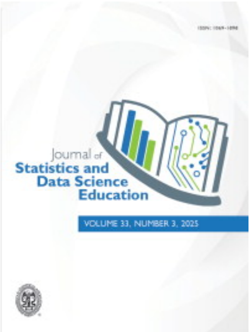

Taylor & Francis
Taylor & Francis Group

Journal of Statistics and Data Science Education

ISSN: 2693-9169 (Online) Journal homepage: www.tandfonline.com/journals/ujse21

The $Q_1Q_2Q_3$ Workflow for Statistics and Data Science Collaborations

Eric A. Vance, Ilana M. Trumble, Jessica L. Alzen & Leanna L. House

To cite this article: Eric A. Vance, Ilana M. Trumble, Jessica L. Alzen & Leanna L. House (05 May 2025): The $Q_1Q_2Q_3$ Workflow for Statistics and Data Science Collaborations, Journal of Statistics and Data Science Education, DOI: 10.1080/26939169.2025.2475775

To link to this article: https://doi.org/10.1080/26939169.2025.2475775

Published online: 05 May 2025.

JOURNAL OF STATISTICS AND DATA SCIENCE EDUCATION
2025, VOL. 00, NO. 0, 1–21
https://doi.org/10.1080/26939169.2025.2475775

The $Q_1Q_2Q_3$ Workflow for Statistics and Data Science Collaborations

Eric A. Vance[a], Ilana M. Trumble[a], Jessica L. Alzen[b], and Leanna L. House[c]

[a]Laboratory for Interdisciplinary Statistical Analysis, Department of Applied Mathematics, University of Colorado Boulder, Boulder, CO; [b]Center for Assessment, Design, Research and Evaluation, University of Colorado Boulder, Boulder, CO; [c]Department of Statistics, Virginia Tech, Blacksburg, VA

ABSTRACT
We develop, advance, and promote a previously existing framework called the Qualitative-Quantitative-Qualitative workflow ($Q_1Q_2Q_3$, pronounced "Q-Q-Q") to systematically guide the content of interdisciplinary collaborations and improve the teaching of statistics and data science. The $Q_1Q_2Q_3$ workflow is designed to help statisticians and data scientists develop skills and techniques for collaboration to work with domain experts across academic fields, industry sectors, and organizations. The $Q_1Q_2Q_3$ workflow explicitly emphasizes the importance of the qualitative context of a project, as well as the qualitative interpretation of quantitative findings. We explain $Q_1Q_2Q_3$ and provide guidance for implementing each stage of the workflow. We describe how we teach $Q_1Q_2Q_3$ within a statistics and data science collaboration course and present data evaluating its effectiveness. We also describe how $Q_1Q_2Q_3$ can be useful for educators teaching introductory, projects-based, and technical statistics and data science courses. We believe that the $Q_1Q_2Q_3$ workflow is an easy-to-implement technique that is beneficial and necessary for statistics and data science education and practice. It can be used to weave ethics into each stage of practice so that statisticians and data scientists can successfully transform evidence into action for the benefit of society.

ARTICLE HISTORY
Received March 2024
Accepted March 2025

KEYWORDS
Data science education; Data science lifecycle; Ethics; Statistical collaboration; Statistical consulting; Statistical practice

1. Introduction

> *The need for interdisciplinary collaboration and a diverse range of skills will become increasingly important for statisticians to remain relevant*—excerpt from the American Statistical Association's (ASA) Statement on *The Role of Statistics in Data Science and Artificial Intelligence* (2023, p. 3).

Statisticians and data scientists have many reasons to collaborate with domain experts across academic fields, industry sectors, and organizations. One reason is to make a deep contribution in a field that could not be created alone (Love et al. 2017). According to Metzger et al. (2023), it is imperative for statistics and data science educators to train students to collaborate across sectors and for individuals to develop their personal collaboration skills. However, interdisciplinary collaboration is multifaceted and requires multiple modes of intelligences (Derr 2000). Collaboration is not simple to teach! To help make sense of the high dimensional space of skills necessary for effective collaboration, Vance and Smith (2019) identified five principal components of collaboration: Attitude, Structure, Content, Communication, and Relationship (ASCCR). They placed Content in the middle of the ASCCR framework to emphasize the deep contribution statisticians and data scientists can make on a consulting/collaboration project.

At a fundamental level, the content or task of most collaborative quantitative projects is to apply statistics and data science to solve a domain problem or make a data-driven decision. We believe that this content of a project requires special attention from practitioners and educators because there is not yet a unifying "Theory of Applied Statistics" that can guide statistical collaborators and investigators to complete a project's tasks. Every project is different, and most require bespoke solutions. Different contexts of the problems and different uses for the solutions require different quantitative methods to be used (Banks 2023). Mallows (1998), in his 1997 ASA Fisher Lecture, called on the field to develop a neat, formal theory so that applied statistics will become easier to teach, and statistics more effectively applied. He stated:

> *The main challenge of applied statistical work is that of taking proper account of contextual issues. Good techniques are not enough; nor are good computer programs, nor powerful theorems. A major intellectual attraction of the discipline is the subtlety of the interplay between the formal statistical procedures and the imperfectly understood substantive questions* (p. 3).

Leman et al. (2015) introduced a clever paradigm for learning and teaching data analytics that we believe can form the basis for a theory of applied statistics. They stated that in a

CONTACT Eric A. Vance *Eric.Vance@Colorado.EDU* Laboratory for Interdisciplinary Statistical Analysis, University of Colorado Boulder, Boulder, CO.

data analysis project, analyzing the Qualitative issues (Q_1) of the context-specific question and data collection method must precede the second layer of Quantitative analysis (Q_2) involving formal mathematics or computations. Then the numbers must be Qualitatively (Q_3) summarized and assessed in a manner consistent with the questions asked in the Q_1 phase. Leman et al. (2015) describe how to use this Q_1-Q_2-Q_3 (QQQ) paradigm in the classroom, not as a tool to teach students, but as a way for instructors to provide a balanced statistics education that provides more focus on Q_1 and Q_3 than in typical statistics courses. Vance and Smith (2019) integrated $Q_1Q_2Q_3$ into the ASCCR framework as a tool for educating and training statisticians and data scientists regarding the Content of collaborative projects.

The $Q_1Q_2Q_3$ workflow has guided our teaching of the content of collaborations since 2013. However, based on our experience, neither Leman et al. (2015) nor Vance and Smith (2019) explained $Q_1Q_2Q_3$ sufficiently for beginning collaborators to implement this workflow. Nor were their explanations sufficient for most educators to infuse it into their pedagogy. Therefore, in this paper, we explain the $Q_1Q_2Q_3$ workflow in detail in Section 2 and how individual practitioners can implement it on collaborative projects in Section 3. Our goal is for readers to be able to implement $Q_1Q_2Q_3$ on future projects and use it to guide their ethical thinking. If a reader is also an educator, our goal is for them to teach and use $Q_1Q_2Q_3$ in the classroom.

This paper has two intended audiences: statistics and data science *collaborators* (including students and educators who also consult/collaborate with domain experts) and statistics and data science *educators* who are curious about how they might apply $Q_1Q_2Q_3$ in various ways in the classroom. In Section 4, we describe how we teach $Q_1Q_2Q_3$ in an interdisciplinary collaboration course, and we evaluate the effectiveness of our pedagogy in Section 5. In Section 6, we describe ways to use and teach $Q_1Q_2Q_3$ in introductory, projects-based, and technical statistics and data science courses. We discuss some of the workflow's benefits and limitations in Section 7, including how it connects with other recommendations in the statistics and data science education literature. In Section 8, we conclude our exploration of the $Q_1Q_2Q_3$ workflow, which we believe is an easy-to-implement technique extremely beneficial and necessary for statistics and data science education and practice.

2. Explanation of $Q_1Q_2Q_3$

In our experience, every technical project—be it in applied statistics, data science, physics, engineering, sociology, or the digital humanities—has three components: Qualitative, Quantitative, and Qualitative. Every effective collaboration must begin with understanding the problem, knowing what data were collected, how and why they were collected, etc. We label the understanding of these first-stage qualitative issues "Q_1." The second stage of a project, "Q_2," is to quantitatively describe, visualize, model, analyze the data. "Q_3" is the end qualitative stage in which the statistician/data scientist translates the statistical results into meaningful qualitative answers or conclusions to the original problem. Put together to mirror the typical flow of a project, $Q_1Q_2Q_3$ becomes a straightforward workflow or lifecycle for statistics and data science practice, a framework for students and practitioners that is understandable, memorable, and easy to implement. Put simply, this workflow is: What is the problem? (Q_1), What did you do? (Q_2), and What does it mean? (Q_3).

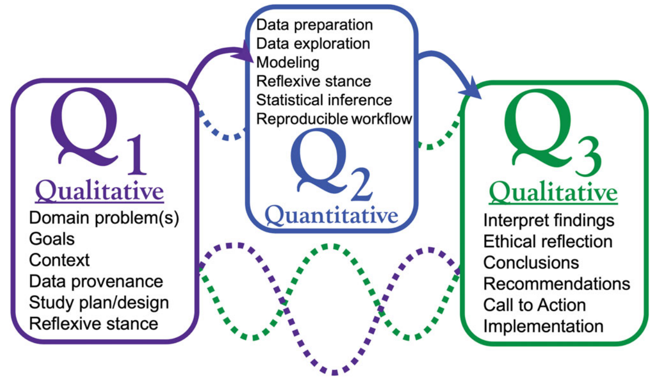

Figure 1. The $Q_1Q_2Q_3$ workflow for effective statistics and data science collaborations iterates between the qualitative and quantitative stages of a project, often non-sequentially.

While statistics and data science collaborations are typically initiated in the order of $Q_1Q_2Q_3$, the stages are interconnected and the process is not always linear, as shown in Figure 1. For example, when interpreting findings (Q_3), the collaboration team may need to revisit the original goals and context of the problem (Q_1) or conduct additional quantitative analyses (Q_2). Similarly, results of exploring data in Q_2 may necessitate updating the research questions in Q_1. We formally define $Q_1Q_2Q_3$ as a three-stage workflow for applied statistics and data science projects, consisting of the interrelated Qualitative (Q_1), Quantitative (Q_2), and Qualitative (Q_3) components of a project.

An often used lifecycle in statistics education is the PPDAC model (Problem-Plan-Data-Analysis-Conclusions) (MacKay and Oldford 1994; Wild and Pfannkuch 1999). Those familiar with it may see straightaway how it maps onto $Q_1Q_2Q_3$: defining the Problem and creating a Plan for data collection belong in Q_1, as do the qualitative aspects of Data collection (e.g., understanding the sources and context of data). Managing, cleaning, and tidying Data and the Analysis of data belong in Q_2. Finally, Conclusions form a part of Q_3.

In Q_1, the statistician/data scientist creates shared understanding with a domain expert (i.e., the "owner" of the original problem) about the goals and context of the project (Vance et al. 2022a), including what the domain problem is, why this problem is important or interesting, how the eventual solution may be implemented ethically in practice, and how the data have been or will be collected. In our experience, statistics and data science collaborators—especially those new to the field—typically rush through or even skip Q_1 to get to Q_2, where they have the most training and feel more comfortable. This is a mistake because the context of the domain problem and how data were collected affects all subsequent outcomes, including (but not limited to) potential ethical infractions and the appropriateness of the statistical methods used (Brown et al. 2018). For example, a statistician who does not give Q_1 its due diligence is at risk of making a Type III error, that is, the error committed by providing the right answer to the wrong question (Kimball 1957).

In Q_2, the statistician/data scientist applies quantitative techniques to analyze the data to answer the Q_1 domain question(s). During Q_2, analysts prepare the data (e.g., clean and tidy the data), explore and summarize data (e.g., visualize the data), formulate models and assess them, perform statistical inference, and summarize quantitative findings in ways that respect ethical guidelines, such as assuring data privacy. Crucially, part of the Q_2 stage is determining analytically and quantitatively whether the data collected are appropriate for solving the problems from Q_1. For example, a Q_2 analysis might suggest that a variable not fully discussed in Q_1 is an important moderating or mediating variable or that necessary modeling conditions have been violated. The analyst may need to clarify additional Q_1 issues with the domain expert before returning to Q_2.

The final stage of our workflow, Q_3, is the concluding qualitative component of a collaborative project. To complete this stage, the statistician/data scientist creates a shared understanding with the domain expert about how the results or findings from Q_2 answer the questions from Q_1 and how the Q_2 findings can be summarized into clear recommendations congruent with the domain expert's language (Gal and Ograjenšek 2016). Statisticians and data scientists who desire to make a positive impact and transform evidence into action for the benefit of society should also work with the domain expert to provide recommendations and a plan for action to implement these recommendations responsibly and ethically based on evidence from Q_2 (Olubusoye et al. 2021; Vance and Smith 2021).

An analogy for the $Q_1Q_2Q_3$ workflow is the "hero's journey" in which an interdisciplinary collaboration is like an epic adventure, and both statistician/data scientist and domain expert are the heroes. The hero's journey consists of three stages, Departure, Initiation, and Return (Campbell 2008), which parallel the three stages of $Q_1Q_2Q_3$:

1. *Departure (Q_1)*.The heroes live in the ordinary world of the status quo and receive a call to adventure. The call to adventure is the problem that the domain expert presents to the statistician/data scientist.
2. *Initiation (Q_2)*.The heroes traverse a threshold to a new world where they face trials. In our analogy, the new world is the world of quantitative reasoning, and the trials are quantitative analyses.
3. *Return (Q_3)*.The heroes again traverse the threshold between worlds, returning to the ordinary world with gained treasure. The treasure is qualitative knowledge that demands a call to action to help others and benefit society.

The heroes are transformed and gain wisdom through their journey. Likewise, in effective collaborations, the statistician/data scientist and domain expert gain knowledge and expertise in both the research domain and statistics/data science.

The $Q_1Q_2Q_3$ stages can be seen in an example collaborative data science research project predicting where enslaved people came from within Africa before they were forcibly shipped across the Atlantic Ocean (Wiens et al. 2022).

1. *Q_1*: Armed conflicts (wars) in the Kingdom of Oyo (modern-day Nigeria) from 1817–1836 resulted in 121,000 enslaved people sent in slave ships to the Americas. While historians have a good record of where the enslaved *went* across the Atlantic, they have no records of where they were from *within* Africa. Led by the project's historian, the team compiled data on historical trade routes and the time, place, and severity of armed conflicts in Oyo.
2. *Q_2*: Led by the statisticians, the team used kriging and a Markov decision process to simulate the capture and transport of the enslaved to ports of departure. They aggregated the simulations to predict the conditional probabilities of the likely origin locations of the enslaved, which were then visualized on maps of the region (see bit.ly/Oyoorigins).
3. *Q_3*: These maps help historians better understand the history of Africa and the entire Atlantic world, whereby the ocean connects, rather than disconnects, Africa, the Americas, and Europe. The paper concludes with a call to action for historians to generate new data about the transatlantic slave trade and to collaborate with data scientists to analyze these data to gain valuable historical insight.

Understanding the Q_1 historical context of how people became enslaved due to internal conflicts/wars and how they were transported to ports of sale was crucial for the statisticians to develop appropriate Q_2 statistical models and algorithms. Similarly, understanding the Q_2 models and their limitations was vital for the team to make appropriate Q_3 conclusions and recommendations. Completing all stages of $Q_1Q_2Q_3$ enabled the collaborators to maximize the potential impact of their work.

3. Implementing $Q_1Q_2Q_3$ in Statistical Collaborations

In our experience, most statistics and data science collaboration projects span multiple meetings, and $Q_1Q_2Q_3$ is a useful workflow for the *project*, not a structure for an individual meeting. To structure individual *meetings*, we recommend implementing the POWER structure (Alzen et al. 2024b; Zahn 2019). The "W" of the five-part POWER meeting structure stands for "Work," and the $Q_1Q_2Q_3$ workflow can guide this work. During initial meetings, the collaborative statistician/data scientist typically works with the domain expert to create shared understanding of the domain expert's goals, problem, and data (Q_1). In a second meeting, they typically work on explaining a statistical analysis (Q_2) or asking great questions to create an appropriate model for the data. Later, or during a third meeting, they may focus on interpreting the results of the statistical analyses, developing recommendations, and initiating a plan to implement their recommendations (Q_3).

Throughout a project, the statistician/data scientist will benefit from adopting the attitude of a collaborative relationship (Halvorsen et al. 2020) in which the statistician considers themselves to be on the same team as the domain expert—just one of many experts in the room—such that the statistician succeeds when the domain expert succeeds. The collaborative statistician believes that $Q_1Q_2Q_3$ is not about showing off their quantitative expertise, rather, it is about helping the domain expert make a good decision and transforming results into action for the benefit of society (Vance et al. 2022c).

With the extraordinary potential of statistics and data science to have positive impact comes a responsibility to engage in ethical practice (Kimball 1957). The $Q_1Q_2Q_3$

workflow can be used to weave ethics into all aspects of data science practice because $Q_1Q_2Q_3$ naturally fosters a reflexive stance, which is self-reflection that considers the effect of the researchers/analysts on what is being investigated and the research outcomes (D'Ignazio and Klein 2020). According to Tanweer et al. (2021, p. 13), "A reflexive stance acknowledges that subjectivity and bias are not aberrations that can ever be fully eradicated from research but inherent aspects of human inquiry that should be acknowledged and accounted for." By placing importance on both qualitative and quantitative aspects of projects, $Q_1Q_2Q_3$ users have designated times to pause and consider at key decision points how their own assumptions, experiences, and relationships influence the collaborative project. A reflexive stance or ethical reflection is built into each stage of $Q_1Q_2Q_3$. We provide details on implementing $Q_1Q_2Q_3$ below.

3.1. Implementing Q_1

To successfully implement Q_1, the statistician/data scientist must create shared understanding about the domain issues that will impact the Q_2 analyses and the Q_3 interpretations, recommendations, and plans for action. That is, the statistician should verify with the domain expert that they both know how these qualitative domain issues may affect the analyses and conclusions (Vance et al. 2022a). Several authors have described what these important qualitative issues are. Peterson et al. (2022) provided eight guidelines, a template, and a table of 23 vital questions for creating a scope of work for a collaborative project, that is, guidance for Q_1 before conducting Q_2 analyses. Cressman and Sharp (2022) detailed how to create a statistical analysis plan by asking 47 primary and secondary questions of the domain expert before proceeding to conduct Q_2 analyses.

Similar in spirit to the Q_1 questions from Peterson et al. (2022) and Cressman and Sharp (2022), and inspired by the "Heilmeier Catechism" (DARPA 2024), Table 1 lists the questions about seven qualitative aspects of a project and related ethical reflection questions we educate our students to ask about at the beginning (Q_1) of their collaborative projects. Statistics and data science collaborators are expected to verify with the domain expert in writing their understanding of the answers (see Appendix A) and engage in personal ethical reflection. We provide further explanation of these questions below.

3.1.1. Questions 1 and 2: Understanding the Problem and Why It's Interesting

In our experience, creating shared understanding around the first two questions of understanding the domain problem and why it is important or interesting and to whom (i.e., who are the stakeholders who care about the problem, and why do they care?) will provide a strong motivation for the statistician/data scientist to become engaged in the project. If the statistician is not interested in this problem, we urge them to invent a plausible reason it is compelling, that is, to find a lens through which to view the problem that makes it attractive. If that is impossible, or if the problem does not align with their moral or ethical values, we suggest referring the domain expert to a more suitable colleague. We find that we can do better statistics and data science and make deeper contributions to the domain when we, too, care about the project. Additionally, conversations around these first two questions can help build strong relationships. According to Vance (2020), making a deep contribution and creating a strong relationship with the domain expert are *the* two ultimate goals of any collaboration. Asking the domain expert three cycles of why they are interested in the project can reveal important context about the project and strengthen the relationship (Vance et al. 2022a).

3.1.2. Question 3: Using the Eventual Solution

Understanding how an eventual solution may be used or implemented in practice and reflecting on who might benefit or be harmed by the solutions can guide the statistician/data scientist to choose the most appropriate quantitative methods analytically and ethically. For example, a very high stakes "life or death" project may demand the most sophisticated, rigorous analysis possible because of the potential risk/harm of getting the anal-

Table 1. Seven qualitative questions for Q_1, with follow-ups and ethical reflection.

Question	Sub-questions or follow-ups	Ethical reflection
1. What is the domain problem?		Does the problem align with my moral values? Who might benefit (from the solution)? Who might be harmed?
2. Why is the problem important or interesting?	To whom is it important or interesting?	
3. How will the eventual solution be used?		
4. What potential data could solve the domain problem?	What data, if it were available and accessible, would help answer the underlying questions?	What are the ethical issues surrounding the potential data (e.g., privacy, consent)?
5. The Five Ws and one H of the actual	a. What data were collected? b. Who or what collected the data? c. Why, and for what purpose, were the data originally collected? d. When were the data collected? e. Where were the data collected? f. How were the data collected?	In which ways might the data be biased (i.e., not representative of
6. What may be the qualitative relationships between variables, for variables both observed and unobserved?		Could we make causal claims? What warnings against causal interpretations might we need?
7. Which types of statistics or data science analyses or techniques might be most useful to the domain expert? Which would not be useful?		

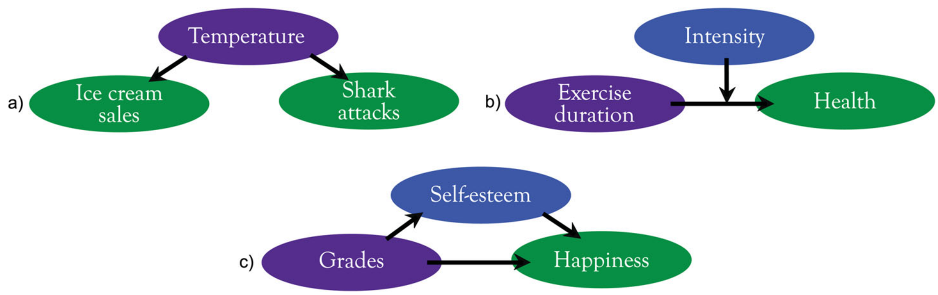

Figure 2. Example causal diagrams with the primary predictor of interest in the bottom left, the dependent variable/outcome in the bottom right, and the (a) confounder, (b) effect modifier, or (c) mediator on top. These diagrams are helpful tools for the Q_1 stage of collaboration and understanding when causal claims can (or cannot) be made.

ysis "wrong." On the other hand, a simple *t*-test or descriptive summary may be the best analysis for other projects. Stallings (2014) wrote about the errors collaborative statisticians make when they conduct an overly complicated analysis that is hard for the domain expert to understand or explain to others and is therefore not used. He concluded that simpler analyses are often more impactful than complex ones. However, an *overly* simple analysis might misstate reality (Scanlan 2014) or disparately benefit or harm certain sub-populations.

3.1.3. Questions 4 and 5: Potential Data and Real Data

Discussing potential data—before discussing the actual data—is an exercise that can help clarify the domain problem (e.g., perhaps the actual underlying domain problem is that the available data are imperfect, inadequate, or unethically collected). It may be that the solution to the domain problem becomes simplified by collecting new data and that the domain expert and data scientist overlook this by focusing only on data already collected. On the other hand, the impossibility of collecting certain data or incorporating existing data into the analyses can illuminate the intricacies of the domain problem and help guide the collection or analysis of alternative data (Keller et al. 2020).

Since most *data* are numbers *with context* (Cobb and Moore 1997), a collaborative statistician/data scientist must learn about the context of the data from the domain expert. As Barrowman (2018) wrote:

> *How data are construed, recorded, and collected is the result of human decisions—decisions about what exactly to measure, when and where to do so, and by what methods. Inevitably, what gets measured and recorded has an impact on the conclusions that are drawn* (p. 130).

Understanding what the actual data are, how they were collected, for what original purpose, and their potential biases (i.e., data provenance) will help the statistician appropriately and ethically explore, visualize, model, and analyze the data and interpret the results (Vance et al. 2022b).

3.1.4. Question 6: Potential Causal Relationships between Variables

Understanding the relationships between observed and unobserved variables is a precursor to the Q_2 activity of modeling the data, as it is essential for identifying confounders, mediators, and effect modifiers (a.k.a. moderators or interactions). A helpful strategy for answering this question is to work with the domain expert to draw a causal diagram with arrows depicting the relationships between variables (Pearl 1995). The three diagrams in Figure 2 depict simple examples of causal diagrams for a confounder, effect modifier, and mediator. In our experience, most domain experts actually want to make causal claims (Hernán 2018). Reflecting on the permissibility of such claims during Q_1 can help guide Q_2 analyses and Q_3 interpretations.

3.1.5. Question 7: Techniques That Will be Most Useful

Finally, discussing at the beginning of a project (Q_1) which types of analyses will be most useful to the domain expert is a key component of creating a scope of work (Peterson et al. 2022) or a statistical analysis sketch (Cressman and Sharp 2022). Are certain techniques frowned upon by the domain expert or their domain? Could new statistics or data science methods be developed to answer the domain question more usefully? Statisticians and data scientists have the potential to drive collaborative projects to develop new technical methods tailored to better answer the particular domain questions (Banks 2023). However, if the domain expert does not understand or cannot defend the choice of a specific statistical technique, that technique may not be useful, and a simpler or more familiar method may need to be used. If the simpler method is not appropriate for the problem and the data, the more advanced method should be used in Q_2 with sufficient time spent explaining the new method to the domain expert in Q_3.

3.1.6. Strategies for Asking These Q_1 Questions

Sometimes a barrage of questions from the statistician can be construed as "impertinent" (Lurie 1958), and the domain expert may resist thoroughly discussing the information relevant to Q_1 because they feel it is an inefficient use of time or unnecessary for the statistician to complete their task. We recommend using the strategies of "asking great questions" (see Vance et al. 2022d) to use these questions as an opportunity to strengthen the interpersonal relationship. One of the strategies is to preface questions with their intent. For example, to start a discussion about the experimental design (Question 5, part f), the statistician may say, "So I can better understand the experiment and appropriately model the data, how were the treatments assigned?" By prefacing questions with the intent, the statistician explains to the domain expert why it is worth their time to discuss the Q_1 components of the problem. Another strategy to encourage discussion about Q_1 aspects of the project is to include the Q_1 questions in the meeting agenda, especially when it is the

Table 2. Six aspects of a Q_2 quantitative analysis, with guiding and reflective questions.

Aspects of Q_2	Guiding questions	Self-reflexive stance
Data preparation	In which ways might my decisions/choices be biased?	
Data exploration	Am I trying to confirm what I think I already know? What evidence is in the data for the opposite of what I believe?	
Modeling	Which type of model best answers the Q_1 question(s)? How successful and sensitive is the model?	What alternative models may be more appropriate?
Assumptions/conditions	What assumptions/conditions are necessary for my model?	Have I tested, reported, and interpreted them honestly and meaningfully?
Inference	What statistical inferences are appropriate?	Am I just p-hacking?
Reproducible workflows	*In what ways am I ensuring that the analysis is reproducible?*	

first meeting between the statistician and domain expert. For an example initial meeting agenda, see bit.ly/gdoccollabtemplate.

3.2. Implementing Q_2

Statisticians and data scientists are likely most familiar and most comfortable with the Q_2 stage of $Q_1Q_2Q_3$ as this stage is often the sole focus of standard statistics or data science courses. Even though Q_2 is a technical, quantitative stage, nuanced choices are needed to decide which methods to apply and how to evaluate their success and interpret their results. Here, we emphasize six aspects of quantitative analyses and highlight strategies and resources to make these choices. In Table 2, we provide questions to guide quantitative analysts through Q_2 and reflect on the ethics of their decisions.

3.2.1. Data Preparation

Data are never raw (Barrowman 2018); many choices go into creating the dataset that an analyst first sees, and these should be discussed with the domain expert in Q_1. Similarly, in the Q_2 phase, choices need to be made to prepare the data since data are rarely in an ideal form for exploration and modeling. Data scientists usually need to clean, tidy, or otherwise wrangle the data (Wickham et al. 2023). Examples of decisions that might affect the conclusions and recommendations later in Q_3 are how to deal with missing data, outlying values, or transformations of variables. Other practical decisions, such as renaming column variables or recoding Likert-scale factors into numbers, might not influence the analysis outcomes, but should still be documented in the analyst's code to aid transparency and reproducibility. During the data preparation tasks, it can be useful to ask oneself some reflective questions: for example, is my treatment of missing data scientifically consistent, that is, is it based on the domain context of the data and how they were collected? In what ways might my choices to include/exclude outliers affect the results? In general, in which ways are my decisions based on my prior beliefs or knowledge about the data? Making our implicit assumptions about the data *explicit* by documenting (in the code) the reasons for our data preparation decisions can serve as a check against unintentionally biasing the data.

3.2.2. Data Exploration

Exploratory data analysis (EDA)—that is, making plots, exploring subgroups, examining variation and covariation of the variables to better understand the data—is a key component of nearly every analysis (Wickham et al. 2023). Sometimes, the analyst will be working on a pre-registered project in which the analysis methods and models are specified before the data have been collected. Even in this case, the analyst will ideally have helped write the pre-registered statistical analysis plan based on their experience with and exploratory analyses of similar, prior data. "Getting our hands dirty with the data" can illuminate possibilities for modeling the data or reveal inadequacies in the data that could be remedied by appropriate cleaning or wrangling of the data. In other words, data exploration is intertwined with data preparation and modeling.

During this part of Q_2, one should be curious about and skeptical of the data, both indicators of a self-reflexive stance. Analysts should ask questions of the data and use EDA to answer them; for example, ask "What do I want to learn more about?" followed-up by "How could this be misleading?" (Wickham et al. 2023, chap. 10.3). Asking oneself these questions and answering them with EDA can ward against a type of confirmation bias in which the questions we ask of the data can lead us to reaffirm what we already think we know (Klayman 1995; Vance and Smith 2021). For example, one can ask oneself: Am I trying to confirm what I think I already know? What evidence is in the data for the opposite of what I believe? In other words, one shouldn't just go looking in the data for what one wants or expects to find.

3.2.3. Choosing and Assessing the Appropriate Model

A guiding question for choosing a model is to ask which type of model is best aligned with the goals and data discussed in Q_1 and best answers the Q_1 question(s), that is, will the model be used for *description*, *prediction*, or *counterfactual prediction* (aka *causal inference*)? (Hernán et al. 2019) Categorizing models this way also aids in explaining models to domain experts and enables analytic pivots when needed. For example, if it has been determined in Q_1 that the data collected are not sufficient to answer the original (predictive or causal) research questions, a useful, reflective question is, "What alternative models may be more appropriate?" The analyst can still help the domain expert describe the data with descriptive models using summary statistics, subgroup comparisons, visualizations, etc. There may still be an important narrative to be told with the existing data, including associations within the data, even if the data are not generalizable to make causal claims or out-of-sample predictions. For a more advanced approach to selecting models during Q_2, we refer readers to Dwivedi (2022) and Dwivedi and Shukla (2020), which contain highly detailed classifications and checklists for selecting models, conducting analyses, and reporting results.

A second guiding question is, "How successful and sensitive (to assumptions/conditions) is the model?" Model success can be quantified based on metrics meaningful to the domain expert and the statistician (e.g., out-of-sample predictive accuracy, log-likelihood fit). Yu and Kumbier (2020) recommend data scientists to consider, measure, and report in their Q_2 analysis the predictive accuracy of the model, the model or algorithm's computational efficiency and scalability, and the stability or sensitivity of the analysis, that is, the likelihood that another researcher making alternative, appropriate decisions would obtain similar conclusions. Again, we advise asking oneself, "What alternative models might be more appropriate to better describe, predict, or counterfactually predict the data and answer the Q_1 research/business/policy questions?"

3.2.4. Assumptions and Conditions Necessary for the Chosen Model

Rossman and Witmer (2019) argue that model assumptions that can be tested should be called *conditions* (not assumptions), and conditions of models should always be tested. If a necessary condition of the model is violated, then another model should be considered. For example, if the relationship between a predictor X_1 and an outcome Y is not linear, then perhaps a spline model should be considered instead of a linear regression. When assumptions about the model cannot be tested with the data, then these *assumptions* (not testable conditions) should be made explicit in the code and subsequent analysis reports. If a key result changes with a small adjustment to the model assumptions/conditions, then the evidence for the result is weak (Anderson 2019). This step of the workflow is a great place to reflect on ethical questions: "What assumptions/conditions are necessary for my model, and have I tested the conditions? Have I reported and interpreted the assumptions and conditions honestly and meaningfully?" Such reflection can help fulfill the ASA Ethical Guidelines for Statistical Practice (Committee on Professional Ethics (COPE) of the American Statistical Association Committee et al. 2022), Specifically, ethical practitioners are "transparent about assumptions made in the execution and interpretation of statistical practices, including methods used, limitations, possible sources of error, and algorithmic biases [and convey] results or applications of statistical practices in ways that are honest and meaningful" (p. 3).

3.2.5. Inference

Statistical inference is not a fourth category of model, but is useful in descriptive, predictive, or counterfactually (causally) predictive models (Hernán et al. 2019). Unfortunately, there are many, many ways to bungle inference; for summaries and examples, see the ASA Statement on Statistical Significance and P-Values (2016) and Greenland et al. (2016). Fortunately, Wasserstein et al. (2019) and associated references provide an abundance of advice on how to do inference right. Our advice is to follow the $Q_1Q_2Q_3$ workflow, that is to say, let the original Q_1 questions guide one's use of inference in Q_2, keeping in mind the entire modeling process (including deviations from study protocols, data preparation, multiple hypothesis tests, etc.) when interpreting, communicating, and using the results in Q_3. Adopting a self-reflexive stance may lead the ethical analyst to question whether one is "just p-hacking," that is, testing for many associations without prior hypotheses or testing hypotheses after the results are already known (Brownstein et al. 2019).

3.2.6. Ensuring That the Analysis is Reproducible

To provide transparency regarding how computational results are produced and to help ensure that our work is reproducible, we recommend statisticians and data scientists habitually create annotated scripts or code notebooks for each project to document choices made during Q_2, reasons for applying the chosen methods, associated computer code used, and the analytical results. These code scripts should be written for ease of re-executability (Krafczyk et al. 2021) and published in a public repository such as GitHub or the Open Science Framework (Horton and Stoudt 2024). Such practice not only helps the research be more reproducible by others—including a domain expert who wants to reuse the code when new data are collected—it can help the individual statistician/data scientist resume work on the project after a hiatus because, as Hadley Wickham (2019) said, "*Every* data science project is a collaboration. At the very least it is a collaboration between you right now and you six months from now."

3.2.7. Note on Designing Studies or Experiments

For collaborative projects that have not yet produced data, statisticians and data scientists can add tremendous value in Q_2 by recommending an efficient study design that ensures sufficient power. In addition, pre-registering the study design, analysis plan, and primary outcome(s) and hypotheses discussed in Q_1 can be a major step toward improving the transparency, reproducibility, and efficiency of scientific research (Munafò et al. 2017).

3.3. Implementing Q_3

Q_3 is the final stage of $Q_1Q_2Q_3$, where a statistician/data scientist can make the most impact on the outcomes of the project. In Q_3, the statistician effectively translates the quantitative evidence from Q_2 into answers to the Q_1 research, business, or policy questions. In our experience, too often the contribution of a statistician ends with the production of a *p*-value. A *p*-value is not itself an answer to a research question nor should *p*-values alone be used to make a decision (American Statistical Association 2016). To achieve greater impact, statisticians or data scientists should summarize the Q_2 findings into recommendations and create a plan with the domain expert to implement these recommendations. As a guide for completing Q_3, we outline seven questions in Table 3 for statisticians and data scientists to ask themselves and the domain expert. Throughout Q_3 and the entire $Q_1Q_2Q_3$ workflow, an ethical data scientist will reflect on the ethical implications of the work and the plan for action. For example, it is important to consider how results can be reported without compromising the privacy or safety of study participants (Thornton et al. 2022).

In the Q_3 stage, it is crucial that the statistician/data scientist and domain expert have clear communication and develop

Table 3. Seven qualitative questions for Q_3.

Question	*Ethical reflection*
1. Qualitatively, what do the results mean?	
2. What are the constraints, limitations, assumptions of the quantitative methods; what conditions are necessary for the results to be valid?	
3. How can we visually display and communicate the results of the analysis in a way the domain expert and their stakeholders will understand?	
4. What are the answers to the domain expert's questions and how are these relevant to the project's goals?	*What are the ethical implications of this work and our plan for action?*
5. What are the practical implications of the answers and Q_2 findings?	
6. What actions do we recommend should be taken?	
7. What is our plan for action to implement these recommendations?	

shared understanding about the results (i.e., questions 1–6 above). To achieve shared understanding, we suggest the following five strategies. First, results should be explained in language that is accessible to the domain expert. Second, the statistician should be intentional about providing the domain expert multiple opportunities to speak and ask questions throughout meetings. For example, periodically checking in with the domain expert by asking "What can I clarify?" or "What can I explain further?" after each explanation can be helpful. Third, the statistician can ask the domain expert how they would explain the results in their own words. For example, the statistician might say, "I want to make sure that I have explained these results clearly. To check that I have done that, could you tell me how *you* would explain these results to your stakeholders?" Fourth, the statistician and domain expert can brainstorm together what actions or implications the results suggest. Lastly, all results, code, findings, conclusions, implications, and recommendations should be written up clearly in a document and shared between the domain expert and statistician.

In our experience, statisticians and data scientists rarely fully complete Q_3, usually stopping short of co-developing a plan for implementing recommendations. This is a missed opportunity to turn evidence into action for the benefit of society. Indeed, former ASA president Phillip Hauser stated that statisticians have an *obligation* to play significant roles in society (Hauser 1963).

To achieve greater impact, we recommend being especially mindful about understanding at the beginning of the project (i.e., question 3 from Q_1) how the eventual solution might be used, and then following through with the domain expert throughout the project to develop a plan for action (i.e., question 7 from Q_3). Even better would be to meet with policy makers/implementers before or during the design of the study to understand what kind of evidence would be persuasive (Hartman et al. 2020), a practice recommended for clinical trial designs by The Global Health Network (2023). Statisticians and data scientists can also choose to pursue collaborative projects that have the most potential to result in positive impact. The ASA's Influencing Discovery, Exploration, and Action (IDEA) Forum is a notable step toward data scientists and statisticians making an impact. The forum helps build strategic partnerships with leaders from academia, government, and industry to grow the influence of statisticians and drive innovative solutions to problems of the world community (Ensor and LaLonde 2023).

4. Teaching $Q_1Q_2Q_3$ for Statistics and Data Science Collaborations

We teach the $Q_1Q_2Q_3$ workflow in a 15-week-semester "Statistics and Data Science Collaboration" course to senior undergraduate statistics and data science (SDS) majors, professional masters SDS students, SDS Ph.D. students, and Ph.D. students in other quantitative fields. We use Team-Based Learning to flip the classroom (Vance 2020) and apply a community of practice approach to teaching and learning (Alzen et al. 2024a). Students are engaged in our five-stage pedagogical process: Prepare, Practice, Do, Reflect, Mentor (LeBlanc et al. 2022). Students *Prepare* by reading about $Q_1Q_2Q_3$ outside of class, *Practice* applying the $Q_1Q_2Q_3$ workflow on homework and in-class exercises, implement (*Do*) $Q_1Q_2Q_3$ on three real collaboration projects, *Reflect* on their learnings, and provide feedback and coaching to peers (*Mentor*) on their use of $Q_1Q_2Q_3$. See Alzen et al. (2024b) for a description of typical projects, which range from straightforward to complex; see Vance and Pruitt (2022) for some guidance on creating a mechanism for generating real projects; see UNKNOWN (in press) for an example of university students consulting on real projects with high school students. Below we present student learning outcomes for our collaboration/capstone students and provide details on how we teach and assess the whole $Q_1Q_2Q_3$ workflow, with references to instructional materials in the appendices.

4.1. Student Learning Outcomes (SLO) for Collaboration Students

A student taught the $Q_1Q_2Q_3$ workflow in a collaboration and/or capstone course should be able to:

1. Understand the $Q_1Q_2Q_3$ workflow (e.g., retrospectively identify the Q_1, Q_2, and Q_3 components of a project)
2. Implement $Q_1Q_2Q_3$ prospectively on their own projects
   (a) With the domain expert, create shared understanding of the context of the problem, questions, and data (Q_1) before beginning the quantitative analysis (Q_2)
   (b) Appropriately prepare, model, and analyze the data (Q_2)
   (c) Translate the findings from their Q_2 analyses into meaningful answers (conclusions) to the original questions (Q_3)
   (d) Use $Q_1Q_2Q_3$ to guide ethical reflection and analysis

(e) [Advanced] Based on the conclusions generated with the domain expert, develop recommendations and a plan for action (Q_3)

3. Self-assess how well they did on the Q_1, Q_2, and Q_3 aspects of their project(s)
4. Assess how well another statistician/student did on the Q_1, Q_2, and Q_3 aspects of a project and provide helpful feedback.

4.2. How We Teach and Assess the $Q_1Q_2Q_3$ Workflow

4.2.1. Prepare

Students prepare to implement $Q_1Q_2Q_3$ in actual collaborations starting in week 2 of the course by reading Vance and Smith (2019), which provides an overview of essential collaboration skills, including a brief introduction to $Q_1Q_2Q_3$. In week 3, students read Section 2, 3, 7, and 8 of this paper (with the option to read the entire paper) to gain a detailed understanding of the $Q_1Q_2Q_3$ workflow. Previous to this paper being drafted, students read a four-page handout about $Q_1Q_2Q_3$ (Vance 2019). Then in week 5, students read about Type III errors (i.e., giving the "right" answer to the wrong question) in the classic paper "Errors of the Third Kind" (Kimball 1957). Near the end of the semester (week 13), students are exposed to a lecture and discussion of an application of $Q_1Q_2Q_3$ for oral presentations we call "QMatrix" (Vance et al. 2024).

4.2.2. Practice

Early in the semester (weeks 2–3), students observe a real collaboration meeting and—as part of a homework assessment—summarize the Q_1, Q_2, and Q_3 aspects they observed (assessing SLO 1). Initial collaboration meetings often only get through Q_1, which is a surprising and valuable lesson for students. During weeks 3–4, students practice asking Q_1 questions during in-class role-plays of meetings.

During week 5, students complete a homework assignment in which they generate an example of a Type III error, ideally one they made or noticed someone else make. Alternatively, the example could be from popular culture, that is, a Type III error committed by someone in a book or movie. In class on the day the assignment is due, students work in their (permanent, 3–5 person) teams to discuss Type III errors and generate strategies to avoid them. See Appendix B for these in-class team exercises. $Q_1Q_2Q_3$ is not explicitly mentioned in the exercises, yet most teams identify "creating shared understanding of Q_1 issues" as a good strategy for avoiding Type III errors. This exercise can be used as another assessment of SLO 1.

In weeks 13 or 14, students attend any quantitative presentation (e.g., a statistics departmental seminar) and complete a homework assignment in which they track how much time the presenter speaks about Q_1, Q_2, and Q_3 (see Appendix C), which is yet another way to assess SLO 1.

4.2.3. Do

During weeks 4–16 (including finals week 16 if necessary), students work in pairs on three real collaboration projects. Students use a meeting notes template (bit.ly/gdoccollabtemplate), which reminds students about asking Q_1 questions and reflecting about ethics throughout $Q_1Q_2Q_3$. Students submit reports about their initial project meetings and final project reports, which are essentially, "summarize Q_1, Q_2, and Q_3; reflect on ethics; self-assess Q_1, Q_2, and Q_3." See Appendix A for the project report prompts. The initial meeting and final reports are used to assess SLO 2 and 3. During weeks 15–16, pairs of students use the $Q_1Q_2Q_3$ workflow to help organize their final project presentations.

4.2.4. Reflect

Reflection is woven into most homework assignments (including final project presentations) and in-class exercises. Specifically, students are asked to reflect on ethical issues of their collaboration projects and self-assess how well they did on Q_1, Q_2, and Q_3 (see Appendix A), which is used for assessing SLO 3. Sprinkled throughout the course are discussions of active projects, which primarily focus on Q_1 and Q_2, and provide opportunities for students to reflect on $Q_1Q_2Q_3$. During a module on ethics in Week 11, students create a personal ethical workflow—a tool to remind them about what ethical issues to think about and when during the $Q_1Q_2Q_3$ workflow. Students provide feedback to each other and then create a team ethical workflow, combining the best parts of the individual workflows (see Appendix D).

4.2.5. Mentor

Students provide peer feedback on many assignments throughout the course. Students typically work in pairs on their collaboration projects, which provides peer mentoring opportunities. In addition to projects and in-class project discussions, students may mentor their peers during occasional (1–2 times in the latter half of the semester) Video Coaching Feedback Sessions (VCFS). During VCFS, students provide feedback and coaching (i.e., mentoring) to their peers about many aspects of the video-recorded collaboration meeting, including $Q_1Q_2Q_3$. Near the end of the semester, students review another team's final project report and submit their feedback to the students and the instructor, who uses this assignment to assess SLO 4.

5. Assessing the Success of the $Q_1Q_2Q_3$ Approach

On a short, outside-of-class survey administered during the last week of the semesters, four cohorts of students in the "Statistics and Data Science Collaboration" course were asked a quantitative, six-point Likert-scale question, "How valuable do you think the $Q_1Q_2Q_3$ approach will be throughout your career?" The most recent three cohorts of students were also asked a qualitative follow-up question, "Has learning about the $Q_1Q_2Q_3$ approach strengthened your ability to successfully complete statistics or data science collaboration projects? Why or why not?" Thirteen responses to this question are used to highlight our findings and, in Section 7, to ground our discussion. The data and code that support the findings of this study are openly available in the Open Science Framework at https://osf.io/69npm.

Most students completed the survey and informed consent forms approved by the University of Colorado Boulder's Institutional Review Board (Protocol 18–0554). The number of students answering the quantitative question was n = 21 of 21 stu-

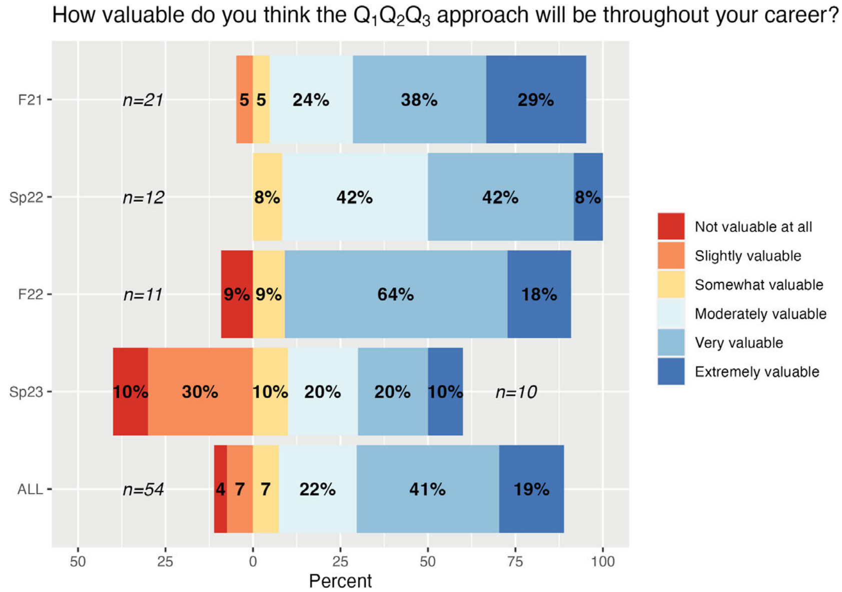

Figure 3. The diverging (stacked) bar chart shows the percentage of responses on a six-point Likert scale for four semesters of students (and the totals) in a collaboration course.

dents in Fall 2021, n = 12 of 12 students in Spring 2022, n = 11 of 15 students in Fall 2022, and n = 10 of 20 students in Spring 2023. Most of these students (85%, 28 of 33 respondents excluding Fall 2021 students) also responded to the qualitative follow-up question. The population of students was 57% undergraduates (predominantly seniors) and 43% graduate students, roughly evenly split between applied masters and MS/Ph.D. students. The demographics of the respondents was not recorded.

The instructor during Fall 2021 and Spring and Fall 2022 was this paper's first author. The instructor during Spring 2023 was not one of this paper's authors. Students in that semester engaged in ostensibly the same pedagogical process as described in Section 4, though they only collaborated on one or two projects (instead of three), did not engage in the "QMatrix" or ethical workflow activities, and did not participate in Video Coaching Feedback Sessions. Figure 3 shows the students' responses to the quantitative question across four semesters, including the total responses. Responses were on a six-point scale of Not valuable at all – Slightly valuable – Somewhat valuable – Moderately valuable – Very valuable – Extremely valuable.

In each semester, at least one student (19%) was enthusiastic about $Q_1Q_2Q_3$, responding that this workflow would be "Extremely valuable" in their careers. A representative qualitative follow-up response from this group from Spring 2023 was: "Learning about the $Q_1Q_2Q_3$ approach has helped me to spend more time understanding the domain problem from a qualitative standpoint and making sure I ask the right questions before moving on with the quantitative portion of the work."[1]

Most students appreciated $Q_1Q_2Q_3$, responding that it will be "Very valuable" (41%) or "Moderately valuable" (22%) in their careers. A representative "Very Valuable" follow-up response from a student in Spring 2023 was:

> *I think that emphasizing the domain problem like in $Q_1Q_2Q_3$ is extremely important because it's hard to remember to focus on the domain problem sometimes when you're more focused on the math and the data. Personally, I'm motivated to do data science because of the domain problem, so I've always placed equal emphasis on both. But I think that this approach was really beneficial for my classmates.*

A student who responded "Moderately valuable" in Spring 2022 wrote: "Yes, I think this approach has strengthened my ability to complete collaboration projects because it has forced me to slow down and fully understand the research goals before trying to interpret data."

One student in each semester responded that they expected $Q_1Q_2Q_3$ to be "Somewhat valuable" (7%) in their careers. A student from Spring 2022 who seemed to forget that Q_1 is Qualitative, Q_2 is Quantitative, and Q_3 is Qualitative responded "Somewhat valuable" and wrote: "Not really – I think for my purposes I would change the order to Q_2, Q_1, Q_3. I think it's more natural to discuss the bigger qualitative picture before diving into the quantitative components of the project."

The rest of the students (11%) did not expect $Q_1Q_2Q_3$ to be especially valuable in their careers, responding that it would be "Slightly valuable" (7%) or "Not valuable at all" (4%). Based on their responses to the qualitative follow-up question, most of these students seemed to have a faulty understanding of $Q_1Q_2Q_3$ and/or confused it with a structure for individual meetings rather than as a workflow or lifecycle for an entire project. A student from Spring 2023 who responded "Slightly valuable" wrote: "It's good to have a structure somewhat, but flexibility is better." A student from Fall 2022 who responded "Not valuable at all" also confused $Q_1Q_2Q_3$ with a meeting structure: "It really didn't, in my opinion how a meeting should be structured should be organized based on the problem at hand. $Q_1Q_2Q_3$ makes me feel like I'm just going through the motions rather than actually learning or doing anything."

[1]Every student quote in this paper comes from a unique student, i.e., the 13 qualitative quotes are from 13 students.

We used Wilcoxon Rank Sum tests in R (R Core Team 2023) to test for pairwise differences in the distributions of student responses by semester. We found that students in Spring 2023 rated $Q_1Q_2Q_3$ about one unit less valuable than students from the other three semesters and that students in Fall 2021 rated $Q_1Q_2Q_3$ about one unit more valuable than students in Spring 2022 (and also Spring 2023). However, only the difference in the distribution of student ratings between Fall 2021 and Spring 2023 was statistically discernible ($p = 0.012$). All other p-values were greater than 0.1. The lower ratings and response rate (50%) for students in Spring 2023 (compared to 92% across the other semesters) may have been due to the new instructor's inexperience with or lack of enthusiasm for teaching $Q_1Q_2Q_3$.

6. Teaching $Q_1Q_2Q_3$ in a Variety of Courses

Among the instructor-facing, "Ten Recommendations for Statistics and Data Science" proposed by the GAISE Revision Committee (2024), we believe that teaching $Q_1Q_2Q_3$ can facilitate the first seven, which are listed below (*in italics*). We briefly describe the relevance of $Q_1Q_2Q_3$ in achieving them.

1. *Teach statistics and data science as iterative processes of gleaning insights from data to inform evidence-based decisions.* This is the main point of the $Q_1Q_2Q_3$ workflow.
2. *Emphasize effective written and oral communication of results from data, with attention to the scope and limitations of conclusions.* This is a focus of Q_3.
3. *Focus on conceptual understanding rather than algebraic manipulation and formulas.* Emphasizing how Q_2 methods fit with Q_1 and Q_3 can help achieve this understanding.
4. *Integrate real data with a context and purpose throughout the course. Select data that are meaningful and engaging to the students.* Understanding the context and purpose of data is the Q_1 stage of every problem.
5. *Encourage multivariable thinking.* Embedded within important Q_2 methods.
6. *Incorporate software/apps to explore concepts and work with data.* This is a great way to teach Q_2 methods.
7. *Emphasize responsible and ethical conduct in the collection and use of data and in their analysis.* Ethical thinking is woven throughout $Q_1Q_2Q_3$.

Because of these attractive features of $Q_1Q_2Q_3$ for statistics and data science education, we advocate for a more general and widespread use of $Q_1Q_2Q_3$ beyond collaborative/capstone courses. In this section, we provide suggestions and examples for how to use $Q_1Q_2Q_3$ to teach and assess a variety of student learning outcomes (SLOs) in introductory, project-based, and technical courses.

6.1. Teaching $Q_1Q_2Q_3$ in Introductory Courses

We believe that $Q_1Q_2Q_3$ should be explicitly taught to introductory students because, if an instructor is teaching statistics or data science as an investigative process as recommended by the GAISE College Report ASA Revision Committee (2016), $Q_1Q_2Q_3$ is a straightforward workflow for that process. Before students learn methods, they should learn that the context of the problem and the provenance of the data informs which methods are most appropriate to use. Teaching $Q_1Q_2Q_3$ early in the course will help students recognize how Q_2 methods fit into the investigative process, that is, Q_2 methods are to be applied to compile quantitative evidence for Q_1 questions to produce relevant Q_3 answers.

Teaching $Q_1Q_2Q_3$ early can also help introduce and discuss ethical issues throughout the course (and in subsequent courses) because of the many opportunities for ethical reflection built into the $Q_1Q_2Q_3$ workflow. Simply asking students, "What ethical issues do you think are pertinent in Q_1 [or Q_2 or Q_3] of this problem?" can lead to a discussion of which ethical principles may apply and what the resolution could or should be.

If the course has weekly lab assignments or a final project with real data and context, we think $Q_1Q_2Q_3$ should be taught as a template and used as a grading rubric for the lab assignments/project reports. See Appendices A and E for examples and the next subsection for details.

We have identified four SLOs that we expect a student to be able to achieve by the end of an intro course:

1. Understand and apply the $Q_1Q_2Q_3$ workflow
2. Analyze and evaluate the interplay of Q_1, Q_2, and Q_3 in a statistics/data science problem
3. Use $Q_1Q_2Q_3$ as a tool for ethical reflection
4. (if applicable) Use $Q_1Q_2Q_3$ to guide their lab assignments or projects (see Section 6.2. for details).

To teach and assess SLO 1, an instructor could ask students to identify the parts of a problem presented in class or in a textbook. For example, ask students to identify: which aspects of Q_1 are present and which are missing; what the Q_2 methods are and what assumptions or conditions are stated or implied; and what the answer is (Q_3) and what the implications of this answer may be. Note that there are not necessarily strict lines dividing these stages. Being able to effectively apply $Q_1Q_2Q_3$ also means that students—when presented with a problem and data—should be able to pose questions about the data or the problem (Q_1), conduct analyses of the data (Q_2), and interpret the results (Q_3), which may, in fact, be the overall learning goal of all statistics and data science courses.

To help teach SLO 2, an instructor could provide examples of how the context of a problem (Q_1) affects the suitability of a method (Q_2) and how the statistical results of an analysis (Q_2) answer the original question (Q_3). To assess this SLO, students could be presented with the details of a statistical analysis (e.g., from a news article with a provocative headline). Students would then be asked to analyze and evaluate how the Q_1 context and the Q_2 quantitative methods used justify—or fail to justify—the Q_3 conclusions and recommendations.

For SLO 3, rather than teaching ethics as an "extra step" by focusing an assignment solely on ethical thinking, ethics could be built into the use of $Q_1Q_2Q_3$. For example, when presented with a provocative news article and statistical analysis, students could be asked to evaluate the interplay of Q_1, Q_2, and Q_3 and also identify and analyze potential ethical issues in each stage.

6.2. Teaching $Q_1Q_2Q_3$ in Projects-Based Courses

In our experience, courses with projects (i.e., any applied assignment ranging from weekly lab assignments to a final project) are ideal venues for teaching $Q_1Q_2Q_3$ and for using the workflow as a template and rubric for project reports. In this type of course, we would expect students to achieve the SLOs in Section 6.1 about learning and applying $Q_1Q_2Q_3$ and the following three SLOs based on their project(s), which could also be used as a basic template for the project assignment. Students in a projects-based course should be able to:

1. Summarize the Q_1, Q_2, and Q_3 aspects of their project, that is, create a report with three sections: What is the problem? [Q_1], What did you do? [Q_2], and What does it mean? [Q_3]
2. Reflect on and discuss how their choices in Q_1 and Q_2 may have affected their conclusions in Q_3
3. Analyze and evaluate at least one ethical quandary, implication, or decision made in each stage of their project (i.e., in Q_1, Q_2, and Q_3).

We recommend teaching students the $Q_1Q_2Q_3$ workflow (see Section 6.1 for suggestions) and then building $Q_1Q_2Q_3$ into students' project assignments. Appendix A provides an example assignment for an open-ended project with domain-expert-generated data. An assignment in which the instructor provides the data may be more appropriate for most introductory and intermediate course projects. In this case, the instructor can choose to also provide Q_1 context or provide a reference for students to learn more about the data (e.g., who collected the data for what original purpose?). Students would then be required to identify or discuss important Q_1 aspects that affect Q_2 and Q_3 and reflect on ethics. We provide an example of a lab assignment in Appendix E and at https://osf.io/69npm in which the instructor provided the data set and the Q_1 context.

One of our lessons learned is that introductory and intermediate students (and, in fact, most statisticians) are not equipped to implement Q_3 recommendations on their own. To effectively transform evidence into action, statisticians and data scientists must almost always collaborate with domain experts. Note that we do not advocate for introductory or intermediate students to collaborate with real domain experts. In our experience, effective interdisciplinary collaboration requires sufficient technical and collaboration skills rarely seen below the level of an upper-division collaboration course.

6.3. Using $Q_1Q_2Q_3$ to Teach Technical Courses

Instructors of technical courses that do not require projects (e.g., "the mathematics of machine learning") may not feel the need to explicitly teach the $Q_1Q_2Q_3$ workflow. Even in these cases, we think that using and talking about $Q_1Q_2Q_3$ in class can help an instructor teach quantitative methods more engagingly. In such a course that focuses on Q_2 methods, we propose three SLOs. Throughout the course, for each Q_2 method presented, students should be able to:

1. Describe contexts for when the method is appropriate (Q_1)
2. Provide (or answer questions about) technical/mathematical details of the method (Q_2)
3. Provide examples of how the results might be useful (Q_3).

Instructors would teach SLOs 1 and 3 by providing Q_1 contexts for each Q_2 method they introduce and describing how the Q_3 results might be useful. This aligns with the conclusion of Ograjenšek and Gal (2016) that the qualitative "need to know" as determined by Q_1 and Q_3 drives all research and is why students are learning the Q_2 methods. The instructor's focus can still be on the intricacies of the Q_2 methods (to teach SLO 2) while providing enough Q_1 and Q_3 context so that no student should ever wonder, 'When are we ever going to use this?'

Instructors could assess SLO 2 through standard quizzes/tests/assignments and all three SLOs together through an assignment that could masquerade as a midterm or final review. Instructors could provide a list of methods that have been covered in class (or ask students to generate such a list). Then students would be asked to summarize the Q_2 aspects of each method (i.e., what the method does and any "important" technical/mathematical details of the method in addition to necessary conditions/assumptions), describe the appropriate Q_1 contexts for the method (i.e., for which types of data are the methods appropriate), and generate a Q_3 example of the usefulness of the method. Note that this type of assignment focused on methods could also be helpful in many technical courses that also have applied projects (e.g., statistical learning).

The following is a brief example of how $Q_1Q_2Q_3$ could be used in a probability course: Students are assigned to answer questions of interest by calculating probabilities from a distribution (e.g., $P(x > 5) =?$). Then they are asked to answer the same questions using a different distribution of their choosing. Finally, they are asked to comment on what about the problem context (Q_1) led them to choose the original distribution, why using the second distribution is wrong (testing their understanding of the Q_2 details about distributions), and how this incorrect Q_2 choice might lead to different Q_3 conclusions.

7. Discussion

Our goal for using $Q_1Q_2Q_3$ is to remind/teach statistics and data science practitioners and students about three high-level concepts:

1. Create shared understanding of the context of the problem, questions, and data (Q_1) *before* beginning the quantitative analysis (Q_2).
2. Be sure to translate the findings from the Q_2 analyses into meaningful answers (conclusions) to the original questions (Q_3).
3. For increased impact of your work, develop recommendations with the domain expert and a plan for action (Q_3).

A quote from a Fall 2022 student illustrates these first two points:

> *[$Q_1Q_2Q_3$] gives a good framework to go about the work. It is beneficial to really flesh out the qualitative goals initially to make sure that subsequent quantitative work is worth the time (i.e., even answers the question). Having a thorough understanding of what the results mean is also beneficial.*

Below we discuss how $Q_1Q_2Q_3$ fits with and is supported by education literature, some benefits of using and teaching $Q_1Q_2Q_3$, and some challenges and future work.

7.1. $Q_1Q_2Q_3$ is Supported by Education Literature

Despite many recommendations and calls for statistics and data science instructors to incorporate qualitative thinking in the first-year course and throughout the statistics major curriculum (Gal and Ograjenšek 2010; Horton 2016; Lee et al. 2022), the modern student's statistics and data science education is still predominantly focused on the theory, methods, and applications of statistics and data science (i.e., Q_2). Ograjenšek and Gal wrote: "Continuing to exclude qualitative methods and thinking from statistical training may hamper our ability to effectively collaborate and communicate with diverse clients and audiences about the contribution and value of statistics and statistical investigations" (2016, p. 176). Using $Q_1Q_2Q_3$ in the classroom and on projects appropriately emphasizes the importance of the *qualitative*, as highlighted by the following two student quotes (from Fall 2022 and Spring 2022, respectively):

> *Yes, [$Q_1Q_2Q_3$ is valuable] because it forced me to ensure I understood the discipline and research problems before starting the code/statistics portion.*

> *[$Q_1Q_2Q_3$] is helpful for figuring out what exactly you want to know before diving into existing data/data collection.*

In addition, we propose that $Q_1Q_2Q_3$ effectively merges the qualitative and quantitative into one coherent framework for statistics and data science as called for in the literature. For example, Tanweer et al. (2021, p. 3) argues that "quantitative and qualitative approaches should be seen as complementary, mutually reinforcing, and co-constitutive of data science when applied to the production of social knowledge." Meng also emphasized this point that "Qualitative and quantitative thinking co-exist and interact at all research stages, and therefore, there should be an on-going emphasis of this interplay in all statistical education and beyond" (2016, p. 187). Meng called for codifying the interplay between the qualitative and quantitative as the "Q-q dynamic, with Q representing the thinking process receiving more emphasis at a particular stage. When and which 'q'–quantitative or qualitative–deserves to be capitalized will depend on the context" (2016, p. 187). A student's comment from Fall 2022 shows how the "Q-q dynamic" can be deduced from study of $Q_1Q_2Q_3$: "I believe the [$Q_1Q_2Q_3$] approach is essential, and it reminds everyone that statistical collaboration demands the marriage of quantitative and qualitative skills."

We believe that Q_1, which encapsulates understanding the context of the problem and data before applying quantitative methods, is very well supported in the statistics and data science education literature (Cobb and Moore 1997; Leonelli 2019; Lee et al. 2022). From the broader education literature comes the sociocultural theory of education, which argues that social and cultural contexts heavily influence what we know and how we learn (Vygotsky and Cole 1978; Alfred 2002). This theory justifies emphasizing the context of problems when teaching statistics and data science because different contexts (Q_1) can induce different analyses (Q_2) and different interpretations of results (Q_3) (Lave 1988). Therefore, emphasizing Q_1 in the classroom and on collaborative projects is fundamental for training in the practice of statistics and data science.

The need to communicate Q_3 conclusions, recommendations, and possible actions is also well-founded in the literature (Lee et al. 2022). For example, on the results of applied statistics projects, Gal and Ograjenšek (2016, p. 204) wrote: "Conclusions have to be presented or reported to clients or stakeholders in ways that they understand and find easy to make sense of, and be congruent with their 'policy language.'" Petocz and Reid (2010) described the work of professional statisticians as almost always involving the communication of the results of statistical procedures and investigations and often requiring the education of the domain expert or user of statistical results. They concluded: "While statistics is essentially a quantitative discipline, it contains a necessary core of qualitative components" (p. 272).

7.2. Benefits of Using and Teaching $Q_1Q_2Q_3$

7.2.1. $Q_1Q_2Q_3$ is Simple Yet Versatile

A student from Fall 2022 commented:

> *"Learning about $Q_1Q_2Q_3$ approach has strengthened my ability to complete data science projects. It is a simple, versatile framework that balances background information, the technical information, and the conclusion. It keeps your project 'grounded' to what is most important to discuss."*

In our view, the value of $Q_1Q_2Q_3$ does indeed stem from its simplicity, ease of use, universal applicability to every data science project, and its versatility. Another student from Fall 2022 commented: "[$Q_1Q_2Q_3$] gives me a nice flow to go about completing a project. The order is very logical and is set up in such a way that it is difficult to make any type of error."

In the classroom, we make use of the versatility of $Q_1Q_2Q_3$ to present concepts such as understanding the provenance of data, adopting a reflexive stance in our data science work, thinking ethically throughout a project, ensuring the reproducibility and replicability of our work, reporting assumptions, and validating conditions of our analyses. $Q_1Q_2Q_3$ also opens up more possibilities and points of references for teaching Q_2 topics to domain experts or students. For example, statistics and data science collaborators and educators can address Q_2 on its own in an exposition of a specific statistical method, in relation to Q_1 while discussing what qualities make a dataset and context appropriate for use with the method, or in relation to Q_3 while discussing the limitations of what could be concluded if that method were used.

The $Q_1Q_2Q_3$ workflow also enables instructors and students to create a shared vocabulary to discuss all aspects of a statistics or data science project. This point was underscored by a Fall 2022 student: "[$Q_1Q_2Q_3$] has given me a named format with which I can identify the stage of a project I might enter at and it gives me a way to describe the progress of a collaboration to the DE [domain expert]."

7.2.2. $Q_1Q_2Q_3$ is Useful for Teaching Ethics in the Classroom

According to a report from the National Academies of Sciences, Engineering, and Medicine (2018), ethical thinking should be woven throughout data science courses and the curriculum. How to integrate ethical thinking into statistics and data science education has been much discussed (Halvorsen et al. 2020). One recent proposal is to teach the Data Science Ethos Lifecycle (Boenig-Liptsin et al. 2022), which is a six-stage data science workflow centered around ethical thinking. The first two stages—Question/Problem Identification and Data Discovery—map onto Q_1, the middle two stages—Exploratory Data Analysis and Modeling—map onto Q_2, and the final two stages—Interpretations/Conclusions/Predictions and Communication/Dissemination/Decision Making—map nicely onto Q_3. We contend that $Q_1Q_2Q_3$ is a simpler alternative to weave ethics into data science practice in the classroom. Teaching $Q_1Q_2Q_3$ in introductory courses, using it for subsequent projects-based courses, and reinforcing it in capstone courses can be a way to integrate ethics into projects at every stage of the investigation (Colando and Hardin 2024) and into the SDS curriculum from top-to-bottom (Baumer et al. 2022).

Specifically, Colando and Hardin (2024) discuss 15 frequently taught ethics topics from a sample of data science ethics courses. As a complete workflow for statistics and data science practice, we believe that $Q_1Q_2Q_3$ would be useful for teaching *all* of those topics. We highlight four of these ethical topics (*in italics*) and briefly discuss how $Q_1Q_2Q_3$ could be used when teaching the topic:

- *Alignment.* Adopting a reflexive stance throughout the $Q_1Q_2Q_3$ workflow can help us determine whether and how our moral values are *aligned* with our data science practices and how our individual views and experiences might be affecting our work.
- *Transparency.* Creating reproducible workflows, explicitly documenting the Q_1 context of a problem and Q_2 model assumptions/conditions, and discussing how decisions made in Q_1 and Q_2 affect the Q_3 interpretations are ways to uphold the ethical tenet of *transparency* in our work.
- *Causation.* Question 6 of Q_1 in Table 1 is really about trying to understand factors that might enable a domain expert to make a *causal claim* or to identify factors to consider in the Q_2 model building and Q_3 model interpretation stages that may warn against making *causal claims*
- *Consequences.* Thinking about the *consequences* of deploying an algorithm in Q_3 is part of the plan of action we advocate for a data science collaborator to co-create with the domain expert.

7.2.3. $Q_1Q_2Q_3$ Encourages and Facilitates Effective Collaboration

A data scientist should never work alone, because, in our experience, statisticians and data scientists typically lack the domain expertise to appropriately understand all of the nuances of the Q_1 context of a complicated problem and lack the professional stature to implement meaningful Q_3 conclusions and recommendations. The $Q_1Q_2Q_3$ workflow encourages, if not requires, data scientists to collaborate with domain experts who do understand Q_1 and could implement Q_3. Hernán et al. (2019) support this view and also advocate for interdisciplinary collaboration by arguing: "Data scientists without subject-matter knowledge cannot conduct causal analyses in isolation: They don't know how to articulate the questions …and they don't know how to answer them …" (p. 48).

In our experience, typical class projects (e.g., Kaggle-style contests) in which students find an "interesting" dataset and develop their own research questions fall short on their desired impact on students' learning and society at large. While these projects can be used to exercise students' technical skills, too often the research questions and results are ultimately meaningless because the students usually lack the domain expertise required to ask novel questions, appropriately interpret the results, and make practical recommendations for action. Instead, educators can teach the $Q_1Q_2Q_3$ workflow and encourage genuine collaborations in undergraduate capstone or graduate courses. In such collaborations, a domain expert originates the problem; answers the students' questions about the Q_1 context and the Q_3 relevance, recommendations, and plans for action; and is in a better position than an SDS student to use the project's results for societal good. For introductory or intermediate students who are not yet ready for real collaborations with domain experts, emphasizing Q_1 and Q_3 aspects of their projects can provide more meaningful experiences and prepare them to have more impact later in their career when they do collaborate with real domain experts.

Perhaps the most practically beneficial result of teaching the $Q_1Q_2Q_3$ workflow is that it reduces pressure on students who mistakenly believe that they must have ready Q_2 answers for the domain expert during their initial collaboration meeting. Understanding that the initial meeting should be primarily focused on Q_1 enables students to think more deeply about the project's context to determine the appropriate statistics and curbs potential coercions from the domain expert to provide a rushed statistical solution. Rather than feeling "put on the spot," students can consult their books, peers, faculty members, and other resources to devise a thoughtful Q_2 approach after the initial meeting has concluded. This empowers students to find a proper solution, not just an expedient one. As Banks (2023) stated in his 2023 ASA Deming Lecture:

> *At this level of detail [of current problems in industrial statistics], there are no general theorems. Every application requires a bespoke solution. And that requires someone with statistical training to sit with domain experts to figure out the particularities of a problem*" (p. 19).

7.3. Challenges and Future Work

The quantitative and qualitative results in Section 5 show that some students might not fully understand or appreciate the usefulness of $Q_1Q_2Q_3$, even when taught by an experienced and enthusiastic instructor. On the other hand, in our experience, many students do immediately grasp the benefits of $Q_1Q_2Q_3$ and instantly connect it with their practice of statistics and data science. Hence, a dual challenge for instructors is to: A) use and present $Q_1Q_2Q_3$ throughout a course in ways that encourage more students to embrace the $Q_1Q_2Q_3$ workflow to guide their applied work and B) accept that $Q_1Q_2Q_3$ might not resonate with some students and provide alternative workflows or lifecycles for these students. Lee et al. (2022) summarize more than a dozen workflows and synthesize them into a six-stage process that could be very useful for K-12 data science education. All of those alternatives to $Q_1Q_2Q_3$ mix the qualitative and quantitative aspects of data science practice.

Another challenge (and future work) is to improve how Q_3 is taught to instill an "evidence-to-action mindset" in students. We posit that emphasizing the evidence-to-action aspect of Q_3 will result in statisticians and data scientists having greater impact on their projects. In our personal experience, statisticians in academia are rarely involved in the "action" resulting from their modeling and analyses. As a result, the full value of an analysis is not realized for most collaborative projects. We hope that this paper will inspire more statistics and data science educators to experiment with innovative pedagogy about how to teach Q_3. We believe that the principles of "evidence communication" in Blastland et al. (2020), which advocate for a clear separation between information and opinion and encourage researchers to adopt a reflexive stance about their intentions to inform or persuade, will be valuable tools for data scientists who want their work to achieve greater impact for societal good.

A limitation of our study of the efficacy of our methods for teaching the $Q_1Q_2Q_3$ workflow is that it relied on students' self-reflection rather than an objective evaluation of students' $Q_1Q_2Q_3$ skills. Also, in most of the semesters we assessed, the instructor was an author of this paper, which potentially biased the results due to the power dynamic. Future work may include designing objective evaluations and investigating the usefulness of $Q_1Q_2Q_3$ at independent validation sites outside of the authors' home institutions with a variety of instructors.

Finally, we believe that the $Q_1Q_2Q_3$ workflow is a useful advancement toward developing a theory of applied statistics, which was called for by Mallows (1998). We hope that others can expand upon $Q_1Q_2Q_3$ in ways to make applied statistics and data science easier to teach and statistics and data science easier to correctly apply.

8. Conclusion

Every statistics or data science project or investigation must mix qualitative and quantitative thinking to succeed. We have designed $Q_1Q_2Q_3$ to be a workflow for statistics and data science practitioners and students to remind them of the important qualitative aspects of their projects and to weave ethical thinking into all aspects of the statistics and data science investigative process. The $Q_1Q_2Q_3$ workflow explicitly emphasizes the importance of the qualitative context of a project at its beginning (Q_1) and the qualitative interpretation (Q_3) of quantitative findings (Q_2) near its end. In our experience, statisticians and data scientists readily understand quantitative thinking, and most need to practice their qualitative thinking. $Q_1Q_2Q_3$ helps to do that. Making students explicitly aware of the important work that comes before and after a quantitative analysis is essential for any course in which students conduct projects, including capstone courses and consulting or collaboration courses.

We provided guidance for students and practitioners to implement the $Q_1Q_2Q_3$ workflow on collaborative projects and strategies for teaching it in collaboration/capstone, introductory, projects-based, and technical courses. We also presented data demonstrating the effectiveness of teaching $Q_1Q_2Q_3$ to beginning collaborators. Teaching this simple yet versatile workflow provides value in the classroom and on collaborative projects. We recommend that $Q_1Q_2Q_3$ be taught early in introductory statistics and data science courses, during projects-based courses, and again in capstone or collaboration courses. $Q_1Q_2Q_3$ can be part of a theory of applied statistics that provides statisticians and data scientists with a framework for successfully contributing to research, policy, and business decisions leading to action for the benefit of society.

Acknowledgments

We thank the domain experts and collaborators—especially Ellery Galvin and Hilary Brumberg—of the Laboratory for Interdisciplinary Statistical Analysis (LISA) and the LISA 2020 Global Network. We also thank the anonymous reviewers and the editors who provided helpful feedback on previous versions of the manuscript.

Data Availability Statement

The deidentified data and code that support the findings of this study are openly available on the Open Science Framework at https://osf.io/69npm/.

Disclosure Statement

The authors report there are no competing interests to declare.

Funding

This material is based upon work supported by the National Science Foundation under Grant No. 1955109, Grant No. 2022138, and Grant No. 2044384 for the projects, "IGE: Transforming the Education and Training of Interdisciplinary Data Scientists (TETRDIS)," "NRT-HDR: Integrated Data Science (Int dS): Teams for Advancing Bioscience Discovery," and "CODE:SWITCH: Integrating Content and Skills from the Humanities." This work was also partially supported by the United States Agency for International Development under Cooperative Agreement #7200AA18CA00022.

ORCID

Eric A. Vance https://orcid.org/0000-0001-5545-1878

References

Alfred, M. V. (2002), "The Promise of Sociocultural Theory in Democratizing Adult Education," *New Directions for Adult and Continuing Education*, 2002, 3–14. *https://doi.org/10.1002/ace.74*.

Alzen, J. L., Cho, K. J., and Vance, E. A. (2024a), "Teaching Statistics and Data Science Collaboration via a Community of Practice," *Journal of Statistics and Data Science Education*, 0, 1–23. *https://doi.org/10.1080/26939169.2024.2422821*.

Alzen, J. L., Trumble, I. M., Cho, K. J., and Vance, E. A. (2024b), "Training Interdisciplinary Data Science Collaborators: A Comparative Case Study," *Journal of Statistics and Data Science Education*, 32, 73–82. *https://doi.org/10.1080/26939169.2023.2191666*.

American Statistical Association (2016), "ASA Statement on Statistical Significance and P-Values," *The American Statistician*, 70, 131–133. *https://doi.org/10.1080/00031305.2016.1154108*.

American Statistical Association (2023), ASA Statement on the Role of Statistics in Data Science and Artificial Intelligence. *American Statistical Association*. Available at *https://www.amstat.org/docs/default-source/amstat-documents/the-role-of-statistics-in-data-science-and-artificial-intelligence.pdf*.

Anderson, A. A. (2019), "Assessing Statistical Results: Magnitude, Precision, and Model Uncertainty," *The American Statistician*, 73, 118–121. *https://doi.org/10.1080/00031305.2018.1537889*.

Banks, D. (2023), "The Future of Industrial Statistics," *Chance*, 36, 19–24. *https://doi.org/10.1080/09332480.2023.2264738*.

Barrowman, N. (2018), "Why Data Is Never Raw," *The New Atlantis, Center for the Study of Technology and Society, 129–135*. Available at *https://www.jstor.org/stable/26498249*

Baumer, B. S., Garcia, R. L., Kim, A. Y., Kinnaird, K. M., and Ott, M. Q. (2022), "Integrating Data Science Ethics Into an Undergraduate Major: A Case Study," *Journal of Statistics and Data Science Education*, 30, 15–28. *https://doi.org/10.1080/26939169.2022.2038041*.

Blastland, M., Freeman, A. L. J., van der Linden, S., Marteau, T. M., and Spiegelhalter, D. (2020), "Five Rules for Evidence Communication," *Nature*, 587, 362–364. *https://doi.org/10.1038/d41586-020-03189-1*.

Boenig-Liptsin, M., Tanweer, A., and Edmundson, A. (2022), "Data Science Ethos Lifecycle: Interplay of Ethical Thinking and Data Science Practice," *Journal of Statistics and Data Science Education*, 30, 228–240. *https://doi.org/10.1080/26939169.2022.2089411*.

Brown, A. W., Kaiser, K. A., and Allison, D. B. (2018), "Issues with Data and Analyses: Errors, Underlying Themes, and Potential Solutions," *Proceedings of the National Academy of Sciences of the United States of America*, 115, 2563–2570. *https://doi.org/10.1073/pnas.1708279115*.

Brownstein, N. C., Louis, T. A., O'Hagan, A., and Pendergast, J. (2019), "The Role of Expert Judgment in Statistical Inference and Evidence-Based Decision-Making," *The American Statistician,* 73, 56–68. *https://doi.org/10.1080/00031305.2018.1529623*.

Campbell, J. (2008), *The Hero with a Thousand Faces* (3rd ed.), Novato, CA: New World Library.

Cobb, G. W., and Moore, D. S. (1997), "Mathematics, Statistics, and Teaching," *The American Mathematical Monthly*, 104, 801–823. *https://doi.org/10.1080/00029890.1997.11990723*.

Colando, S., and Hardin, J. (2024), "Philosophy within Data Science Ethics Courses," *Journal of Statistics and Data Science Education,* 0, 1–22. *https://doi.org/10.1080/26939169.2024.2394542*.

Committee on Professional Ethics (COPE) of the American Statistical Association (2022), "Ethical Guidelines for Statistical Practice," *American Statistical Association*. Available at *https://www.amstat.org/docs/default-source/amstat-documents/ethicalguidelines.pdf*

Cressman, K. A., and Sharp, J. L. (2022), "Crafting Statistical Analysis Plans: A Cross-Discipline Approach," *Stat,* 11, e528. *https://doi.org/10.1002/sta4.528*.

D'Ignazio, C., and Klein, L. F. (2020), *Data Feminism, Strong Ideas*, Cambridge, MA: MIT Press.

DARPA (2024), "The Heilmeier Catechism," *Defense Advanced Research Projects Agency*. Available at at *https://www.darpa.mil/work-with-us/heilmeier-catechism*.

Derr, J. (2000), *Statistical Consulting: A Guide to Effective Communication*, Pacific Grove, CA Duxbury Press.

Dwivedi, A. K. (2022), "How to Write Statistical Analysis Section in Medical Research," *Journal of Investigative Medicine*, 70, 1759–1770. *https://doi.org/10.1136/jim-2022-002479*.

Dwivedi, A. K., and Shukla, R. (2020), "Evidence-Based Statistical Analysis and Methods in Biomedical Research (SAMBR) Checklists According to Design Features," *Cancer Reports (Hoboken, N.J.)*, 3, e1211. *https://doi.org/10.1002/cnr2.1211*.

Ensor, K. B., and LaLonde, D. (2023), "The Influencing Discovery Exploration & Action Forum: An Initiative of the American Statistical Association," *CHANCE*, 36, 4–8. *https://doi.org/10.1080/09332480.2023.2179258*.

GAISE College Report ASA Revision Committee (2016), Guidelines for Assessment and Instruction in Statistics Education College Report 2016, *American Statistical Association*. Available at *https://www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege_Full.pdf*.

GAISE Revision Committee (2024), "Ten Recommendations for Statistics and Data Science," *College GAISE Revision*. Available at at *https://community.amstat.org/collegegaiserevision/home*.

Gal, I., and Ograjenšek, I. (2010), "Qualitative Research in the Service of Understanding Learners and Users of Statistics," *International Statistical Review*, 78, 287–296. *https://doi.org/10.1111/j.1751-5823.2010.00104.x*.

Gal, I., and Ograjenšek, I. (2016), "Rejoinder: More on Enhancing Statistics Education with Qualitative Ideas," *International Statistical Review*, 84, 202–209. *https://doi.org/10.1111/insr.12157*.

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., and Altman, D. G. (2016), "Statistical Tests, P Values, Confidence Intervals, and Power: A Guide to Misinterpretations," *European Journal of Epidemiology*, 31, 337–350. *https://doi.org/10.1007/s10654-016-0149-3*.

Halvorsen, K. T., Hanford, K. J., Vance, E. A., Wilson, J., and Zahn, D. (2020), "Transforming Your Stumbling Blocks into Stepping Stones," in *JSM Proceedings*, Alexandria, VA: American Statistical Association: Statistical Consulting Section, pp. 2523–2541.

Hartman, D., Heaton, P., Cammack, N., Hudson, I., Dolley, S., Netsi, E., Norman, T., and Mundel, T. (2020), "Clinical Trials in the Pandemic Age: What Is Fit for Purpose?," *Gates Open Research*, 4, 58. *https://doi.org/10.12688/gatesopenres.13146.1*.

Hauser, P. M. (1963), "Statistics and Society: Presidential Address," *Journal of the American Statistical Association*, 58, 1–12. *https://doi.org/10.1080/01621459.1963.10500829*.

Hernán, M. A. (2018), "The C-Word: Scientific Euphemisms Do Not Improve Causal Inference From Observational Data," *American Journal of Public Health*, 108, 616–619. *https://doi.org/10.2105/AJPH.2018.304337*.

Hernán, M. A., Hsu, J., and Healy, B. (2019), "A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks," *CHANCE*, 32, 42–49. *https://doi.org/10.1080/09332480.2019.1579578*.

Horton, N. J. (2016), "Discussion: Making Progress in a Crowded Market," *International Statistical Review*, 84, 179–181. *https://doi.org/10.1111/insr.12150*.

Horton, N. J., and Stoudt, S. (2024), "Editorial: Guidelines and Best Practices to Share Deidentified Data and Code," *Journal of Statistics and Data Science Education*, 32, 227–231. *https://doi.org/10.1080/26939169.2024.2364737*.

Keller, S. A., Shipp, S. S., Schroeder, A. D., and Korkmaz, G. (2020), "Doing Data Science: A Framework and Case Study," *Harvard Data Science Review*, 2. *https://doi.org/10.1162/99608f92.2d83f7f5*.

Kimball, A. W. (1957), "Errors of the Third Kind in Statistical Consulting," *Journal of the American Statistical Association*, 52, 133–142. *https://doi.org/10.2307/2280840*.

Klayman, J. (1995), "Varieties of Confirmation Bias," in *Psychology of Learning and Motivation*, eds. J. Busemeyer, R. Hastie, and D. L. Medin, New York: Academic Press. pp. 385–418. *https://doi.org/10.1016/S0079-7421(08)60315-1*.

Krafczyk, M. S., Shi, A., Bhaskar, A., Marinov, D., and Stodden, V. (2021), "Learning from Reproducing Computational Results: introducing Three Principles and the Reproduction Package," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379, 20200069. *https://doi.org/10.1098/rsta.2020.0069*.

Lave, J. (1988), *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life*, Cambridge, UK: Cambridge University Press.

LeBlanc, M., Rueegg, C. S., Bekiroğlu, N., Esterhuizen, T. M., Fagerland, M. W., Falk, R. S., Frøslie, K. F., Graf, E., Heinze, G., Held, U., Holst, R., Lange, T., Mazumdar, M., Myrberg, I. H., Posch, M., Sergeant, J. C., Vach, W., Vance, E. A., Weedon-Fekjaer, H., and Zucknick, M. (2022), "Statistical Advising: Professional Development Opportunities for the Biostatistician," *Statistics in Medicine*, 41, 847–859. *https://doi.org/10.1002/sim.9290*.

Lee, H., Mojica, G., Thrasher, E., and Baumgartner, P. (2022), "Investigating Data like A Data Scientist: Key Practices and Processes," *Statistics Education Research Journal*, 21, 3. *https://doi.org/10.52041/serj.v21i2.41*.

Leman, S., House, L., and Hoegh, A. (2015), "Developing a New Interdisciplinary Computational Analytics Undergraduate Program: A Qualitative-Quantitative-Qualitative Approach," *American Statistician*, 69, 397–408. *https://doi.org/10.1080/00031305.2015.1090337*.

Leonelli, S. (2019), "What Distinguishes Data From Models?," *European Journal for Philosophy of Science*, 9, 22. *https://doi.org/10.1007/s13194-018-0246-0*.

Love, K., Vance, E. A., Harrell, F. E., Johnson, D. E., Kutner, M. H., Snee, R. D., and Zahn, D. (2017), "Developing a Career in the Practice of Statistics: The Mentor's Perspective," *The American Statistician*, 71, 38–46. *https://doi.org/10.1080/00031305.2016.1255257*.

Lurie, W. (1958), "The Impertinent Questioner: The Scientist's Guide to the Statistician's Mind," *American Scientist*, 46, 57–61.

MacKay, R. J., and Oldford, W. (1994), *Stat 231 Course Notes Fall 1994*. Waterloo: University of Waterloo.

Mallows, C. (1998), "The Zeroth Problem," *The American Statistician*, 52, 1–9. *https://doi.org/10.1080/00031305.1998.10480528*.

Meng, X.-L. (2016), "Discussion: The Q-q Dynamic for Deeper Learning and Research," *International Statistical Review*, 84, 181–189. *https://doi.org/10.1111/insr.12151*.

Metzger, T. A., Pruitt, T. R., Alzen, J. L., Goshu, A. T., and Vance, E. A. (2023), "The Impact of Working in A Statistical Laboratory on the Statistics Student Collaborators," *Statistics Education Research Journal*, 22, 9. *https://doi.org/10.52041/serj.v22i3.584*.

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie Du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., and Ioannidis, J. P. A. (2017), "A Manifesto for Reproducible Science," *Nature Human Behaviour*, 1, 1–9. *https://doi.org/10.1038/s41562-016-0021*.

National Academies of Sciences, Engineering, and Medicine (2018), *Data Science for Undergraduates: Opportunities and Options*, Washington, DC: The National Academies Press.

Ograjenšek, I., and Gal, I. (2016), "Enhancing Statistics Education by Including Qualitative Research," *International Statistical Review*, 84, 165–178. *https://doi.org/10.1111/insr.12158*.

Olubusoye, O. E., Akintande, O. J., and Vance, E. A. (2021), "Transforming Evidence to Action: The Case of Election Participation in Nigeria," *CHANCE*, 34, 13–23. *https://doi.org/10.1080/09332480.2021.1979807*.

Pearl, J. (1995), "Causal Diagrams for Empirical Research," *Biometrika*, 82, 669–688. *https://doi.org/10.2307/2337329*.

Peterson, R. A., Hochheimer, C. J., Grunwald, G. K., Johnson, R. L., Wood, C., and Sammel, M. D. (2022), "Reaping What You SOW: Guidelines and Strategies for Writing Scopes of Work for Statistical Consulting," *Stat*, 11, e496. *https://doi.org/10.1002/sta4.496*.

Petocz, P., and Reid, A. (2010), "On Becoming a Statistician—A Qualitative View," *International Statistical Review*, 78, 271–286. *https://doi.org/10.1111/j.1751-5823.2010.00101.x*.

R Core Team (2023), *R: A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing.

Rossman, A. J., and Witmer, J. (2019), "Interview With Jeff Witmer," *Journal of Statistics Education*, 27, 48–57. *https://doi.org/10.1080/10691898.2019.1603506*.

Scanlan, J. P. (2014), "Race and Mortality Revisited," *Society*, 51, 328–346. *https://doi.org/10.1007/s12115-014-9790-1*.

Stallings, J. (2014), "Type IV Errors: How Collaboration Can Lead to Simpler Analyses," *Amstat News,* 24–25.

Tanweer, A., Gade, E. K., Krafft, P. M., and Dreier, S. (2021), "Why the Data Revolution Needs Qualitative Thinking," *Harvard Data Science Review*, 3. *https://doi.org/10.1162/99608f92.eee0b0da*.

The Global Health Network (2023), "DAC: Design, Analyze, Communicate," *DAC Trials*. Available at *https://dac-trials.tghn.org/*.

Thornton, S., Roy, D., Parry, S., LaLonde, D., Martinez, W., Ellis, R., and Corliss, D. (2022), "Towards Statistical Best Practices for Gender and Sex Data," *Significance*, 19, 40–45. https://doi.org/10.1111/1740-9713.01614.

Vance, E. A. (2019), "Content of Collaborations QQQ," Public Files for Collaboration in a Bag. Available at: *https://osf.io/z6wdg*.

Vance, E. A. (2020), "Goals for Statistics and Data Science Collaborations," in *JSM Proceedings*, Alexandria, VA: American Statistical Association: Statistical Consulting Section, pp. 2198–2209.

Vance, E. A., Alzen, J. L., and Smith, H. S. (2022a), "Creating Shared Understanding in Statistics and Data Science Collaborations," *Journal of Statistics and Data Science Education*, 30, 54–64. *https://doi.org/10.1080/26939169.2022.2035286*.

Vance, E. A., and Pruitt, T. R. (2022), "Statistics and Data Science Collaboration Laboratories: Engines for Development," in *Promoting Statistical Practice and Collaboration in Developing Countries*, Boca Raton, FL: Chapman and Hall/CRC, pp. 3–26.

Vance, E. A., and Smith, H. S. (2019), "The ASCCR Frame for Learning Essential Collaboration Skills," *Journal of Statistics Education*, 27, 265–274. *https://doi.org/10.1080/10691898.2019.1687370*.

Vance, E. A., and Smith, H. S. (2021), "Asking Great Questions: Part of a Theory of Communication in Interdisciplinary Collaborations," in *JSM Proceedings*, Alexandria, VA: American Statistical Association: Statistical Consulting Section, pp. 995–1008.

Vance, E. A., Glimp, D. R., Pieplow, N. D., Garrity, J. M., and Melbourne, B. A. (2022b), "Integrating the Humanities into Data Science Education," *Statistics Education Research Journal*, 21, 9. *https://doi.org/10.52041/serj.v21i2.42*.

Vance, E. A., Love, K., Awe, O. O., and Pruitt, T. R. (2022c), "LISA 2020: Promoting Statistical Practice and Collaboration in Developing Countries," in *Promoting Statistical Practice and Collaboration in Developing Countries*, Boca Raton, FL: Chapman and Hall/CRC, pp. 27–46.

Vance, E. A., Trumble, I. M., Alzen, J. L., and Smith, H. S. (2022d), "Asking Great Questions," *Stat*, 11. e471 *https://doi.org/10.1002/sta4.471*.

Vance, E., Trumble, I., and Macaluso, G. (2024), "Communicating Quantitative Information to the C-Suite." OSF. Available at *https://osf.io/wstvb/*.

Vygotsky, L. S., and Cole, M. (1978), *Mind in Society: Development of Higher Psychological Processes*, Cambridge, MA: Harvard University Press.

Wasserstein, R. L., Schirm, A. L., and Lazar, N. A. (2019), "Moving to a World Beyond '$p < 0.05$,'" *The American Statistician*, 73, 1–19. *https://doi.org/10.1080/00031305.2019.1583913*.

Wickham, H. (2019), "Keynote: What is Data Science?," in *ASA Conference on Statistical Practice*, NewOrleans, LA.

Wickham, H., Çetinkaya-Rundel, M., and Grolemund, G. (2023), *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*, Sebastopol, CA: O'Reilly Media.

Wiens, A., Lovejoy, H. B., Mullen, Z., and Vance, E. A. (2022), "A Modelling Strategy to Estimate Conditional Probabilities of African Origins: The Collapse of the Oyo Empire and the Transatlantic Slave Trade, 1817–1836," *Journal of the Royal Statistical Society Series A: Statistics in Society*, 185, 1247–1270. *https://doi.org/10.1111/rssa.12833*.

Wild, C. J., and Pfannkuch, M. (1999), "Statistical Thinking in Empirical Enquiry," *International Statistical Review*, 67, 223–248.

Yu, B., and Kumbier, K. (2020), "Veridical Data Science," *Proceedings of the National Academy of Sciences of the United States of America*, 117, 3920–3929. *https://doi.org/10.1073/pnas.1901326117*.

Zahn, D. (2019), *Stumbling Blocks to Stepping Stones*, Bloomington, Indiana: iUniverse.

Appendix A. Initial Collaboration Meeting and Final Project Report Prompt

Individual Assignment 7 A&B: LISA Collaboration Meetings Project #1

(1) Conduct an initial meeting and subsequent follow-up activities with your domain expert to help them solve their research problem(s) or answer their business or policy question(s).
(2) Please record your Zoom meeting after getting verbal permission from the Domain Expert. Upload the video to "Submit Videos Here" folder on OSF: Assignments/Submit Assignments Here/IA7 projects/Submit Videos Here
(3) Complete the write-up as instructed below. All students involved in the project should collaborate to create one IA7A document that is sent to the Domain Expert (via email) and to the course instructor (via OSF). The Domain Expert should comment on the document—making changes as necessary—and indicating that he or she agrees with what is written.
(4) Append your initial meeting notes to this document, either by copying and pasting or providing a link to your notes.
(5) IA7B will be due shortly after the completion of the project or by the end of the semester, whichever comes first. IA7B will be a final summary of the project, not just the initial meeting and subsequent follow-up.
(6) IA7AB is to be sent to the Domain Expert for comments and a statement that shared understanding has been achieved.

LISA Collaboration Project Report

Collaborators

Domain expert name:

Domain expert affiliation: (PhD student from AeroSpace Engineering)

Domain expert contact info:

Date of initial meeting:

Date(s) of additional meetings (for IA7B):

Date of report:

Video recording of initial meeting uploaded to OSF?: Yes/No. Filename:

Q_1: Qualitative

Q_1 is the project's initial Qualitative component, which sets the foundation for the Quantitative component of the project (Q_2) and the implementation of the solution (Q_3). Specifically, there are seven aspects of Q_1 relevant for every project.

1: What is the domain problem? Be sure to write down their overall research/business/policy goals and their specific scientific questions.
2: Why is this problem important or interesting? This should be answered individually by each collaborator on the project: Why is your domain expert's research/business/policy question interesting to you? If it's not interesting, make up a plausible reason.
3: How will the eventual solution be used? How they will use the answers to their research questions (i.e., what is their intended outcome of the research) and how will this help achieve the overall goal of the project?
4: What potential data could solve the domain problem? That is, what data, if it were available and accessible, would help answer the underlying research questions or guide the business or policy decisions? This is an important hypothetical exercise.
5: The actual data (only if data have already been collected)
   - a: What data have been collected?
   - b: Why were the data collected originally? (For what purpose?)
   - c: and d: When and where were the data collected?
   - e: Who or what collected the data?
   - f: How were the data collected? With what instrumentation/methods?
6: What may be the qualitative relationships between variables, for those observed and unobserved?
7: Which types of statistical analyses or techniques would be most useful to the domain expert? Which would not be useful?

Ethical Reflection questions (pick at least one question to reflect on in writing for IA7A):

(A) Does the problem align with my moral values?
(B) Who might benefit (from the solution)? Who might be harmed?
(C) What are the ethical issues surrounding the potential data (e.g., privacy, consent)?
(D) In which ways might the data be biased (i.e., not representative of our population of interest)? Are the data usable (e.g., obtained with appropriate consent, low risk of violating subjects' privacy)? How well do we understand the context of the data?
(E) Could we make causal claims? What warnings against causal interpretations might we need?

Q_2: Quantitative

Summarize the statistical collaborators' quantitative contribution/advice. Did the domain expert understand the statistics? This can be whatever (if any) quantitative contribution/advice you provided during the initial meeting or in a subsequent follow-up email. If there has not been any Q_2 advice so far, indicate your thoughts of potential directions for Q_2.

Q_3: Qualitative

Did the contribution/advice/solution answer the researchers' questions? Will it help the domain expert achieve his or her overall research goal(s)? Are there any practical constraints limiting the effectiveness of the proposed Q_2 statistical solution? What is the answer to the research question(s)? Note: it is uncommon for an initial meeting and follow-up activities to result in Q_3 conclusions or recommendations. If these have already occurred, please detail them here. If they have not occurred, just state that Q_3 has not yet occurred.

Individual Reflections (to be Sent to the Instructor but Not to the Domain Expert)

What do you think went well in the project? What do you think could have been improved? What might you have done differently if you had the chance to do it over again? What is your overall impression of participating in your first LISA collaboration project?

(For IA7B)

(1) What did you learn from collaborating on this project?
(2) Ask yourself the questions from Tables 1–3 in the QQQ manuscript reading (Reading 1.2). What did you do well? What were opportunities to improve? Provide a summary statement of your self-assessment. Be sure to include a self-assessment of your ethical thinking and reflections in this summary statement.

Shared Understanding Statement

This has been reviewed by the Domain Expert (Yes/No)

The Domain Expert made edits or additions (Yes/No)

The Domain Expert agrees that shared understanding has been created on the topics above (Yes/No)

Initial Meeting Notes: Append your initial meeting notes here or provide a link.

Appendix B. In-Class Exercise about Kimball's Type III Errors

Team Application Exercises Based on Kimball's 1957 "Type III Errors" Paper

(1) If you were to explain Type III errors to a non-statistician, which of the below would be the BEST example of a Type III error? [these answer choices for questions 1 and 2 are copied and pasted from the students' homework assignments due that day and may change from semester to semester]
(2) Which would be the LEAST EFFECTIVE example of a Type III error?
   - (A) During office hours a student came to me asking help on a homework problem. That person was complaining that they do not get the same answer as indicated. I was trying to solve the problem for 10mins and was unable to solve it. Then I checked the original problem on the text book and the student has made an error when copying down the problem, hence getting a different solution. Had I checked the problem on the text book first time, would have been able to save few minutes of my time.
   - (B) At work, we calculate the conversion rate of our retail store by utilizing a counter at the front doors where customers enter and exit. Also, the number of transactions at the register is recorded. The conversion is calculated by dividing the number of transactions over a time period by the number of customers who enter the store resulting in the percentage of customers who made a purchase. The district manager asked us for ideas on how to increase our conversion and one coworker came up with the solution of reducing the number of people that is counted at the front of the store (by a number of ways, one of which included employees ducking under the counter so that they are not counted.) This was a type III error. While taking these measures may result in a higher calculated conversion percentage, the point of increasing conversion was to cause a higher percentage of shoppers to make a purchase which these measures would not achieve. This error was a result of a misunderstanding between my coworker and the district manager
   - (C) In the movie *Se7en*, when David Mills played by Brad Pitt decides to go with the serial killer to find the location of the final body,

he is searching for the answer to the wrong question. He is more concerned with answering the question of where the final body is than he is concerned with figuring out the killer's endgame. If he were able to figure that out, then he might have realized that the killer was baiting him into making a point. However, he found the answer to the wrong question and ultimately made the killer's point for him rather than stopping his plan.

(D) This is a classic joke: In a restaurant, a guest asked the waiter, "Can you taste this soup?" The waiter was very nervous, and changed the soup for him immediately. And the guest said, "Waiter, please, can you taste this soup?" This time the waiter said, "OK. I'll taste it. Where's your spoon?" This is like a type III error because replacing the soup was the right answer to the wrong question. The real question was, "Waiter, can you give me a spoon?"

(E) I was once helping a friend out with a probability homework. I read over their question and seeing it was similar to another problem they had shown me early in the year, I quickly produced a solution and felt good about helping a friend. However, had I actually read the question deeply I would have found that it differed from the old problem. The new problem focused on sampling without replacement. The old one which I had solved was about sampling with replacement. Needless to say my solution wasn't of much help and I had to redo it. This is a classic type-III error as I gave the right answer to the wrong problem.

(3) "Errors of the Third Kind in Statistical Consulting" by A.W. Kimball was written in 1957. Do you think Type III errors are more or less common today than in 1957?

A. More common
B. Less common

(4) Do you think the rate of committing Type III errors is higher or lower today than in 1957?

A. Higher rate of Type III errors today compared to 1957
B. Lower rate of Type III errors today compared to 1957

(5) In which era do you think it would be easier to avoid Type III errors?

A. 1957
B. Today

(6) According to Dr. Kimball, what is the best remedy for Type III errors?

A. Spend more time in the classroom.
B. Teach students to be more reflective.
C. Talk about collaboration (consulting) in more classes.
D. Contemplate what you would do differently under similar situations in the future.
E. Integrate experience collaborating (consulting) into the graduate curricula in statistics programs.

(7) What does your team think would be the best remedy for avoiding Type III errors in LISA collaboration meetings? Think of this individually then discuss as a team. Write your individual tips and best team consensus tip on the class Google sheet.

Appendix C. Homework Assignment Applying $Q_1Q_2Q_3$ to Quantitative Seminars/Presentations

Individual Assignment 10

(1) Attend a live (can be on Zoom) quantitative presentation at least 40 minutes long
(2) Summarize the presentation
(3) Assess the structure of the presentation
(4) Complete the following…

Quantitative Presentation Review and Assessment

0. Presenter, title, date, venue (what kind of talk was it?)

(1) Using the "QQQ talk assessment sheet.pdf" (on OSF: *https://osf.io/sh2nu*), track how much time the presenter spent on Q_1, Q_2, Q_3, and X—where X means anything not Q_1, Q_2, or Q_3 (such as audience questions and post-talk Q&A). Turn in your sheet with your assignment.
(2) QQQ: Briefly summarize the Q_1, Q_2, and Q_3 of the project the presenter presented.
(3) Paragraph of overall comments on the presentation content/structure. What was the structure? Did the presenter employ the "what is/what could be" structure? Did they use the QMatrix $Q_1Q_3 \ldots Q_3$ structure?
$Q_2 \ldots Q_2$
$Q_3 \ldots Q_3$.
(4) Approximately what percentage of the presenter's time was spent on Q_1, Q_2, Q_3, and X?
(5) What was the longest (approximate) continuous stretch of Q_2 (in minutes)?
(6) Using the analogy of snorkeling (diving down for Q_2 and coming up for Q_3 "air") v. scuba diving (staying deep in Q_2 for long stretches without coming up for "air"), was the speaker more of a snorkeler or a scuba diver?
(7) Presentation goals: What do you think the presenter's goals were for the presentation? Did he or she accomplish them? Why or why not?
(8) Feedback on the slides. Did they effectively add to the presentation? Did they use the Assertion Evidence structure? How could they be improved?
(9) Presentation style: What behaviors or actions did the presenter do that facilitated or detracted from the accomplishment of the goals of the presentation?
(10) How well did the presenter use the remote environment? How do you think the presentation changed because it was not in person? (For in-person talks, what did the presenter do to leverage the in-person nature of the talk compared to Zoom?)
(11) On a scale from 0 to 10, where 0 means a "complete waste of time," 10 means a "fantastic use of time," and 5 is "about average," how useful was the hour or so you spent watching this talk? (whole numbers only; this is meant to be your subjective rating of the *talk*, not of the assessment exercise)
(12) (optional) Overall feedback for the presenter. What did they do well? What are opportunities for improvement? Note: the presenter will never see this feedback unless you send it to him or her or give me permission to bundle the feedback anonymously.

Learning objectives for this assignment include:

- Identifying the (QQQ) structure of a quantitative talk
- Reflecting on the usefulness of this structure
- Synthesizing class learnings to evaluate the effectiveness of the speaker's slides
- Reflecting on how remote presentations differ from in person presentations
- (Optional) Practice providing helpful feedback on a presentation
- Making connections between behaviors/actions and the goals of the presentation

Appendix D. Homework Assignment on Ethics

Individual Assignment 6: IRB Certification for Research on Human Subjects and Ethical Workflow

(1) Complete the CITI IRB training.
(2) Register your training with CU Boulder.
(3) Upload a copy of your certification onto OSF.
(4) Write a few sentences reflecting on your IRB certification experience.
(5) Create a personal Ethical Workflow
(6) Provide feedback to your teammates on their Ethical Workflow.
(7) For Team Assignment 4 (TA4) you will create a Team Ethical Workflow

Details

(1) Go to <link>. Read some of the text then use CITI's Single Sign On (SSO) to sign on to the CITI site.

Click the option to automatically link your CITI account to your myCU SSO account. From there "add course" on compliance training for "Human Research for the IRB." Then choose either the training for Biomedical Research Investigators and Key Personnel or the training for Social Behavioral Research Investigators and Key Personnel.

Read through the modules and score at least 80% on the quizzes. I recommend immediately retaking any quiz in which you score less than 80%.

(2) After completing the training, your result might automatically get transferred to CU's system. If not, make sure your training gets registered with CU.
(3) Save your certificate as a PDF, rename the file using the following conventions: YYYY-MM_First.Last.IRB.Field.citiCompletionReport xxxxxxx.pdf, where Field is "SocSci" or "BioMed" or "SocSci+BioMed" if you did both.

Upload this file to the IRB folder on OSF.

(4) Did you learn anything new? What do you think of human subject research? Did anything you learned affect the way you see an ethical statistician behaving? What is the importance of acting ethically for a collaborative statistician/data scientist? Write this paragraph reflection within your IA6 document.
(5) Based on the principles from the Belmont Report (as you learned through your CITI IRB training), the principles from the ASA Ethical Guidelines 2022 document, and your knowledge of the $Q_1Q_2Q_3$ Workflow, create your own personal Ethical Workflow. What questions should you ask yourself about ethics (ethical reflection) and when during the workflow of project? Which questions do you think you need to ask the Domain Expert?
(6) Provide feedback to your teammates (by 2:00PM on Wednesday, November 6) on their personal ethical workflows. Was anything surprising? Are there critical components missing from their workflow? Were there points your teammates made that you will incorporate into your personal ethical workflow?
(7) TA4 (mostly in class on Wed, Nov 6) will be to create a Team Ethical Workflow. Combine the best parts of your teammates' personal workflows. This assignment is part of creating a Community of Practice (CoP) in which we co-create tools to help us move from beginner to master collaborators. Your team will present the main components of your team's workflow to the other team and the professor at the end of class.

Learning objectives for this assignment include:

- Learning about the history of human subjects experiments and the need for ethics training
- Understanding the key points for protecting human subjects in research
- Gaining certification to be able to work on human subjects research projects in LISA
- Reflecting on how thinking about ethics is important for a collaborative statistician
- Putting this reflection into practice by synthesizing the key points of ethical practice with the intention to implement them in current and future collaboration projects
- Working with your team to co-create a tool to help you advance along the path from a beginner to a master collaborator.

Appendix E. Example Project Using $Q_1Q_2Q_3$

Intro to Data Science Lab 11: Baby Names [Modified for QQQ]

Overall Description of Lab 11 (Also Available at Https://Osf.io/69npm)

Each teammate should individually try to answer all of the questions in this lab and create all of the plots. If you get stuck, ask your teammates for help. Compare your results with your teammates and reconcile any discrepancies.

In this lab you will practice:

- downloading and importing a new dataset
- merging two datasets
- calculating percentiles and probabilities
- making comparisons between a variable of interest and other "background" variables you don't really care about
- using regular expressions to manipulate strings
- plotting histograms and determining what constitutes an "extreme" value in the distribution
- determining if an observed change is "significant"
- reflecting on ethical issues
- applying the QQQ workflow and identifying how the stages interact

Analyzing Baby Names

Q1 Context Provided to the Students

Your data science team has been hired by a pregnant couple to analyze the names of babies in the US (appearing 5 or more times) from 1880to2018. The dataset for 1880–2017 is called babynames from the library(babynames). The data from 2018 is available from the USA Social Security Administration. Google it to find the 2018 data and add it to the 1880–2017 dataset.

Your clients' baby is due on December 18, 2019. The clients want the sex of the baby to be a surprise and so are looking for names for both boys and girls. They want names to satisfy their idiosyncratic criteria of:

- Clearly for either males or females (no gender-ambiguous names)
- "Trending" but not "trendy," that is, on the upslope but not superpopular today. Maybe very popular in 10–20 years.
- Cannot be easily (and commonly) shortened to a nickname (like "Benjamin" to "Ben")

The first step for this type of data science problem is gaining a better understanding of what the client or domain expert really wants and understanding the data we are working with. The best way to do that is to ask questions of the domain expert with whom you are collaborating. Ask them qualitative questions like:

- "What problem do you want to solve?"
- "Why is this important or interesting?"
- "How will the eventual solution be used?" and "How is this relevant to the business?"
- "In a perfect world, what data could be analyzed to solve this problem?"
- "What data have been collected? Why were the data collected originally? For what purpose? When and where were the data collected? Who or what collected the data? How were the data collected, that is, with what instrumentation/methods?"
- "Which types of data science analyses or techniques would be most useful to the domain expert? Which would not be useful?"

Since you were just handed a dataset and some problems/questions to answer, you can't actually get answers to these questions. You can, however, learn more about the dataset by looking at the help file? babynames. For your Q1 summary in the project report, list the Q1 questions you think would be most important to ask and what your clients' hypothetical answers were. Also, identify at least one aspect of Q1 that will impact Q2 and Q3.

Q1 Ethical Reflection

Reflect on the ethics of engaging in this project. For example, discuss how the ethical principles of Privacy and Consent may or may not be relevant for this project. [Note to reader: Because the data only includes names appearing 5 or more times in a year, privacy might not be an issue.] Is this an ethical use of data science? Why or why not?

Q2: Creating Individual Plots of Your Name's Popularity Over Time

Create a plot over time (1880–2018) of the proportion of names in a given year that start with the first three letters of your name (i.e., a time series of the popularity of the first three letters of your name). On the same plot, show the proportion of your name over time. Use a regular expression to match most of the various spellings of your name.

Show what regular expression you used to find your name in the dataset. How popular is your name now compared to when you were born? How popular is your name now relative to the other names that start with the same three letters?

Using Q2 to Answer Questions About Your Own Name

How did the popularity of your name (in the US) change from the year you were born to 2018? How did the popularity of your name change from X years before your birth year to your birth year (X is 2018 minus your birth year)? Use a regex to match most of the versions of your name. If your name was not a popular name in the US, you should use a similar name (or a nickname) that does appear in the dataset.

Answer these questions for all your team's names by describing A) the actual change in proportions and B) the relative change, expressed as a percentile of all changes in baby name proportions over those timeframes.

How did your team decide to deal with gender and the fact that you probably are not all born in the same year?

Show the changes in your names' popularity on two histograms. Make a special annotation on the plot (maybe by using a special color or adding text to the plot) for changes in names that you believe are in the "tails" of the histogram, that is, very far from the middle of the distribution. The names in the extremes of the distributions are the names that show "significant" or noticeable changes over the years.

Which of your team members' names show a "significant" change over the years?

Self-Reflexive Stance in Q2

Answer and briefly discuss some of the Q2 self-reflective questions in Table 2.

Your Q3 Recommendation for a Boy and Girl Name for Your Clients

Individually come up with two candidate names (one for a boy and one for a girl) that satisfy the clients' desires. As a team, pick one boy name and one girl name to present to your clients. State why you recommend these names. Produce a plot(s) providing visual evidence for your recommendations.

Q3 Ethical Reflection

Reflect on any ethical implications of your work. Some things to consider are what, if any, are the risks of your analyses? Who might benefit or be harmed by your work? Is there something you could have done in Q2 that would have led to potential harms?

Format and Order for Team Lab Report

Create a team Quarto (or R Markdown) document knitted to html and submitted to OSF. Put your team name into the filename of the submitted document.

Your document should include:

- Summaries of the Q1, Q2, and Q3 aspects of this lab assignment/project
- A reflection and discussion of how your choices in Q1 and Q2 may have affected your conclusions/recommendations in Q3.
- Embedded within your QQQ summaries, analyze and evaluate at least one ethical quandary, implication, or decision made in each stage of your project (i.e., in Q1, Q2, and Q3).
- The final section of your report should be a summary of which team member did what.

Rubric for Assessment

(1) How well did the team summarize the Q1, Q2, and Q3 aspects of their project?
(2) To what degree did the teams or individuals engage in ethical reflection?
(3) To what degree were Q1 aspects with important downstream effects identified?
(4) To what degree are the Q2 analyses reproducible?
(5) To what degree are the Q3 recommendations based on their Q2 analyses and appropriate for the Q1 context?