch speakers give students an overview of how the workshop content is related to college and career options. E.g., in 2022, two PhD biologists walked students through innovative case studies of how they used DS to solve problems that save lives and improve health. Other speakers explained their own career trajectories to make the steps for choosing college majors and careers more transparent, and they overviewed where and how to access college and career information.

# Activities & Learning Objectives

The following graphics shows how we expect the workshop activities to relate to short-term learning objectives. It also summarizes what we invested to produce the activities.

These objectives include 1) a broader understanding and appeal of DS, 2) improved confidence in programming, statistics, math and reasoning skills, and 3) greater college and career preparedness. The [background section spells out in more detail](#) how each activity is expected to impact each outcome in the short and [long term](#).

[Students were surveyed](#) before and after the workshop to assess in how far these short-term objectives were reached or not after 8 weeks.

The learning objectives are referenced on each page of the respective evaluation section.

## Resources

### Curriculum

- Developed by 4 team members (includes instructor guide, Jupyter/Colab Notebooks, Data Science Reasoning Framework (DSRF), and scoping of two real projects with data)

### Staffing

- 1 project manager
- 4 instructors
- 6 Mentors
- 8 lunch speakers

### Partners

- Data Science Institute
- Argonne National Laboratory
- Center for Spatial Data Science
- Office of Civic Engagement

### Funding

- Hymen Milgrom Supporting Organization (student stipends)
- Data Science Institute (catering)

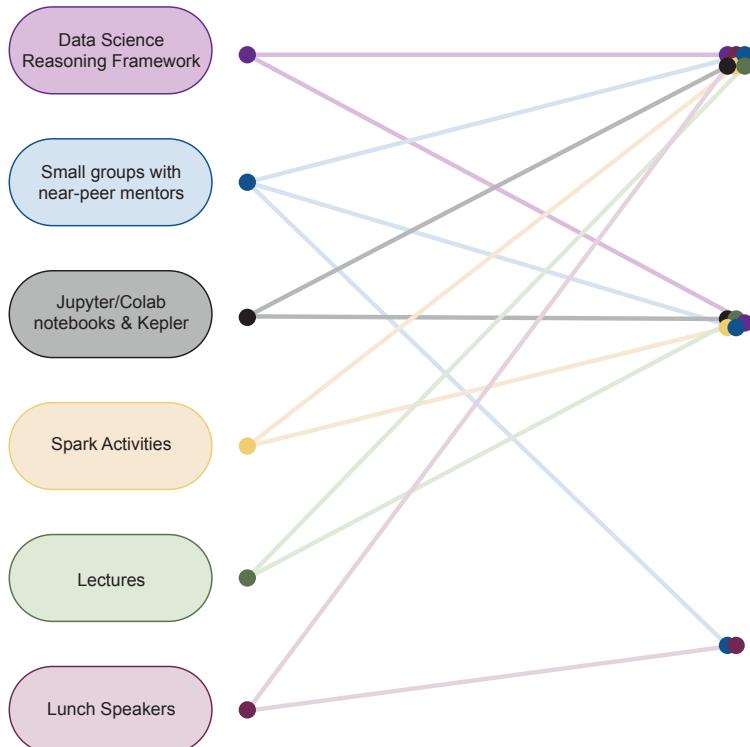
### Tech Infrastructure

- Jupyter Hub/Google Colab
- Chromebook Loaners

### Meeting Space

- Crerar Library @ UChicago

## Activities



## Learning Objectives

### 1. Understanding: What is Data Science?

- 1.1 Students see computer science and data science in a larger context of scientific reasoning and problem solving with data
- 1.2 Students who might not major in computer science find new relevance and broader appeal in computer science and data science
- 1.3 Students realize broader connection of computer science and statistics to STEM fields (e.g., biology, environmental science, social science, etc.)

### 2. Skills: Programming, Statistics, Mapping, and Reasoning

- 2.1 Programmers add skills in reasoning with data while others (like debaters) add programming skills
- 2.2 Students' confidence increases in:
  - analyzing data with Python and Kepler (mapping tool)
  - understanding foundational statistical concepts
  - using both to critically explore causes of problems

### 3. College and Career Preparedness

- 3.1 Students realize that data science can be used in various fields that can match their interests
- 3.2 Students find Data4All useful for clarifying data science-related college and career goals

# EVALUATION



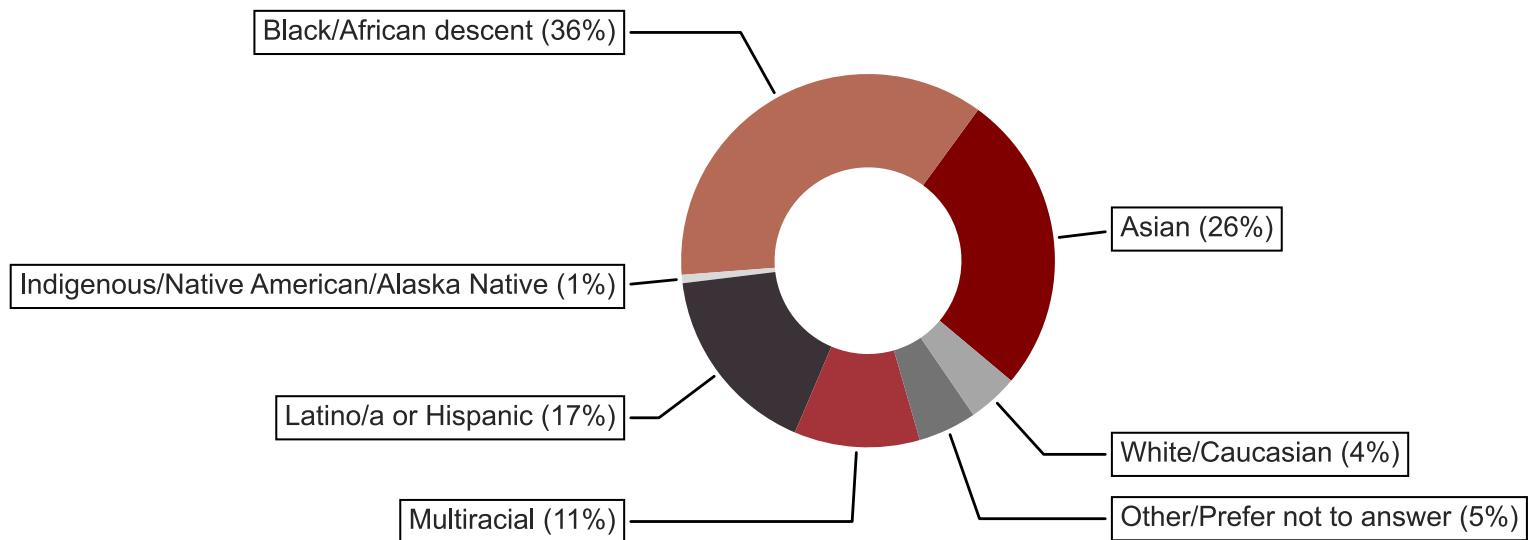
# 1. DID DATA4ALL REACH ITS TARGET AUDIENCE? YES



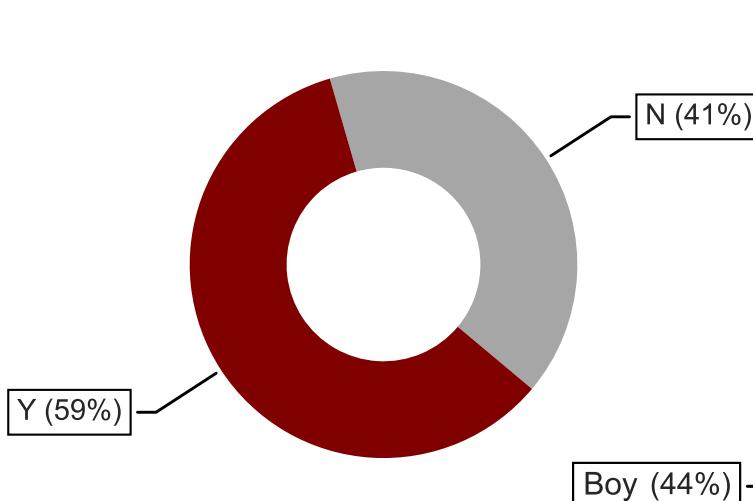
# Data4All students largely include gender and racial/ethnic minorities.

Fall 2022 – Fall 2024

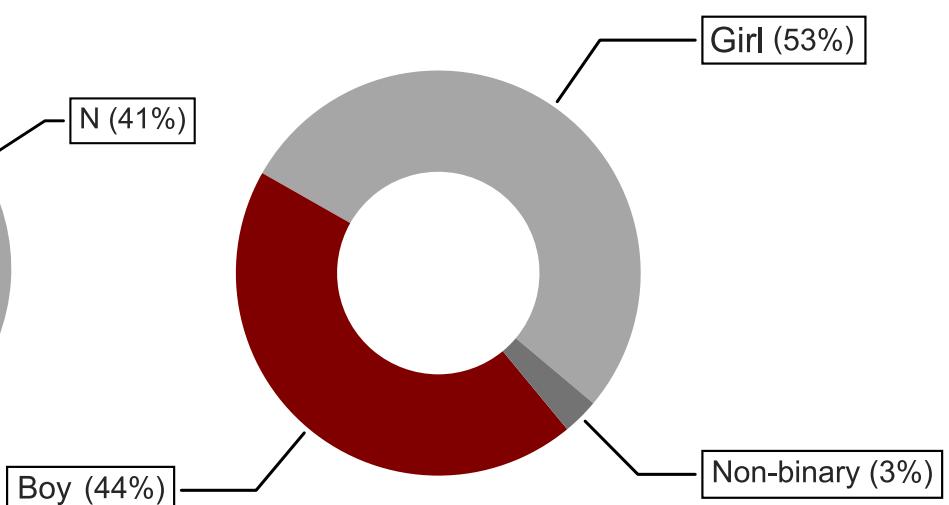
Racial/Ethnic Demographics



Underrepresented Minority\* Proportion



Gender Demographics

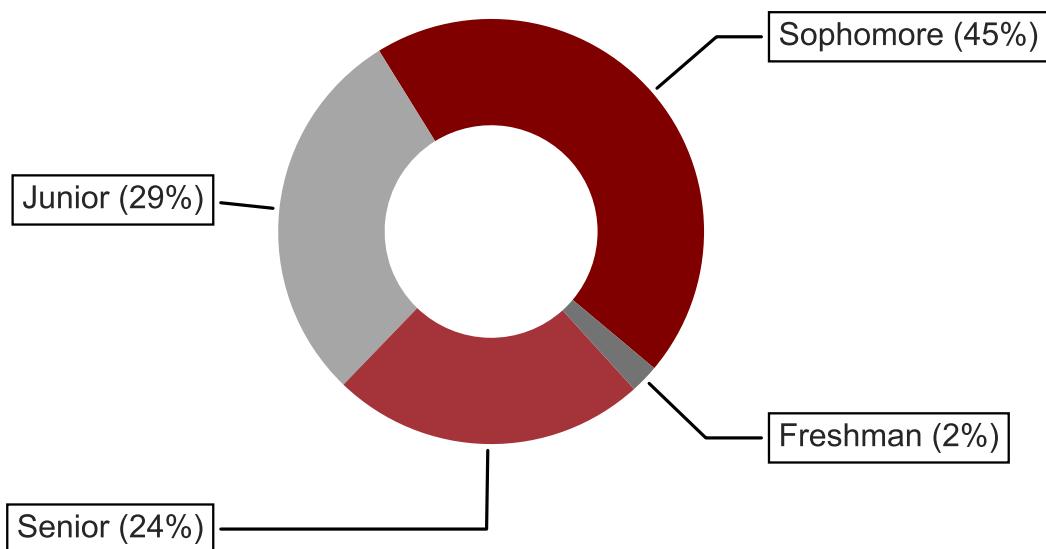


\*Per the [National Science Foundation definition](#), **underrepresented minorities** are races or ethnicities whose representation in STEM employment and S&E education is smaller than their representation in the U.S. population. This includes those that are Black/African descent, Hispanics/Latinos, and Native American or Alaska Natives.

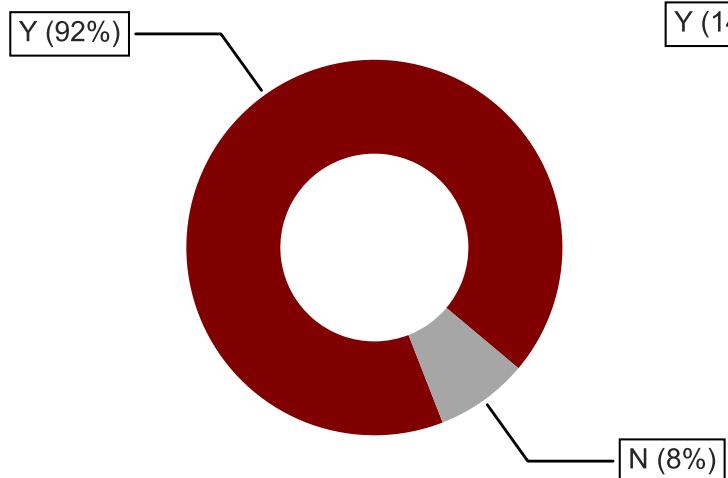
# A majority of Data4All students attend public schools.

Fall 2022 – Fall 2024

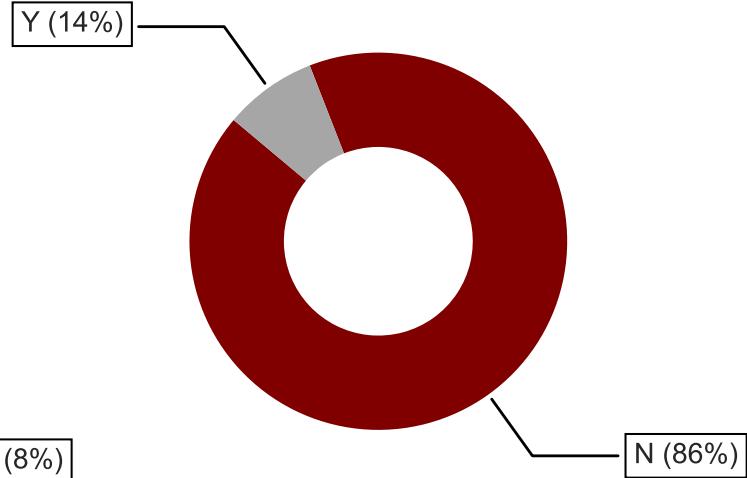
Grade Levels



Chicago Public School Proportion

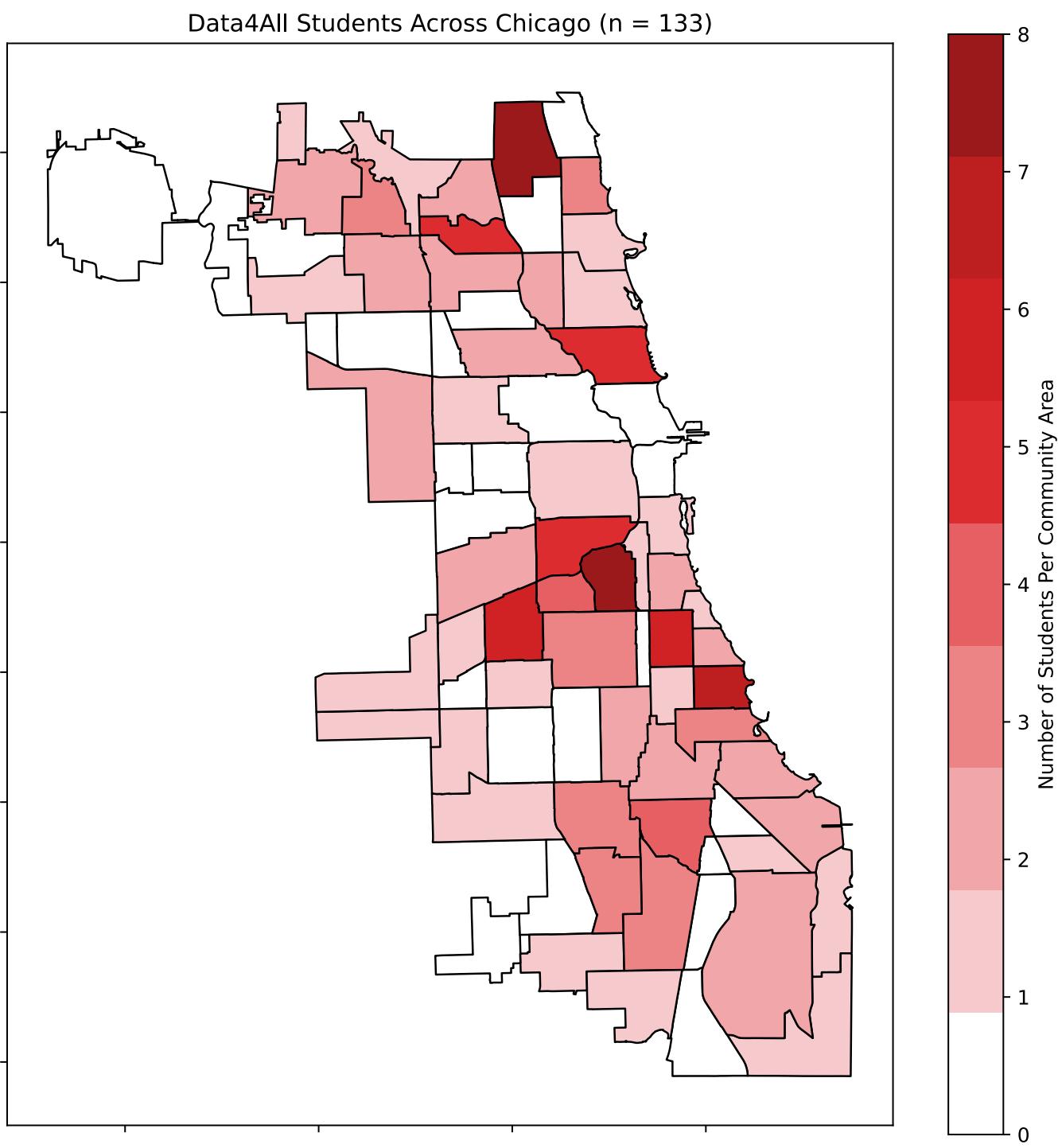


Charter/Private School Proportions



Students lived in 54 of the 77 community areas in within the City of Chicago.

Fall 2022 – Fall 2024



# Students came from 47\* high schools across Chicagoland, majority being part of the Chicago Public School district.

Fall 2022 – Fall 2024

## District Schools (29)

Carl Schurz High School  
Chicago Bulls College Prep  
Disney II Magnet High School  
Dr. Martin Luther King Jr. College Preparatory High School  
Englewood STEM High School  
Frederick Von Steuben Metropolitan Science Center  
George Washington High School  
George Westinghouse College Prep  
Gwendolyn Brooks College Prep  
Hyde Park Academy High School  
John Hancock College Prep  
Kenwood Academy High School  
Lane Tech College Prep High School  
Lincoln Park High School  
Lindblom Math and Science Academy  
Marie Skłodowska Curie Metropolitan High School  
Northside College Prep  
Ogden International High School  
Richard T. Crane Medical Preparatory High School  
Nicholas Senn High School  
Eric Solorio Academy High School  
South Shore International College Prep  
Thomas Kelly College Preparatory High School  
Von Steuben Metropolitan High School  
Walter Payton College Preparatory High School  
Whitney M. Young Magnet High School  
William Howard Taft High school  
William Jones College Preparatory High School  
Marist High School

## Charter Schools (9)

Gary Comer College Prep  
Butler College Prep  
Johnson College Prep  
Mansueto High School  
Muchin College Prep  
Pritzker College Prep  
UIC College Prep  
Urban Prep Academy  
Urban Prep Academy

## Private Schools (7)

De LaSalle Institute  
DePaul College Prep  
Holy Trinity High School  
Josephinum Academy of the Sacred Heart  
Leo Catholic High School  
University of Chicago Laboratory School  
Willows Academy

## Chicago Suburbs Schools (2)

Homewood-Flossmoor High School  
Thornwood High School

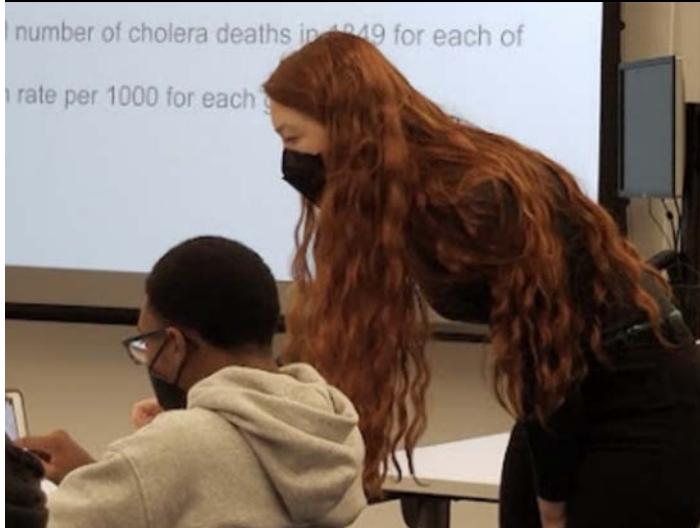
## Alternative Schooling (2)

Homeschooled  
James Madison Online High School

\*One student attended online high school, for a total of 48 high schools represented in the program.

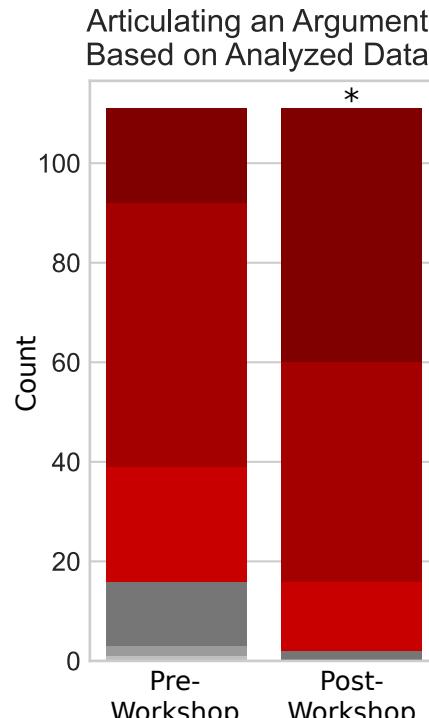
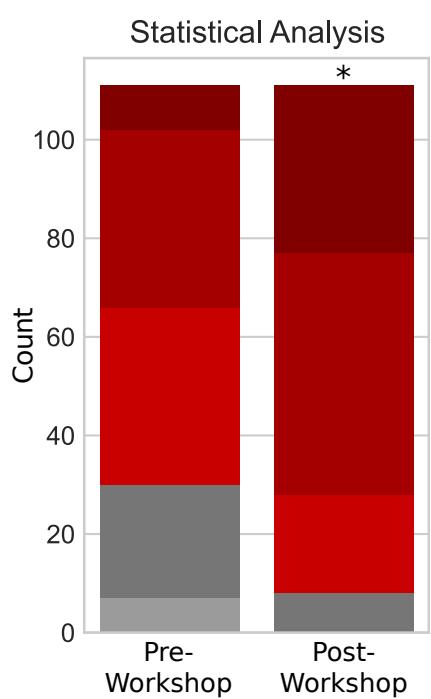
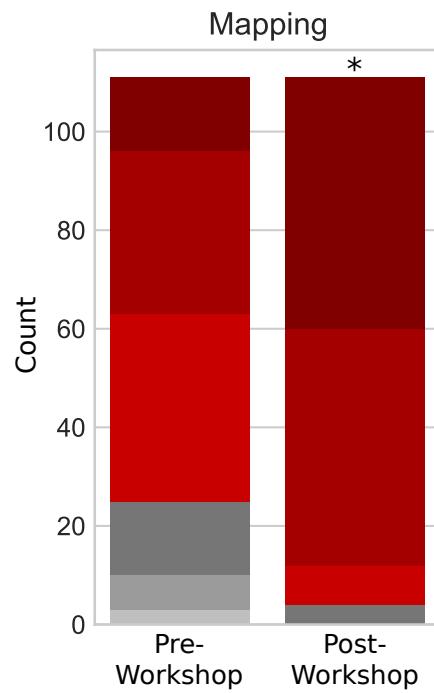
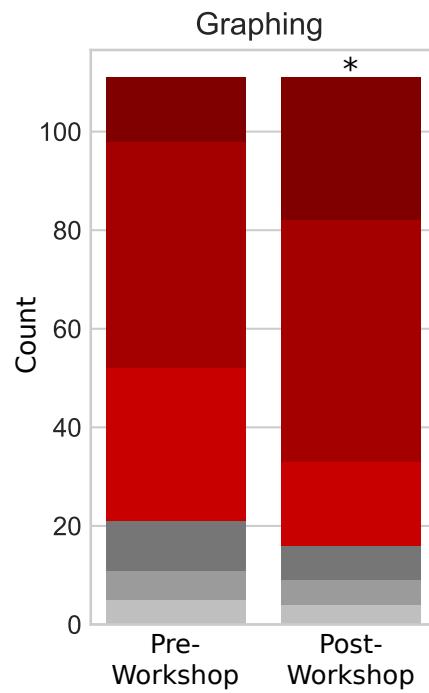
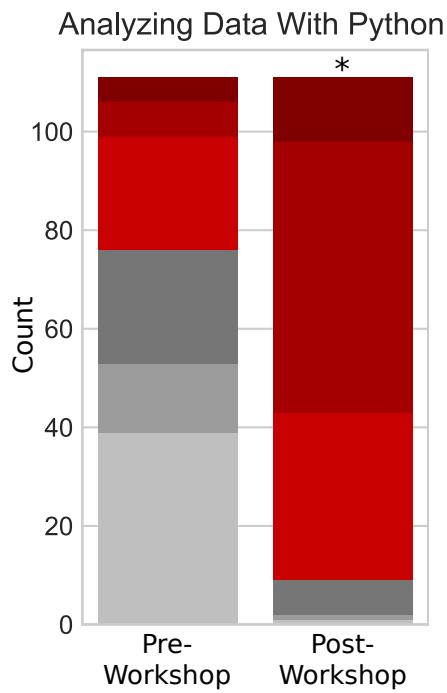


# 2. DID DATA4ALL REACH ITS LEARNING OBJECTIVES? YES



# Students reported increased comfortability in several DS skills after completing Data4All.

## Learning Objective 2.2



 Not comfortable at all  
Uncomfortable  
Slightly uncomfortable  
Slightly comfortable  
Comfortable  
Extremely comfortable

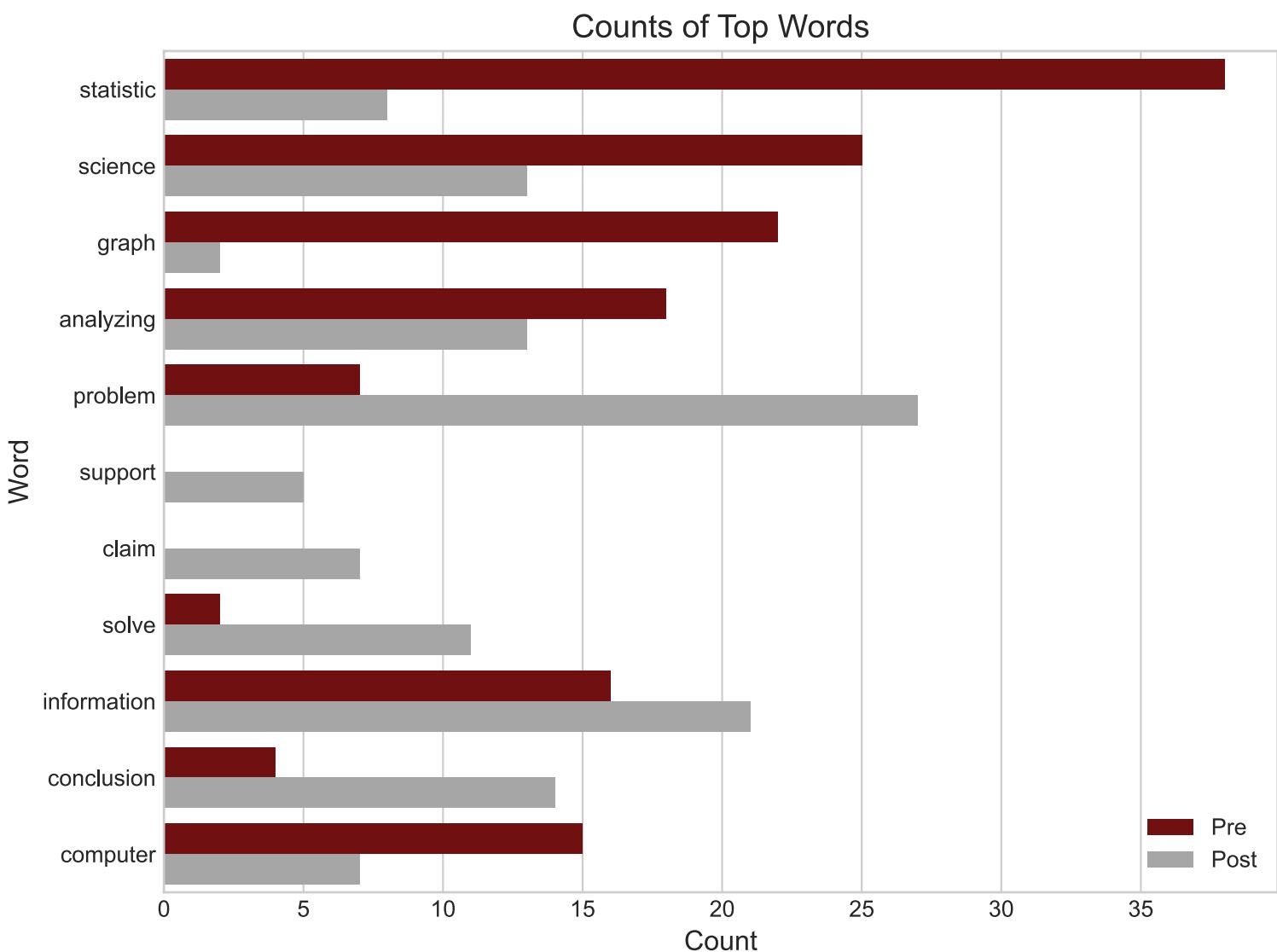
Data was converted from ordinal categories to corresponding numerical values, ranging from 21 (Not comfortable at all) to 26 (Extremely comfortable). The Wilcoxon signed-rank test was used to statistically analyze differences in comfortability in a non-parametric fashion. Data was assessed for normality before analysis by using the Shapiro-Wilk test with a significance threshold of 0.05. \* $p < 0.05$  corresponding to Wilcoxon signed-rank test.

# Data4All helped students see DS in a larger context of scientific reasoning.

## Learning Objective 1.1

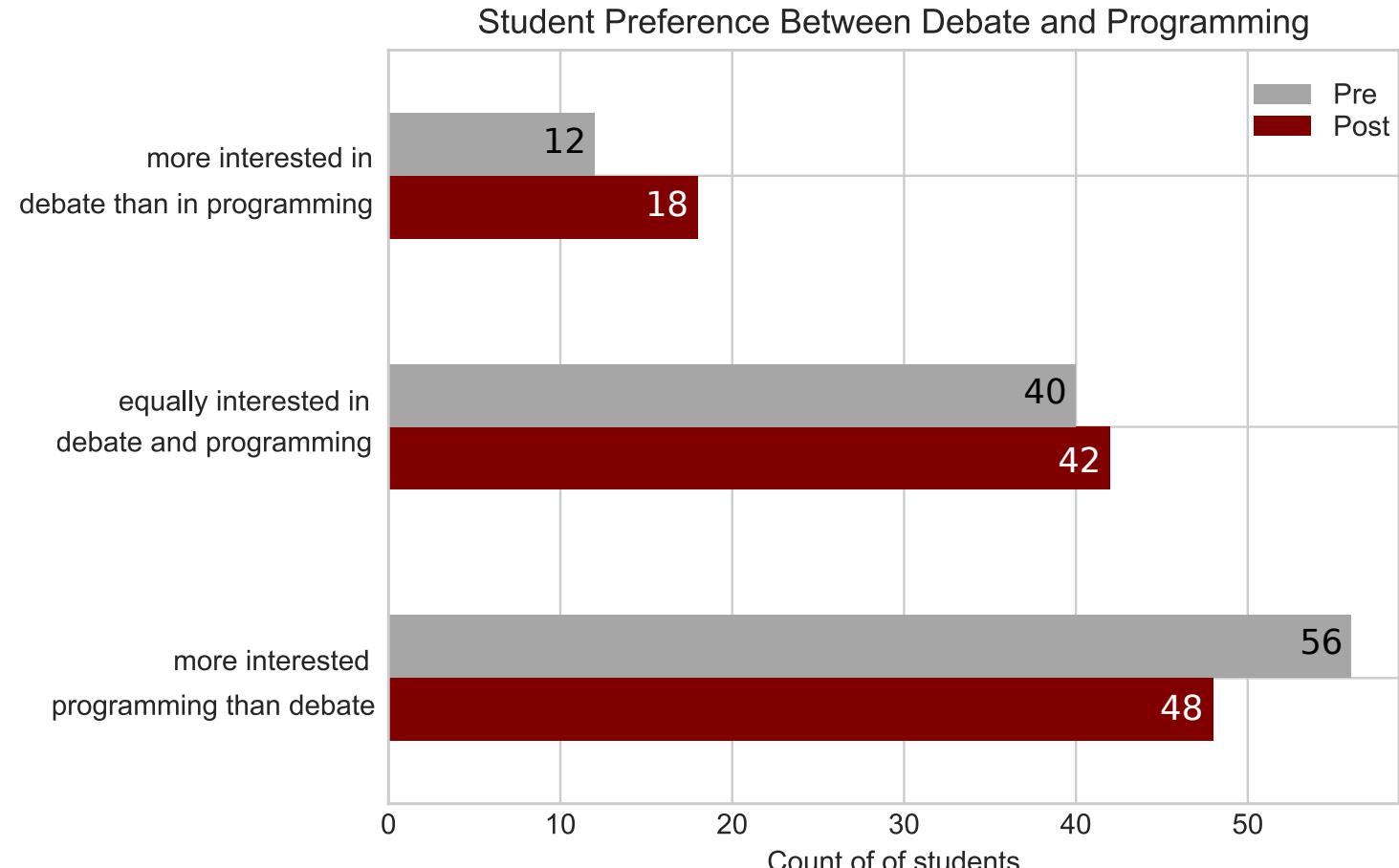
Before and after Data4All, students were asked the question, “**When you hear the phrase ‘DS’, what comes to mind?**” Below is a visualization of the occurrence of the frequent words mentioned in the pre- and post-responses. Artifact terms such as “DS,” “data,” and “think” were removed from the answers before conducting the analysis. A total of 114 responses from students that participated in Data4All iterations from spring 2023 to fall 2024 were analyzed.

After Data4All, students saw DS in a larger context of scientific reasoning. Prior to the program, words frequently included in student responses were “statistic,” “science,” “graph,” “analyzing,” and “information,” suggesting that students largely attributed DS to more technical tasks. After the program, students expanded their perspectives to include processes of inquiry, as denoted by increased frequency of the words “problem,” “support,” “claim,” “information,” and “conclusion.”



# Integration of scientific inquiry broadened the appeal of debate based on evidence-based reasoning after participation in Data4All

## Learning Objective 1.2



**Two seniors from Englewood STEM and Homewood- Flossmoor who are interested in exploring fields beyond CS and DS:**

"My experience in DS workshops have been very influential, i love science, debate, and mysteries, and DS is a great way to bring all of my interests together into one very influential field of study."

"It's expanded my view of DS so much, which is so helpful because I feel like I can use this in my future."

**Northside College Prep junior without prior programming experience:**

"I enjoyed the notebooks and the Cholera problems had me more interested. I also enjoyed learning more about UChicago from all the Mentors, they helped a lot with telling us about the school and showing us around during lunches."

**Marist HS sophomore, on what excites them most about learning more about DS:**

"The debates and the interaction with other like-minded coders is what excites me the most."

# Through the DSRF and speakers, students realized the broader connection of computer science & statistics to STEM

## Learning Objective 1.3

### In the Words of Our Students:

**Jones College Prep senior who is interested in exploring fields beyond computation and DS:** “My experience in the workshop has made me even more inclined to pursue computer science more than before. Mainly because I was in between pursuing public health and computer science but now I am sure that I can use both my skills in data and computer science to also pursue my passion for solving health and environmental issues.”

**Lane Tech senior with AP CS experience:** “It made me realize of new career possibilities I was not aware of before. For example, combining DS with medicine (Microbiology).”

**Crane Medical Prep sophomore with programming experience:** “It has shown me a new side of DS by seeing that it can come with many different factors that i haven't thought about. … “The highlight was this group activity of showing and proving our hypothesis/claim.”

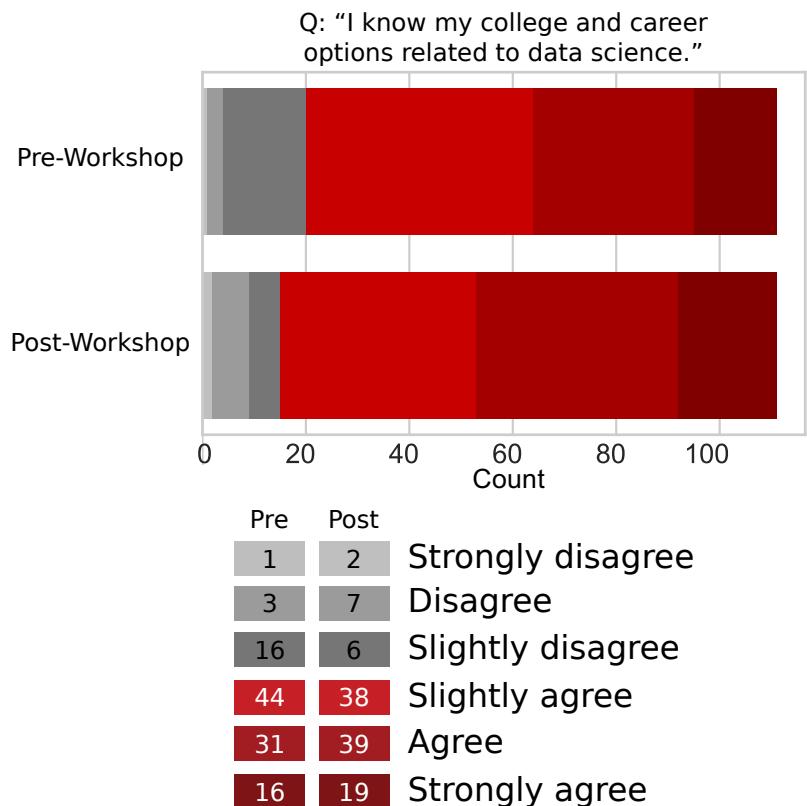
**Walter Payton senior with AP CS experience:** “I really enjoyed being able to talk with my teammates about our theories for our case-studies it really feels like we are scientists conducting and testing real-life experiments.”



# Students gained some clarity around college and career options related to DS and realized its broader relevance through guest speakers.

Learning Objective 3.1

Learning Objective 3.2



Students ranked their agreement with the statement "*I know my college and career options related to DS*" before and after the program. While the number of students that were in some form of agreement with this statement increased after the workshop, there was also an increase in those that "strongly disagreed" or "disagreed." This may be due to students not remembering their response to the question from the pre-workshop survey or students wanting further clarity of their career options.

#### **Weekly feedback from anonymous students (2023):**

"I enjoyed learning about the different game labs and careers within them. I think Ashley did a really good job explaining the different aspects of the lab and the connections within them, like how data scientists team up with others to help them create visuals with their data."

"Today's highlight for me was the tour we took of the computer science labs downstairs. It was really exciting to learn more about specific careers and majors in computer science, and what you could do in them."

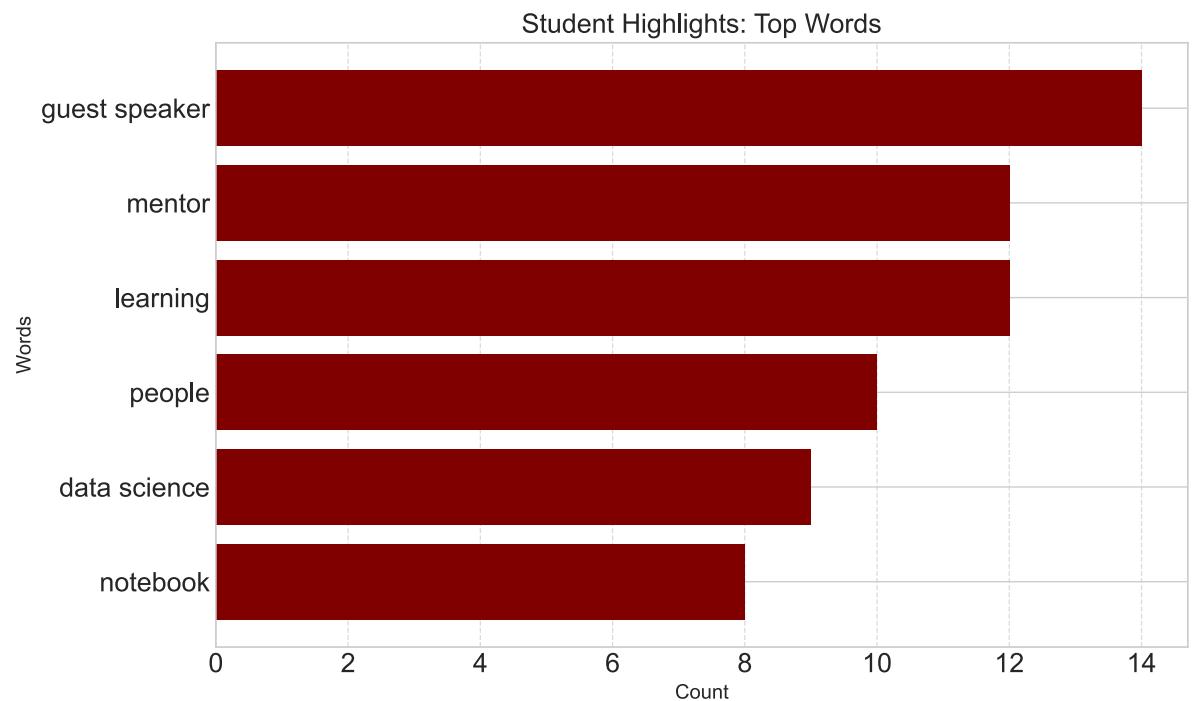
"A highlight today was seeing how much terminology was used from biology and U.S History. It was very cool to see."

### 3. EVALUATION OF DATA4ALL PROGRAM

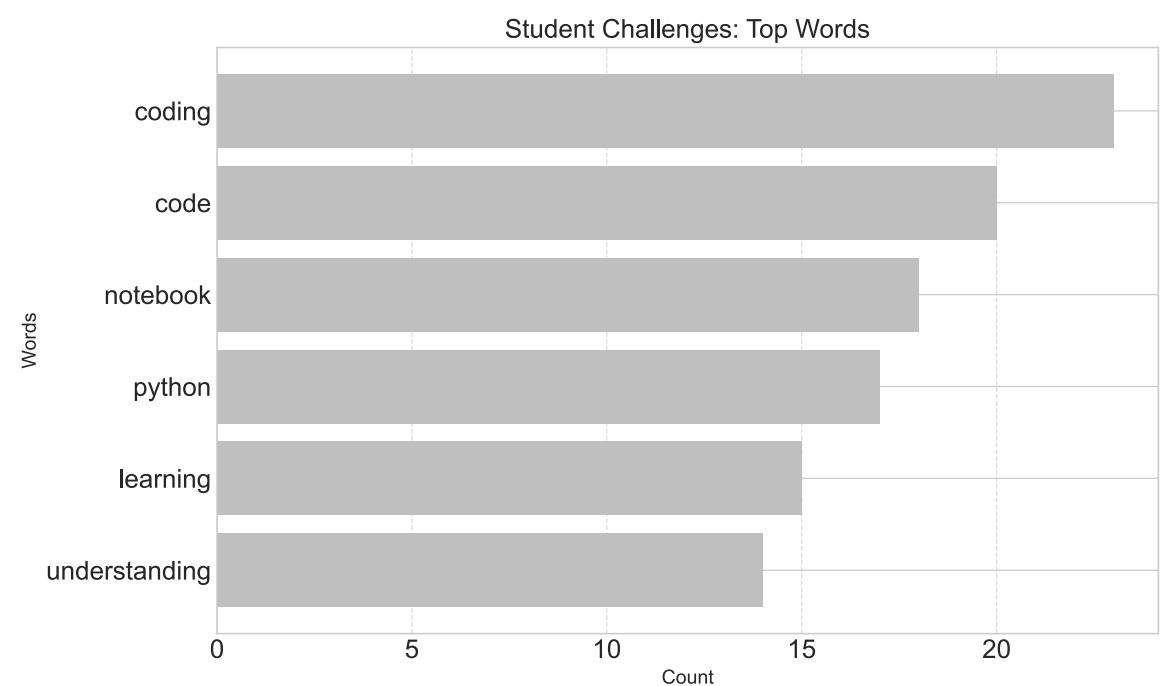


**Students emphasize social aspects as highlights of the program even though coding in the Jupyter/Colab notebooks was the most challenging.**

Following completion of Data4All, students were asked:



What was the biggest challenge in the program?



# SMALL GROUPS, MENTORS, & GUEST SPEAKERS

Students worked in small groups of 4-6 with one near-peer mentor each. The near-peer Mentors were UChicago undergraduate students with expertise in DS and scientific reasoning. Mentors helped students work through the reasoning, programming, mapping, and statistical concepts and addressed their questions about college.

Most weeks, students heard from guest speakers that utilize DS in various professions, including academic research, tech industry, and social policy.

As the feedback illustrates, students found their Mentors and small groups especially helpful for learning how to program and debug their code. Students also expressed positive sentiments about small groups, mentors, and having guest speakers, with evaluation data showing 98%, 98%, and 97% satisfaction, respectively.

## In the Words of Our Students:

**Englewood STEM senior with extensive Python skills:** “I adore my group, we always have a fun time while getting our work done. We share thoughts, ideas, and just feed off of each other's creativity.”

**Mansueto HS senior with less programming experience:** “Sometimes when we would be working on the jupyter hub I would get stuck on some parts but then I would ask for help and I would have my mentor or one of the people in my group help me.”

**Jones College Prep senior with some Python skills:** “My biggest challenge has been learning how to connect the pieces of data so far. My teammates really help me see the connections through our discussions but I hope to also develop those skills on my own as well.”



# SPARK ACTIVITIES & LECTURES

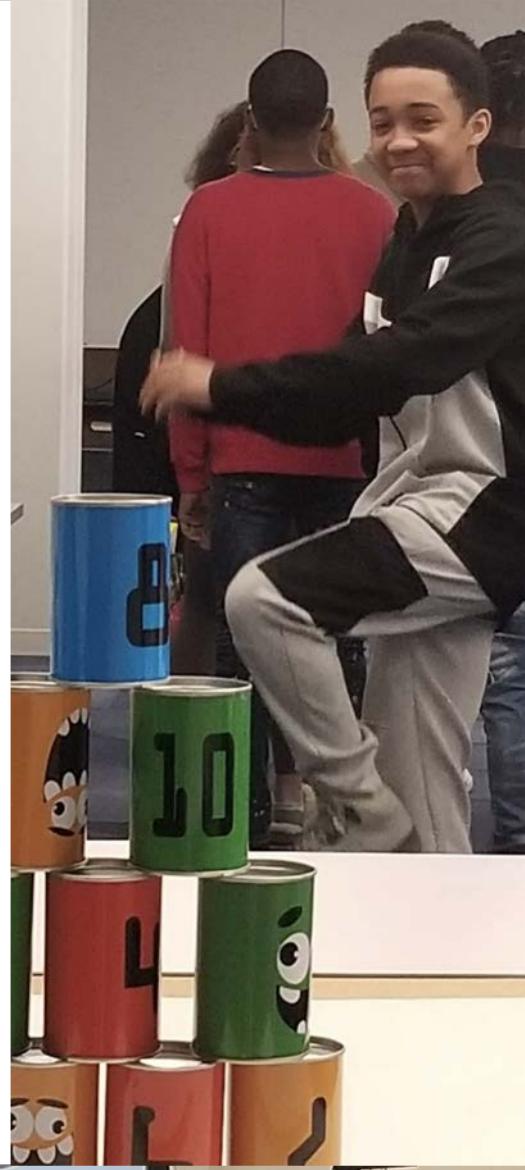
Spark activities and lectures primed students to help orient their thinking on approaching relevant topics discussed in Jupyter/Google Colab notebooks. For example, one spark activity discusses the use of different variables of data that can be used to determine who is the GOAT (Greatest Of All Time) in basketball. Students then apply this reasoning to find appropriate variables to determine how cholera was impacting London in the 1800s.

Students expressed positive sentiments toward lectures and spark activities as didactic tools to supplement learning, reporting 95% and 94% satisfaction, respectively.

## In the Words of Our Students:

### **Northside College Prep Junior with prior programming experience:**

**experience:** “The highlight of my time in the program was all the fun games we played and the friends made both within my table and with the rest of the group. I really enjoyed how we were able to learn in an exciting environment and how we were able to move around a lot instead of sitting in a chair all day.”



# NOTEBOOKS & KEPLER ACTIVITIES

Jupyter and Google Colab notebooks were used to teach students Python programming in an interactive and integrative environment where they can visualize output and errors for debugging. In addition to providing structure coding exercises, notebooks also included reflective writing prompts to reinforce statistical concepts and critical thinking skills covered in lectures and throughout the notebooks.

Kepler was used as an accessible online platform where students can create maps and examine geospatial patterns. Using spatial projections as a basis for inquiry, students were able to formulate testable questions and explore these questions using data in the notebooks.

Students really liked the programming and mapping exercises even though they also found the work with Jupyter/Colab notebooks the most challenging. Of 109 post-workshop responses, 94% and 96% of students expressed positive sentiments about notebooks and kepler activities, respectively.

## In the Words of Our Students:

**Walter Payton senior who took AP CS:** "I usually don't enjoy coding, but the way things were broken down in the notebook and how extensive things were explained (either in the notebook or by the mentor) really made it more enjoyable and manageable. It's also nice to know that I can now analyze both tables and graphs to draw conclusions. "

**Walter Payton sophomore:** [A highlight of your time in the program ] :"Using Kepler to analyze data"



# CASE STUDIES

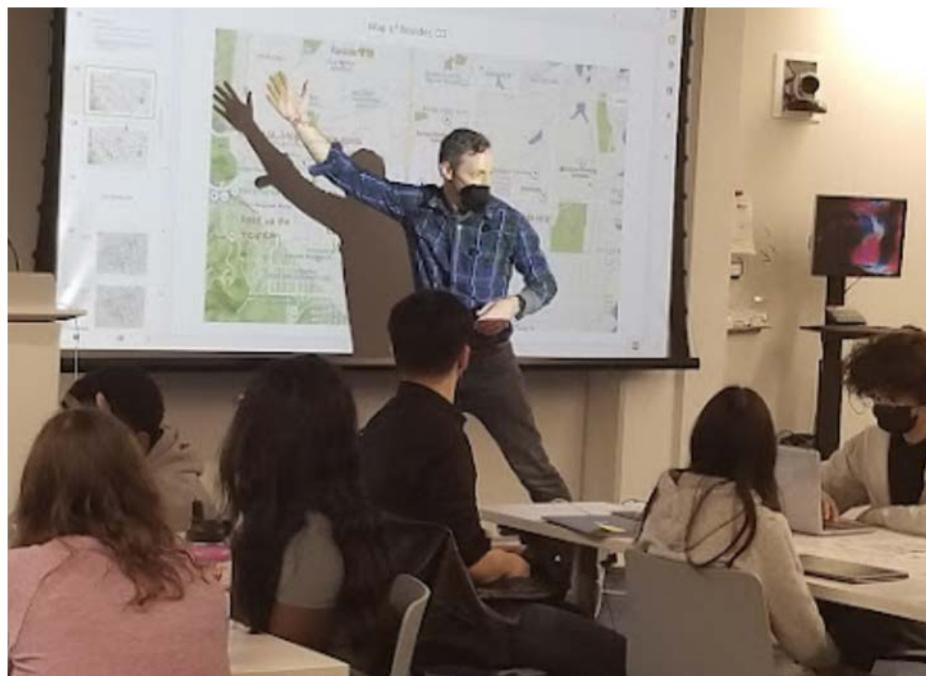
Data4All applies its reasoning framework to two real-world problems that students address with programming, mapping, and statistical tools: 1) **Why cholera spread** in 19th century London and 2) **how COVID-19 impacted Chicago neighborhoods differently** in the early 2020s. Since the first case is historical, it is more scaffolded, while the 2nd case allows students to apply the same reasoning to a more open-ended problem that was relevant to their lived experience.

Students enjoyed working through case studies as a way of learning DS, with 93% of students reporting positive sentiments about the John Snow case and 98% of students reporting positive sentiments about the COVID-19 case.

## In the Words of Our Students:

**Von Steuben Metropolitan Science Center junior:** “My highlight of this program would most likely be the activity where we mapped out the positivity rates of covid in Chicago. We put stickers and analyzed where each rating would be.”

**Thomas Kelly College Prep senior:** [A highlight of your time in the program ] :“Getting an introduction to python and figuring out the answer or important variables that lead to figuring out how the cholera outbreak started or who is most impacted by Covid.”



# BACKGROUND



# About the Hosts



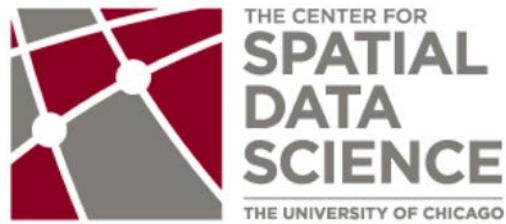
The Data Science Institute (DSI) executes the University of Chicago's bold, innovative vision of Data Science and AI as a new discipline. The DSI seeds research on the interdisciplinary frontiers of this emerging field, forms partnerships with industry, government, and social impact organizations, and supports holistic DS education. It partners with the UChicago Department of Computer Science, where it is located, and the UChicago Department of Statistics.



Argonne is a multidisciplinary science and engineering research center, where talented scientists and engineers work together to answer the biggest questions facing humanity, from how to obtain affordable clean energy to protecting ourselves and our environment. Through its middle school to graduate programming, Argonne Education connects with over 30,000 youth and families at Outreach events, hosts over 4,400 middle and high school students at the Learning Center, and employs over 900 undergraduate & graduate students each year.



# About the Hosts



The team at the Center for Spatial Data Science (CSDS) thinks spatially about research problems. CSDS develops state-of-the-art methods for geospatial analysis; and applies them to policy-relevant research in the social sciences. It developed the open spatial software GeoDa, which has been downloaded over half a million times, and hosts spatial analytics lectures with over 670,000 views on its YouTube channel. It hosts a larger research project on integrating scientific reasoning with spatial DS: <https://puttingscienceintodatascience.org/>



The Office of Civic Engagement (OCE) connects the University to the city and the South Side. OCE stewards the University's commitment to the city, supporting other academic and administrative units as they develop and advance their distinct civic priorities. It also leads areas of strategic work that extend the University's reach and impact to the city of Chicago and the South Side.



# About the Surveys

This report is based on the results of 3 student surveys conducted with Google forms in each of the 2022, 2023, and 2024. These surveys were deployed before the beginning of the workshop, in the middle, and after the workshop. It also draws on feedback from near-peer Mentors and instructors obtained in focus groups after the workshop. Because we revised the survey questions for 2023 to take the logic model into account that we developed after the 2022 workshop, in some cases there is no direct comparison between the 2022 and later results.

**1) Baseline survey before workshop (N=24 in fall 2022; N=24 in spring 2023; N=30 in fall 2023; N=30 in spring 2024; N=30 in fall 2024)**

All students completed this survey before the start of the workshop. The survey asked questions about their demographics, schools, and experience with and confidence in computing, data analysis, and scientific reasoning.

**2) Midpoint survey (N=20 in fall 2022; N=23 in spring 2023; N=28 in fall 2023; N=30 in spring 2024; N=30 in fall 2024)**

The midpoint survey contains questions about Data4All's Mentors, instructors, activities and levels of confidence in workshop-related skills. Since a larger number of students completed the midpoint than final survey in 2022, many of the 5-point scale results in this report are based on midpoint results for 2022 (which tended to be consistent with the final results).

**3) Survey at workshop end (N=13 in fall 2022; N=21 in spring 2023; N=30 in fall 2023; N=30 in spring 2024; N=30 in fall 2024)**

The final survey had many questions that were the same as in the midpoint and pre-workshop surveys, so that pre-post and mid-post analyses could be performed. The smallest number of students typically answered the final survey's questions in the last workshop (week 8). We saw some attrition on the last workshop day in both workshops (an issue attempted to address by tethering student stipends to attendance).

# Open Teaching Materials

All of the teaching materials were peer reviewed and revised by Damaris Hernandez in the summer of 2023. She earned a B.S. and M.Ed. from UC Berkeley, has extensive expertise in math and DS, and worked as a math middle school teacher at Sayre Language Academy in Chicago.

The revised materials will be accessible openly and for free in fall 2024 here:

<https://github.com/uchicago-dsi/2023-data4all>

Please contact us at [spatial@uchicago.edu](mailto:spatial@uchicago.edu) with any questions.



# Acknowledgements

**Funding:** Hymen Milgrom Supporting Organization, UChicago Office of Civic Engagement, Inclusive Innovation Fund, Data Science Institute

**In-kind contributions:** Data Science Institute (DSI), Argonne National Lab (ANL), Center for Spatial Data Science (CSDS)

**Collaborators:** UChicago Office of Civic Engagement, Peter Vinten-Johansen (historian | cholera case) and Kevin Credit (data and research | covid case), Jeanne Century (program & logic model), Shaz Rasul (recruitment)

**Curriculum Developers:** John Domyancich (ANL), Tyler Skluzacek, Julia Koschinsky (CSDS), Bethany Frank (ANL)

**Instructors:** John Domyancich, Kelly Sturner, Jacqueline Otmanski, Azucena Rodriguez (all ANL), Joshua McKeever, Bradley Stoller, and Evelyn Campbell (DSI)

**Mentors:** Anna Poon, Arnav BrahmaSandra, Christine Park, Cielo Flores Martinez, Disha Mohta, Douglas Williams, Eliza Migdal, Emma Ricks, Erin Ku, Jenny Lee, Kennedy Coleman, Moturi Ondeyo, Obi Obetta, Osinachi Nwosu, Pratham Srivastav Gandhi, Ryan Michaud, Sebastian Plaza, Victoria Kielb, Yannick Tanyi, Yurou Li, Ziyu Ren, Amanda Kube Jotte, Susanna Lange, and Evelyn Campbell

**Logistics:** Jessica Sweeney, Katherine Rosengarten, Maria Victoria Fernandez, Shwetha Srinivasan, Charisma Lambert, and Evelyn Campbell

**Speakers:** Dr. Amy Kirby (CDC), Ashlyn Sparrow (UChicago Weston Game Lab), Dom McKoy (To & Through), Veronica Hauad (UChicago Admissions), Brandon McCallister (UChicago Admissions), Dr. Brandeis Marshall (DataEdX), Dr. Chantel Perry (theseen.ai), Joshua McKeever (BSD), Kevin Bryson (CS), Jared Bryson, Dr. Kriti Seghal (DSI), Dr. Meghan Hutch (DSI), Dr. Jordan Kemp, Noma Shields, Alexandra Usher, Chuks C-Madu (Netflix), Torie Bates (Google), and Evelyn Campbell

**Photography:** Joe Sterbenc, Julia Koschinsky, Charisma Lambert, and Evelyn Campbell

**Teaching Material Revisions:** Damaris Hernandez (peer review), Julia Koschinsky, and Evelyn Campbell

**Evaluation and Analysis:** Julia Koschinsky (report content and design), Evelyn Campbell 34 (report content and design, data collection, analysis, and visualization)

Mentor Yannick Tanyi guiding students through a geospatial analysis of Chicago COVID-19 data.

