# Application of Variance Analysis to Measure the Impact of STEM Methodology Components in an Educational Intervention

**Alfredo García Suárez**
Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la Computación, Puebla, México, 72570
*alfredo.garciasuarez@viep.com.mx,*

**Juan Manuel González Calleros**
Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la Computación, Puebla, México, 72570
*jumagoca78@gmail.com*

and

**Josefina Guerrero García**
Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la Electrónica, Puebla, México, 72570
*Joseguga01@gmail.com*

**Abstract**

This study presents the application of variance analysis (ANOVA) to evaluate the effectiveness of the STEM (Science, Technology, Engineering, and Mathematics) methodology in an educational intervention. The primary objective was to measure the learning impact per session and assess the influence of three key components: preparation time, teacher performance, and social interaction. The analysis revealed no statistically significant differences in the average learning rate per session, suggesting a consistent learning outcome above 75% throughout the intervention. However, notable differences emerged when examining the influence of specific components. Regarding preparation time, courses with extensive preparation achieved an average learning rate of 94.5%, in contrast to only 13.64% for sessions with immediate preparation. In terms of teacher performance, instructors who exceeded teaching expectations reached an average learning rate of 92.74%, while those demonstrating deficiencies resulted in only 11.25%. Finally, classrooms characterized by dynamic social interaction reported a 95.14% learning rate, compared to just 10.19% under minimal interaction conditions. These findings emphasize that the effectiveness of STEM methodology is significantly enhanced when supported by thorough preparation, high teaching quality, and active classroom engagement. The study highlights the value of pedagogical planning and collaborative learning environments in optimizing STEM-based education.

**Keywords:** Data Science, STEM Education, Educational Data Analysis, Learning Impact Assessment.

## 1 Introduction

In recent years, Artificial Intelligence (AI) has advanced significantly, driven by the development of algorithms and computational tools that facilitate the analysis of large volumes of data and the creation of predictive models. In this context, the Py-thon ecosystem has established itself as one of the most widely used languages for AI research and development due to its simplicity, flexibility and the robustness of its libraries [1]. Among the most notable and widely used tools for the data analysis described in this work are Scikit-learn [2] and SciPy [3], which allow addressing a wide variety of machine learning, data processing and data analysis.

Another of the tools used for this data analysis was Numpy, which focuses on the efficient processing of large matrices and multidimensional arrays and is the cornerstone for high-performance numerical calculations. NumPy provides optimized data structures and advanced mathematical functions that allow working with data quickly and efficiently, which is essential in the implementation of machine learning models and other AI approaches [4]. Data analysis and results presented in this work were carried out from the data collected in an educational intervention where the STEM methodology was applied. The data collection instruments used in the educational intervention were interview formats that evaluated the direct contribution of various elements in the students' learning process, among which the following stand out: methodology, social interaction within the classroom, planning of practices, guidance or orientation in the development of practices, work combining theory and practice, preparation time for each session, resources used in the intervention, proposed bibliography, organization in the intervention, teacher performance and the level of understanding achieved by students on the topics taught in the educational intervention.

Based on the data obtained, the combined use of Scikit-learn and SciPy was implemented in the data analysis, which allowed us not only to build machine learning models [5], but also to address fundamental mathematical and statistical tasks that support these models. This article explores how these two libraries can be used together to solve real data analysis problems, from data cleaning and preprocessing to model interpretation and evaluation. This process is carried out with the purpose of providing accurate and efficient solutions [6].

Chapter 2 presents the state of the art on the approaches and techniques that have been used to analyze the STEM methodology.

Chapter 3 describes the data analysis methodology used in this research work.

Chapter 4 illustrates the results obtained using simple linear regression and analysis of variance (ANOVA).

Finally, Chapter 5 contains a discussion of the results and final conclusions.

## 2 Learning Analysis Techniques in STEM Educational Interventions

Over time, STEM methodology has gained increasing relevance in the educational field due to its comprehensive and multidisciplinary approach that fosters critical thinking, problem solving, and creativity among students. Its application in educational interventions seeks not only to improve the learning of technical concepts, but also to develop key skills for the 21st century, such as collaboration, innovation, and adaptability.

The success of applying STEM as a teaching methodology depends largely on an accurate and detailed evaluation of the results obtained, which has led to an increase in the use of data analysis strategies to measure the effectiveness of STEM pedagogical approaches. Through data analysis, educators can gain valuable insights into academic performance, student engagement, and the impact of the methodologies used in their learning process.

This article reviews the main analysis strategies and techniques used to evaluate the results of STEM educational interventions, exploring quantitative and qualitative approaches that allow for a deeper understanding of educational effects. This paper addresses methods such as statistical analysis, network analysis, machine learning, and data mining, which are essential for identifying patterns, predicting future performance, and improving teaching personalization. It also examines the challenges associated with implementing these techniques, including the data preprocessing stage, interpretation of results, and application of findings to adjust and improve pedagogical interventions.

One of the most widely used techniques for the evaluation and analysis of results derived from the application of the STEM methodology are Machine learning algorithms, as described in the analysis presented in [7], this study aimed to evaluate and predict factors affecting STEM students' future intention to enroll in chemistry-related courses. Results showed that attitude toward chemistry and perceived behavioral control represent the most influential factors, followed by autonomy and affective behavior. Another case of analysis was applying the Multivariate analysis of variance (MANOVA) technique, this multivariate analysis was used to determine possible differences between male and female math anxiety and math self-efficacy levels. Male engineering students reported higher self-efficacy and lower math anxiety levels, and this difference was shown to be significant according to the MANOVA results [8]. In the research work presented in [9], the technique called Error Analysis (EA) was used to a wide corpus of exercises and essays written by third year students of mechanical engineering, with the main purpose of achieving a precise diagnosis of the students' strengths and weaknesses in writing skills. In the educational intervention described in [10], a Questionnaire developed from the TPB (Theory of Planned Behavior) model was applied, this study analyzed STEM education's effect on students' learning satisfaction. Extending the planned behavior theory, this study aimed to predict high school students' learning satisfaction with STEM education. Attitude is the most important factor influencing student satisfaction and acceptance toward STEM education.

There are other works that focus on student satisfaction as de-scribed in [11], where it indicates that the National Student Survey (NSS) has become one of the most important metrics to assess student satisfaction that directly influences the university league tables and the Teaching Excellence Framework (TEF). The NSS results and it was concluded that there is strong evidence, that "teaching" and "Organization and management" are the vital influential factors on the overall satisfaction of students.

Among the classification algorithms is the Classification algorithm CNN (Convolutional Neural Network). This technique was used in [12]. In this study, automatic assessment methods for piano performances were developed. The results show that the CNN (Convolutional Neural Network) approach is superior to the other approaches (Support Vector Machine (SVM), Naive Bayes (NB) and Long Short-Term Memory (LSTM), with a classification accuracy of more than eighty percent.

Analyses on STEM methodology have not only focused on students, but also on assessment processes as described in [13], where the Display and Evaluate (DE) technique was applied. This technique consists of self-evaluation, group evaluation, and teacher evaluation. Self-evaluation focuses on finding and filling new knowledge and skills in the learning process; group peer assessments can draw on peers' opinions and suggestions to provide a learning atmosphere for mutual learning and improvement; teacher evaluation focuses on process evaluation rather than just result.

Regarding correlation algorithms, the study presented in [14] describes the application of CT (Computational Thinking) assessment, this tool sheds light on the correlation between algorithmic thinking skills and age in early childhood, revealing that age is a predictor factor for algorithmic thinking and, therefore, for CT.

Another of the factors analyzed is the mood and mental health of the teachers. In the study presented in [15], a mixed-method approach (cross-sectional study and a bibliometric analysis) is described, this study analyzes the factors caused anxiety or depression in teachers during pandemic COVID-19 combining a cross-sectional study of university teachers from Ecuador and Spain with a medium of twenty years of working experience and a bibliometric analysis carried out in three databases. The combination of the results indicated that mental health in STEM teachers at universities is related to diverse factors, from training to the family and working balance.

In the review of the literature, various qualitative studies were also found, as illustrated in [16] where they carried out an analysis based on interests, prestige, and sex type, this study investigates career profiles of German university freshmen based on interests, prestige, and sex type. Of note, especially female students in STEM subjects with a low proportion of females distributed widely across the 11 profiles identified. Another qualitative investigation is presented in [17], where an analysis is carried out over the explanations and tasks set out in different science textbooks, findings showed inconsistencies detected in the distribution of volume measurement-related content, as well as in the strategies, units' tools used in the two areas.

Table 1 presents a summary of the studies that have been found in the state of the art on some Techniques for evaluating approaches to STEM methodology such as Machine learning algorithms, Multivariate analysis of variance, Classification algorithm CNN, Error Analysis, Computational Thinking, Mixed-method approach, Cor-relation analysis and others.

**Table 1:** Techniques for evaluating approaches to STEM methodology: state of the art

| Assessment | Data Collection Instrument | Target Variable | Reference |
|---|---|---|---|
| Machine learning algorithms | AI Quantitative | Students' intention | [7] |
| Multivariate analysis of variance (MANOVA) | Quantitative | Anxiety | [8] |
| Error Analysis (EA) | Quantitative | Writing skills | [9] |
| TPB (Theory of Planned Behavior) model | Quantitative | Students' learning satisfaction | [10] |
| National Student Survey (NSS) | Quantitative | Satisfaction of students | [11] |
| Classification algorithm CNN (Convolutional Neural Network) | Quantitative | Performance of students | [12] |
| Display and Evaluate (DE) | Quantitative (diversified) | Process evaluation | [13] |

| | | | |
|---|---|---|---|
| CT (Computational Thinking) assessment | Quantitative | Correlation between algorithmic thinking skills and age | [14] |
| Mixed-method approach (cross-sectional study and a bibliometric analysis) | Quantitative and Qualitative | Anxiety or depression in teachers | [15] |
| Analysis based on interests, prestige, and sex type | Qualitative | Career profiles of university students | [16] |
| Analysis over the explanations and tasks set out in different science textbooks | Qualitative | Inconsistencies in science textbooks | [17] |

## 3 Methodology

The data analysis presented in this paper used the data collected from an educational intervention with high school students. The data collection instruments used were interviews and class diaries. The topics covered during this intervention were software development and hardware implementation. The STEM methodology was the teaching methodology applied during the educational intervention. In research papers [18] and [19] the experimental development of this teaching-learning experience is described. The data analysis was developed in three stages: Data preprocessing (stage 1), Linear correlation analysis (stage 2) and Analysis of Variance ANOVA (stage 3). Fig. 1 shows the general structure of the data analysis.

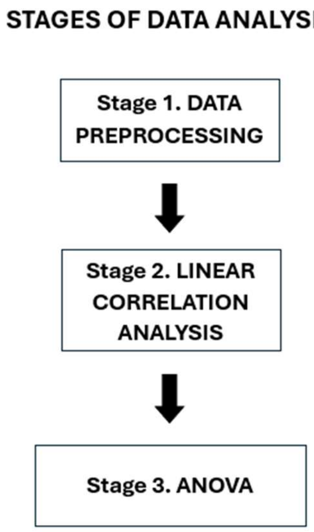


**Figure 1**: Stages of data analysis performed

### 3.1 Description of the educational intervention

For this research, a practical teaching-learning case was developed using the STEM Maker methodology. A diverse group of students was formed, with educational levels ranging from elementary school to postgraduate studies. The participants in this course-workshop were between 13 and 49 years old. Table 2 presents the general age distribution.

**Table 2:** Frequency table of ages of the participants

| Age (years) | Participants (Number) |
|---|---|
| 13 | 1 |
| 15 | 4 |
| 16 | 8 |
| 17 | 8 |
| 18 | 1 |
| 19 | 1 |
| 20 | 1 |
| 21 | 10 |
| 22 | 14 |
| 23 | 1 |
| 24 | 2 |
| 28 | 1 |
| 32 | 1 |
| 40 | 1 |
| 42 | 1 |
| 44 | 1 |
| 45 | 1 |
| 49 | 1 |

The STEM Maker workshop was conducted over 16 sessions, held once a week. Each session lasted 2 hours. The total session time was structured into specific stages, divided as follows:

Stage 1 (Introduction): Students were welcomed, and key concepts and theories relevant to the session were explained (30 minutes).

Stage 2 (Instruction): The instructor provided a step-by-step explanation of practical activity, demonstrating how to apply both hardware and software tools, integrating the concepts introduced in Stage 1 (30 minutes).

Stage 3 (Hands-on Practice): Students engaged in practical exercises, applying the necessary software and setting up the appropriate hardware.

Stage 4 (Challenge): A problem-solving activity was presented for students to complete, requiring them to apply both theoretical and technical knowledge. This stage aimed to strengthen skills aligned with the intended learning outcomes.

At the end of Stage 3 in each session, the following assessment tools were used to evaluate students: Class diary, course quality interview, and individual progress scale.

These tools were used to gather feedback on the workshop and to collect data for future comparative analysis with results from other instructional strategies.

### 3.2 Data preprocessing

In the data preprocessing stage, the following actions were performed on the datasets collected from the educational intervention: Data extraction, imputation of null values, treatment of outliers, conversion of variables and ordinal coding of the independent variables and the dependent variable.

Data extraction. In this process, a database was created that included the column of the target variable (Level Comprehension) and the columns of the independent variables (Methodology, Social Interaction, Planning Practice, Practical Guidance, Theoretical-Practical Work, Preparation Time, Resources, Bibliography, Organization, Performance Teacher).

Imputation of Null Values. This action was performed using the "backward fill" technique.

Outliers Treatment. The variables were expressed in text categories, where no outliers were found.

Cordinal coding of variables. Coding was necessary to be able to process the numerical algorithms, this was because the data variables were originally of the "string" type.

The diagram in Fig. 2 illustrates the actions performed in the data preprocessing stage.

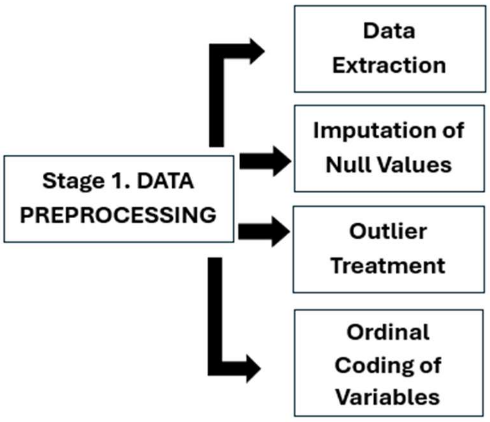


**Figure 2**: Preprocessing Stage Actions

### 3.3 Analysis of Linear Regression

In the linear correlation analysis stage, the following actions were performed on the dataset obtained from the previous stage (preprocessing): Definition of the dependent variable (target) and the independent variables, analysis of simple correlations and identification of the multiple linear regression model with the highest correlation with respect to the target variable.

Definition of variables. In this process, the dependent variable is assigned to a vector ("y"), and the independent variables are assigned to a matrix array ("X").

Simple linear correlation analysis. In this action, the matrix of all correlations between the intersections of all numerical variables is obtained and the variables with the highest correlation with the target variable are identified using a heatmap.

Implementation of the multiple regression model. The variables identified in the previous step are used as inputs to the model to predict the target variable. Finally, the model is evaluated with the correlation ($\rho$) and determination (R2) coefficients obtained. The actions performed in the linear correlation analysis stage are described in Fig. 3.

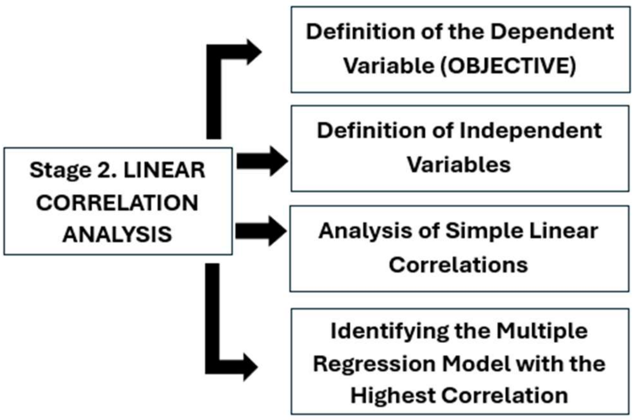


**Figure 3**: Description of linear regression analysis

### 3.4 Analysis of variance ANOVA

In the ANOVA analysis stage, the following actions were performed with the preprocessed dataset: The first ANOVA, the target variable was the course session dates, and the percentage of learning achieved through the application of the STEM methodology was chosen as the independent variable.

In the second ANOVA, the target variable was course preparation time, and the percentage of learning achieved through the application of the STEM methodology was chosen as the independent variable.

In the third ANOVA, the target variable was the instructor's performance in the course, and the percentage of learning achieved through the application of the STEM methodology was chosen as the independent variable.

Finally, in the fourth ANOVA, social interaction within the classroom was chosen as the target variable, and the percentage of learning achieved through the application of the STEM methodology was chosen as the independent variable. The actions performed in the ANOVA analysis stage are described in Fig. 4.

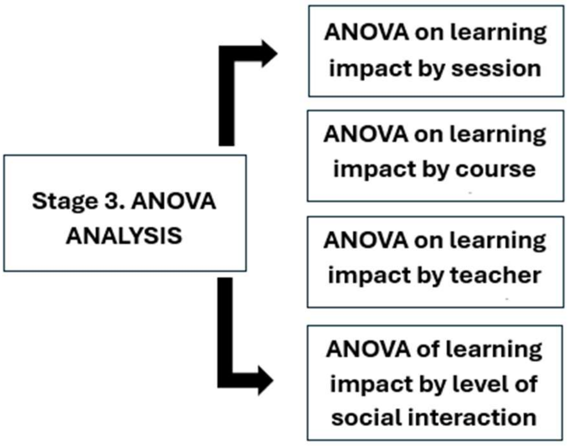


**Figure 4**: Description of ANOVA analysis

## 4 Results

Results derived from the data analysis are presented below in three sections: results of the simple linear regression analysis, results of the multiple linear regression analysis and results of the ANOVA analysis.

### 4.1 Results of the data analysis with simple linear regression

The simple linear regression analysis fits the model with the dependent variable (Level Comprehension) and the independent variables (Methodology, Social Interaction, Planning Practice, Practical Guidance, Theoretical-Practical Work Preparation Time, Resources, Bibliography, Organization and Performance Teacher).

In the first step, Pearson correlation coefficients (ρ) were obtained for all combinations of variables, as illustrated in the heatmap in Fig. 5. In the second step, only the correlations related to the target variable were extracted and ordered from highest to lowest as shown in Table 2, where it can be observed that all independent variables have a correlation greater than ρ=0.67, which indicates that they are in a range of considerable correlation to strong correlation. In addition, the three variables that have the highest correlation with respect to the students' Level Comprehension are "Resources (ρ= 0.796273)," "Bibliography (ρ= 0.793834)" and "Preparation Time (ρ= 0.793679)".

The graph in Fig. 6 illustrates the original values (blue dots) of the variable "Level Comprehension" compared to the predicted values (red dots) by the model with the highest correlation described in equation 1. The independent variable "Resources" represents the x-axis, while the dependent variable "Level Comprehension" represents the y-axis.

$$y = (0.798)x + 0.402 \tag{1}$$

Where:
y= Level Comprehension Predicted
x= Resources

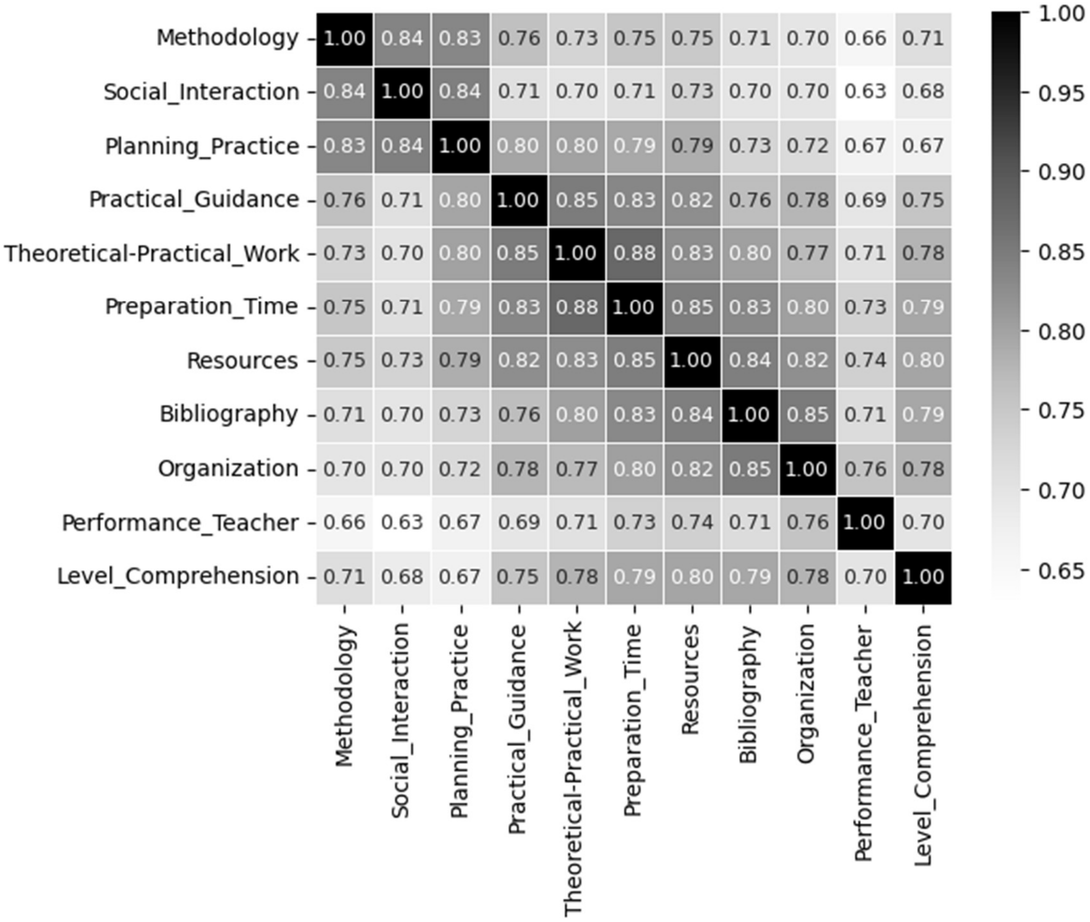


**Figure 5**: Heatmap of simple linear correlations

**Table 3:** Correlation table with respect to the target variable (Level Comprehension).

| Independent variable | Correlation coefficient (ρ) | Coefficient of determination ($R^2$) |
|---|---|---|
| Resources | 0.796273 | 0.63405 |
| Bibliography | 0.793834 | 0.63017 |
| Preparation Time | 0.793679 | 0.62992 |
| Organization | 0.778737 | 0.60643 |
| Theoretical-Practical Work | 0.778311 | 0.60576 |
| Practical Guidance | 0.753706 | 0.56807 |
| Methodology | 0.712076 | 0.50705 |
| Performance Teacher | 0.699197 | 0.48887 |
| Social Interaction | 0.682951 | 0.46642 |
| Planning Practice | 0.671273 | 0.45060 |

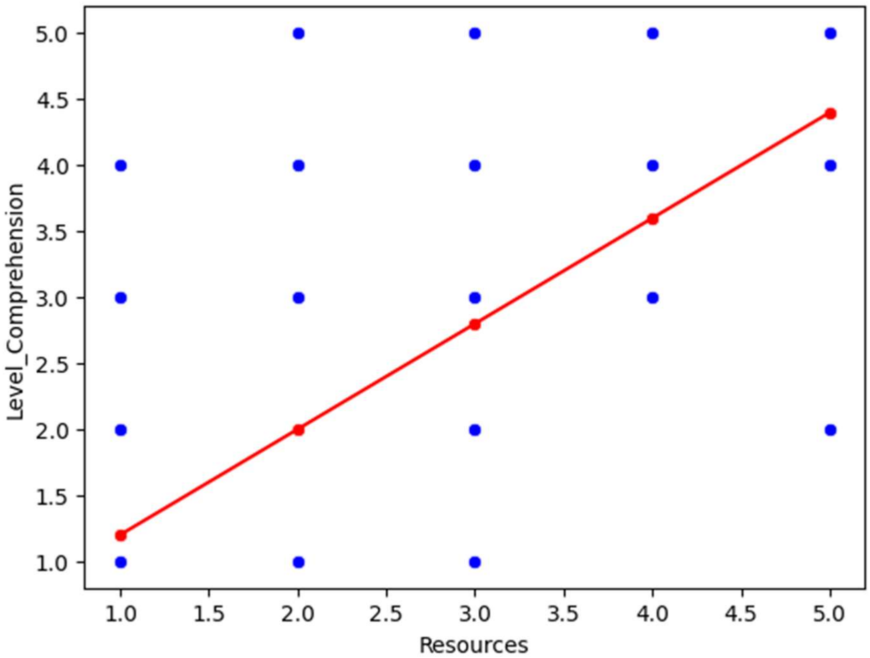


**Figure 6**: Original values of level comprehension (blue dots) Predicted values with the simple linear regression model (red dots)

In the multiple linear regression analysis, the 5 independent variables with the highest correlation (Resources, Bibliography, Preparation Time, Organization, Theoretical-Practical Work) with respect to the target variable (Level Comprehension) were chosen to adjust the prediction model.

The correlation obtained with the multiple linear regression model (described in equation 2), was ($\rho = 0.8493$) and a coefficient of determination of ($R^2 = 0.7213$). The correlation achieved with the multiple regression model was greater than any correlation obtained with the simple linear regression models.

$$y = 0.169x_1 + 0.171x_2 + 0.193x_3 + 0.196x_4 + 0.197x_5 + 0.158 \quad (2)$$

Where:
$y$= Level Comprehension Predicted
$x_1$= Theoretical-Practical Work
$x_2$= Preparation Time
$x_3$= Bibliography
$x_4$= Resources
$x_5$= Organization

**4.2 Results of the data analysis with ANOVA**

An Analysis of Variance (ANOVA) was conducted to compare the average learning rates across different dates, with the objective of evaluating the impact of the STEM methodology on learning outcomes throughout the sessions.

The ANOVA results yielded a p-value of 0.731, indicating that there was no statistically significant difference in the average learning rates between sessions. Despite this, the average learning rates remained consistently high, ranging from 75.6% to 100%. This suggests that, while the STEM methodology did not lead to significant variations in learning progression across sessions, the overall learning performance remained consistently above 75% throughout the study period.

These findings (showed in Fig. 7) imply that the methodology may have contributed to maintaining a high level of learning effectiveness, even if no significant temporal changes were observed.

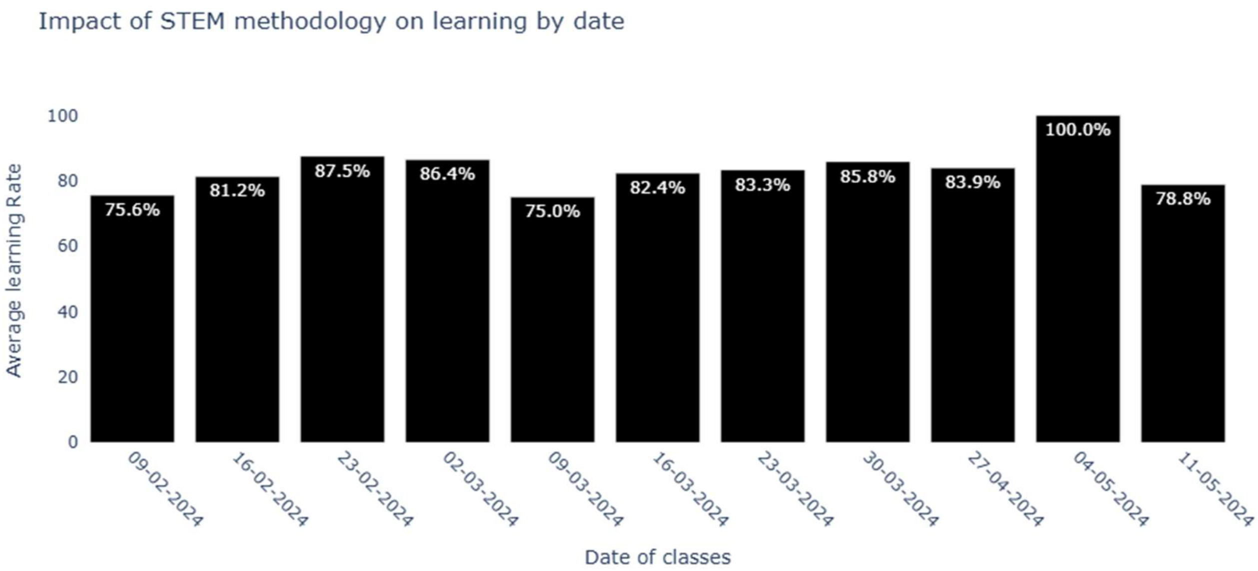


**Figure 7**: Average learning rate by session

A second Analysis of Variance (ANOVA) was conducted to compare the average learning rates based on the course preparation time, with the aim of evaluating the impact of the STEM methodology depending on the amount of time dedicated to course preparation.

The ANOVA results revealed a highly significant difference between the groups, with a p-value of 1.132e-59, indicating that the time spent on preparation had a profound effect on the learning outcomes.

The mean learning rates for each preparation time were as follows: 94.59% for extensive preparation, 83.43% for extended preparation, 73.15% for moderate preparation, 21.43% for short preparation, and 13.64% for immediate preparation.

These findings (illustrated in Fig. 8) suggest that the amount of preparation time significantly influenced the effectiveness of the STEM methodology. The highest learning rates were observed in the groups with more extensive preparation, while those with minimal or no preparation showed notably lower learning rates. This emphasizes the importance of adequate preparation time in maximizing the impact of the STEM approach on student learning outcomes.

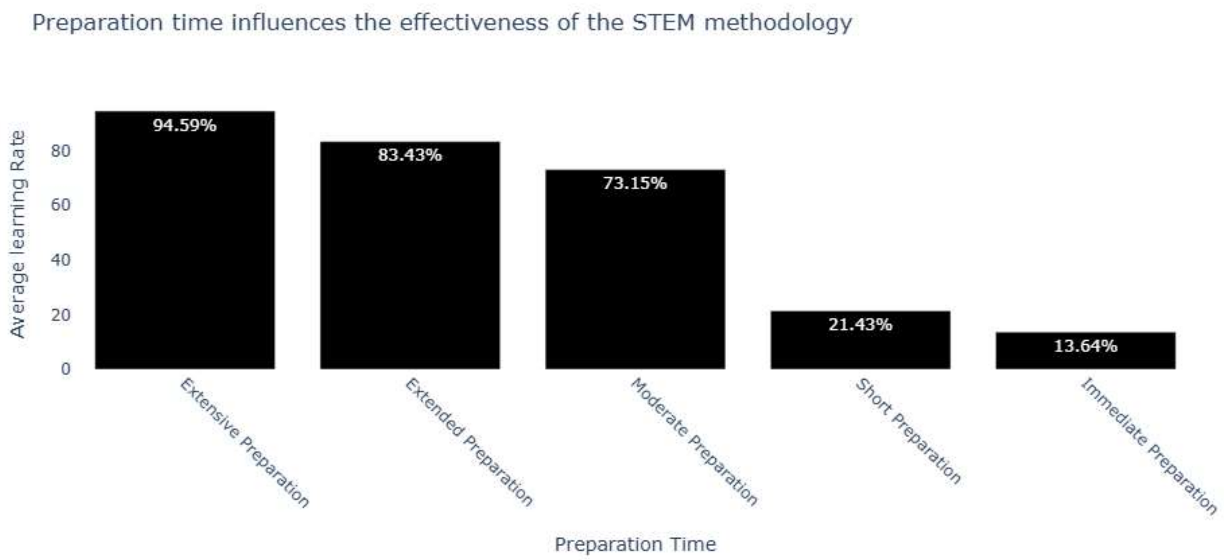


**Figure 8**: Average learning rate by preparation time of the course

A third Analysis of Variance (ANOVA) was conducted to compare the average learning rates based on the performance of the instructor, with the aim of assessing the impact of the STEM methodology in relation to the instructor's effectiveness in the classroom.

The ANOVA results showed a highly significant difference between the groups, with a p-value of 2.857e-46, indicating that the instructor's performance had a substantial effect on the learning outcomes of the students.

The mean learning rates for each level of instructor performance were as follows: 92.74% for "exceeds expectations", 83.09% for "performs effectively", 82.59% for "demonstrates strong teaching skills", 45.45% for "meets minimum requirements", and 11.25% for "shows deficiencies".

These findings (illustrated in Fig. 9) highlight the critical role of instructor performance in the success of the STEM methodology. The highest learning rates were observed when the instructor exceeded expectations or performed effectively, whereas much lower learning rates were seen when the instructor showed deficiencies or only met the minimum requirements. This suggests that the teacher's ability to engage students and effectively deliver the STEM curriculum is a key factor in maximizing learning outcomes.

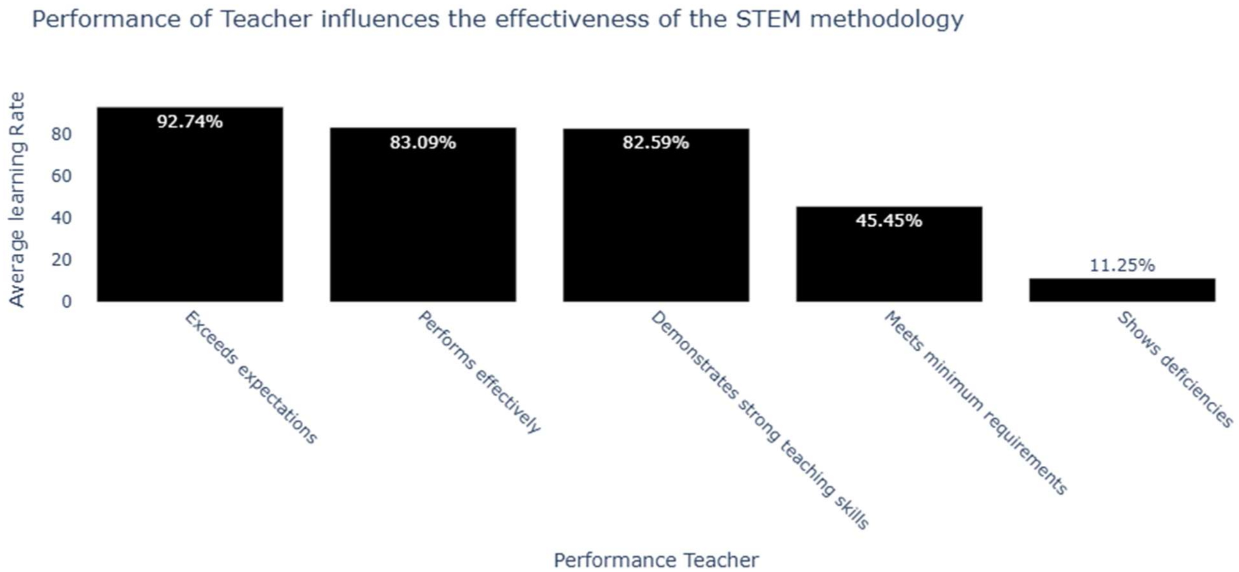


**Figure 9**: Average learning rate by performance of teacher in classes

A fourth Analysis of Variance (ANOVA) was conducted to compare the average learning rates based on the level of social interaction in the classroom, with the objective of assessing the impact of the STEM methodology depending on the extent of student engagement and interaction during lessons.

The ANOVA results revealed a highly significant difference between the groups, with a p-value of 8.890e-82, indicating that the level of social interaction had a profound effect on the learning outcomes.

The mean learning rates for each level of social interaction were as follows: 95.14% for dynamic interaction, 81.62% for active interaction, 67.86% for moderate interaction, 25.00% for limited interaction, and 10.19% for minimal interaction.

These findings (showed in Fig. 10) suggest that greater social interaction within the classroom plays a crucial role in enhancing the effectiveness of the STEM methodology. The highest learning rates were observed in environments with dynamic interaction, while lower learning rates were associated with limited or minimal interaction. This highlights the importance of fostering an interactive and engaging classroom environment to optimize student learning outcomes in STEM-based instruction.

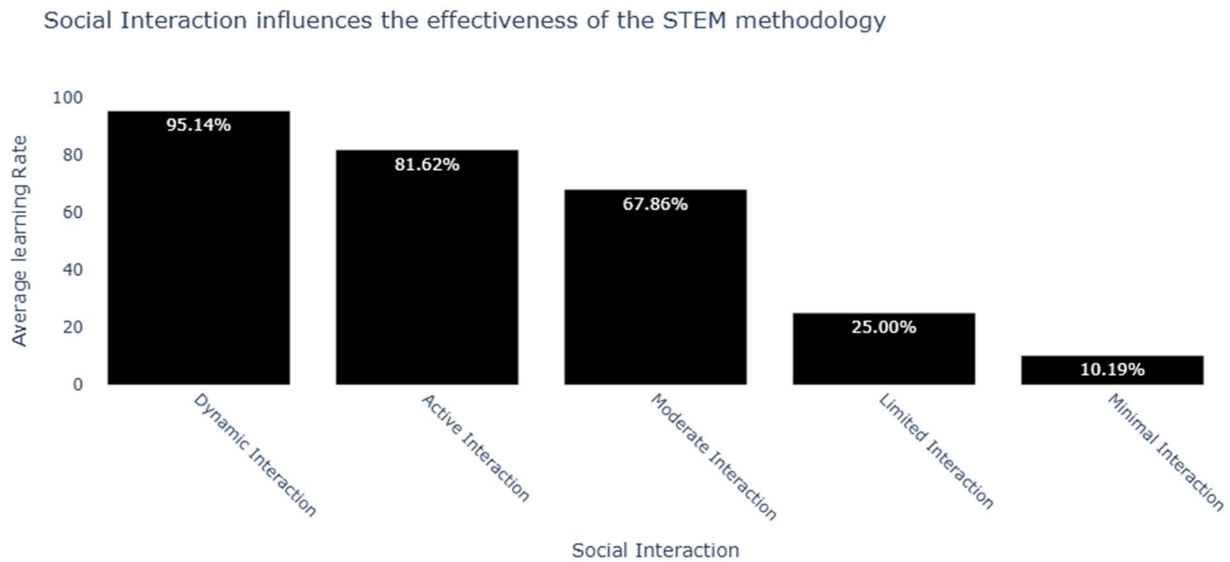


**Figure 10**: Average learning rate by social interaction in classroom

## 5 Conclusions

The results of the four ANOVA analyses provide valuable insights into the factors influencing the effectiveness of the STEM methodology in fostering student learning outcomes.

The first analysis (Impact of Date on Learning Rates) revealed that there were no significant changes in learning rates across different sessions, suggesting that the STEM methodology led to consistent learning outcomes. Although the p-value indicated no statistical significance, the consistent performance above 75% in all sessions suggests that the approach maintained a high level of student engagement and knowledge retention throughout the study period.

The second analysis (Impact of Course Preparation Time) highlighted the critical role of preparation time in enhancing learning outcomes. The results showed that as the preparation time increased, so did the average learning rate, with the highest rates observed in the extensive preparation group (94.59%). This demonstrates that a well-prepared course structure is essential for maximizing the effectiveness of STEM teaching strategies.

The third analysis (Impact of Instructor Performance) emphasized the importance of the instructor's performance in shaping student learning outcomes. Instructors who exceeded expectations or performed effectively led to significantly higher learning rates, while those who showed deficiencies had notably lower rates. This underscores the pivotal role that instructor effectiveness plays in the successful implementation of the STEM methodology.

The fourth analysis (Impact of Social Interaction) demonstrated that the level of social interaction in the classroom had a profound effect on student learning. High levels of interaction, such as dynamic or active engagement, were associated with significantly higher learning rates (95.14% and 81.62%, respectively). In contrast, minimal or limited interaction resulted in much lower learning outcomes. This indicates that fostering an interactive and collaborative learning environment is crucial for enhancing the impact of the STEM methodology.

These findings suggest that the STEM methodology's success is influenced by several key factors: preparation time, instructor performance, and social interaction in the classroom. Future research will focus on exploring how these elements can be optimized to further improve the effectiveness of STEM teaching strategies. Additionally, it may be valuable to investigate how these factors interact with other variables, such as student motivation and prior knowledge, to provide a more comprehensive understanding of the methodology's impact.

One of the challenges encountered during the Maker workshop was the limited physical space available for hands-on activities, which made it difficult to work comfortably with physical components. This issue will be considered when planning a second edition of the STEM Maker workshop. Another challenge was the number of participants relative to the presence of only one facilitator, which made it difficult to provide timely feedback. Reducing the number of attendees could enhance the effectiveness of support and monitoring.

Despite these challenges, the workshop had a positive impact across all participant contexts. Some of the experiences shared included receiving awards in STEM-related fields, helping high school students define their academic focus, and the development of more ambitious projects, among others.

For future iterations, the goal is to address the identified limitations to provide a more meaningful learning experience, particularly aimed at high school and university educators.

## References


[1] T. E. Oliphant, “Python for scientific computing,” Computing in Science & Engineering, vol. 9, no. 3, pp. 10-20, May–Jun. 2007. [Online]. Available: https://doi.org/10.1109/MCSE.2007.58

[2] S. Raschka, Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow, 1st ed. Packt Publishing Ltd., 2015.

[3] E. Jones, T. Oliphant, and P. Peterson, “SciPy: Open-source scientific tools for Python,” 2001. [Online]. Available: https://www.scipy.org/

[4] S. van der Walt, S. C. Colbert, and G. Varoquaux, “The NumPy array: A structure for efficient numerical computation,” Computing in Science & Engineering, vol. 13, no. 2, pp. 22-30, Mar–Apr. 2011. [Online]. Available: https://doi.org/10.1109/MCSE.2011.37

[5] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, … and E. Duchesnay, “Scikit-learn: Machine learning in Python,” Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011. [Online]. Available: https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf

[6] L. Buitinck, G. Louppe, A. Joly, and J. Vanderplas, “API design for machine learning software: Experiences from the scikit-learn project,” in ECML PKDD Workshop: Languages for Data Mining and Machine Learning, 2013, pp. 108-122. [Online]. Available: https://scikit-learn.org/stable/about.html

[7] A. K. S. Ong, “A machine learning ensemble approach for predicting factors affecting STEM students’ future intention to enroll in chemistry-related courses,” Sustainability, vol. 14, no. 23, pp. 16041-16058, Nov. 2022. [Online]. Available: https://doi.org/10.3390/su142316041

[8] G. Morán-Soto and O. I. González-Peña, “Mathematics anxiety and self-efficacy of Mexican engineering students: Is there gender gap?,” Education Sciences, vol. 12, no. 6, pp. 391-403, Jun. 2022. [Online]. Available: https://doi.org/10.3390/educsci12060391

[9] M.-J. Arévalo, M. A. Cantera, V. García-Marina, and M. Alves-Castro, “Analysis of university STEM students’ mathematical, linguistic, rhetorical–organizational assignment errors,” Education Sciences, vol. 11, no. 4, pp. 173-190, Apr. 2021. [Online]. Available: https://doi.org/10.3390/educsci11040173

[10] A. Sofroniou, B. Premnath, and K. Poutos, “Capturing student satisfaction: A case study on the National Student Survey results to identify the needs of students in STEM related courses for a better learning experience,” Education Sciences, vol. 10, no. 12, pp. 378-400, Dec. 2020. [Online]. Available: https://doi.org/10.3390/educsci10120378

[11] J. Zhao, T. T. Wijaya, M. Mailizar, and A. Habibi, “Factors influencing student satisfaction toward STEM education: Exploratory study using structural equation modeling,” Applied Sciences, vol. 12, no. 19, pp. 9717-9737, Sep. 2022. [Online]. Available: https://doi.org/10.3390/app12199717

[12] V. Phanichraksaphong and W.-H. Tsai, “Automatic evaluation of piano performances for STEAM education,” Applied Sciences, vol. 11, no. 24, pp. 11783-11805, Dec. 2021. [Online]. Available: https://doi.org/10.3390/app112411783

[13] M. S. Domínguez and D. Mocencahua, “Propuesta educativa del movimiento maker como herramienta para generar estrategias de aprendizaje de matemáticas,” in Tecnologías aplicables a la educación: innovación educativa, XI Encuentro Iberoamericano de Educación, vol. 1, no. 1, 2016. [Online]. Available: https://bit.ly/3lvuRs2

[14] K. Kanaki and M. Kalogiannakis, “Assessing algorithmic thinking skills in relation to age in early childhood STEM education,” Education Sciences, vol. 12, no. 6, pp. 380-401, Jun. 2022. [Online]. Available: https://doi.org/10.3390/educsci12060380

[15] J. A. Navarro-Espinosa, M. Vaquero-Abellán, A.-J. Perea-Moreno, G. Pedrós-Pérez, P. Aparicio-Martínez, and M. P. Martínez-Jiménez, “The influence of technology on mental well-being of STEM teachers at university level: COVID-19 as a stressor,” International Journal of Environmental Research and Public Health, vol. 18, no. 18, pp. 9605-9628, Sep. 2021. [Online]. Available: https://doi.org/10.3390/ijerph18189605

[16] D. Mouton, F. G. Hartmann, and B. Ertl, “Career profiles of university students: How STEM students distinguish regarding interests, prestige and sextype,” Education Sciences, vol. 13, no. 3, pp. 324-349, Mar. 2023. [Online]. Available: https://doi.org/10.3390/educsci13030324

[17] A. Hasani, D. E. Juansah, I. J. Sari, and R. A. Z. El Islami, “Conceptual frameworks on how to teach STEM concepts in Bahasa Indonesia subject as integrated learning in Grades 1–3 at elementary school in the Curriculum 2013 to contribute to sustainability education,” Sustainability, vol. 13, no. 1, pp. 173-188, Jan. 2020. [Online]. Available: https://doi.org/10.3390/su13010173

[18] A. García, J. M. Gonzalez, and J. Guerrero, “Application of the STEM methodology in the development of a closed-loop control system,” in Proceedings of the 13th International Conference on Virtual Campus (JICV 2023), 2023. [Online]. Available: https://doi.org/10.1109/JICV59748.2023.10565656

[19] A. Garcia Suárez, J. M. Gonzalez Calleros, and J. Guerrero Garcia, “Measurement of efficiency parameters from the application of STEM Maker teaching: A practical case,” in Proc. 2023 Conference, 2023. [Online]. Available: https://doi.org/10.1145/3630970.3631049