Project-based learning in life sciences statistics courses: Dynamics, pitfalls, and educational gains

Anna Khalemsky[1] and Yelena Stukalin[2]
[1]Jerusalem Multidisciplinary College, Israel; [2]The Academic College of Tel-Aviv-Yaffo, Israel
anaha@edu.jmc.ac.il

This study explores how project-based learning (PBL) enhances statistics education for students in the two departments—biotechnology and medical laboratory sciences. Semester-long practical projects enable students to simulate real-world data analysis and develop research skills by selecting topics, building databases, reviewing literature, applying various methods using analytical software, interpreting results, and drawing conclusions. MedLab students, who take the course early in their studies, express strong motivation to take initiative and work independently throughout the project. In contrast. Biotechnology students, who are at a more advanced stage in their academic studies, rate communication with peers and lecturer very high. The study draws on two surveys: a Likert-scale questionnaire completed by 64 students to assess PBL experiences, and an open-ended survey of 27 students analyzed using natural language processing to identify sentiment and themes. Findings reveal key differences and similarities, informing how PBL can be tailored to enhance learning in STEM education.

INTRODUCTION

In recent years, the delineation between STEM and non-STEM disciplines has become increasingly blurred in the context of applied statistics education. On one hand, technological advancements and the integration of artificial intelligence have enabled social science students, traditionally outside the STEM domain, to demonstrate strong analytical capabilities (Parviz, 2024). On the other hand, STEM students, once primarily focused on mathematically rigorous statistical content, are now encouraged to engage with the "softer" dimensions of learning by incorporating social science concepts and interpretative analysis (Pacella et al., 2022).

As applied statistics courses become more closely aligned with graduates' professional trajectories, academic institutions have placed greater emphasis on cultivating the key skills required for future career success (Kurczewska et al., 2020). In response, active learning—particularly project-based learning (PBL)—has emerged as an important pedagogical strategy across data science curricula, including courses in descriptive and inferential statistics, data mining, and data visualization (Kokotsaki et al., 2016). Contemporary analytical work is no longer divided into discrete stages handled by separate experts. Instead, professionals are expected to understand the problem, select appropriate methods, perform the analysis, and interpret the results. PBL provides a powerful framework for developing these end-to-end analytical competencies (Granado-Alcón et al., 2020). The literature highlights that PBL enhances students' understanding of core statistical concepts, builds practical skills with analytical tools, fosters critical thinking through authentic problem-solving, and promotes motivation by encouraging active engagement and ownership of learning, especially when projects align with their interests and career goals (Kokotsaki et al., 2016).

STEM-oriented statistics courses provide a comprehensive foundation for advanced study and future careers, covering core topics such as descriptive statistics, correlation analysis, prediction models, and statistical inference, which underpin more complex, application-specific models (Horton & Hardin, 2015; Watson, 2017). Each of these topics should be reflected in project-based learning.

In life sciences disciplines, research plays a central role—from understanding the interconnections between biology, chemistry, and physics and their real-world applications, to seeking solutions for unresolved problems (Gibson & Mourad, 2018). These elements are fundamental to students' academic development and continue to shape their professional paths. Project-based learning (PBL) offers valuable opportunities for students to engage directly with these challenges (Movahedzadeh et al., 2012). By working with real datasets from biology, medicine, or environmental science, students transform abstract statistical concepts into practical tools for analyzing complex biological systems (Gibson & Mourad, 2018). As life sciences increasingly rely on data-driven research, PBL helps students develop essential research skills, including experimental design, hypothesis testing, data analysis, and scientific reporting—all critical for evidence-based practice (Robeva et al., 2020). Through project work, students learn to integrate statistical reasoning with domain-specific knowledge, enhancing their ability to interpret data in areas such as genetics, epidemiology, and clinical trials. Finally, PBL fosters collaboration and communication skills that are vital for multidisciplinary teamwork in labs, hospitals, and research institutions, where life scientists must interact effectively with data analysts, clinicians, and other professionals (Li et al., 2020).

This study builds on previous research exploring the significance of project-based learning across different fields and students' perspectives on it (Khalemsky et al., 2024; Khalemsky & Stukalin, 2024, 2023). We focus specifically on students in STEM disciplines—biotechnology and medical laboratory sciences—to examine both the shared characteristics and differences between these two areas. This paper brings together findings from two studies conducted at an academic college offering statistics courses across both STEM and non-STEM departments. In the first study, students completed a survey that included demographic information and ten questions assessing their attitudes toward project-based learning. The second study focused on open-ended responses, inviting students to reflect in depth on their experiences with project-based learning. As the research is still ongoing, we present here a subset of the questions.

METHODS

In both STEM departments, students completed practical projects integrated throughout the course, accounting for up to 30% of the final grade. At the semester's start, they selected a topic in their field, built or sourced a database, defined variables, and formulated research questions and hypotheses. They then conducted descriptive statistics with visualizations using SPSS and Microsoft Excel, followed by correlation analysis and multivariate regression. Next, they performed statistical inference, including hypothesis testing and confidence intervals for single, paired, and independent samples, as well as one-way ANOVA. Finally, they related their findings to the research hypotheses.

Two surveys were administered during the 2023–2024 academic year. The first survey, completed by 64 students, consisted primarily of Likert-scale items designed to evaluate perceptions of the project-based learning (PBL) methodology and to identify differences across sub-groups. The second survey, involving 27 students, included open-ended questions aimed at capturing students' reflections and attitudes in a more exploratory and nuanced format. Open-ended responses were preprocessed using standard natural language processing (NLP) techniques, including lowercasing, stopword removal, and lemmatization (Chowdhary, 2020). Sentiment analysis was performed using TextBlob to compute polarity, that measures the emotional orientation of a text, indicating whether the expressed sentiment is positive, negative, or neutral (ranging from -1 to 1) and subjectivity, that refers to the degree to which a text expresses personal opinions, feelings, or judgments rather than objective facts (ranging from 0 to 1) scores (Montoyo et al., 2012). Thematic coding combined keyword frequency analysis with manual review to uncover common patterns in students' attitudes, challenges, and reflections regarding project-based learning (Vaughn & Turner, 2016). The choice of analysis method is guided by the survey design and aims to capture nuances in students' responses.

RESULTS

Table 1 presents the distribution of the research variables in the first survey. For categorical variables, each cell contains the simple count and the percentage, and the p-value for the Chi-square test of independence. For numerical variables, each cell contains mean (std.dev), and median with IQR. In addition, we conducted the following tests for the comparison of two sub-groups: T-test for mean differences (the 1st p-value) and Mann-Whitney U test for median differences (the 2nd p-value).

Table 1.

	Biotechnology	Medical Laboratory	p-value
Number of students	N=50	N=14	---
Gender	Female 39 (78%) Male 11 (22%)	Female 4 (28.6%) Male 10 (71.4%)	0.635
Age	23.250 (4.160) 22 (22,25.75)	28.091 (5.991) 27 (23,33)	0.026* 0.01*

In: Birk, L., Loth, G., Jotzo, L., Binder, K., & Frischemeier, D. (Eds.) (2026). *Statistics and Data Science Education in STEAM. Proceedings of the Satellite Conference of the International Association for Statistical Education (IASE), September/October 2025, Münster, Germany*. © ISI/IASE.

Mathematical background	High 18 (45%) Medium19 (47.5%) Low 13 (32.5%)	High 4 (28.6%) Medium 5 (35.7%) Low 5 (35.7%)	0.757
Working status	Yes 33 (66%) No 17 (34%)	Yes 12 (85.7%) No 2 (14.3%)	0.154
ADHD diagnosis	Yes 15 (30%) No 35 (70%)	Yes 3 (21.4%) No 11 (78.6%)	0.528
Learning hours per week	3.510 (1.845) 3 (2,4)	3.846 (1.951) 3 [2,5]	0.585 0.552
PP1 - Do you think that the combination of the practical project that simulates real research enriches the course?	6.14 (2.807) 7 (4,8)	5.50 (3.653) 4 (2.75,10)	0.551 0.578
PP2 - Do you feel more engaged when there is a practical project in the course?	5.46 (2.873) 5 (3.75,8)	5.71 (3.361) 5 (3.25,9.25)	0.779 0.774
PP3 - Do you prefer a more structured data mining course?	6.48 (2.613) 6.5 (4.75,9)	6.50 (3.276) 7 (3.75,10)	0.981 0.831
PP4 - Do you like the autonomy of choosing your own topic and data?	5.82 (2.883) 6 (3.75,8)	7.43 (2.503) 8 (5.75,10)	0.063 0.065
PP5 - Do you like the autonomy in decision-making?	6.22 (2.951) 7 (4,9)	7.43 (2.652) 8 (5,10)	0.172 0.148
PP6 - How much do you like teamwork during practical projects?	6.38 (3.257) 7 (4,9.25)	5.71 (2.972) 6 (4,8)	0.494 0.358
PP7 - Do you feel free to contact the lecturer for help?	7.48 (2.837) 8.5 (5,10)	7.57 (2.681) 9 (6.25,9.25)	0.915 0.751
PP8 - Do you feel free to ask other students for help?	7.38 (2.955) 8 (5.75,10)	6.14 (2.931) 6 (4,8.25)	0.170 0.102
PP9 - Are you using external information sources for learning goals?	6.04 (3.220) 7 (3,9)	5.50 (3.156) 5 (2,8.25)	0.580 0.595
PP10 - Do you think you will use the tools you acquired in the project in your further studies or future work?	6.60 (2.777) 7 (5,9)	6.57 (2.954) 7 (4.25,9)	0.973 0.954
SE (achievement)	5.379 (0.930) 5.429 (4.857,6.071)	5.367 (1.556) 3.586 (3.821,6.893)	0.979 0.567
SE (learning)	4.824 (0.907) 4.682 (4.091,5.477)	5.24 (1.296) 5.591 (4.023,6.25)	0.275 0.160

* Significant for 5% significance level

Natural language processing methodology was used in the second survey analysis to understand and interpret the nuances of students' attitudes towards a practical project. Each open-ended question was analysed using polarity and subjectivity values. We compared these values for biotechnology and medical laboratory students using an independent samples T-test with a 5% significance level. Means and standard deviations are shown in Table 2.

Table 2. NLP analysis – comparison of polarity and subjectivity values of students' attitudes towards the importance of practical projects.

	Biotechnology	Medical laboratory	t-test p-value
Number of students	20	7	---
Polarity	0.147 (0.169)	0.155 (0.169)	0.916
Subjectivity	0.365 (0.119)	0.613 (0.314)	0.035*

We used thematic coding to divide all answers to the open-ended question into three separate homogeneous groups: strongly supportive, neutral, and strongly unsupportive. Table 3 demonstrates the joint distribution of the department and the obtained label.

Table 3. Chi-squared test for independence.

Label	Biotechnology	Medical laboratory	Cramer's V
Strongly supportive	5 (25%)	3 (42.9%)	0.196 p-value=0.596
Neutral	14 (70%)	4 (57.1%)	
Strongly unsupportive	1 (5%)	0 (0%)	

Finally, we used the same technique to find topics/themes in students' answers. Four topics were revealed: skill development, students' engagement, skepticism or criticism towards PBL, and the importance of real-world applications. Table 4 demonstrates several examples of each type.

Table 4. Examples of students' answers for different topics.

Theme	Description	Example quote
Skill development	Emphasizes the projects' role in building applied skills and real-world competence.	#We have the shovel thinking that teaches me about the project we did is what points in this test sharpens the material. #The project helped us to understand the material, especially for the outputs, if we didn't have projects we wouldn't be messing with the stats and the outputs properly, and it will help us with the final test. #In general, for me, practical projects contribute more to understanding the material than just practicing and memorizing in a dry way of solving exercises.
Students' engagement	Students' active involvement, interest, and emotional connection to the learning process.	#Students' engagement with writing the practical project contributes to a better understanding of the theoretical material studied. #Basically, the topic we studied was relatively difficult. Things that I was supposed to come up with really encouraged me and my project friends to understand the meaning of the calculations in general. #People like me tend to isolate themselves, so a project like this really gave me the option to connect with new people. It's not always possible to embrace those close to you. Sometimes it's difficult to form friendships, and this is a beautiful way for me. #I'm asking myself, maybe this is a really good way to pass the course in general and to cancel tests and instead of giving assignments but ones that also give real data, maybe it will encourage people to look for the meaning of learning in real life. Thank you very much for the experience, I will probably use it.
Skepticism or Critique	Expresses doubt or criticism about the PBL value or implementation.	#I would prefer to submit exercises instead of the project. #It was not felt that it contributed too much to my understanding of the material. I felt it was a bit of a waste of time. Projects help to understand the material better. I think that projects need to have a higher weight in the final grade. Students will know the material better because they will invest more time.
Real world applications	Importance of using statistics in real-life and professional contexts.	#Prepares students to use statistics in the real world, teaches students about samples, and teaches how to use Excel. #Integrating a practical project into a statistics course is important: 1. Allows students to apply theory in reality. 2. Helps develop skills such as data analysis and programming. 3. It provides practical experience that is important for the labor market.

DISCUSSION AND CONCLUSIONS

The analysis of the first survey offers a fresh perspective on assessing students' attitudes toward project-based learning in statistics courses within two life sciences departments. It captures multiple facets of the learning experience—including strengths, challenges, and nuanced personal insights. The questions explore not only the perceived value and engagement stemming from practical projects but also students' preferences regarding autonomy (self-direction), course structure, and collaboration with peers and the instructor, which are core components of effective PBL environments.

Medlab students tended to report greater appreciation for autonomy in both choosing a topic and decision-making. In contrast, biotechnology students reported greater motivation in seeking collaboration with peers and the instructor. These findings align with the program structure: medical laboratory students take the statistics course during the first semester of their first academic year, while biotechnology students typically take it in the second semester of their second year. During the first academic year, most of the academic programs are very structured and do not provide students the autonomy and decision-making. No significant differences were noted in learning self-efficacy, along with a large but non-significant gap in achievement self-efficacy. These trends, though not always statistically significant due to small sample sizes, suggest that differences in age, academic maturity, and program progression may influence learning preferences and experiences.

Further supporting these findings, text analysis of open-ended responses revealed that medical laboratory students expressed themselves with significantly greater subjectivity compared to biotechnology students, although the polarity of responses did not differ between the groups. This fact may reflect the earlier academic stage and increased emotional investment of medlab students, who take the course during their first semester. Categorical labeling of responses as strongly supportive, neutral, or unsupportive showed no significant difference between the groups, aligning with the quantitative findings that attitudes toward practical projects were generally moderate and similar across both populations. Together, these results suggest that while students share broadly positive perceptions of PBL, important differences in autonomy preferences, self-efficacy, and emotional framing of experiences are influenced by age, academic progression, and disciplinary context. Analysis of open-ended responses in the second survey highlights several important topics regarding the integration of practical projects into statistics education. Many students emphasized skill development, describing projects as very important in building applied competencies and deepening their understanding of statistical outputs beyond rote memorization. Additionally, a topic of student engagement emerged, with responses illustrating how practical projects fostered emotional investment, collaborative learning, and even new social connections among students who might otherwise remain isolated. Notably, some students expressed skepticism or critics, questioning the added value of the projects and suggesting alternative formats such as traditional exercises. Finally, students frequently recognized the importance of real-world applications, noting that practical projects bridge theoretical learning with essential professional skills such as data analysis, programming, and statistical reasoning for real-world contexts. Together, these findings reinforce the multifaceted benefits of PBL while also highlighting areas for further refinement to maximize the impact.

The findings highlight that integrating comprehensive practical projects into statistics education enhances skill development, real-world application, and student engagement. Differences in autonomy preferences, self-efficacy, and subjectivity between groups suggest that teaching strategies should be adapted to students' academic stage and disciplinary background. Providing structured support while allowing for independent work can better meet diverse needs. Addressing students' critiques, such as the desire for greater project relevance and assessment weight, can further improve motivation. Overall, thoughtful project design is essential for deepening understanding and building applied competencies in statistics.

RESEARCH LIMITATIONS

The main limitation of this study is its sample size. Although it spans several departments, both STEM and non-STEM, this analysis focused solely on STEM departments, resulting in a smaller sample. Additionally, the intercultural composition of the two departments posed challenges, as some students struggled to express their views in open-ended questionnaires.

REFERENCES

Chowdhary, K. R. (2020). Natural Language Processing. In K. R. Chowdhary (Ed.), *Fundamentals of Artificial Intelligence* (pp. 603–649). Springer. https://doi.org/10.1007/978-81-322-3972-7_19

Gibson, J. P., & Mourad, T. (2018). The growing importance of data literacy in life science education. *American Journal of Botany*, *105*(12), 1953–1956. https://www.jstor.org/stable/26617166

Granado-Alcón, M. del C., Gómez-Baya, D., Herrera-Gutiérrez, E., Vélez-Toral, M., Alonso-Martín, P., & Martínez-Frutos, M. T. (2020). Project-Based Learning and the Acquisition of Competencies and Knowledge Transfer in Higher Education. *Sustainability*, *12*(23). https://doi.org/10.3390/su122310062

Horton, N. J., & Hardin, J. S. (2015). Teaching the Next Generation of Statistics Students to "Think With Data": Special Issue on Statistics and the Undergraduate Curriculum. *The American Statistician*, *69*(4), 259–265. https://doi.org/10.1080/00031305.2015.1094283

Khalemsky, A., Gelbard, R., & Stukalin, Y. (2025). Constructing a Course on Classification Methods for Undergraduate Non-STEM Students: Striving to Reach Knowledge Discovery. *Journal of Statistics and Data Science Education*, 33(1), 68–76. https://doi.org/10.1080/26939169.2024.2320218

Khalemsky, A., & Stukalin, Y. (2024). Effect of students´ ethnic and cultural backgrounds on the success of practical projects within statistics and data science courses. *Statistics Education Research Journal*, *23*(2), Article 2. https://doi.org/10.52041/serj.v23i2.721

Khalemsky, A., & Stukalin, Y. (2023). Teaching classification module in data science course for undergraduate non-STEM students using project based learning methodology. In E.M. Jones (Ed.), *Fostering Learning of Statistics and Data Science. Proceedings of the Satellite conference of the International Association for Statistical Education (IASE)*. ISI/IASE. https://doi.org/10.52041/iase2023.104

Kokotsaki, D., Menzies, V., & Wiggins, A. (2016). Project-based learning: A review of the literature. *Improving Schools*, *19*(3), 267–277. https://doi.org/10.1177/1365480216659733

Kurczewska, A., Doryń, W., & Wawrzyniak, D. (2020). An Everlasting Battle between Theoretical Knowledge and Practical Skills? The Joint Impact of Education and Professional Experience on Entrepreneurial Success. *Entrepreneurial Business and Economics Review*, *8*(2), 219–237. https://doi.org/10.15678/EBER.2020.080212

Li, B., Jia,X., Chi, Y., Liu,X, & Jia, B. (2020). Project-based learning in a collaborative group can enhance student skill and ability in the biochemical laboratory: A case study. *Journal of Biological Education*, *54*(4), 404–418. https://doi.org/10.1080/00219266.2019.1600570

Montoyo, A., Martínez-Barco, P., & Balahur, A. (2012). Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments. *Decision Support Systems*, *53*(4), 675–679. https://doi.org/10.1016/j.dss.2012.05.022

Movahedzadeh, F., Patwell, R., Rieker, J. E., & Gonzalez, T. (2012). Project-Based Learning to Promote Effective Learning in Biotechnology Courses. *Education Research International*, *2012*, Article 536024. https://doi.org/10.1155/2012/536024

Pacella, D., Fabbricatore, R., D'Enza, A. I., Galluccio, C., & Palumbo, F. (2022). Teaching STEM Subjects in Non-STEM Degrees: An Adaptive Learning Model for Teaching Statistics. In F. Ouyand, P. Jiao, B. M. McLaren & A. H. Alavi (Eds.), *Artificial Intelligence in STEM Education* (chapter 5). CRC Press.

Parviz, M. (2024). AI in education: Comparative perspectives from STEM and Non-STEM instructors. *Computers and Education Open*, 6, Article 100190. https://doi.org/10.1016/j.caeo.2024.100190

Robeva, R. S., Jungck, J. R., & Gross, L. J. (2020). Changing the Nature of Quantitative Biology Education: Data Science as a Driver. *Bulletin of Mathematical Biology*, *82*(10), Article 127. https://doi.org/10.1007/s11538-020-00785-0

Vaughn, P., & Turner, C. (2016). Decoding via Coding: Analyzing Qualitative Text Data Through Thematic Coding and Survey Methodologies. *Journal of Library Administration*, *56*(1), 41–51. https://doi.org/10.1080/01930826.2015.1105035

Watson, J. M. (2017). Linking Science and Statistics: Curriculum Expectations in Three Countries. *International Journal of Science and Mathematics Education*, *15*(6), 1057–1073. https://doi.org/10.1007/s10763-015-9673-y