A Scaffolded Approach to Data Literacy Skills in 2YC Students using Authentic Data from OOI and other Sources

Denise L. Bristol1, Anna S. Pfeiffer-Herbert2, & Jessica L. Olney3

1Hillsborough Community College, Tampa, Florida and 2Stockton University, Galloway, New Jersey  Contact: dbristol@hcfl.edu

What are students’ perceptions of, & skills with large, real-world oceanographic data visualizations in 2YC students?

Introduction:
Up to 80% of 2YC students have math anxiety & a survey of undergraduate biology students found that 60-80% had not previously worked with authentic data.1, 3 In addition, students may struggle to work with data and visualizations due to limited previous experiences and exposure.2, 5 Data literacy skills are needed by undergraduates to become fluent in working with authentic data and align with expected on-the-job skills.

The Ocean Data Labs lab manual: Exploring the Ocean with OOI Data, was created to provide high quality open access (OER) ocean data literacy activities that allow undergraduates to use real-world oceanographic data from the OOI in accessible ways, while being easy for professors to integrate into their teaching. It supports undergraduate students by enhancing graph interpretation skills, recognition of patterns, critical thinking skills and provide opportunity to work with large, professionally collected data sets. Within the collection Lab 2: Building Data Skills - The Display of Oceanographic Data6 orientds students to different types of data using a scaffolded learning cycle approach. The activities are carefully crafted to teach students how to orient themselves to various oceanographic data visualizations, recognize trend patterns and correlations, interpret ‘messy’ data and apply their knowledge.

Methods:
A pilot study was conducted to determine if the lab activity changed student perception and skills working with data visualizations. Pre and post surveys were administered to introductory 2YC oceanography students in asynchronous online courses during the Fall 2023 and Spring 2024 before and after completing the Ocean Data Lab 2: Building Data Skills - The Display of Oceanographic Data. Students were asked about each data visualization type: simple scatter plot, time series, bubble chart, messy scatter plot, vertical profile.2 Questions were about: • Opinions and feelings (Table 1) • Familiarity (Table 2) • identifying data points • map/chart data interpretation • identifying patterns in data to determine if students’ perceptions and familiarity changed after the activity and if there was increased understanding of the data visualizations.

Results and Discussion:
Sixty-four (n=64) students completed the pre & post surveys across 5 different classes after 16 students were removed from the survey because they did not complete one of the surveys. A paired proportion test was used to determine change in pre & post feelings and correct answers to data interpretation questions.

Student Perceptions: The pre-survey indicated that students were most familiar with a simple scatter plot and least familiar with vertical sections (p<). Percentage of students ‘not familiar at all’ with the other data visualizations ranged from 12 to 36%. Post survey indicated that students had increased familiarity with all graphics and 78% of the students reported they were ‘somewhat’ or ‘much more’ confident reading & interpreting data visualizations while 8% became less ‘f’. There was statistical significance for each question for total reduction of negative and increase in positive perceptions between pre and post, except for the positive expressions for the vertical profile & all neutral. Students’ familiarity with graphics increased across all types as did their comfort and confidence with reading and interpreting the data after completing the activity (p<). Student's overall negative perceptions decreased for every graphic, most notably they were less confused, challenged, anxious and scared.

Data Interpretation: As expected, introductory students were most familiar (87% pre; 91% post) with a simple linear scatter plot and were correct reading (90%) and identifying the (93%) post for it. Students also did very well on the three bubble chart questions both pre & post ranging from 81-93.8%. Statistically significant change in correct answers were: the question on the dual Y-axis time series for maximum temperature increased from 55% to 76.6%; recognition of trend type on the messy scatter plot increased from 79.7% to 98.4% and the dual Y-axis time series scatter plot and were correct reading (90%) and identifying the trend (93% post) for it. Students also did very well on the three bubble chart questions both pre & post ranging from 81-93.8%. Statistically significant change in correct answers were: the question on the dual Y-axis time series for maximum temperature increased from 55% to 76.6%; recognition of trend type on the messy scatter plot increased from 79.7% to 98.4% and the dual Y-axis time series scatter plot and were correct reading (90%) and identifying the trend (93% post) for it. Students also did very well on the three bubble chart questions both pre & post ranging from 81-93.8%. Statistically significant change in correct answers were: the question on the dual Y-axis time series for maximum temperature increased from 55% to 76.6%; recognition of trend type on the messy scatter plot increased from 79.7% to 98.4% and the dual Y-axis time series scatter plot and were correct reading (90%) and identifying the trend (93% post) for it. Students also did very well on the three bubble chart questions both pre & post ranging from 81-93.8%. Statistically significant change in correct answers were: the question on the dual Y-axis time series for maximum temperature increased from 55% to 76.6%; recognition of trend type on the messy scatter plot increased from 79.7% to 98.4% and the dual Y-axis time series scatter plot and were correct reading (90%) and identifying the trend (93% post) for it.