Paper ID #53905

Teaching contextual reasoning as a key skill in data science education

Cathryn Carson, University of California, Berkeley

Cathryn Carson is a historian of contemporary science and technology and scholar of STS (Science, Technology, and Society) at UC Berkeley. She has been active in interdisciplinary collaborations in undergraduate and graduate education, including nuclear engineering and data science, since the 2000s. She currently co-leads the Human Technology Futures group, a proposed department in the College of Computing, Data Science, and Society (CDSS), and chairs the Department of History. Trained as a historian of 20th-century physics, she co-chaired Berkeley's Data Science Education Rapid Action Team, which framed the broad-based, integrative vision its undergraduate data science curriculum; and chaired the Faculty Advisory Board that developed the blueprint for Berkeley's new College of CDSS. She is part of the teaching team for Data C104, Human Contexts and Ethics of Data, which is the largest humanities or social science class at UC Berkeley, at a scale of about 700 students each semester. Her current research is on disciplines and interdisciplinarity in contemporary universities and on the intellectual history of the philosophy of technology.

Lisa Yan, University of California, Berkeley

Lisa Yan is an Assistant Teaching Professor of Electrical Engineering and Computer Sciences at the University of California, Berkeley. She is a computer science and data science educator and researcher with expertise in data analysis of student learning in large courses, computing ethics education, and teaching assistant education. She received her Bachelor of Science degree in Electrical Engineering and Computer Sciences from UC Berkeley and her Master of Science and PhD degrees in Electrical Engineering from Stanford University.

Ari Edmundson, University of California, Berkeley

Ari Edmundson is a lecturer in the Department of History at the University of California, Berkeley. He teaches and develops Human Contexts and Ethics curriculum for data science majors and minors.

Joshua Grossman, University of California, Berkeley

Josh Grossman is assistant teaching professor of statistics and data science at UC Berkeley. He has conducted research on discrimination law, college admissions, pretrial detention, and policing. His current research focus is AI-assisted grading in data science courses. He is also interested in how computerized testing centers can be used to provide in-person, low-stakes, and high-frequency quizzing at-scale in large courses. Prior to his academic career, Josh worked as a product manager at IXL Learning, a K–12 education technology company. Josh completed his Ph.D. in Computational Social Science at Stanford University's Department of Management Science & Engineering (MS&E) and his bachelor's degree at Harvard University.

Teaching contextual reasoning as a key skill in data science education

Contextualization is a theme of recurring significance in engineering education. It is highly relevant to the emerging interdisciplinary field of data science/data analytics, which brings together computing, statistics, data systems engineering, and real-world applications. Data analytics and algorithmic decision-making systems have major impacts in our world today akin to other engineering practices. The models, algorithms, and other products created by data science practitioners derive from data and assumptions that inherently originate in particular contexts. Yet such tools abstract from those circumstances to create generalizable mechanisms and analyses that are meant to be cut loose from particularity and travel across contexts. In fact, this kind of abstraction is arguably intrinsic to technical training and specifically to computational and inferential thinking, which are defined as core skills in the data science education literature. Data science is thus a good setting in which to examine the kinds of work that "context" can do in technical education, including integrating key strategies into pedagogy, examining impact in students' learning, building scaffolded curriculum to achieve key goals toward robust and responsible professional practice, and fleshing out a substantial place for a form of ethics that itself goes beyond abstraction.

"Contextual reasoning," as used in this paper, is a cognitive skill that moves counter to abstracting from particular circumstances. In teaching data science, contextual reasoning centers skills of systematically analyzing where data comes from, how it is made, and what assumptions it may harbor; how the intermediate and end products of technical practice (including data and algorithms) respond to or carry traces of the human contexts from which they are drawn; and how the actual impact of data-driven solutions in the real world depends on how they end up embedded in concrete settings and contexts. As implemented in a data science curriculum that from the start has included "human contexts and ethics," contextual reasoning has proven to be essential in advanced data science courses that consider real-world problems. Developing contextual reasoning is an ongoing project within this curriculum. This paper provides a mid-project examination of how contextual reasoning serves technical students and how it can be taught as a core skill across an undergraduate curriculum in ways that are responsive to growing discussions about contextualization and sociotechnical integration as broader strategies in technical education.

1 Introduction

Recent scholarship in engineering education has highlighted the importance of contextualization as a core element of students' preparation for professional practice. Contextualization is intended to help students grasp key aspects of a problem situation, be alert to social situations and dynamics, and be responsible and ethical in engaging with them. In pedagogy that centers contextualization, students are supported and encouraged to recognize context, respect it, judge in alignment with it, and act skillfully and appropriately within it—several distinct aspects that come together in one term. Recent studies have explored some of the different meanings and functions of contextualization in technical education, and the ways that it intersects with other disciplines such as STS (Science, Technology, and Society). In some cases contextualization can mean giving students 'richer,' more concrete real-world cases to engage with. In other uses, contextualization can highlight design considerations, direct attention to social impact of technologies, or bring into view social structures that enable and constrain technical work. [1,2,3]. Context continues to be an inviting way for a range of educators to 'put the social (back) in' to technical work [6,7,8]. Like sociotechnical integration, which supports students in learning to see how social and technical aspects of real-world situations are intrinsically bound together in a way that both recalls STS and opens the door to questions of politics, ethics, and justice [20,21], contextualization blurs the boundary between—and intrinsically constructs interactions across—what is "content" and what is "context," and what is "technical" and what is "social." Contextualization and sociotechnical integration together present alternatives to decontextualized technical practice, with the goal of helping students orient themselves toward robust, responsible, and ethical real-world work [1,5,22,23,24].

Data science is a relatively new field whose disciplinary boundaries and practices are still fluid. Undergraduate programs in data science have been emerging for about a decade. Conventional definitions of data science often depict it as grounded in statistics and computing, the latter often including a component of data systems engineering, and widely oriented toward real-world application that is taken to involve "domain knowledge." While different programs emphasize different approaches—ranging from a focus on the cognitive skills of "computational" and "inferential thinking" to the embeddedness of each data science problem in a distinct domain of application requiring domain-specific knowledge—the field as a whole is constructed around the back-and-forth between abstraction and particularity. Students are taught foundations, principles, and techniques that are then meant to be applied to specific situations in order to understand the world and guide action in it [4,9]. To many data science educators, "context" will sound familiar: it can be taken to mean domain knowledge.

Beyond domain knowledge, however, undergraduate data science education can also be considered to include another form of context. Within our own institution and its undergraduate data science program, "Human Contexts and Ethics" (HCE) forms a crucial fourth pillar of robust and responsible data practice, along with mastery of the three pillars of inferential thinking (statistics), computational thinking (computing), and domain knowledge. These pillars become familiar to students through the program's degree requirements, which include "computational and inferential depth," "domain emphasis," and "human contexts and ethics," and are frequently discussed in and across classes. The phrase "human contexts and ethics" highlights that contexts are essential and integral to making sense of the ethics of data (or any technology). Parallel to the emphasis in the technical data science curriculum on teaching students the skills of computational and inferential thinking, HCE teaches students contextual reasoning. The cognitive skill of contextual reasoning is a complement and sometimes a counterweight to the abstraction dominant in computer science and statistics pedagogy.

This paper explains what we mean by "contextual reasoning," as a set of cognitive practices that are rooted in liberal arts disciplines to infuse and enrich data science technical education. Next, it illustrates the use of contextual reasoning with specific examples from three existing courses: a dedicated "human contexts and ethics" class (Data 104: Human Contexts and Ethics of Data) and technical classes at the junior (third-year) and senior (fourth-year) level (Data 100: Principles and Techniques of Data Science, and Data 102: Data, Inference, and Decisions).

It then provides a mid-project report exploring the factors involved in cultivating contextual reasoning skills earlier in the curriculum, via developing a new first- or second-year class (Course A) to scaffold follow-on contextual reasoning objectives in third- and fourth-year courses. When faced with a dataset or a visualization, what do you need to know about the world to make sense of it? To communicate about it to others? To deploy it responsibly? Given the overwhelming complexity of possible contexts, where should a data scientist begin? This part of the paper outlines the design of a class that teaches students where to look, how to look, and what to look for.

Positionality

Our author team represents three disciplines and several institutional positionalities. One of us is a computer scientist in a department of electrical engineering and computer sciences, one is a computational social scientist in a department of statistics, and two are historians (an intellectual historian and a historian of science and technology with STS leanings) in a department of history. All of us have some degree of technical education, including undergraduate coursework, a masters, and two PhDs, and all have some interest in data science, computing, society, and justice. Two of us are assistant teaching professors, one is an early/mid-career lecturer with a continuing contract, and one is a full professor. We all currently have primary teaching responsibilities in our institution's data science program, and we have worked together in the program for several years. One of us was on the design team for the data science program in the mid-2010s, helping set its parameters including defining the concept of HCE, and subsequently directed it for several years.

2 What is contextual reasoning?

Contextual reasoning encompasses a suite of skills that return the practitioner's gaze back to the concrete world from which their data originated and through which it circulates. Much of the power of computational and statistical forms of reasoning stems from their ability to abstract from the messiness of the world. Data is always a selective representation of a phenomenon of the world—something that "stands for" something else—and thereby leaves out most of the rest of the world. Although this abstraction makes the data more mobile and reconfigurable, allowing it to circulate and be recombined in ways that enable new kinds of patterns and relationships to become visible, it can also make it easier to forget the significance of what's left out, or even what's left in. Students who work only with data points as abstract, fungible tokens without learning to think about the *making* of data may imagine that their data is an unproblematic, transparent, one-to-one representation of a coherent underlying reality rather than a particular interpretation of it.

Contextual reasoning is about more than understanding where data comes from and what it represents. It is also about reflecting on the embodied practices of those working with data (e.g., how the tech workplace is set up and functions) and the social, political, economic, environmental worlds that shape and are shaped by data and data-driven sociotechnical systems. For example, if a practitioner works with census data that categorizes people by "race" or ethnicity, then it is not enough to be aware that this is fraught and complicated. A practitioner who is familiar with the complex history and politics of racial categories and the multiplicity of race will draw dramatically different inferences from the data and carry a different sense of responsibility to the people in the data.

Which contexts matter in any given case is an essential question. As every "text" has innumerable contexts, the direction to "contextualize your data" is empty without some clear normative orientation to the contexts that matter most. Further, as pointed out in an astute anthropological analysis of data science practice [10], it is not the case that data scientists are just bereft of awareness of context. Rather, professions like data science coalesce around different notions of what counts as a relevant context and what "context" even means.

Although all disciplines in practice presuppose some underlying normative practices for handling "contexts," this does not mean that existing norms and practices for reasoning about contexts in data science are cognitively and ethically adequate. To build contextual reasoning into data science curriculum as a distinct skill goes beyond addressing a perceived deficiency in students' exposure to contextual reasoning in the classroom, and it intervenes in the existing norms of contextual reasoning in data science, which simply highlight factual and theoretical knowledge of application domains. There are powerful practices for analyzing other human contexts developed in the social sciences and humanities that can broaden data science students' horizons for what counts as a context worth addressing.

In our curriculum we have developed what we call the *Human Contexts and Ethics (HCE) Toolkit*, a set of concepts distilled from various disciplines in the humanities and social sciences that orient students to the most vital and often-overlooked contexts of data and technologies. Ranging from general social-theoretical concepts like *power* and *identity/positionality* to more specialized concepts such as *sociotechnical systems* and *co-production*, the HCE Toolkit serves as a kind of scaffolding or equipment for contextual reasoning by orienting students to the most salient dynamics and structures of the social life of data. Rather than emphasizing abstract definitions or theoretical statements, the HCE Toolkit gives students practice asking questions: what narrative animates a particular data collection project, and who does it serve? Where did it come from? What histories, institutions, and people contributed? How does an artifact shift power or link up with particular political projects? Who are the people in the data, the data subjects, and how does the data speak for them?

Teaching contextual reasoning in both dedicated ethics courses and in modules integrated into existing technical courses has taught us a further lesson. It is easy to take for granted the underlying *skills* that make robust contextual reasoning practices possible. The skills that provide the "infrastructure" of contextual reasoning certainly surface in liberal arts education. Since contextual reasoning involves structured ways of thinking about how data is entangled with social power, or how contemporary data practices are extensions of deep seated historical processes, then it necessarily depends on students possessing basic skills for learning about those contexts in the first place: by knowing where to look (*research*), how to parse the sometimes difficult academic texts about them (*literacy*), how to apply what they read to data practice (*synthesis*), and how to communicate about their reasoning clearly (*writing*). However, these skills are not well-represented in existing data science curricula; at present, student preparation for contextual reasoning is uneven and assumes students have acquired the core skills in their secondary education or through other coursework. In short, giving students access to the benefits of a liberal arts education within their data science education is essential to the cultivation of contextual reasoning abilities.

3 Experiences teaching contextual reasoning

The concept of contextual reasoning emerged from about a decade of developing Human Contexts and Ethics (HCE) curriculum across multiple levels of the data science curriculum at our institution. This curriculum includes a dedicated course (Data 104) that fulfills the program's HCE course requirement, commonly taken in students' third or fourth year of university study, and integrated HCE content in third- or fourth-year technical courses (Data 100 and Data 102) [25,26,27].

The data science program at our institution was first offered in 2018. It is an interdisciplinary major and minor program between faculty in Statistics, Electrical Engineering and Computer Sciences, and several dozen other departments across our campus (e.g., Economics, Molecular and Cell Biology, Psychology, History). Run as an interdisciplinary program without being lodged in a department, it is the largest undergraduate program at UC Berkeley, which is a large public R1 institution, and graduated 1,074 majors and 450 minors in academic year 2024-25. It is a B.A. degree and is not ABET accredited, as our institution's computer science program has likewise opted out of ABET. HCE is thus anchored in and responds to our institution's own sense of mission and purpose. Data 104, 100, and 102 are offered every semester with enrollments each semester typically reaching 700+ (Data 104), 1000+ (Data 100), and about 250 (Data 102). The authors are the instructors of Data 104 and 100 and collaborate closely with the instructors of Data 102.

Data 104

Contextual reasoning in the third-/fourth-year course Data 104 is crystallized in the HCE Toolkit. As mentioned above, these tools act like pointers, orienting students to specific kinds of contexts. For example, students practice asking how data is entangled in specific relationships of *power*, or how data always *represent* the world selectively according to *narratives* about how the world works and specific viewpoints that get shaped by the *positionality* and *identity* of its makers.

For example, through a close examination of a case study about pollution data from the local air quality management district, the third lecture of the course, "Making Data," teaches students to recognize how data is always made, not found; and made for specific reasons, guided by particular assumptions that frame how the world can become legible. What does it mean to measure air pollution at a regional level as measured by concentrations of standardized 'criteria pollutants' defined by the U.S. federal government, versus specific, hyper-local fenceline measurements of compounds released in particular industrial processes, such as steel production or refinery flares? How do these ways of making data shape what we see and think about pollution's impact on human health? Who decides what measurements are made, and made public? How do their availability or non-availability open the door to different kinds of individual, societal, and regulatory action? The key concepts from the HCE Toolkit here are *representation* and *classification*.

Another lesson in the course emphasizes how historical context powerfully inflects how one interprets the social significance of data visualizations. It examines a popular visualization from Hans Rosling's Gapminder project plotting the relationship between women's fertility and infant mortality by country [11]. Along with other Gapminder visualizations, this plot appears to reveal a narrative of progress, a natural law of economic development and modernization, with countries following a normative path of development. Students that encounter this plot without context tend to find it unproblematic, and accept the basic narrative that declining fertility is a necessary component for progress and modernization (even if they're careful not to confuse correlation with causation). By putting these visualizations into the context of the history of global population statistics, however, we teach students to see this data and the narratives that animates it in a different light. This data on women's fertility was collected in the second half of the twentieth century as a part of a project of Global North and elite-led international development in the Global South. Together with the new discipline of demography, and animated by a racialized narrative of a ticking population time bomb, international organizations from the Global North aimed to understand population trends in order to manage them. Women's fertility in the Global South became a focal point for population management through family planning, abortion, and sterilization [12,13,14].

Data 100

In the third-year (junior) level data science class, we first developed an approach to integrating contextual reasoning into multi-week course projects, where students apply data science techniques to solve real world problems. This approach took its most complete form in the course's first major project (Table 1). Students are asked to assume the role of the Cook County, Illinois Assessor's Office (CCAO) and predict the value of residential properties in Cook County Illinois for property tax assessment. Superficially this project is a familiar statistics lesson: fitting a linear model. But whereas a typical project gives students an unequivocal quantitative measure of success, by placing this task in a particular social context, we problematize the meaning of success by showing how careful attention to the proper political contexts of the problem reveals the essential ambiguity and contestability of every measure of error.

Craft at least two questions about housing in Cook County that can be answered with this dataset and provide the type of analytical tool you would use to answer it. For one of the questions that you asked, draw by hand what a plot resulting from this analysis might look like.
Consider people selling homes, Cook County Assessor's Office (CCAO) appraisers, and people buying homes. When their interests conflict, whose perspective should the CCAO prioritize in valuing homes, and why?
When evaluating your model that predicts the price of homes, we used root mean squared error (RMSE). In the context of estimating the value of houses, what does the residual mean for an individual homeowner? How does a positive or negative residual affect them in terms of property taxes?
In your own words, what makes a model's predictions of property values for tax assessment purposes "fair"?

Table 1: Sample of abridged open-ended questions from Data 100's first project on building a model to assess the value of residential properties. See projA1 and projA2 [28] for the full assignment.

Students begin the project primed to create a model that minimizes root mean square error, which they take to be a reliable measure of the accuracy of their model. But the larger context of this task throws this assumption into doubt: students work with a dataset that was made public as part of an attempt by the CCAO to re-establish its legitimacy after a scandal revealed its tax assessments to have been sharply regressive, leaving a disproportionate tax burden on low income and non-white homeowners. By building a public data repository and appealing to the public perception that more sophisticated data practices and algorithms could produce accurate and therefore fair assessments, the CCAO aimed to rebuild public trust. With this context in mind, we ask students to consider how different measures of accuracy (error) capture different intuitions about what makes for a "fair" model, and to consider how the positionality of different constituencies might entail different and competing expectations about fairness. Root mean square error, they realize, doesn't account for regressivity, so minimizing it wouldn't necessarily address the underlying problem. Is a model that minimizes overall error more fair than one that is less accurate but minimizes regressivity or eliminates it completely? Would a model that errs on the side of a progressive tax rate be more fair than one which distributes errors more evenly across the socioeconomic spectrum? Who decides?

The project then pushes students to move past well-trodden questions of fairness and to consider the very framing of the problem itself: how does accuracy come to stand in for fairness in the first place? Why is an accurate algorithm supposed to engender social trust? Asking these questions and getting students to think about the specific constellation of political forces in Cook County pushes them to question their assumption that publishing accurate numbers would automatically foster trust.

Following on previous work in Data 100 [25], we extended our approach to integrate contextual reasoning across the course via weekly homework assignments, where students already explore real world data through the abstract lens of data science technical skills. New contextual reasoning exercises are inserted as open-ended questions toward the end of each Jupyter Notebook assignment (Table 2). We evaluated initial student responses to these questions. Among 1,874 respondents to a mandatory mid-semester survey conducted in both Fall 2025 and Spring 2026 (>90% response rate), approximately 15% of respondents found the open-ended questions "Not at all valuable," with 65% responding "Somewhat valuable" and 20% responding "Very valuable." These three proportions did not deviate substantially across term (Fall 2025 or Spring 2026), class year (first, second, third, fourth, fifth or higher, graduate, or exchange), gender, or status as having learned English as a first language. Additional results derived from free text student feedback is reported in Table 3.

Data 102

Finally, working with colleagues teaching Data 102, a senior level course on statistical decision-making, we developed a new component of the final course project that requires students to research reflect on the human contexts of their projects, and a new lecture on algorithmic fairness that modeled the kinds of reasoning expected of students on their final project.

SETUP: You are a data scientist working for the San Francisco Department of Public Health. Your manager has been allocated a small budget to address the health inspection scores of the restaurants in one neighborhood in San Francisco. TASK: Recommend one neighborhood where funding would have the biggest impact.
SETUP: You are a data scientist working for the New York Times (NYT). Your manager is concerned that NYT exhibits bias in its articles and asks you to look into this. TASK: Provide evidence for or against the hypothesis that NYT exhibits bias in its articles.
SETUP: Suppose you are interested in studying the effect of raising the minimum wage on prices and the unemployment rate. Assume you have infinite time and infinite resources, you can observe multiple different universes at the same time, and you can collect data about anything in any universe. TASK: Design one hypothetical plot that demonstrates the effect of changing the minimum wage on prices and/or unemployment.
SETUP: You are a data scientist working for Streaming Co., a new movie streaming company. Your manager wants to understand the changes in movie preferences over time to help them come up with the next blockbuster movie. TASK: Your task is to create one visualization to help your manager understand the changes in movie trends over time and suggest at least one thing you might want to include in this blockbuster movie.

Table 2: Abridged open-ended homework questions from Data 100. See hw02b, hw03, hw04, and hw07 [29] for the full questions.

Theme	Count (%)	Theme	Count (%)
Enjoyable	428 (23)	Grading concerns	84 (5)
Time-consuming	311 (17)	Belief that nothing was learned	83 (4)
Allows for creativity	246 (13)	Not enjoyable	79 (4)
Need more guidance to start	182 (10)	Overreliance on AI	58 (3)
Provides real-world preparation	178 (10)	Want examples of student work	53 (3)
Difficulty with open-ended thinking	161 (9)	Appreciate AI support for code	38 (2)

Table 3: Common themes from free text student feedback on open-ended questions in Data 100, among 1,874 respondents from the 2025–2026 academic year.

The Data 102 lecture worked off of a well-known case study [15] about the Allegheny Family Screening Tool (AFST), a predictive algorithmic system developed and deployed by the public child welfare office in Allegheny County, Pennsylvania to screen intake calls about potential cases of child neglect and abuse by generating a risk score for the concerned families. Despite the good intentions behind the system, designed to help protect children and support families, the system ultimately reinforced the association of poverty with poor parenting, and acted as an apparatus of surveillance and class discipline.

The lectures scaffolded contextual reasoning by gradually moving away from the kinds of abstract technical questions familiar from the rest of the course to progressively more concrete and holistic social contexts. Starting with an already abstract framing of a real world problem ("we want our tool to use data to perform its assigned task while treating protected classes fairly"), and then operationalizing it as a purely technical problem (fixing an accuracy metric, dataset, model, measure of fairness, etc.), the lectures then worked outwards, questioning the assumptions that went into the original framing of the problem and its technical operationalization by examining progressively more expansive contexts, from the most immediate (e.g., what do your variables represent, what proxies are you using,what are you really trying to predict, which measure of fairness should we use, what does fairness mean) to the more holistic (how does the history of race and class in the United States shape the meaning of racial categories in the data, what political histories made it thinkable or justifiable to use predictive analytics to manage poverty and child welfare).

While the lecture offered an exceptional model of contextual reasoning, there was no space earlier in the semester to give students practice employing this style of thinking. This resulted in suboptimal performance on the contextual reasoning components of the final project in the first two semesters since they were introduced [25].

4 New course design

We are currently midway through a design process for a new course: *Course A*, a new first- or second-year course dedicated to giving students foundational skills in contextual reasoning, as a way to support its higher-level integration into already existing third-/fourth-year technical courses such as Data 100 and Data 102. This section provides a mid-project report on the design work. It serves as both a reflection on experiences elsewhere in the curriculum and a way of concretely fleshing out how contextual reasoning can be taught as its own set of skills by starting from data and moving outward to contexts.

The motivation for developing Course A comes not only from having developed a sense of the skill-based meaning of contextual reasoning across our iterative interventions, but also from having encountered challenges in teaching contextual reasoning later. When students have been socialized in a data science curriculum that mostly elides context in courses taken before the third/junior year, students in third- and fourth-year technical courses (such as Data 100 and Data 102) have sometimes reported being confused by assignments that engage their contextual reasoning, or they downplay their importance. Even more, many students lacked the basic literacy skills that provide the essential infrastructure of contextual reasoning—in some cases because of the limits of their secondary education or because skills had atrophied across several years of coursework requiring only minimal reading and writing.

Course A will start from data and move outward to contexts organized around the data science lifecycle, which is a standard conceptual framework in the field that has provided useful scaffolding for introducing HCE concepts in courses like Data 100 [16]. The data science lifecycle describes the path of a typical data science project across a sequence of stages. Although differing in their details, most versions of the lifecycle include problem framing, data acquisition, exploratory data analysis, modeling, drawing conclusions and reporting findings. We put a strong emphasis on the part of the lifecycle most neglected in current data science courses at our institution: *making data*. Technical data science classes may neglect or downplay sustained inquiry into the creation of datasets, instead prioritizing teaching the computational and statistical techniques for analyzing off-the-shelf datasets. Although some work has been done to stress how contexts are indispensable for making sense of these datasets, this rarely involves putting pressure on the social and historical origins of that data: where it comes from, who made it, for what purposes, with what assumptions about the world, etc.

Course A is designed to scaffold increasingly sophisticated contextual reasoning skills throughout the course, operating at two levels as students move through it. On the first level, students will learn to move past the most "proximate" context of every dataset, visualization, or model, to the larger, more structural and historical contexts, working "outwards" from the given "text" of the data and the more technical operations through progressively more complex layers of context (as in the Data 100 property tax assessment and Data 102 fairness example). The second level works in parallel to expanding the range and scope of contextual reasoning stepwise. We gradually scaffold students with the basic, infrastructural skills of contextual reasoning, asking them to contextualize with greater degrees of independence as the semester progresses. While the first level involves teaching students *where* to look for context, the second level focuses on cultivating the skills necessary for parsing those contexts.

For example, at the start of the semester, students will be given a dataset or visualization with no accompanying descriptions other than the data labels and will be asked to speculate about it: What does the data represent? Who collected it? For what purposes? What assumptions about the world frame the data? What relationship do you expect would obtain between two categorical variables plotted on x- and y-axes? In a second step, with a new data dataset or visualization, students will be given a brief curated description of a context by the instructors, and then will be asked to explain how that context shapes their interpretation of their data. What did they think the data meant drawing only on their common sense? How did knowing this bit of historical and social context change how they viewed that data or a particular way of using that data? In a third step, students will be given a piece of sophisticated academic writing (e.g., a history or sociology article or chapter) to use to contextualize the data text. In a fourth and final step, students will be presented with a given dataset or visualization and will be asked to do their own independent research using academic databases to discover relevant contexts, connect their research to their dataset, and effectively communicate about the data-in-context: what it says (and does not say) about the world; who it represents, how its use in certain contexts shifts power, etc.

The first and most extensive unit of Course A will expand on and give students practice applying the key insights of the "Making Data" lecture in Data 104, drawing on further theoretical literature and ethnographic work on the micropolitics of making data [17,18,19]. The purpose is not only to instruct students didactically in the relevance of contexts in understanding data, but to give them repeated practice and first-hand experience with these contexts so asking about them becomes an intuitive and systematic part of their data practice. Among the earliest exercises in Course A will ask students to collect their own data in a variety of different contexts and using multiple methods (observation, survey, designing a form, etc.) so they get hands-on experience with the contexts that shape data making.

Many students at our large university campus superstitiously avoid walking over a large metallic campus seal embedded in a busy pedestrian walkway. As a tangible exercise in making data, students will design and carry out a data collection plan to answer the following: What proportion of students intentionally avoid walking over this campus seal? While simple on its face, this question challenges students to make fuzzy categorizations on the fly: Which pedestrians are students? Who is actively avoiding the seal, and who misses the seal just by chance? The exercise challenges students to consider the downstream effects of the assumptions embedded in their data collection plan: Is walking behavior the same at all times of the day, at different points in the school year, or when the weather changes? Are there avoidant students who choose to walk on a different pathway altogether, and are thus excluded from the dataset? Does the presence of an observant student affect pedestrian behavior? Students will also have to consider how their own decisions and states of mind affect the data they produce: If I decide to alter the criteria for my categorizations, do I discard existing data? To what extent does the quality of my categorizations change as I become increasingly fatigued? Finally, to emphasize how different data collection approaches can lead to different results, we will visualize the distribution of student outcomes as part of a follow-up lecture and solicit possible explanations for the largest discrepancies in the estimated proportion of avoiders. This exercise provides students with an important lesson not typically learned until after graduation: The most time-consuming and influential components of data science projects are often collecting, validating, and interpreting raw data, as opposed to heavy computation and modeling.

In a different exercise, students will consider how best to represent a real-world phenomenon with a tabular dataset, and then compare and contrast their idealized dataset with a public dataset that proxies a similar phenomenon. For example, students will be asked to consider how to construct tabular data that represents an interaction between a police officer and a stopped driver. What is the downstream purpose of this data? Should all stops be recorded, even if they do not end in a citation? Which demographic information should be collected, and should it be officially verified by a third party? How can body-worn camera footage be converted into a set of numeric features? At what point are we asking officers to record too much information? How does the design of a data entry form influence the interpretation of stop events? In what ways can recollection of stop events be distorted? Which relevant stop events are practically impossible to record, like changes in facial expressions or intonation of voice? After students design their idealized dataset, they will write a mock README for the dataset that explains the context and limitations of each feature in the dataset. Finally, we will present students with real examples of tabular policing data from the Stanford Open Policing Project. Students will compare and contrast the contents of these public datasets with their own idealized dataset. Students will consider the kinds of questions for which public datasets provide sufficient context to answer, and which questions need additional context provided by a domain expert.

The exercises are designed to help students pay attention to recurring, high-level matters through concrete examples and experiences. These include: How data represents the world; what labor goes into making data, including how that labor gets organized and standardized; which classifications get used, and what gets left out when different kinds of classifications are used; how difficult it can be in practice to make decisions about how to classify phenomena and take them out of the ambiguous flux of everyday experience; and how the specific historical and institutional setting of data collection shapes how we evaluate its responsible use.

The second main unit of Course A will emphasize the next stage in the data science lifecycle, *Exploratory Data Analysis* (EDA). EDA reveals patterns, anomalies, and assumptions hidden in datasets, complementing and enriching the context provided by written descriptions of a dataset's features and observations. This unit will focus largely on interpreting visualizations of raw data and basic summary statistics, as opposed to fitted statistical models, and will continue to develop and extend many of the core contextual reasoning skills introduced in the first unit. This unit addresses a common limitation of many data science courses: While students spend a lot of time learning how to mechanically produce visualizations, they often spend little to no time thinking deeply about the larger story and context behind published plots of real data.

In one exercise, students will be provided with a dataset and a dataset description. They will draw, by hand, their predictions of the distributions of important features and their predictions of the shape of bivariate plots of related features. Before the true plots are revealed, students will discuss discrepancies between their own predictions and their peers' predictions, and consider how differences in assumptions drive these discrepancies. After the true plots are revealed, students will be asked to describe the extent to which their intuition and assumptions about the dataset have changed, state any new questions they have about the dataset, and plan additional EDA needed to answer those questions. This exercise intentionally slows the EDA process: Rather than just mechanically plotting all individual features and all possible comparisons without considering dataset context, students carefully consider why certain features or relationships should be plotted, and how the resulting plots align with their intuition or change it.

In a different exercise, students will play "telephone" with data. Pairs of students are provided a textual conclusion of a published study. Each pair hand-draws candidate plots to communicate that conclusion. After deciding on a single best candidate plot, each pair of students exchanges their single best plot with a different pair of students. Each receiving pair attempts to decipher the conclusion of the hand-drawn plot, and returns their guess to the original pair. Finally, each pair discusses any discrepancies between the true conclusion of their assigned study and the deciphered conclusion. This exercise provides students with practical experience thinking through how subjective plot design choices can influence the perception of purportedly objective results.

5 Conclusions

This paper describes a multi-level approach to teaching contextual reasoning in a technical undergraduate major. Beyond encouraging students to *put their technical work in contex*t, and in addition to helping them *see the significance of context*, we propose that contextual reasoning equips students with cognitive skills to *identify particularly consequential contexts and trace their impacts* on outcomes. A key instrument of teaching contextual reasoning is a "human contexts and ethics" toolkit drawn from the humanities and social sciences for the purpose of teaching ethics and responsible practice. The approaches built into the toolkit can be successfully used in technical classes to give students practice in contextual reasoning as an integral part of robust technical practice.

The focus of this paper is data science education, which in some ways is particularly amenable to fleshing out the work of contextual reasoning. All data is made through human actions and incorporates human decisions. Their representational consequences may stay with the data, visibly or not, as it travels—across different contexts and into data technologies such as automated decision systems and predictive algorithms. Practitioners' choices about how to represent and shape the world through data are the subject not only of careful attention in the discipline of statistics, but also in other disciplines including STS (Science, Technology, and Society) and critical data studies. Thus data science education is a fruitful site to explore the work of contextualization, and doubly so because the field is already interdisciplinary and still relatively new in formation, making it a valuable setting in which to engage.

Data science may not be a wholly singular case, however. At a higher level, it may give useful pointers for other technical and engineering majors. Contextualization and contextual reasoning make a difference in data science in part because *context* runs counter to moves of *abstraction*. Abstraction is built into the technical disciplines of computing and statistics, and each engineering discipline likely has its own set of moves of abstraction—framing what can be put to the side when moving from concrete particulars to a more general treatment. Helping students grapple with the back-and-forth between particulars and abstraction seems to be a critical piece of contextual technical education.

With focused effort, close collaboration, and a commitment to repeated practice across the curriculum, it is possible to teach students how to see and assess the consequences of the human contexts of data across a technical major. However, this work is challenging in multiple ways. In existing courses, it requires adding material to already full syllabi, which intrinsically calls on instructors to make judgment calls about priorities. More than that, it requires working together to smoothly integrate technical and contextual learning objectives and preparing instructors (including teaching assistants) to handle them. Unless students have enough practice early in the curriculum, they will arrive in third- or fourth-year (junior- and senior-level) courses without enough experience or skills. Thus an ecosystem approach is needed that works across courses and specifically embeds contextual reasoning early in the curriculum.

References

[1] M. S. Kleine, K. Zacharias, and D. Ozkan. “Contextualization in engineering education: A scoping literature review.” Journal of Engineering Education 113 (2024): 894-918. 10.1002/jee.20570.

[2] M. S. Kleine, K. Zakarias, and D. S. Ozkan. “Contextualization as virtue in engineering education.” Paper presented at 2021 ASEE Virtual Annual Conference Content Access, Virtual Conference. 10.18260/1-2--36844.

[3] E. A. Mosyjowksi, S. R. Daly, and S. J. Skerlos. “Peripheral or fundamental? Instructor perspectives on integrating societal content in engineering courses.” Studies in Engineering Education 6(2) (2025): 74-93. 10.21061/see.201.

[4] A. Adhikari, J. DeNero, and M. I Jordan. “Interleaving computational and inferential thinking: Data science for undergraduates at Berkeley.” Harvard Data Science Review 3(2) (2021). 10.1162/99608f92.cb0fa8d2

[5] E. A. Reddy, M. S. Kleine, M. Parsons, and D. Nieusma. “Sociotechnical integration: What is it? Why do we need it? How do we do it?” Paper presented at 2023 ASEE Annual Conference & Exposition, Baltimore, Maryland. 10.18260/1-2--44239.

[6] R. Koh, “Putting the (social and environmental) ‘context’ in context-rich problems,” Paper presented at ASEE Zone 1 Conference - Spring 2023, State College,, Pennsylvania. 10.18260/1-2--45081

[7] R. H. Sangree. “Engagement in practice: A cornerstone design course fusing the engineering design process with community engagement in context-based design.” Paper presented at 2025 ASEE Annual Conference & Exposition, Montreal, Quebec, Canada. 10.18260/1-2--56350.

[8] M. L. Matthews, S. Koloutsou-Vakakis, and A. S. Stillwell. “CASE STUDY: Project-based integration of societal context with engineering communication and computational thinking in an upper-level civil and environmental engineering course.” Paper presented at 2025 ASEE Annual Conference & Exposition, Montreal, Quebec, Canada. 10.18260/1-2--56070

[9] S. J. Qin. “Data science education with domain knowledge and system principles.” Harvard Data Science Review 3(2) (2021). 10.1162/99608f92.6f037267

[10] N. Seaver. “The nice thing about context is that everyone has it.” Media, culture, and society 37(7): 1101-1109. 10.1177/0163443715594102

[11] https://www.gapminder.org/tools/#$chart-type=bubbles&url=v2

[12] E. K. Merchant. Building the population bomb. Oxford: Oxford University Press, 2021.

[13] M. Murphy. The economization of life. Duke: Duke University Press, 2017. 10.1215/9780822373216

[14] A. Bandarage. Women, population, and global crisis: A political-economic analysis. New York: Bloomsbury, 1997.

[15] V. Eubanks. Automating inequality. How high-tech tools profile, police, and punish the poor. St. Martin’s Press, 2018. Chapter 4, “The Allegheny Algorithm.”

[16] M. Boenig-Liptsin, A. Tanweer, and A. Edmundson. “Data science ethos lifecycle: Interplay of ethical thinking and data science practice.” Journal of statistics and data science education 30(3) (2022): 228-240. 10.1080/26939169.2022.2089411

[17] L. Gitelman, ed. “Raw Data” is an oxymoron. Cambridge, MA: MIT Press, 2013. 10.7551/mitpress/9302.001.0001

[18] C. D’Ignazio and L. F. Klein. Data feminism. Cambridge, MA: MIT Press, 2020. 10.7551/mitpress/11805.001.0001

[19] C. Biruk. Cooking data: Culture and politics in an African research world. Durham, NC: Duke University Press, 2018. 10.1215/9780822371823

[20] S. Sismondo. An introduction to science and technology studies. 2nd ed. Chichester: Wiley-Blackwell, 2010.

[21] E. A. Cech. “The (mis)framing of social justice: Why ideologies of depoliticization and meritocracy hinder engineers’ ability to think about social injustices.” In Engineering education for social justice, ed. J. Lucena, 67-84. Dordrecht: Springer, 2013. 10.1007/978-94-007-6350-0_4

[22] D. Nieusma. “Analyzing context by design: Engineering education reform via social-technical integration.” In International perspectives on engineering education: Engineering education and practice in context, eds. S. H. Christensen et al., vol. 1, 415-434. Cham: Springer, 2015. 10.1007/978-3-319-16169-3_20

[23] D. S. Özkan. “Sociotechnical curricular approaches for a culture of engagement in engineering: Students navigating power and equity.” Engineering Studies, 2025, 10.1080/19378629.2025.2556386.

[24] L. Czerwionka, S. S. Joshi, G. O. Rios-Rojas, and K. A. Davis. “Developing students’ sociotechnical thinking in the humanities-engineering classroom: A treatment versus control study using a scenario-based assessment.” J. Eng. Educ., 114(3) (2025), e70004. 10.1002/jee.70004

[25] C. Carson, L. Yan, A. Edmundson, and A. Strang. “Sociotechnical integration in data science education.” Paper presented at 2025 ASEE Annual Conference & Exposition, Montreal, Quebec, Canada. 10.18260/1-2--57660

[26] A. Brooks-Ramirez, R. Dang, B. A. V. Benitez, and L. Yan. “Scaling responsible data science education: The role of a teaching assistant in bridging the sociotechnical divide.” Paper presented at 2025 ASEE Annual Conference & Exposition, Montreal, Quebec, Canada. 10.18260/1-2--57629

[27] L. Yan. “Teaching our teacher assistants to thrive: A reflexive, inclusive approach to scalable undergraduate education.” In Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 1 (SIGCSETS 2025). Association for Computing Machinery, New York, NY, USA, 1281-1287. 10.1145/3641554.3701950

[28] http://github.com/DS-100/sp26-student/tree/main/proj

[29] http://github.com/DS-100/sp26-student/tree/main/hw