Identifying Core Formal Assessment Competencies and Informal Learning Outcomes of Data Science and Data Literacy

Yale Quan: University of Washington, College of Education
Marc Sager: Southern Methodist University, Simmons School of Education and Human Development

Introduction

Over the past decade, the field of Data Science (DS) has enjoyed explosive growth as both a career path (Bureau of Labor Statistics, 2024) as well as a program of study at the undergraduate and graduate levels (National Academies of Sciences, Engineering, and Medicine, 2018). In response, formal and informal development of curriculum in the areas of Data Science education and Data Literacy (DL) education increased (Jiang et al., 2022; Shreiner, 2018).

Definitions

Data Literacy encompasses comprehension of the data production process, the capacity to craft pertinent inquiries addressable with data, the skill to choose and apply suitable analyses, and the ability to draw data-driven inferences, and arguments while considering the contextual factors influencing data generation (Louie, 2022; Stornaiuolo, 2020). Further, critical DL revolves around comprehending the potential uses of data, its impact on individuals and society, and its role in shaping our world (Louie, 2022; Stornaiuolo, 2020).

Data Science is an interdisciplinary field in which data are used to investigate and answer consequential questions using a collection of scientific methods from statistics, computer science, programming, and mathematics (Burrill & Pfannkuch, 2024; Pennington et al., 2020).

Research Questions

Assessments that formally measure students' mastery of these subjects provide a snapshot view within a specific school or classroom. However, comparing these measures across multiple schools and states would be challenging, if not impossible. Additionally, DS and DL can be taught and assessed both in formal classroom and out-of-school learning environments. This naturally leads us to two questions:

RQ1)

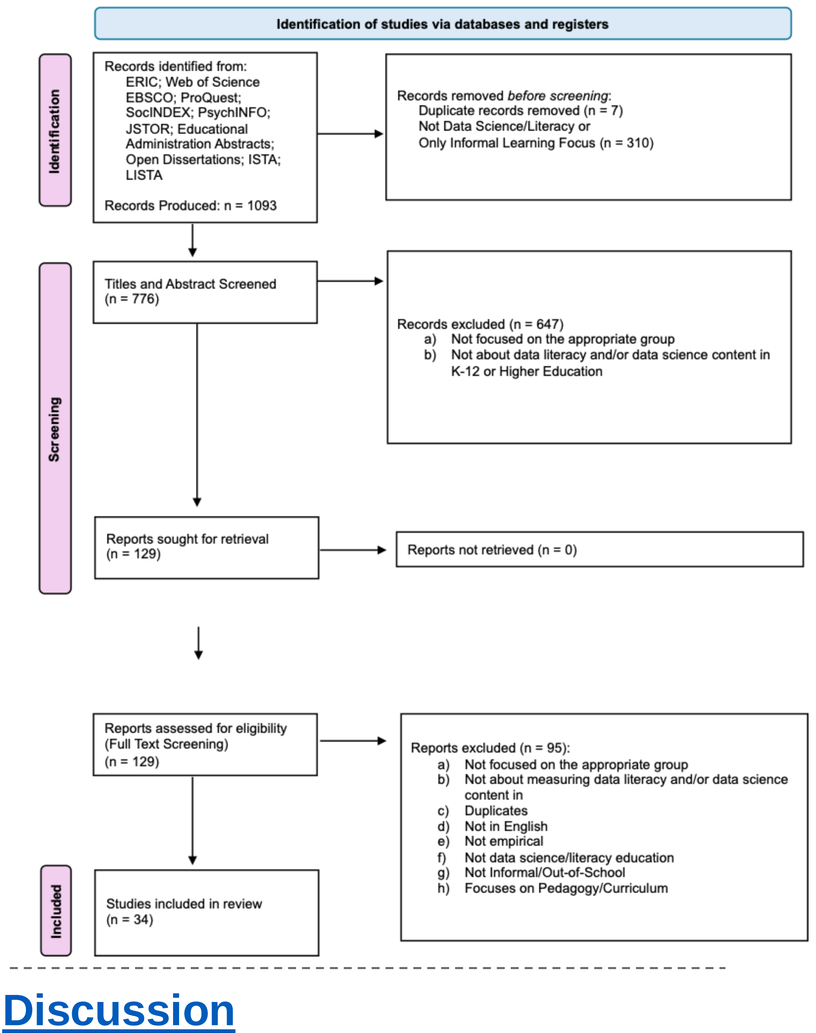
- What core competencies are associated with existing DS and DL formal assessment and curriculum?

RQ2)

- What are the primary learning outcomes associated with DS in informal learning environments?

Methodology

To answer out research questions we performed a systematic literature review using the PRISMA method.



Discussion

To effectively integrate DS education, practitioners should consider best practices highlighted in the review. These include

- Integrating connected learning principles to make data exploration relevant to youths' lives.
- Designing accessible tools to support hands-on activities, and involving families and communities to create a supportive learning environment
- Emphasizing ethics and personalization
- Ensuring that activities are responsible, relevant, and accessible to all learners

To effectively measure DS and DL competency we recommend the assessments should:

- Be positioned within a specific content area and focus on measuring competency in targeted sub-domains.
- Focus on the development of a scientific approach

and,

- Require the use of data to explore and answer research questions. Along with emphasizing the interpretation and communication of results both visually and numerically

By adopting these strategies, both formal and informal learning environments can foster an engaging and impactful learning experience, as they prepare youth to navigate and contribute to a data-driven world.

Detailed Results

RQ1)

- DL and DS are complex constructs measured indirectly through related components
- Formal classroom DL assessments focus on measuring competency in datasets and visualizations.
  - Datasets are measured through three sub-domains: data source quality, relevance to a specific content area, and result communication
  - Data visualizations are assessed by understanding graphical conventions, specific data representations, and data-related context.
  - An essential sub-domain of DL involves awareness of data privacy concerns and ethical considerations.
- DS competencies include:
  - Applying technology for data exploration and analysis, requiring knowledge of statistical concepts, methodologies, and relevant content areas
  - DS competency is interdisciplinary and requires understanding of both the methodology used to conduct the analysis and content knowledge relevant to the data

RQ2)

- Youth engagement and empowerment are key outcomes of DS in informal learning settings.
- Relating data to real-life contexts, such as daily life, community issues, and historical events enhances learner engagement and motivation.
- Hands-on activities like asking questions, producing, analyzing, and presenting data promote scientific inquiry and reinforce learning.
- Practical data skills development, including data wrangling and modeling, enables learners to clean, organize, and create meaningful narratives from data.
- Critical DL is emphasized, fostering inquiry, reasoning, comparisons, and critical evaluation of data to derive well-founded conclusions.

Acknowledgements

This research was conducted with support from Data Science 4 Everyone (DS4E). We would like to thank the DS4E staff and the other Summer 2024 Graduate Research Assistants for their support.

University of Washington; Measurement & Statistics, College of Education

Southern Methodist University; Simmons School of Education and Human Development