



Cary Institute
of Ecosystem Studies

Hudson River and Baltimore Data Jams: Creatively Engaging Students with Large Datasets

Cornelia Harris, Alan Berkowitz, Bess
Caplan, Shelley Doster

ESA Meeting – August 11, 2015

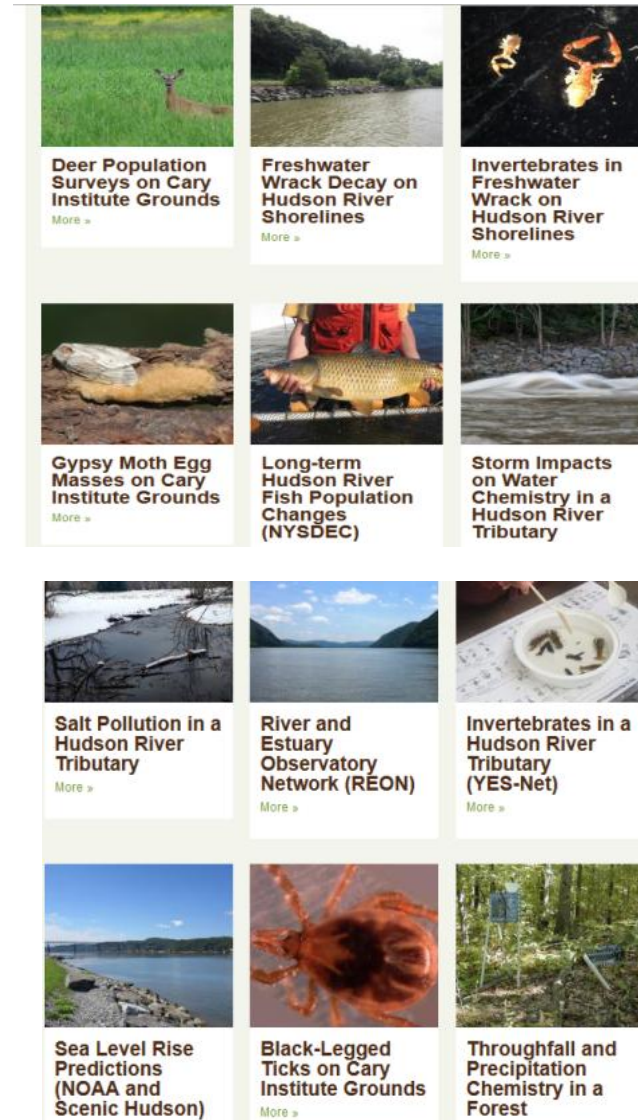
What is a Data Jam?

- A Data Jam is a competition where students are challenged to creatively tell a data-based story (using expert data sets)
- Modeled after the Desert Data Jam at Asombro Institute (started their competition in 2011)



Data Jam Overview

- There are many long-term, large ecological data sets which are intimidating and often confusing for the non-scientific audience (Bestlemeyer, et al. 2015)
- Students choose a data set that is interesting to them and come up with a creative way to share their understanding with others
 - Baltimore, as an LTER site, can use the EcoTrends website
 - In the Hudson Valley, we created a library of data sets (we currently have 42)





Why STEAM?



- Participating in the arts improves spatial-temporal reasoning, creative thinking, originality and abstract thought (Minton, 2002)
- Making visualizations is central to scientific thinking – through drawing, students make their thinking explicit and specific (Schwartz, 1995)
- Publicly sharing representations helps students learn by critiquing their own work and others' (Linn et al, 2000)
- Drawing, as opposed to writing summaries or providing oral explanations, can help students organize their knowledge in new ways and integrate new understandings (Chi et al., 1989, Kombartzky et al, 2010, Ainsworth et al, 2003)

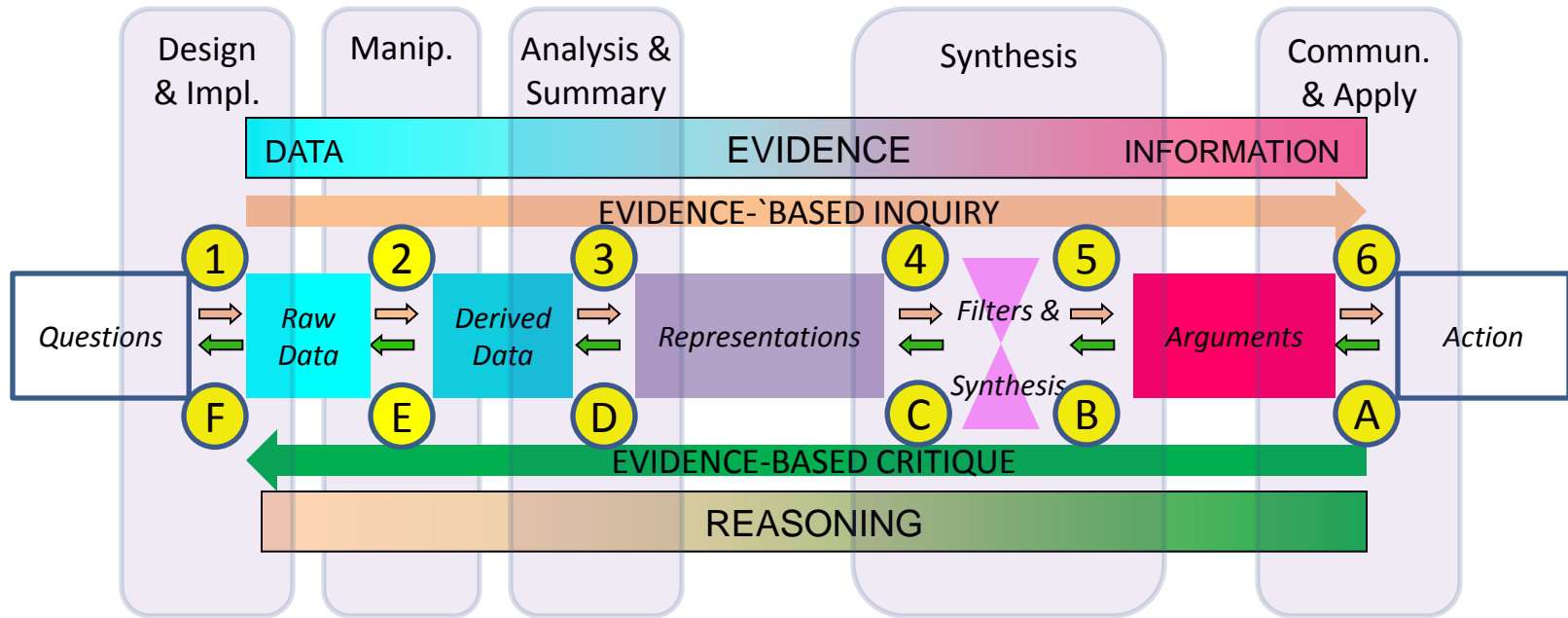
Our specific questions:

1. Does a Data Jam motivate students to work with secondary data?
2. Does participating in a Data Jam increase students' ecological knowledge and interest?
3. What can a Data Jam tell us about students' data literacy skills, creativity, and interest in ecology?

Our Sources of Evidence:

1. Data Jam Projects
2. Student Surveys : voluntary feedback from participants
3. Student Assessments : students who did and did not participate
4. Informal interviews

A Framework of Critique and Inquiry Practices



Inquiry Practices:

1. Primary research – question/hypotheses, study design, data collection
2. Data manipulation – descriptive statistics, sub-setting data, indices
3. Summarizing results – graph, diagrams, tables, bottom line, statistical tests
4. Filtering results – selecting salient, relevant, and reliable results
5. Synthesizing – combining, integrating, meta-analysis
6. Communicating and recommending

Critique Practices:

- A. Media Literacy – critically evaluating media claims based on the argument
- B. Evaluating the synthesis in the argument
- C. Evaluating the filtering of evidence used in the argument
- D. Critiquing the representation – right graph or summary? Adequate information about it?
- E. Evaluating manipulated data – right descriptive statistics, sub-setting data of, indices?
- F. Evaluating questions/hypotheses, study design, data collection

Data Jam Entries Include:

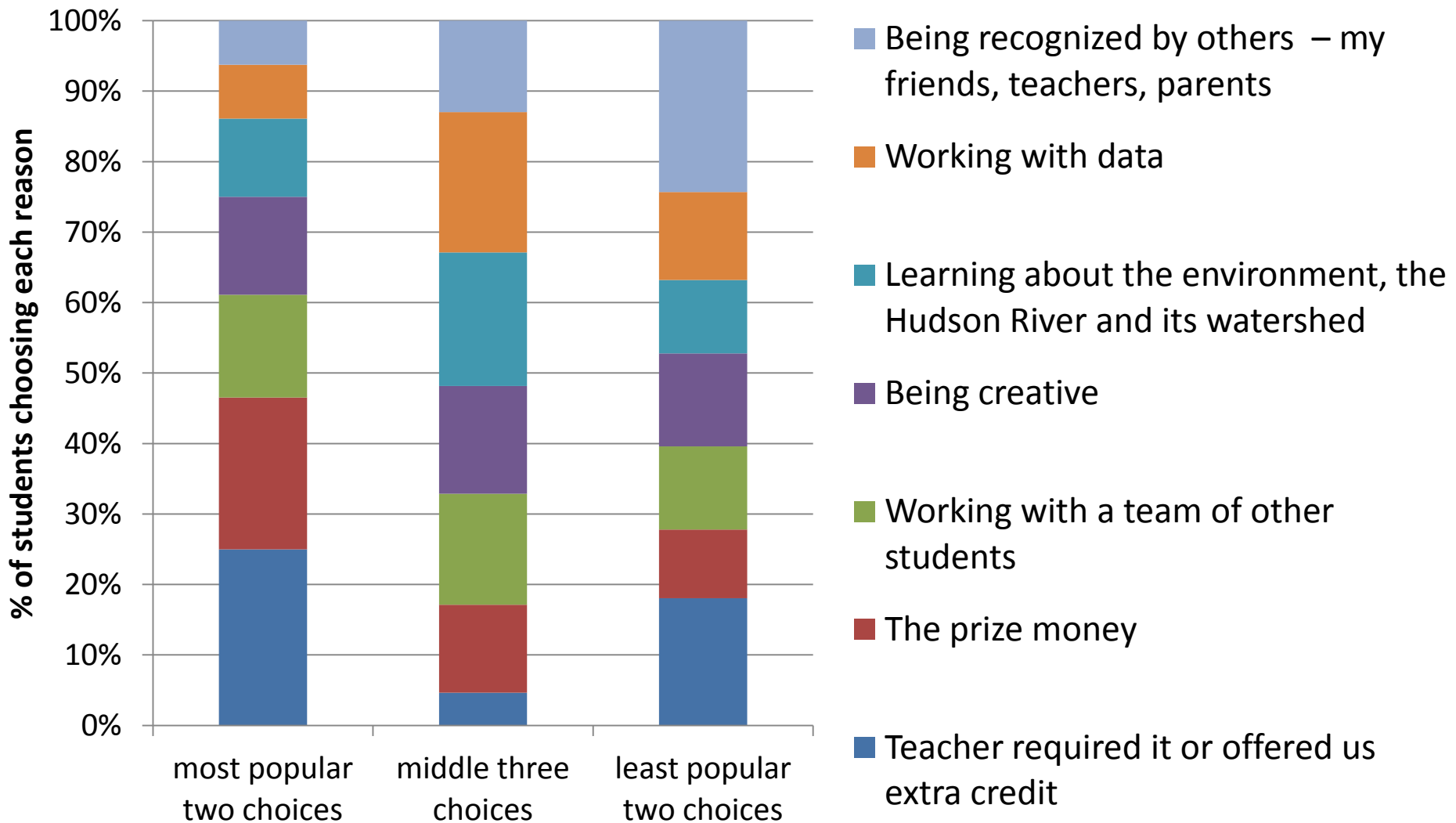
- A report or poster
 - Includes metadata, references, logistics
 - 30 points in rubric
- A creative expression piece
 - Visual arts, videos, sculptures, songs, poems, stories, etc.
 - 30 points in rubric
- Scientific merit (included in poster or report)
 - Data summary, interpretation, representation, synthesis, interpretation and reasoning
 - 40 points in rubric

1. Does a Data Jam motivate students to work with secondary data?

	Baltimore		Hudson Valley	
	2014	2015	2014	2015
Number of student projects	10	31	26	100
Number of students participating	25	76	89	240



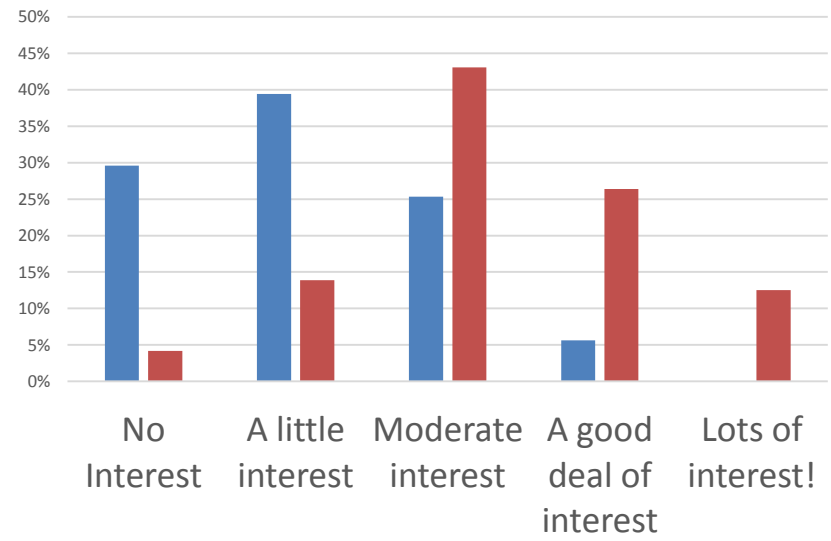
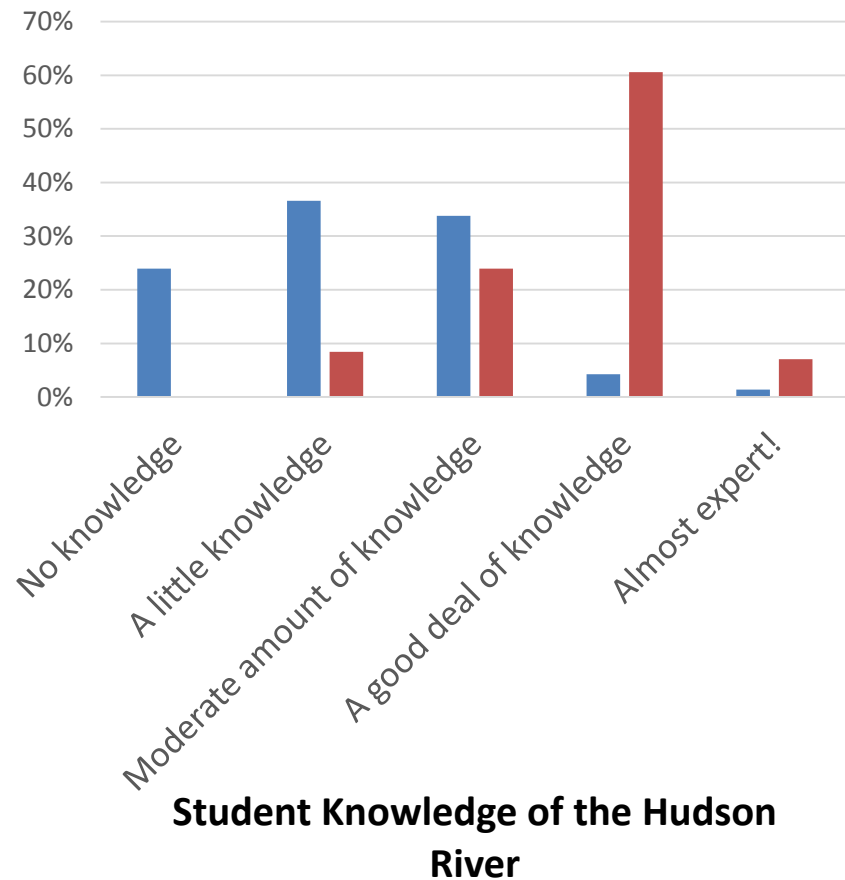
Why did students participate?



Why did students participate?

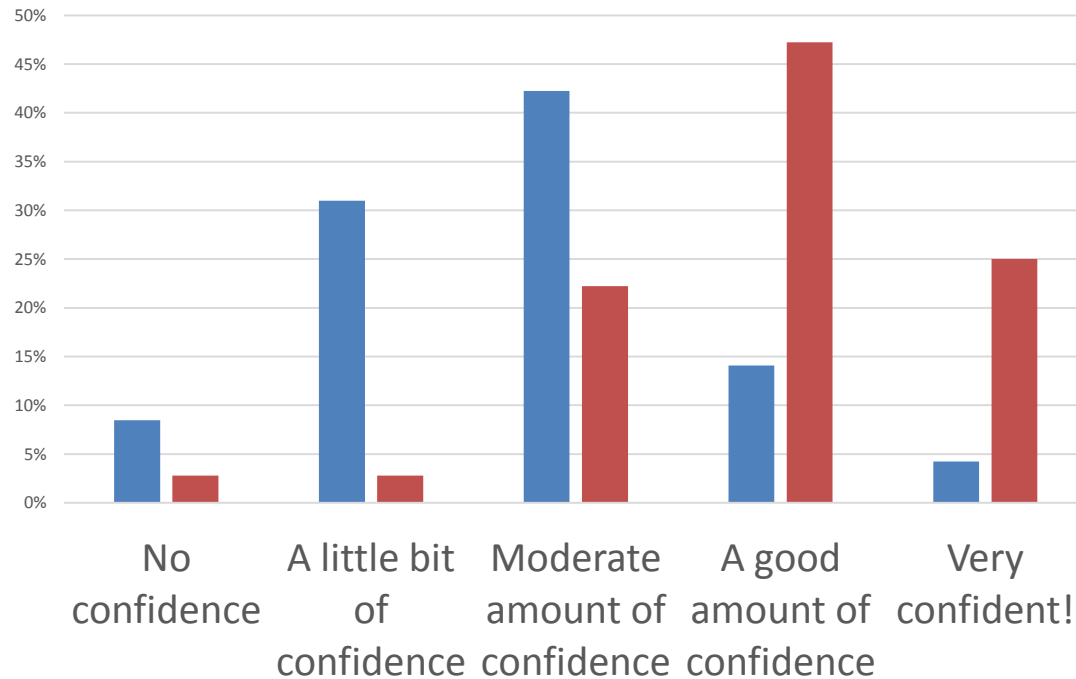
- Most important reason:
 - teacher required it or offered extra credit (50%)
 - although this was least important reason for many (36%)
- Second most important reason:
 - the prize money (43%)
- Least important reason:
 - Recognition by others
- Other reasons were of high or moderate import:
 - Data, Learning, Creativity, Group Work
- 73% would do it again next year

2. Does participating increase students' ecological knowledge and interest? *



Before After

2. Does participating increase students' ecological knowledge and interest?



Student confidence in working with data

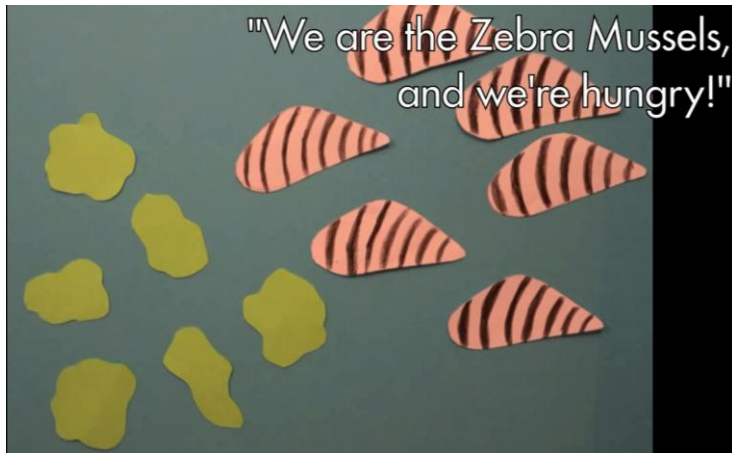
■ Before ■ After

2. Does participating increase students' ecological knowledge and interest?

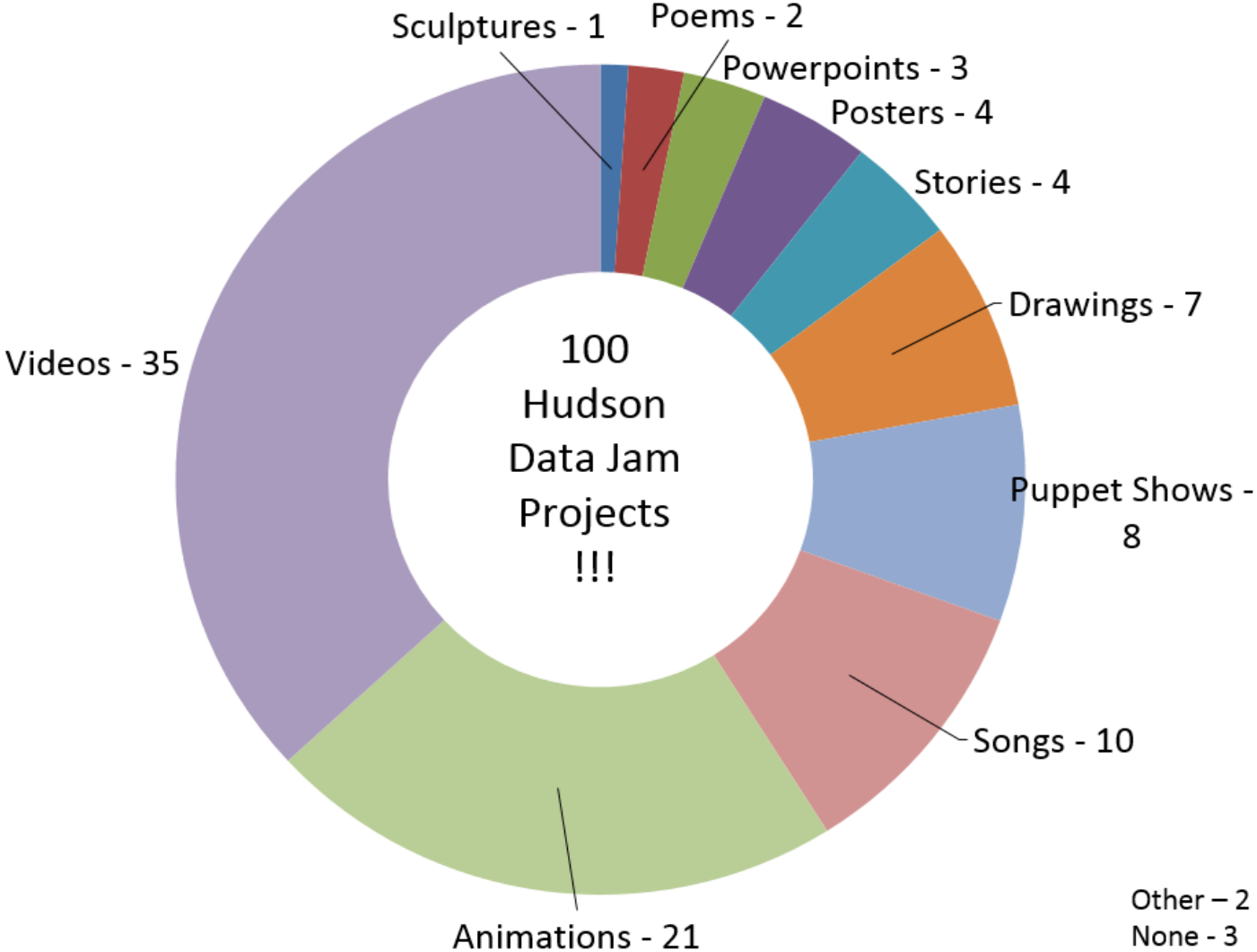
- Student knowledge and interest in the Hudson River and its watershed both increased in 2014 & 2015
- Student confidence in working with data increased in 2015, but not in 2014



3. What can a Data Jam tell us about students' data literacy skills, creativity, and interest in ecology?



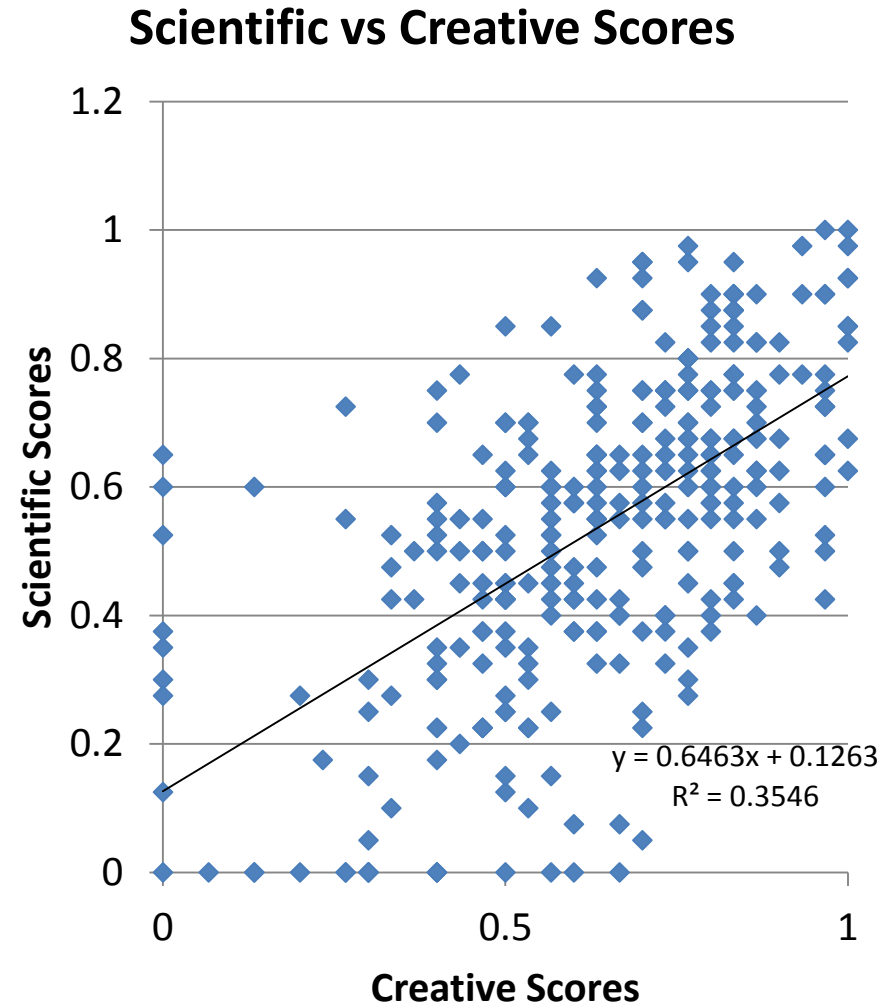
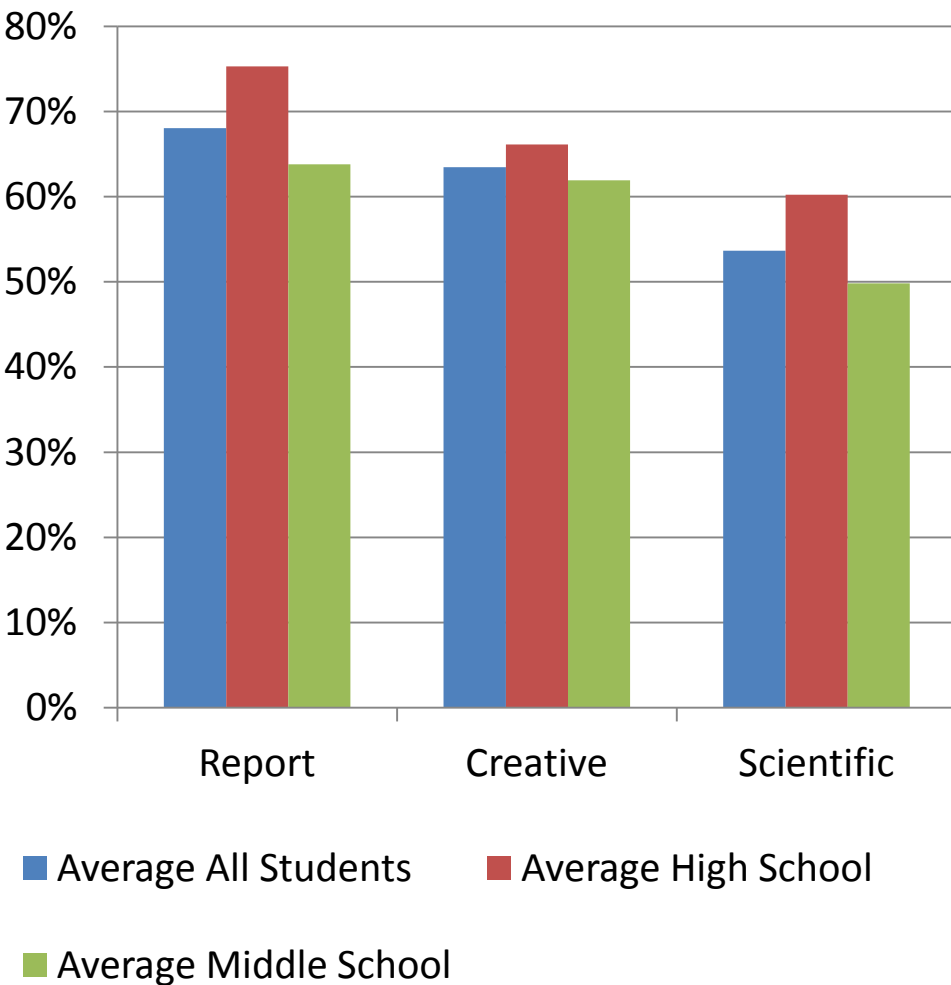
What kinds of creative projects did students do? (HV)



What ecology topics were students interested in?

	# Middle School Projects	# High School Projects
Macroinvertebrates	29 * (27 required)	2
Zebra mussels	4	8
Hudson River fish	5	4
Deer populations	5	4
Historic pollution	5	1
Salinity pollution	2	3
Dissolved oxygen	0	3
Hurricanes	2	3

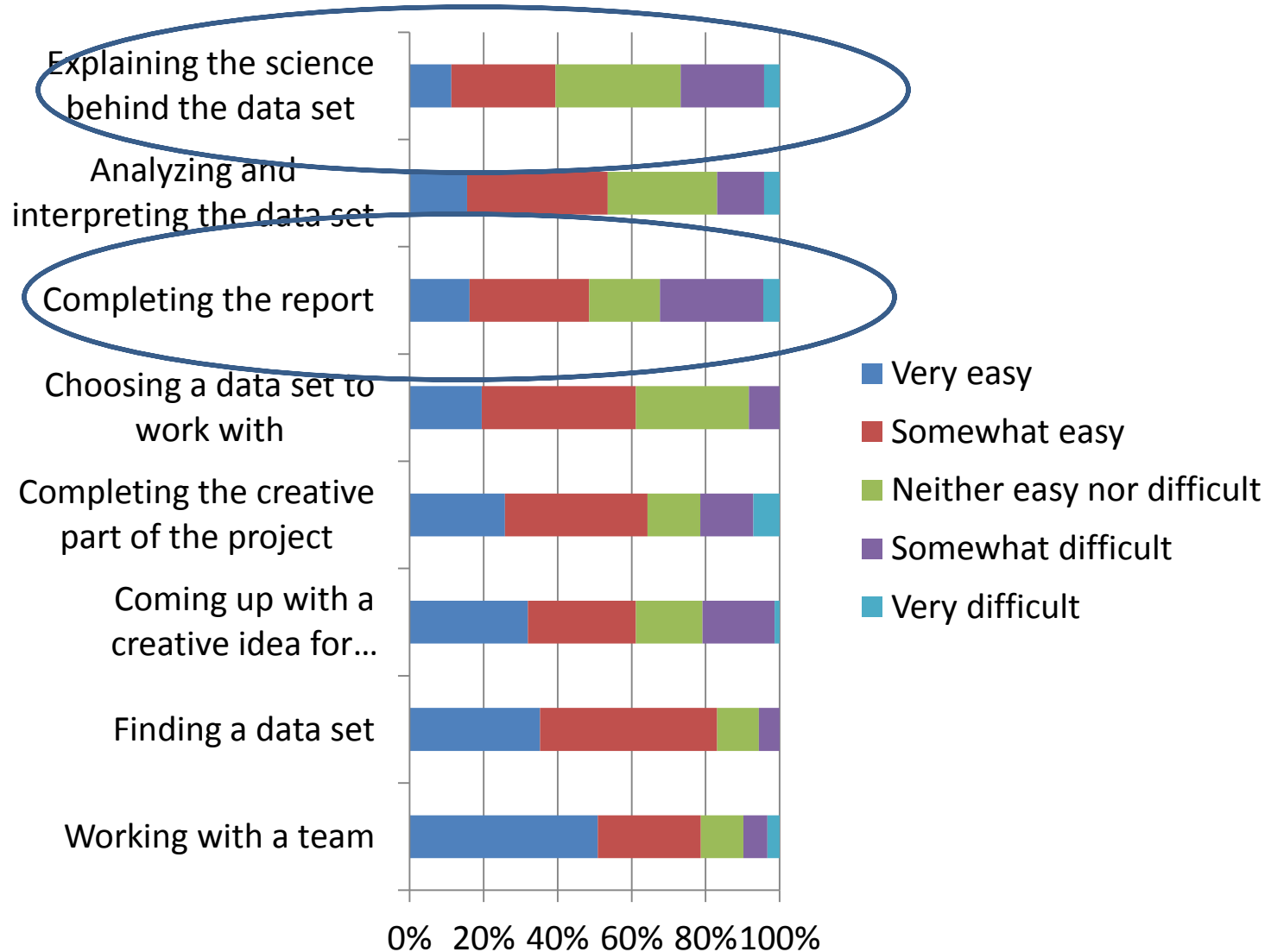
3. What can a Data Jam tell us about students' data literacy skills, creativity, and interest in ecology?

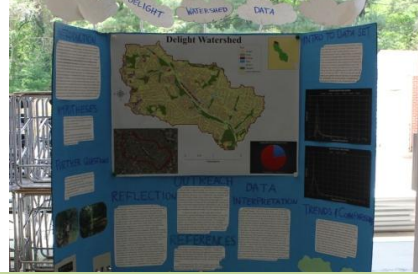


3. What can a Data Jam tell us about students' data literacy skills, creativity, and interest in ecology?

- High school students scored higher than middle school students in all components
- Scores for report > creativity > scientific merit
- Positive relationship between scores for creativity and science with much variability

What did students find difficult?





Demonstrated Data Literacy Skills

Baltimore		Hudson Valley	
2014	2015	2014	2015
80% of students made their own graphs	Winning & runner up middle school teams integrated different types of social science data	All middle school students used prepared graphs and looked at only one data set	Five middle school projects used more than one data set
Three groups compared water quality across two different streams	One high school group used STELLA to model stream inputs/outputs	Two student groups (both the top prize winners in the HS category) created their own graphs using provided data	Four high school teams used more than one dataset in their projects
	Two high school groups compared multiple provided data sets	One student group incorporated a large online data set (HRECOS)	Two high school student teams used outside data sets

Scientific Merit Scores – 2015 Hudson Data Jam Student Projects

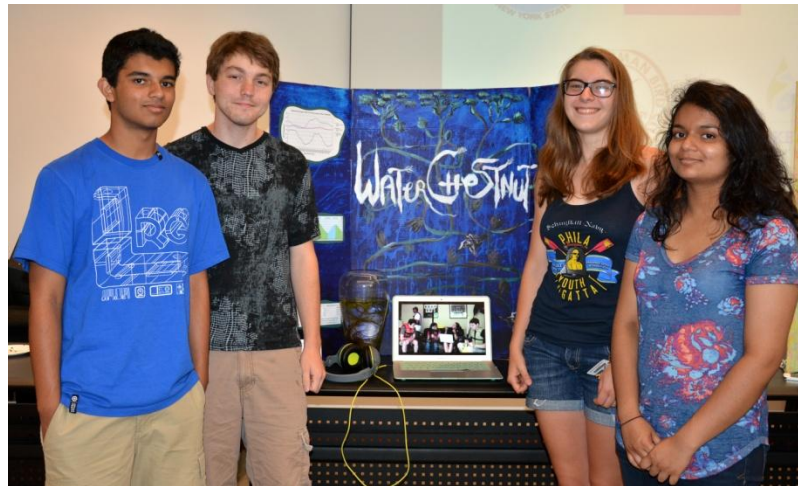
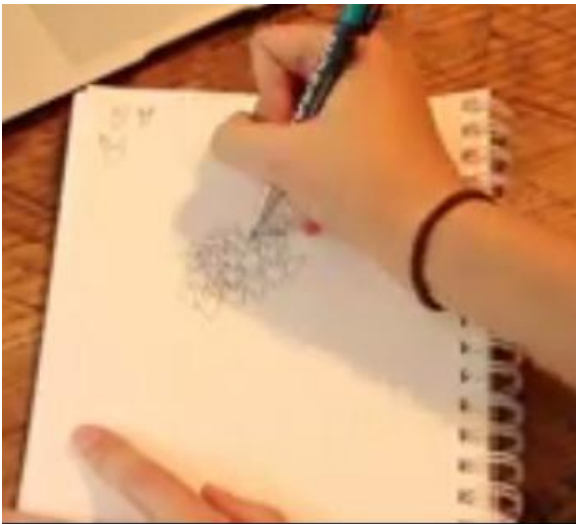
Scientific Component	Our criteria (summarized)	Average score
Trend or comparison	Project accurately describes the trend(s) or comparisons of the data set	63%
Representation	Project includes a representation that correctly displays the data used.	62%
Data interpretation	Project provides the reasoning about the topic to explain the trend or comparison that was discovered	60%
New hypotheses & questions	Project gives at least two additional ideas about future scientific research.	57%
Describing data	Projects explains the data accurately using basic descriptive statistics (mean, standard deviation, t-tests, or other appropriate measures of significance) and describes variability (range, variance).	42%
Data set complexity	Project synthesizes additional data, either from outside sources or provided data sets	40%

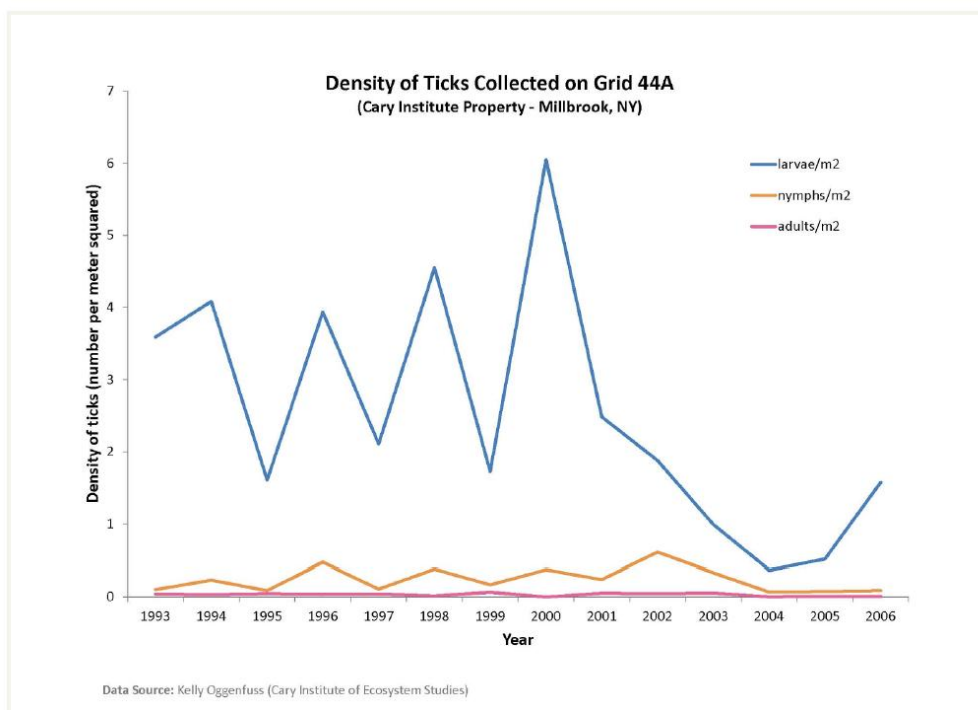
Conclusions

- Students enjoy doing the Data Jam for a variety of reasons
- They believe that they learn something
- Students *think* they are doing a good job with the science but are not prepared, nor supported, to really evaluate and analyze data
- Students need more content knowledge and experience to be able to
 - Do their own analyses of data
 - Synthesize multiple sources of data

Example projects

- A project that didn't incorporate multiple data sets (but was still a lot of fun!)
- A project that incorporated additional data and/or used data in a new way





Sampling is performed
at least once
during the peak time of each stage
It takes place
on days,
when the ground is dry and there's a nice clear sky

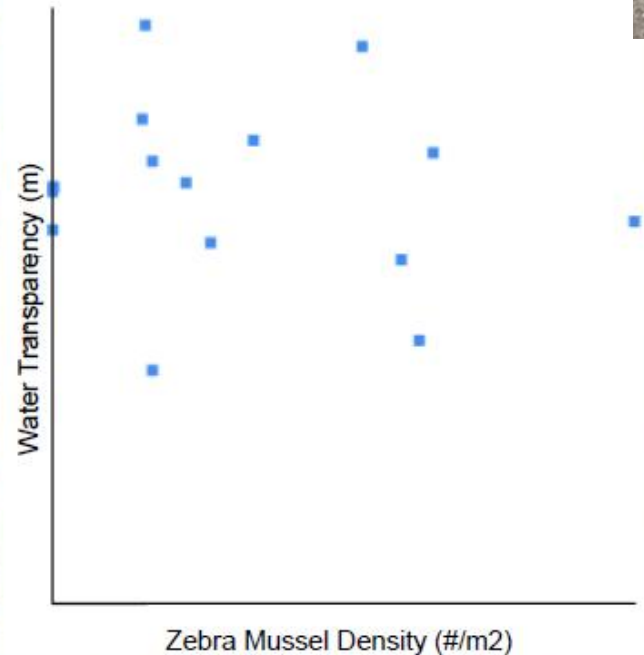
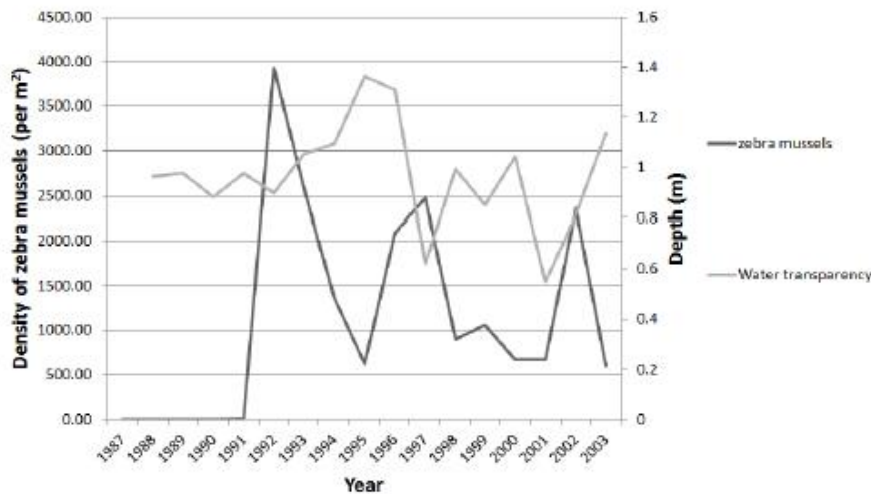
[Videos](#)

Data Trends



r-value: $-.11$
N: 16
p-value: $.68$

Zebra Mussels and Water Transparency in the Hudson River



Data obtained from scientists at the Cary Institute of Ecosystem Studies

- We observed that zebra mussel density and water transparency did **not** have a strong correlation with each other
- **Water transparency** appears to be generally **higher** in the **absence** of **zebra mussels**

www.caryinstitute.org/data-jam



2015 Hudson Data Jam Competition

Making data "sing" through creative expression

Now in its second year, the Hudson Data Jam Competition challenges students to creatively tell stories for a general audience using data from the Hudson River watershed.

 **for Students**

Youth Camps & Programs

Graduate Opportunities

REU Program

Data Jam Competition

» 2014 Winning Projects

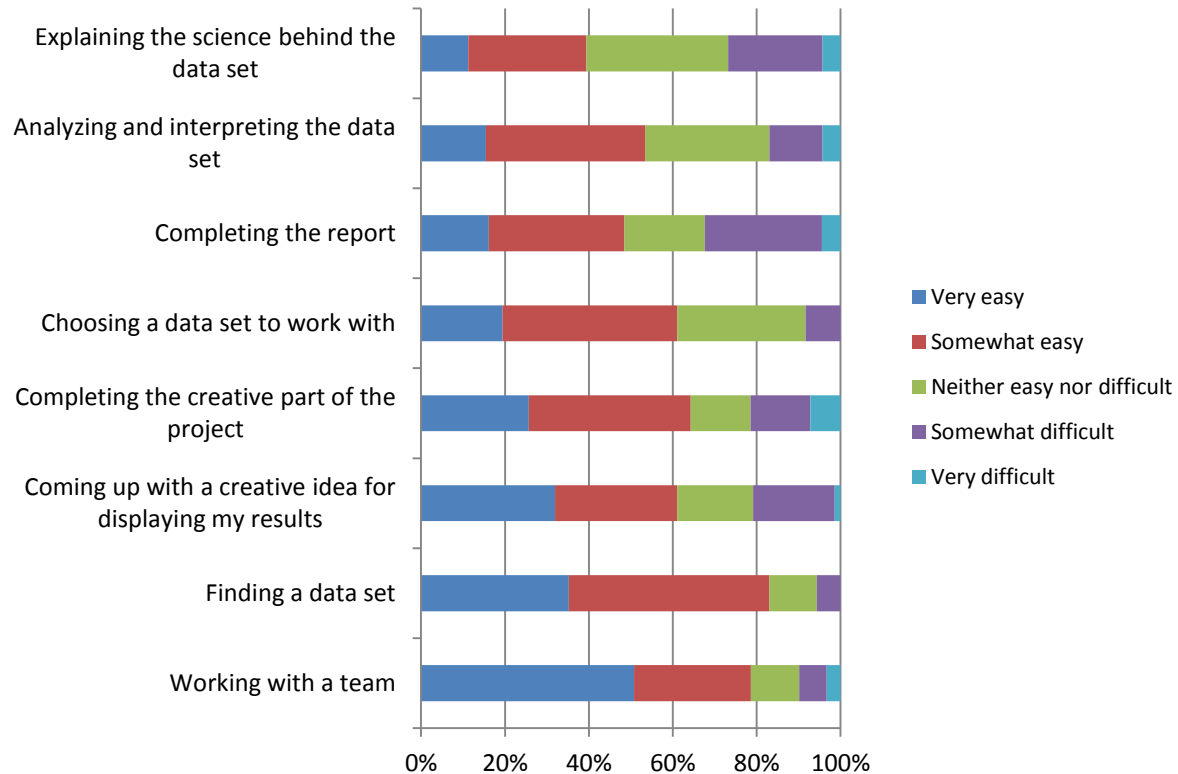
» For Students

» For Teachers

» Rules & Regulations



What did students find difficult?



- Most challenging: Completing the creative part of the project
- “Somewhat difficult”: Completing the report