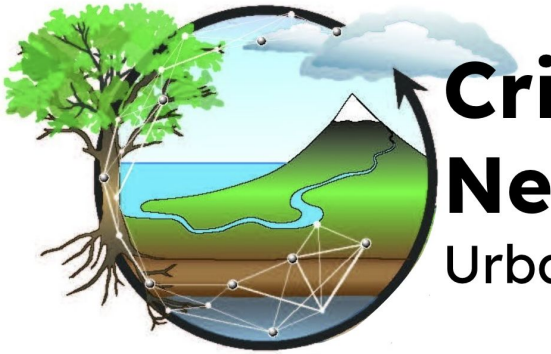
Urban Critical Zone Data Jam

Webinar #2: Analyzing and Visualizing Data
Thursday, March 13, 2025 6:00-7:30 p.m. EDT



Critical Zone Network

Urban Cluster

Hello & Welcome!

The Urban Critical Zone Education and Outreach Team.

Angie Hood
Education Program Leader
Cary Institute of Ecosystem Studies

Alan Berkowitz
Head of Education
Cary Institute of Ecosystem Studies

Tonight's Agenda

What to expect tonight

1. Welcomes and objectives
2. Quick review of Webinar #1
3. Objective #1 - Creating a data visualization
4. Objective #2 - Unpacking, Interpreting, and Analyzing
5. Objective #3 - Creating the written report
6. What's next?
7. Adjourn
8. Optional Q&A

Tonight’s Objectives

Goal of the Urban Critical Zone (UCZ) Data Jam

Promote data literacy using real, local datasets collected by researchers along the Fall Zone.

Objectives of tonight’s webinar

Participants learn about resources designed to help them support their students:

1. Determine the best way to create a visual representation of their selected data.
2. Interpret and analyze the information presented in their visualization.
3. Create a written report to communicate their data findings, which will serve as the basis of their creative component.

What is Data Literacy?

NCES defines data literacy as:

> "[...] the practice of examining and understanding data to draw and communicate conclusions and make decisions."

National Forum on Education Statistics. (2024). *Forum Guide to Data Literacy* (NFES 2024-079). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

What is Data Literacy?

A data-literate individual can...

1. **Read data** (interpretation and analysis)
2. **Work with data** (data collection and compilation)
3. **Communicate information found within the data** (effectively sharing the “data story” in order to inform others of important information found in the data)
4. **Reason with data** (use the data to make decisions)
5. **Critique data-based arguments** (evaluate claims that are, or should be, supported by sound data analysis and interpretation)

(Crabtree & Nehme, 2024)

Why Focus on Data Literacy?

Data literacy can lead to simple and effective solutions to real world problems. Data provides a common language through which those who are data literate can hold effective, data-driven conversations that will allow them to make informed decisions and take action based on cause and and effect relationships evident in and explained by the data.

National Forum on Education Statistics. (2024). *Forum Guide to Data Literacy* (NFES 2024-079). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

The UCZ Data Jam Webinars

Urban Critical Zone Data Jam Competition

Getting Started: UCZ Advisor Resources

UCZ Data Jam: Essentials & Registration

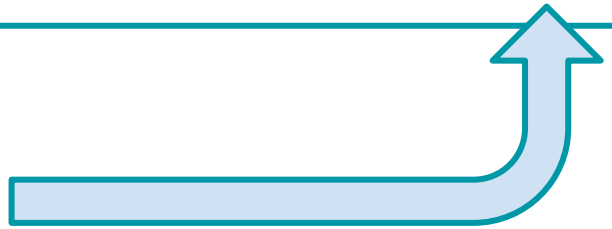
UCZ Datasets

UCZ Data Jam: Important Documents

Data Jam FAQs

We've gathered the following resources to help you on your journey.

- **Data Jam Essentials** is an overview of the contest and complete contest information.
- We have a special "advisors-only" Google Drive Folder with sample graphs for most of the datasets. Please email us to receive access.
- We asked veteran Data Jam coaches to share their advice. Read their **tips** on how to Jam.
- We will be hosting a three-part webinar series designed to support you and your students as they embark on their Data Jam journey. If you are unable to attend a scheduled webinar, a recorded version will be linked below. Additionally, you may find **recorded webinars** from our sister Data Jam–the Hudson Data Jam–useful.
  - **UCZ Webinar #1**: Working with the data & introduction to CODAP
    - Supporting **slide deck** used in webinar
  - **UCZ Webinar #2**: Data analysis & visualization (Recording available March 17)
  - **UCZ Webinar #3**: Effective communication of a data story (Recording available March 24)



bit.ly/UCZDataJam

The UCZ Datasets

Urban Critical Zone Data Jam Competition

Getting Started: UCZ Advisor Resources

UCZ Data Jam: Essentials & Registration

UCZ Datasets

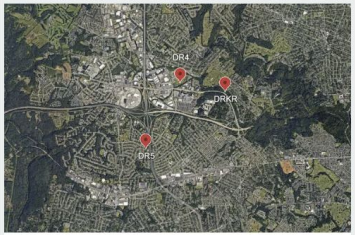
UCZ Data Jam: Important Documents

Data Jam FAQs

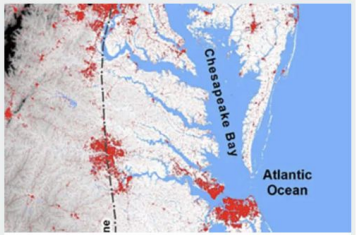
bit.ly/UCZDataJam

Datasets

Anions in the Pennypack



Greenhouse Gases in Dead Run



Stream Chemistry Across the Fall Zone

Each dataset includes:

1. A brief introduction to the UCZ project and specific dataset
2. A metadata page in the workbook
3. The dataset
4. Maps of the watersheds (satellite, MRLC, USGS watershed basin)

Objective #1: Choosing your data visualization.

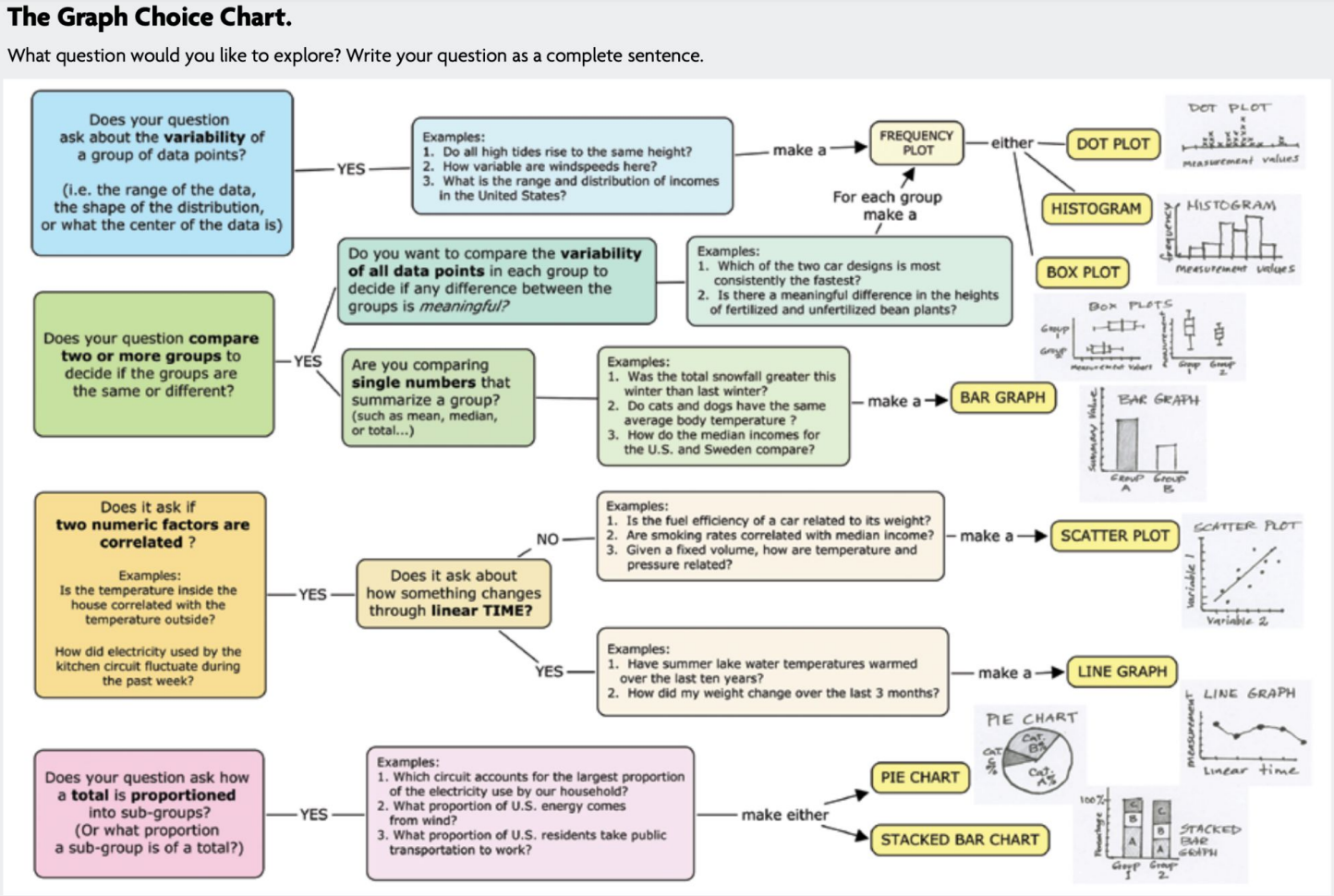
The first step to any good dataquest is a solid **research question**.

We will explore two options that you can offer your students to help them formulate their research question as they embark on their Data Jam Journey.

Data Visualization Option #1: Graph Choice Chart (GCC)

The GCC is a student-facing decision making tree that will help your students choose the appropriate method of data visualization for their selected **research question**.

(Webber, *et al*. 2014)
Full article available from NSTA.



Data Visualization Option #1: Graph Choice Chart (GCC)

To utilize the GCC, students must first have a precisely worded **research question**.

Students should use the information provided in the About the Data tab to consider the following questions:

Where is the data being collected?

Who is collecting the data?

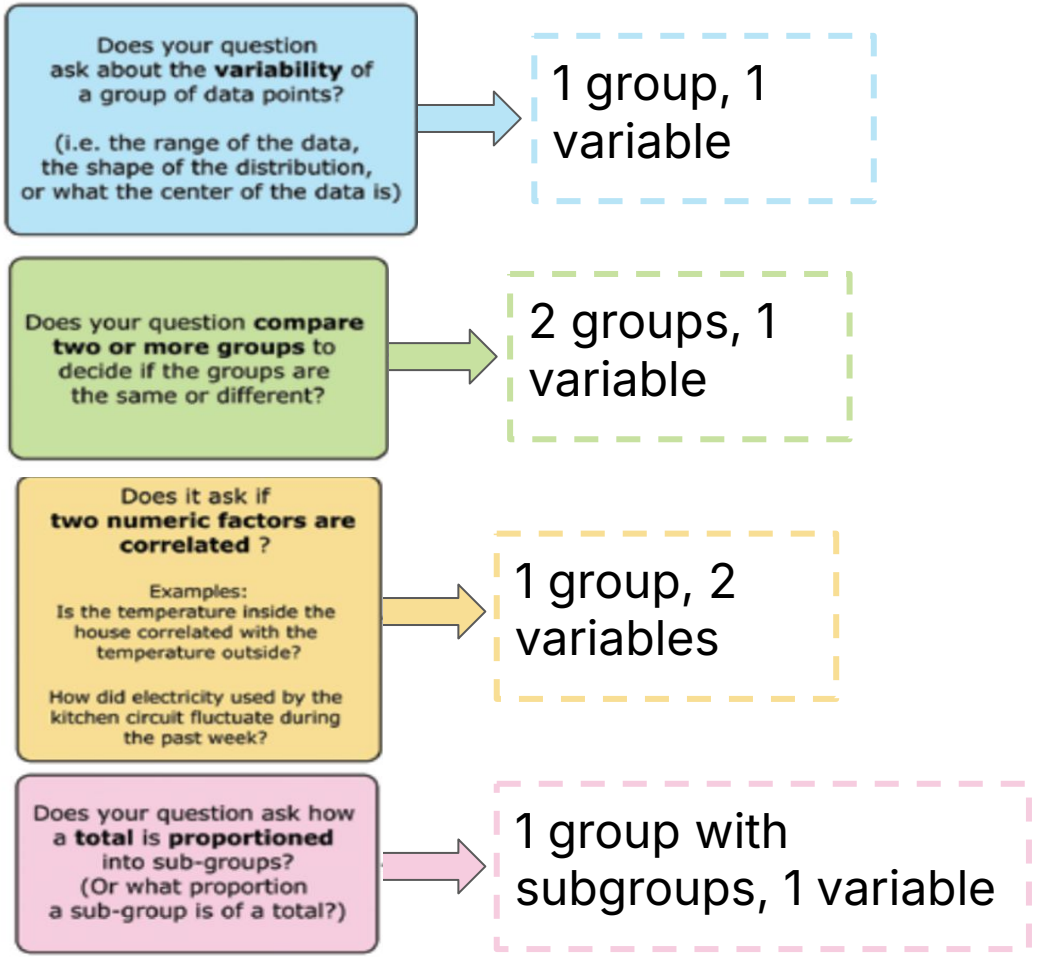
Why is this data being collected?

What are the independent variables? In what units are they being measured?

What are the dependent variables? In what units are they being measured?

Once students have completed this preliminary investigation of the dataset, ask them to formulate a **research question** that they think they might be able to find an answer to within their dataset.

Data Visualization Option #1: Graph Choice Chart (GCC)



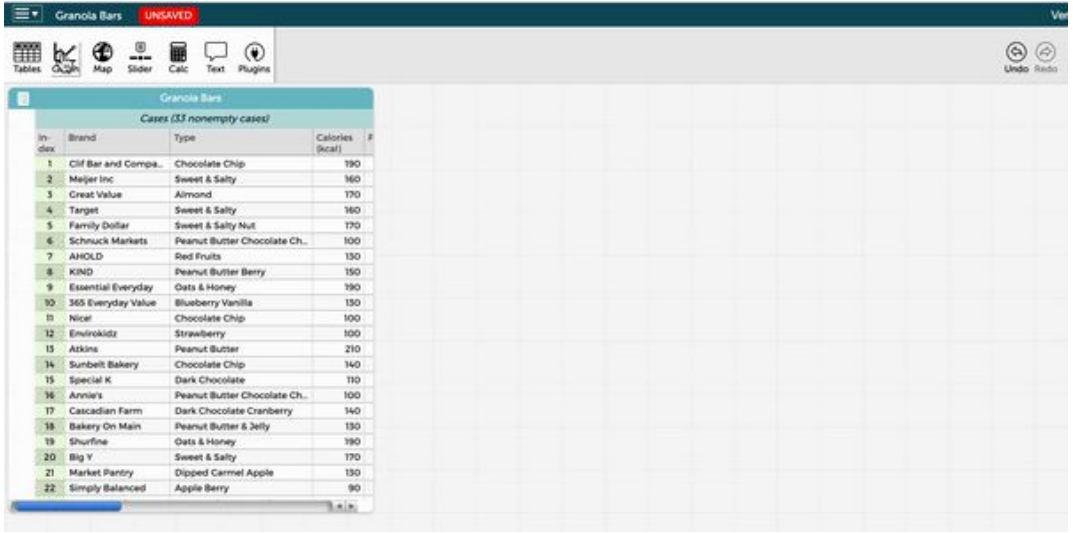
With their newly formulated their **research questions**, students can use the GCC to determine which graph type will best fit their inquiry needs.

Data Visualization Option #2: Drag & Drop Exploration

Alternatively, students may choose to use a drag-and-drop data analysis platform like CODAP to explore the dataset to “see what they see.”

Through drag-and-drop exploration, students have the opportunity to “mess around with” the data until they see something interesting pop up in their chart.

Using this method, students will develop their **research question** after noticing an interesting trend, pattern, or comparison in the data



CODAP

The Concord Consortium

Comparing the Data Visualization Options

Let's use our "Stream Chemistry Along the Fall Zone" dataset as an example.

Following the procedure included in the GCC option, the formulation of a student's research question may look like this:

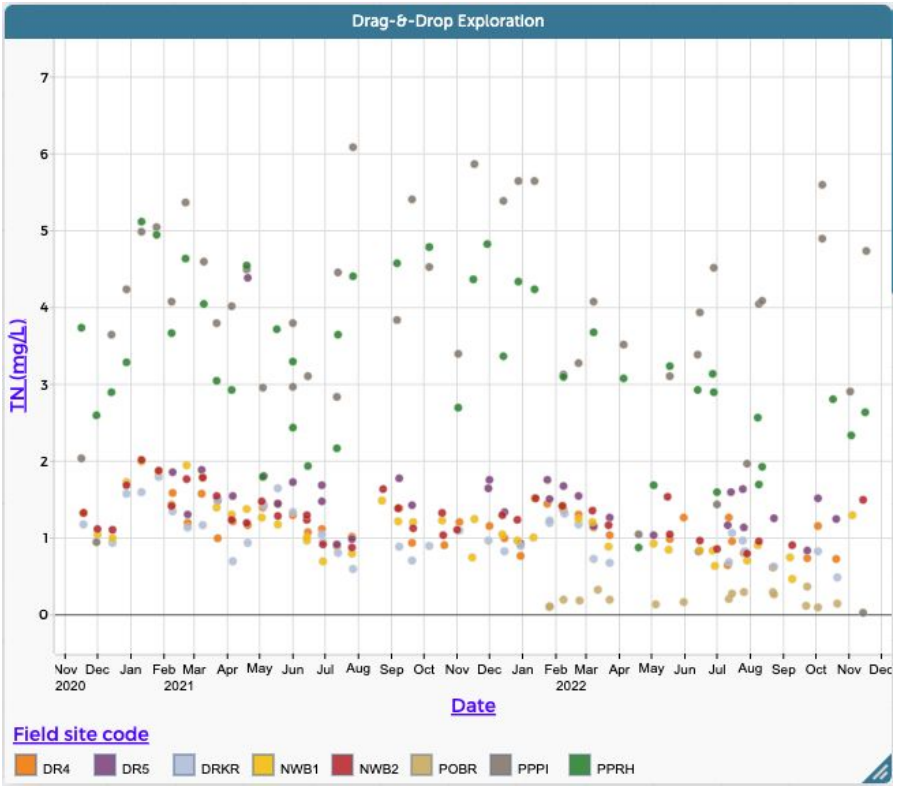
"I see that the data offers TN concentrations across all 8 research sites. I wonder '**Does TN vary significantly between the research sites?'** Now that I have my research question, I can 1) make my hypothesis, and 2) use the GCC to determine which graph will be the best fit for investigating my research question."

Four Watersheds Bi-Weekly Stream Chemistry 2020-2024 dataset	
❖ Background Information	The chemistry of streams can reveal a great deal of information about the precipitation and watershed feeding the stream, and about the stream system itself. Each chemical constituent can tell a different story. In the Urban Critical Zone project, scientists have been measuring the chemistry of streams in the Philadelphia PA, Baltimore MD, Washington DC and Raleigh NC metropolitan areas to understand how the geology, topography, ecology, urban structures and human activities influence water quality. Each city is on or near the Fall Zone - the transition between the hilly, rocky Piedmont region and flatter, sandy Coastal Plain region - along the Eastern seaboard. In this way, they are helping urban residents and managers make choices that will improve the quality of the water in their neighborhoods and leaving their regions to feed important coastal waterways such as the Delaware, Chesapeake and North Carolina coastal bays.
❖ Data Collection Site and Methods:	Stream water samples were collected bi-weekly (every two weeks) from February 2021-March 2024 in four watersheds as part of the Urban Critical Zone research project. Three of the watersheds were sampled at two or more locations nested within one another. Specifically, from North to South: Pennypack Creek watershed starts in suburban Philadelphia and runs through the Northeast part of the City before flowing into the Delaware River. A sewage treatment plant discharges into the stream. Sampling is done at two USGS gaging stations, one midway through the watershed at Pine Hill (PPPI) and the other near the base of the watershed at Lower Rhawn Rd (PPRH). Pond Branch watershed is mostly forested and is northwest of Baltimore in Oregon Ridge County Park. It is a first order watershed sampled a just one location near its base (POBR). Dead Run watershed on the west side of Baltimore, MD. The sites are "nested", starting with Dead Run 5 (DR5), a small headwater sub-watershed that feeds into Dead Run 4 (DR4), which in turn feeds into the sampling site at the base of the whole watershed where the stream flows into the Gwynns Falls Creek, Dead Run at Franklintown (DRKR). Northwest Branch Anacostia Creek watershed in suburban Maryland and Washington DC. It is sampled near this midpoint of the watershed (NWB1) and near its based (NWB2). Water samples are collected and then analyzed for the concentration a number of chemical constituents using standard analytical procedures.
❖ Source of Dataset:	This data was collected by Urban Critical Zone researchers, scientists, and support team. The data were then analyzed and provided for use by the Urban Critical Zone Education & Outreach team by Dr. Joel Moore, Towson University
❖ Dataset comments:	1. There are many missing values in this very large dataset.
❖ Dataset Variables	
Field Site Code	Site identification code: DRKR - Dead Run at Franklintown. This is at the base of the entire Dead Run watershed, including DR5 and DR4 DR4 - A sub-watershed in the south and west part of the Dead Run watershed that receives water from DR5, and flows into DRKR. DR5 - A sub-watershed in the southwest corner of the Dead Run watershed. It feeds into DR4. NWB1 - Northwest Branch Anacostia River upper portion of the watershed, flowing into NWB2. NWB2 - Northwest Branch Anacostia River lower (whole watershed) portion, receiving water from NWB1. POBR - Pond Branch PPPI - Pennypack Creek at Pine Road, midway through watershed, flowing into PPRH. PPRH - Pennypack Creek at Lower Rhawn Road, near the base of the watershed, receiving water from PPPI.
Date	Date the sample was collected, in format MONTH-DAY-YEAR
Time	Time stream sample collected (Eastern STANDARD Time)
Alkalinity as HCO3 (mg/L)	
TN (mg/L)	Total nitrogen, includes dissolve (nitrate, ammonium, urea, etc.) and particulate N
TOC (mg/L)	Total organic carbon (includes DOC and particular C).
TDN (mg/L)	Total dissolved nitrogen
DOC (mg/L)	Dissolved organic carbon (excludes particulate organic carbon)
Cl (mg/L)	Chloride, -1 charge
Nitrite as NO2 (mg/L)	Nitrite (note, nitrite is harmful to people)
Nitrate-N (mg/L)	Nitrate, expressed as the mass of N only
SO4 (mg/L)	sulfate, -2 charge
Na (mg/L)	sodium, +1 charge
K (mg/L)	potassium, +1 charge
Mg (mg/L)	magnesium, +2 charge
Ca (mg/L)	calcium, +2 charge
Nitrate-NO3 (mg/L)	Nitrate concentration - this is used to calculate the [Nitrate-N (mg/L)] concentration; students who are interested in investigating the nitrate concentrations in the Fall Zone watersheds should focus on Nitrate-N (mg/L) as their variable.

Data Visualization Option #1: Graph Choice Chart (GCC)

Alternatively, a student who is exploring the dataset using a drag-and-drop data analysis platform may come across a data visualization like the one shown here. The formulation of this student's research question may look like this:

"Weird! Both of the Pennypack watershed sites have significantly higher TN concentration than the rest of the research sites. I think my research question will be '**Why are the TN concentrations significantly higher at the Pennypack research sites, compared to the others?**'"



Objective #2: Interpreting and Analyzing the Data

Now that we have a data visualization, what do we do with it?

Students can sometimes struggle with the task of interpreting and analyzing data. We've prepared a check-list of data unpacking, interpreting, and analyzing questions to help guide your students.

Step #1: Unpack the Data

The students need to take time to get to know the data, to answer the question "*What information do I have here?*" With the information provided in the About the Data tab, students should answer the following questions:

Where is the data being collected?

Who is collecting the data?

Why is this data being collected?

What are the independent variables? In what units are they being measured?

What are the dependent variables? In what units are they being measured?

If these questions seem familiar - you're right! These are the same questions that students using the GCC method will ask themselves. If your students have already completed these tasks, they can move on to Step #2: Interpret the Data.

Step #2: Interpret the Data

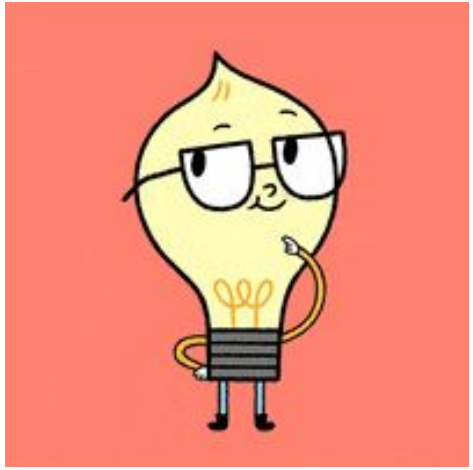
Now that students have gotten to know their datasets a little better, they can start to make observations about what the data visualization has to offer. In this step, students want to answer the question "*What do I see in the graph?*"

- What are the ranges of the data in our visualization?
- What types of outliers exist?
- How are the variables changing?
- What patterns/ trends do we see in the data?

A similar procedure is followed in the Identify and Interpret (I^2) Strategy.

Step #3: Analyze the Data

Now students are ready to start making connections between the information presented in their data visualization and the scientific phenomenon behind their observations, in order to explain the observed patterns, trends, and comparisons. In this final step, students want to answer the question "*What does it all mean?*"



Analyzing the Data

This step tends to be the most difficult for students, as it involves making connections from prior learning to their datasets, and often seeking out and utilizing additional resources to support their explanations.

As part of the UCZ Data Jam Competition judging rubric, students are expected to support their scientific explanations with at least two additional, credible sources beyond the information provided in with the dataset.

We recommend providing students with additional resources and support to guide them through their research.

Is that online source credible?

It is important to remind your students that on the internet, anyone anywhere is allowed to post “information”! How to spot a credible source and reliable information is a must-know for any good science report. The Source Credibility Guide is a great resource for students searching online for science to support their Data findings.

The Source Credibility Guide

Origin

From where was the source retrieved?

- For Digital: What is the domain?
  - » com, .org, .gov, .net?
- For Print: Who is the publisher?
  - » Is the source printed by a well-known publishing press, or a university press, or other?
- Is the source primary or secondary?

Author

Who is the author of the source?

- How qualified is the author to write on this topic?
- Is the author sponsored by an organization?

Purpose

What is the purpose of the source?

- Who is the intended audience?
- Why was the source written?
  - » Inform? Opinion? Entertain? Persuade? Sell?

Perspective

From what perspective is the source written?

- What is the author’s point of view?
  - » What is the tone or voice of the writing?
  - » Is the writing biased/unbiased? Does there seem to be an agenda?

Academic

How scholarly is the source?

- Is the content supported by evidence?
  - » Are there references? Does the author cite credible sources?
- Is the source peer reviewed?

Relevance

How relevant is the source?

- When was the content published?
  - » How current is the source and/or when was it last updated?
- Does the source fit the needs of the assignment?
  - » Does the assignment require an overview, or something specific?
  - » Does the assignment require primary sources?

The Source Credibility Guide

Objective #3: Communicating the story

Your students are now ready to tackle the written report!

We have provided educators with additional tools that they can use with their students to help them plan and execute their Data Jam data analysis and report.

Additional resources

These worksheets and lessons may help students work through understanding Data Jam, getting to know their datasets, and graphing their data.

Data support

- **Getting to Know your Dataset**: This student-facing worksheet guides a student through getting to know their dataset, from prior knowledge to understanding the metadata to graphing the data and analyzing the graph.
- **Graph Choice Chart**: This paper from the November 2014 issue of The Science Teacher is a great teacher-facing resource, which includes a chart that helps students choose the best type of graph for their particular science question.
- **Student Planning Sheet**: This planning tool can help your classes break Data Jam down into small digestible steps with concrete deadlines.

Urban Critical Zone Data Jam Competition

Getting Started: UCZ Advisor Resources

UCZ Data Jam: Essentials & Registration

UCZ Datasets

UCZ Data Jam: Important Documents

Data Jam FAQs

This one is good, too!

bit.ly/UCZDataJam

Coming Up!

- **Webinar #3 - Effective Communication of a Data Story (March 20)**
  - Transforming the results of data interpretation and analysis into creative expressions.
  - Keeping the data story at the forefront of the creative interpretation.
- **UCZ Data Jam Competition Registration - April 4, 2025**
- **UCZ Data Jam Competition Project Submission - ~~May 6, 2025~~**

UPDATE! Shift in deadline under consideration due to the start of AP Exams. Decision TBA soon.

Thank you!

Angela Hood

Education Program Leader - Cary Institute of Ecosystem Studies

hooda@caryinstitute.org

UCZ Education Email List↓↓↓

bit.ly/3Ww31uU

Cary Institute of Ecosystem Studies

Critical Zone Network
Urban Cluster

EAR-2012344