2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte
Paper ID #52269

# Transdisciplinary Data Science Course for Graduate Students Interested in Community Research

**Dr. Hua Li, Texas A&M University - Kingsville**

Dr. Hua Li, a Professor in Mechanical and Industrial Engineering at Texas A&M University-Kingsville, is interested in data science, sustainable manufacturing, renewable energy, sustainability assessment, and engineering education. Dr. Li has served as P.I. and Co-P.I. in various projects funded by NSF, NASA, USDA, etc.

**Dr. Michael Preuss, Exquiri Consulting, LLC**

Michael Preuss, EdD, is the Co-founder and Lead Consultant for Exquiri Consulting, LLC. His primary focus is providing assistance to grant project teams in planning and development, through external evaluation, and as publication support. Most of his work is completed for Minority-Serving Institutions and he publishes regarding findings on a regular basis.

**jianhong Ren, Texas A&M University - Kingsville**
**Dr. David Ramirez**

Dr. David Ramirez is a tenured Associate Professor of the Department of Environmental Engineering at Texas A&M University-Kingsville (TAMUK). He is the graduate coordinator of the doctoral program in environmental engineering. He has served as the Directo

**Christine Reiser Robbins, Texas A&M University - Kingsville**
**Tushar Sinha, Texas A&M University - Kingsville**

# Transdisciplinary Data Science Course for Graduate Students Interested in Community Research


## Abstract

Coastal communities face unprecedented environmental challenges such as frequent and intensified floods and droughts, uncertain availability of water resources, harmful algal blooms, and deteriorating water and air quality. The management of air and water resources (AWR) is becoming more complex due to ever-changing socio-environmental factors such as increasing demand caused by rapid population growth, industrial development, urbanization, and climate change. Finding convergent solutions to these issues demands well-coordinated teamwork and a new outlook toward education and research across the engineering, sciences, and humanities disciplines along with active engagement of communities. To effectively address these challenges, future environmental professionals should have in-depth knowledge of the dynamic relationships among natural, social, and engineering systems and an ability to transcend traditional disciplinary boundaries. As part of an NSF Research Traineeship project, a transdisciplinary data science course was planned and implemented for a group of master's and doctoral students who are interested in community research at Texas A&M University-Kingsville. The course was designed for graduate students in STEM majors with or without a data science background. This course focused on providing hands-on experience and fundamental knowledge of data science processes, data analysis, and rapid ethnographic assessment to develop transdisciplinary solutions to environmental issues and challenges. It has two parts, data science fundamentals and qualitative informatics. Data science fundamentals encompass data and simulation-driven research and problem-solving, statistical and exploratory data analysis using R and Python, data formats, data acquisition and cleaning, cyberinfrastructure systems for data integration, data visualization, machine learning, and ethical considerations in data science. Qualitative informatics encompasses research methods central to stakeholder engagement, mixed-methods research design, and qualitative data analysis, including maintaining fieldnotes, memoing, coding, and interpretation. In this paper, the course description, the need for it, and first-year outcomes are presented. The course was first offered in fall 2024, where all students had engineering majors and a research focus in AWR. Pre- and post-participation self-reports of knowledge and skills were collected relevant to 30 topics identified as learning objectives in the course materials. Learning and skill advancement were reported across all areas surveyed, and despite a small group size (n = 12), six of the pre- to post-participation comparisons were statistically significant. Implications of transdisciplinary challenges and the findings for the future of engineering education are discussed.



Keywords: Data Science, Transdisciplinary, Graduate Education.


## Project Background

The NSF Research Traineeship (NRT) project at Texas A&M University-Kingsville envisions future environmental professionals as global leaders in effectively integrating knowledge from different disciplines to address the burgeoning challenges to sustain and enhance the quality of life across regions. It is developing and institutionalizing a transformative graduate training model that can educate a new generation of environmental professionals to master skills in data science, be equipped with sufficient professional and communication skills, apply those skills in a transdisciplinary framework to address environmental sustainability issues, and engage these actions within culturally, environmentally, and sustainable humanitarian practices to achieve solution convergence. The entire NRT project builds around three components: (1) academic skills preparation; (2) professional skills development; and (3) humanitarian service-learning experience.

To solve increasingly complex problems surrounding air and water resources (AWR), future scientists/engineers will need to adopt new lines of inquiry involving transdisciplinarity [1]. Transdisciplinarity is an approach to human communication, meaning-making, and collaborative decisions that not only supports interdisciplinary thinking but also supports the co-construction and soft assembly of complex models and paradigms that value the input and design of structures across disciplines from the very beginning [2], [3], [4]. By adopting a transdisciplinary approach, this NRT program eschews traditional compartmentalized discipline-specific approaches because they are often part of the problem, rather than the solution.

Since data science itself is a transdisciplinary field, most existing graduate-level data science courses are designed specifically for certain field or major [5],[6]. The transdisciplinary data science course is a new course developed as part of the academic skills preparation component of the NRT project. It was first offered in fall 2024 with 12 students. This paper introduces the detailed design of this new course and students' feedback and survey responses.

## Course Design

This transdisciplinary data science course, titled Data Science for Next Generation of Community Researchers, is a three-credit-hour graduate-level course open to students in STEM majors at Texas A&M University-Kingsville. It focused on providing hands-on experience and fundamental knowledge of data science processes, data analysis, and rapid ethnographic assessment to support transdisciplinary solutions to environmental issues and challenges. It had two component parts, data science fundamentals and qualitative informatics. Data science fundamentals included data and simulation-driven research and problem-solving, statistical and exploratory data analysis using R and Python, data formats, acquisition and cleaning, cyberinfrastructure systems for data integration, data visualization, machine learning, and ethical issues in data science. Qualitative informatics included research methods central to stakeholder engagement, mixed methodological research design, and qualitative data analysis, including maintaining fieldnotes, memoing, coding, and interpretation.

The course materials were developed by a group of faculty members from Environmental Engineering, Industrial Engineering, Anthropology, and Bilingual Education. The weekly course topics included the following.

(1) Introduction to data science and basic statistics.
(2) Statistical and exploratory data analysis.
(3) Distributions and non-parametric statistics.
(4) Basic programing in R.
(5) Data analysis in R.
(6) Basics in Python programing.

(7) Data analysis and quality checking using Python.
(8) Data visualization using Python.
(9) Introduction to machine learning and deep learning.
(10) Rapid ethnographic assessment and ethnographic observation.
(11) Community engagement goals.
(12) Research methods central to stakeholder engagement.
(13) Techniques for completing interviews.
(14) Fieldnotes and memoing.
(15) Coding qualitative information and visualization.
(16) Qualitative data analysis.
(17) Data ethics and data sharing.
(18) Protocols for conducting focus groups.
(19) Mixed methodology survey design and administration.

In addition to homework and two exams, a team project was required for all students. Two students, or three with the approval of the course instructor, did the project together, which was one of the major assignments. The team project involved obtaining a dataset, performing statistical analysis, creating visualization tools or high quality spatial and/or temporal plots of the data. Students selected their team project topics, which needed approval from the instructors by submitting a one-page objective statement (per team) by week eight of the course, within the overarching theme of the course. In the one-page objective statement, the students were asked to include two items: an abstract up to five lines that will explain the theme/objective of the project. And a detailed outline of the expected project report. The students were asked to cite the data sources in the report and consult with the instructor (e.g., email or meetings) regarding their idea. The final project report was limited to no more than 12 double-spaced pages, excluding references. In addition, each team gave a 15-minute presentation to introduce their final project results, which was graded based on objectives, content, organization, individual presentation skills, and technical merit.

After successfully completing this course, students are expected to achieve the following student learning outcomes (SLOs).
1. Apply data-driven simulation techniques in formulating and solving problems.
2. Develop scripts in different software to perform statistical data analysis.
3. Work with distinct types of data formats and perform data quality checking and cleaning.
4. Visualize data and create high-quality figures and maps.
5. Analyze data and apply machine learning techniques.
6. Design transdisciplinary qualitative research solutions to pressing environmental issues.
7. Analytically memo germane documents, interviews, photographs, and focus group data to code and consolidate codes into themes and processes.
8. Evaluate the differences between qualitative and quantitative research designs.
9. Describe fundamental ethical issues in data science and stakeholder engagement.

**Student Feedback and Survey Results**

As mentioned before, the data science course was first offered in fall of 2024 and co-taught by two faculty members from Environmental Engineering and Industrial Engineering. The

Environmental Engineering faculty with data science background covered topics related to data science fundamentals (about 2/3 of the course). The Industrial Engineering faculty member with experience in qualitative data analysis covered topics related to qualitative informatics (about 1/3 of the course). A total of 12 graduate students, including both M.S. and Ph.D. students, took the course. It should be noted that this data science course is required for NRT-funded students, while four out of the 12 students were non-NRT-funded students.

Pre- and post-surveys, which include questions specific to the course's instructional objectives, were conducted to collect students' feedback. In both surveys, students were asked to provide ratings of their understanding or skill in specific areas relevant to SLOs and related topic areas in the data science course. The queries used in both surveys to measure learning in topic areas related to the data science course were devised by the NRT project's external evaluator. Statements from the proposal and the course syllabus identifying proposed learning outcomes were utilized as the starting point. Many of the learning objectives in the syllabus were sufficiently specific to be easily converted to operational statements. Several other related ideas were added based on outcomes of the process employed with the course topics as listed in Table 1. Prompts that are restatements of course SLOs are marked by the SLO number. When a SLO for the course included two topics, they were split out as part a and part b. Other labels used are "course description" for topics noted in the course description but grouped under a broader category in the list of learning objectives. Details of the queries used in the surveys, the instructional topics and SLOs on which they were based are listed in Table 1.

Table 1: Course Skill and Associated Operational Statements used in Surveys.

| **Source** | **Skill to be Taught** | **Queries Used in Surveys** |
|---|---|---|
| Course description | Mixed method research design | Develop a plan for a mixed methods investigation. |
| Course schedule | Rapid ethnographic assessment | Conduct a rapid ethnographic assessment. |
| Course schedule | Case study | Conduct a case study. |
| Course schedule | Focus group protocol | Appropriately plan and conduct a research focus group. |
| Course schedule | Ethnographic observations | Plan and complete a set of observations and interviews intended to lead to conclusions about people's or society's actions and commitments. |
| Course schedule | Interview techniques | Plan the format and volume of interviews needed to achieve a research objective. |
| Course schedule | Cyberinfrastructure systems for data integration | I can explain how the combination of distributed hardware, software, networks, and inputs from people facilitate processing of large data sets. |
| Course schedule | Exploratory data analysis | Analyze a data set to summarize its main characteristics and determine methods of analysis that are appropriate and workable. |
| Course description | Use of R in data analysis | Use R in data analysis. |
| Course description | Use of Python in data analysis | Use Python in data analysis. |
| Course description | Fieldnotes | Compile field notes as research data. |
| Course description | Interpretation of qualitative data | Use concepts derived from things people have said, done, or produced to construct a theoretical argument. |
| Course schedule | Ethical issues in data sharing | Describe ethical issues related to sharing research data sets. |
| SLO 1 | Data-driven simulation in problem definition/solving | Repeatedly mimic real-world performance of a system using a computer to arrive at a result or conclusion. |
| SLO 2 | Stat anal using different software | Develop scripts in different software to perform statistical analysis of data. |

| | | |
|---|---|---|
| SLOs 3a/5a | Distinct data formats; analyze data | Analyze data that appears in different formats (e.g., numeric, text, video, audio, binary, bitmap, etc.). |
| SLOs 3b, 5a | Data quality check/clean in different formats; analyze data | Perform data quality checking and cleaning in different formats (e.g., numeric, text, video, audio, binary, bitmap, etc.). |
| SLO 4 | Data visualization as high-quality figures and maps | Create images (e.g., diagrams, figures, maps) that represent outcomes of data analysis. |
| SLO 5b | Use of machine learning in data analysis | Use machine learning techniques in data analysis. |
| SLO 6 | Design transdisciplinary qualitative research solutions to pressing environmental issues | Design transdisciplinary qualitative investigations regarding pressing environmental issues. |
| SLO 5a, 7a | Analyze data; analyze qual data | Code qualitative data. |
| SLO 5a/7b | Analyze data; consolidate codes as themes and processes | Consolidate qualitative codes into themes and processes. |
| SLOs 8/7a | Diffs quant/qual research design; what is germane | Determine what data is needed to answer a research question. |
| SLO 9a | Ethical issues in data science | Describe ethical issues in data science. |
| SLO 9b | Ethical issues in stakeholder engagement | Describe ethical issues in stakeholder engagement. |
| SLO 7a; evaluator suggestion | Investigative scope; power analysis; what is germane | Determine how much data is needed to answer a research question. |
| Evaluator suggestion | Data organization, storage, retrieval | Organize data sets for use, storage, and future access. |

An item-by-item accounting of the responses is provided in Table 2. A 10-point scale was used on the survey with higher values indicating greater confidence. The numbers of students taking pre- and post- surveys are different at several points because informants elected to skip questions. Statistical analysis was completed as a t-test as the underlying assumptions were met, independent samples that are normally distributed with homogeneity in variance [7], [8].

Table 2: Students' Pre- and Post-survey results on a 10-point scale

| **Prompt** | **Period (Pre/Post)** | | | |
|---|---|---|---|---|
| | **N** | **Mean** | **Mode** | **SD** |
| Develop a plan for a mixed methods investigation. | 11/12 | 5.82/7.71 | 5/8 | 2.29/1.21 |
| Repeatedly mimic real-world performance of a system using a computer to arrive at a result or conclusion. | 12/12 | 5.33/6.50 | 9/8 | 3.06/2.03 |
| Determine what data is needed to answer a research question. | 12/12 | 6.75/7.67 | 8/8 | 2.17/1.37 |
| Determine how much data is needed to answer a research question. | 12/12 | 6.25/7.08 | 7/7 | 2.20/1.71 |
| Conduct a rapid ethnographic assessment. | 10/12 | 5.70/5.83 | 6/7 | 2.24/2.41 |
| **Conduct a case study.** | 12/12 | 4.92/6.92* | 2/7 | 2.63/1.93 |
| Appropriately plan and conduct a research focus group. | 12/12 | 5.33/7.17 | 6/10 | 2.59/2.19 |
| Plan and complete a set of observations and interviews intended to lead to conclusions about people's or society's actions and commitments. | 11/12 | 5.55/6.83 | 8/7 | 2.57/2.11 |
| Plan the format and volume of interviews needed to achieve a research objective. | 10/12 | 5.60/6.75 | 8/8 | 2.87/2.09 |
| I can explain how the combination of distributed hardware, software, networks, and inputs from people facilitate processing of large datasets. | 9/12 | 6.89/6.92 | 7/7 | 2.42/2.02 |
| Analyze a data set to summarize its main characteristics and determine methods of analysis that are appropriate and workable. | 11/12 | 6.27/7.33 | 7/9 | 2.45/1.97 |
| Organize data sets for use, storage, and future access. | 11/12 | 6.45/7.83 | 8/8 | 2.06/1.52 |

| | | | | |
|---|---|---|---|---|
| Develop scripts in different software to perform statistical analysis of data. | 10/12 | 5.40/7.00 | 8/7 | 2.65/2.16 |
| Analyze data that appears in different formats (e.g., numeric, text, video, audio, binary, bitmap, etc.). | 12/12 | 5.83/6.75 | 8/9 | 2.67/2.20 |
| Perform data quality checking and cleaning in different formats (e.g., numeric, text, video, audio, binary, bitmap, etc.). | 10/12 | 5.50/6.42 | 7/7 | 2.29/2.02 |
| Use R in data analysis. | 11/12 | 4.27/6.08 | 7/9 | 3.11/2.06 |
| Use Python in data analysis. | 10/12 | 5.00/6.83 | 7/7 | 2.97/2.37 |
| Use machine learning techniques in data analysis. | 9/12 | 5.11/5.92 | 5/5 | 1.97/2.33 |
| **Create images (e.g., diagrams, figures, maps) that represent outcomes of data analysis.** | 12/12 | 5.75/8.17** | 7/8 | 2.42/1.14 |
| **Design transdisciplinary qualitative investigations regarding pressing environmental issues.** | 10/11 | 4.80/6.82* | 7/7 | 2.18/1.59 |
| Compile field notes as research data. | 11/12 | 6.36/6.92 | 8/10 | 2.06/2.22 |
| Code qualitative data. | 9/12 | 5.78/6.00 | 7/8 | 2.25/2.74 |
| Consolidate qualitative codes into themes and processes. | 9/12 | 5.78/6.17 | 6/6 | 1.87/1.95 |
| **Use concepts derived from things people have said, done, or produced to construct a theoretical argument.** | 12/12 | 5.17/7.08* | 7/6 | 2.23/1.66 |
| **Describe ethical issues in data science.** | 12/12 | 5.08/7.00* | 7/6 | 2.22/1.44 |
| Describe ethical issues in stakeholder engagement. | 12/12 | 5.25/7.00 | 7/6 | 2.45/1.53 |
| Describe ethical issues related to sharing research data sets. | 11/12 | 5.63/7.42 | 7/8 | 2.27/1.32 |
| Applying data science techniques in real-world investigations of environmental issues. | 11/12 | 6.27/7.50 | 8/7 | 1.77/1.26 |
| **Planning a transdisciplinary investigation of an environmental issue in collaboration with stakeholder groups.** | 11/12 | 5.09/7.00* | 7/7 | 2.27/1.47 |
| Conducting transdisciplinary research regarding an environmental issue in collaboration with stakeholder groups. | 11/12 | 5.27/6.75 | 7/7 | 2.42/1.53 |
| Note: **bold text marks significant differences;** * represents p < 0.05, ** p < 0.01, and *** p < 0.001. | | | | |

Figure 1 below illustrates the patterns as percentage of change in students' confidence on the survey queries from pre-survey to post-survey. The outcomes were strongly positive and therefore encouraging but represented one offering of the program so cannot support generalizable interpretation of specific points without replication. Students reported substantial learning in respect to all the items. The increases in six of the areas were large enough to be statistically significant even though the informant count was small, 12 persons, limiting statistical power. The following is a summary of the findings.

- All the measures had increases in mean and decreases in standard deviation pre- to post-instruction.
- A limited number, six, had significant to strongly significant positive differences pre- to post-instruction even though the n, 11 or 12 persons, for the group was small. This indicates trainee self-perception of skill was higher post-instruction and the associated decreases in standard deviation that it was more consistent across the group.
- The significant findings in the data science course were in the following topic areas: (1) case studies, (2) data visualization, (3) transdisciplinary qualitative research, (4) interpretation of qualitative data, (5) ethical issues in data science, and (6) planning transdisciplinary research with stakeholders. These are also topics in which graduate students in engineering and environmental science are unlikely to have experience.

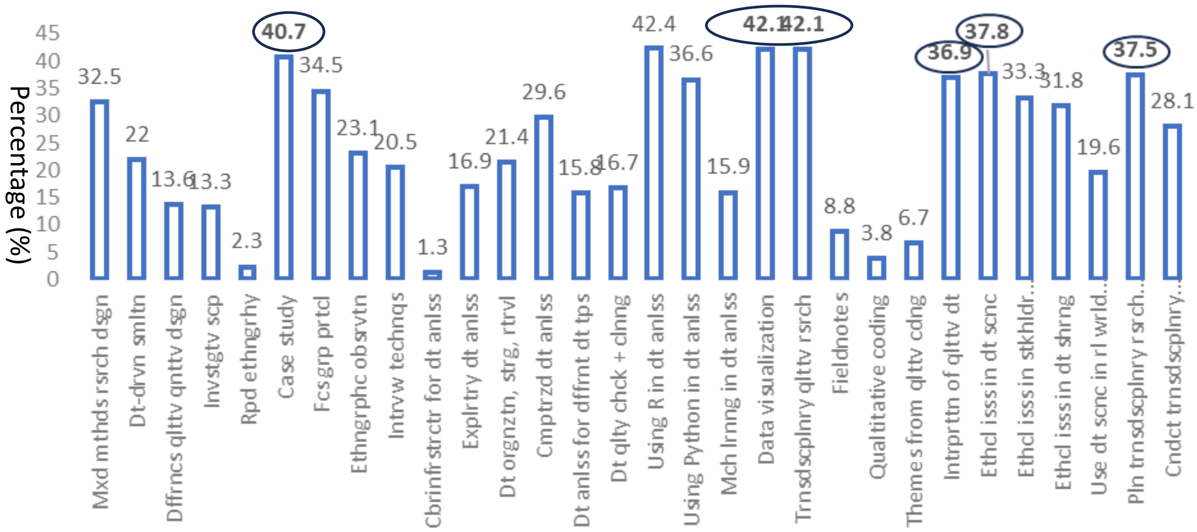


Figure 1: Percent Change in Knowledge related to the Data Science Course Topics (**bold**/circled text indicates statistically significant change).

Qualitative data was also gathered from the project participants following completion of their first year in the NRT cohort. Themes identified relevant to the course and the outcomes achieved are as follows.

- Students communicated that working on projects with their peers resulted in valued relationships, a consistent support network, and academic assistance.
- The patterns of interaction built into the NRT programming blossomed into a network of personal and professional relationships that the students valued and worked to maintain.
- Students noted collaboration, sharing knowledge and ideas, asking questions of each other, talking about challenges, providing encouragement and feedback to each other, brainstorming about how they could collaborate in research, and troubleshooting with and for each other as patterns they practiced.
- Students stated that they held each other accountable, and through their interactions, reinforced a general commitment to excellence.
- NRT participants derived motivation to persist and advance through the interactions they had with their peers in the program.

The commitment to and realization of a community within the cohort of participating students is a valuable outcome of the project. As has just been noted, students attributed some of their motivation, persistence, pursuit of high standards, and success to being part of a close-knit NRT student community at Texas A&M University-Kingsville. That this outcome was coupled with the transdisciplinary learning achieved moved the students toward the project goal of producing future environmental professionals who will be global leaders in effectively integrating knowledge from different disciplines to meet the burgeoning challenges to sustain and enhance the quality of life across regions.

## Lessons Learned and Continuous Improvements

Most students in fall 2024 were from Environmental Engineering, Industrial Engineering, and Mechanical Engineering backgrounds, who may not have much prior experience with or knowledge regarding data visualization and qualitative data analysis. So, they tended to gain more knowledge in these areas, as reflected in their pre- and post-surveys. In Fall 2024, the course was taught in a hybrid format by offering both in-person and online sessions to meet the students' needs. It should be noted that students who were in the classroom did better overall in the team projects, which was an important way to strengthen the skills and knowledge learned in the lectures.

In Fall 2025, the course was offered to a second cohort of students from different engineering majors plus a computer science major. More hands-on practices of qualitative data analysis were added in Fall 2025. Although the assessment data from the second cohort is still bring gathered, some early observations suggest that most students liked the added hands-on practices. In Fall 2025, the course instructors also emphasized the connection between qualitative and quantitative research and their applications in community research. To enhance students' skills on transdisciplinary research, the students were asked to focus on the generation of research questions and selection of research methods in their projects.

Since the data science course will be offered every fall semester, more students will complete the course in coming years. The course instructors will keep improving the course contents based on students' feedback and academic background. Multiple waves of individuals who experience the same programming but at different times with differing external circumstances will contribute to the data set. Thus, there is potential to verify the initial findings and, should instructional circumstances remain static, combine data from two or more offerings, increasing the statistical power. There is also an opportunity to gather additional feedback that may confirm the efficacy of the intra-cohort network-building elements of the undertaking and provide further insight into how those relationships form, what might challenge such formation, what strengthens it, and whether the relationships and interaction patterns persist as the students take up professional commitments. Since the course is open to all graduate students in STEM majors, the project team is actively promoting this new course to students at different colleges in the university. This will facilitate collection of learning data and feedback from students outside engineering or computer science majors.

## Acknowledgments

The authors are thankful for support from National Science Foundation. The material presented is based upon work supported by the National Science Foundation under Awards #2225109 and #2244523. Any opinions, findings, and conclusions, and recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## References


[1] National Research Council, Division on Earth, Life Studies, Board on Life Sciences, & Committee on Key Challenge Areas for Convergence. (2014). Convergence: Facilitating transdisciplinary integration of life sciences, physical sciences, engineering, and beyond.

[2] Lawrence, R. J. (2015). Advances in transdisciplinarity: Epistemologies, methodologies and processes, Futures, 65, 1–9.

[3] Pohl, C. (2010). From transdisciplinarity to transdisciplinary research. Transdisciplinary Journal of Engineering & Science, 1, 65-73.

[4] Darbellay, F. (2015). Rethinking inter-and transdisciplinarity: Undisciplined knowledge and the emergence of a new thought style. Futures, 65, 163-174.

[5] Msweli, N. T., Mawela, T., & Twinomurinzi, H. (2023). Transdisciplinary teaching practices for data science education: A comprehensive framework for integrating disciplines. Social Sciences & Humanities Open, 8(1), 100628.

[6] Holderieath, J. J., Crosby, M. K., Jacques, L. A., & Chowriapp, P. (2025). The Case for Transdisciplinary Data Science Education in Agricultural Economics. Applied Economics Teaching Resources (AETR), 7(4), 1-11.

[7] Bobbitt, Z. (March 16, 2021). The four assumptions made in a t-test. https://www.statology.org/t-test-assumptions/

[8] Kim, T. K. & Park, J. H. (April 1, 2019). More about the basic assumptions of the t-test: normality and sample size. Korean Journal of Anesthesiology, 72(4), 331-335. https://pmc.ncbi.nlm.nih.gov/articles/PMC6676026/https://pmc.ncbi.nlm.nih.gov/articles/PMC6676026/