*International Journal of Basic and Applied Sciences*, 14 (6) (2025) 115-122

International Journal of Basic and Applied Sciences

*Website: www.sciencepubco.com/index.php/IJBAS*
*https://doi.org/10.14419/y9ynz887*
***Research paper***

# Uncovering The Impact of COVID-19 Disruptions on Students Mathematics Achievement: A CART Analysis of Selected PISA 2022 Data

**Pei-Ching Chao [1], Gregory S. Ching [2] ***

*[1] Center for Teacher Education, Fu Jen Catholic University, Taiwan*
*[2] Graduate Institute of Educational Administration and Policy, National ChengChi University, Taiwan*
**Corresponding author E-mail: gching@nccu.edu.tw*

*Received: September 24, 2025, Accepted: October 2, 2025, Published: October 8, 2025*

**Abstract**

The COVID-19 pandemic disrupted education worldwide, with mathematics learning particularly affected due to its reliance on cumulative knowledge and structured instruction. This study investigates the influence of socioeconomic background and pandemic-related disruptions on mathematics achievement across countries using data from the Program for International Student Assessment (PISA) 2022 COVID-19 module. The dataset included 109,097 secondary students from 17 participating countries, with mathematics performance measured as the average of ten plausible values. Predictor variables included socioeconomic status, emotional impact, perceived learning loss, family support, and access to digital resources. Multiple linear regression analysis was applied to identify independent contributions of each predictor, while Classification and Regression Tree (CART) modeling captured non-linear interactions and threshold effects. Results showed that socioeconomic status was the strongest positive factor, followed by digital access as a modest contributor, whereas perceived learning loss and emotional impact emerged as strong negative influences. Family support showed limited predictive power when modeled together with other variables. CART analysis further demonstrated that students with high socioeconomic status and low learning loss were most likely to achieve higher mathematics scores, while students with low socioeconomic status were consistently classified as low achievers regardless of other conditions. These findings highlight how COVID-19 amplified pre-existing inequalities in mathematics education, revealing that disadvantage and disruption interact to magnify vulnerability. The study underscores the need for equity-focused recovery policies that address both structural socioeconomic gaps and targeted interventions for learning recovery in mathematics.

***Keywords***: *COVID-19 Educational Disruption; Mathematics Achievement; Socioeconomic Status; Learning Loss; Digital Access; Regression Analysis; Decision Tree Modeling; Educational Equity.*

## 1. Introduction

The COVID-19 pandemic has caused unprecedented disruptions to education systems around the globe (Chin et al., 2022; d'Orville, 2020). As schools closed and transitioned to remote learning, students' academic experiences were dramatically altered, raising urgent concerns about the long-term impacts of these disruptions on learning outcomes, particularly in foundational subjects such as mathematics (Contini et al., 2022; Schult et al., 2022). While research has begun to document the general effects of the pandemic on student achievement, less is known about how specific types of COVID-19-related disruptions, such as emotional strain, perceived learning loss, family support, and access to digital resources, which jointly shape academic performance, more specifically within mathematics achievement across diverse international contexts. Understanding such dynamics requires not only descriptive statistics but also formal mathematical modeling to quantify, predict, and classify student outcomes under pandemic conditions.

Emotional disruption, often tied to fear, uncertainty, and isolation during the pandemic, has been shown to interfere with cognitive processes vital to learning, such as attention, memory, and problem-solving (Zaccoletti et al., 2020). Students experiencing heightened emotional distress have reported lower motivation and engagement in remote settings (Bubb & Jones, 2020; Camacho et al., 2021), potentially leading to decreased academic performance. In addition, learning loss, especially in mathematics, emerged as a central concern for policymakers and educators during school closures (Donnelly & Patrinos, 2022; Tashtoush et al., 2023). Unlike literacy, which can often be reinforced through everyday life, mathematics learning relies heavily on structured instruction and cumulative knowledge (Nesher, 1986). Studies have shown that when students returned to school during the Fall of 2020, only around 37 to 50% of the typical gains in mathematics learning were recorded (Kuhfeld et al., 2020). Perceived learning loss, even when not directly measured, has also been linked to lower academic self-concept and achievement (Kania & Juandi, 2023).

Similarly, family support during remote learning played a protective role, but its effectiveness was unevenly distributed (Sosa Díaz, 2021). Parents from more advantaged backgrounds were more likely to provide cognitive and emotional support, supervise learning activities, and offer supplementary materials (Gaxiola Romero et al., 2022; Scrimin et al., 2022; Xian et al., 2023). In contrast, socioeconomically disadvantaged families often lacked conducive home learning environments (Easterbrook et al., 2023). Access to digital resources, such as devices, internet connectivity, and online platforms, was also a major determinant of participation in remote learning (Mathrani et al., 2022; Ramsetty & Adams, 2020). The so-called "digital divide" compounded pre-existing inequalities and placed marginalized students at greater risk of academic setbacks (Di Pietro, 2021; Di Pietro et al., 2020). Research consistently shows that mathematics learning was more negatively affected than reading, with slower and more uneven recovery (Hammerstein et al., 2021).

Critical synthesis and gaps - Evidence on pandemic learning loss remains heterogeneous across contexts and methods. For example, Engzell et al. (2021) estimate large losses concentrated among disadvantaged students, whereas Donnelly and Patrinos (2022) document wide variability depending on grade level, subject, and assessment period. Similarly, findings on digital access diverge: while Ramsetty and Adams (2020) emphasize how the lack of connectivity directly excludes students, Mathrani et al. (2022) argue that access alone was insufficient without pedagogical orchestration and home routines. These inconsistencies highlight two key gaps: 1) the need to jointly model emotional, instructional, familial, and technological disruptions within a large multi-country dataset; and 2) the value of applying complementary parametric (regression) and non-parametric (CART) approaches to capture both average effects and threshold-based interaction patterns.

This study conceptualizes these pandemic-related factors as predictor variables $X_i$ within two complementary models: multiple linear regression and CART analysis. In regression, mathematics achievement $Y$ is modeled as a linear function of predictors,

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \epsilon$$

where $X_1$ represents socioeconomic status (ESCS), $X_2$ emotional impact, $X_3$ perceived learning loss, $X_4$ family support, and $X_5$ digital access. This framework estimates the independent contribution of each factor to achievement. In contrast, CART analysis models non-linear and interaction effects by recursively partitioning the predictor space. At each node $t$, the impurity is computed as

$$G(t) = 1 - \sum_{i=1}^{c} p(i|t)^2$$

and the split that minimizes weighted impurity across child nodes is selected. This produces an interpretable decision function $f(X)$ that classifies students into high or low achievers based on conditional thresholds (e.g., high ESCS and low learning loss).

Overall, this study is conceptually grounded in Bronfenbrenner's (2000) Ecological Systems Theory which posits that individual development and learning are shaped by interactions across multiple nested systems, from immediate contexts such as family and school (micro-systems) to broader societal structures (macro-systems). The COVID-19 pandemic disrupted several of these systems simultaneously, affecting students' emotional well-being, access to educational resources, home learning environments, and perceptions of academic continuity. By adopting this ecological perspective, the study recognizes that mathematics achievement during the pandemic cannot be attributed to isolated factors alone, but rather to the complex interplay of personal, familial, and contextual influences.

Despite the growing literature on COVID-19 and education, few large-scale studies have simultaneously examined the combined effects of emotional, learning, familial, and technological disruptions on student performance using both linear and non-linear approaches. Moreover, most studies have focused on single-country samples or limited variable sets. With these having been said, the current study addresses this gap by analyzing data from the Program for International Student Assessment (PISA) 2022 COVID-19 module, encompassing more than 109,000 students from multiple countries, and applying CART analysis to explore complex interaction effects. CART is especially well-suited for uncovering decision patterns in large, heterogeneous datasets, as it allows for interpretable, hierarchical segmentation of predictors (Mehenni & Moussaoui, 2012). This method complements traditional regression by revealing how different combinations of conditions (e.g., low socioeconomic status and high learning loss) map onto student outcomes in ways that are both intuitive and actionable.

Considering these issues, this study pursues the following objectives:

- To describe the patterns of COVID-19-related disruptions (emotional impact, learning loss, family support, and digital access) among secondary students from countries that participated in the PISA 2022 COVID-19 module.
- To examine the bivariate relationships between mathematics achievement and key COVID-19 disruption variables using Pearson's correlation coefficients to quantify linear associations.
- To determine the relative contributions of background and COVID-related variables in predicting students' mathematics performance through a multiple linear regression model.
- To explore complex and non-linear interaction effects among predictors using a CART model.
- To compare and synthesize findings from linear regression and decision-tree classification models to generate actionable insights for post-pandemic educational interventions and support strategies.

## 2. Method

### 2.1. Study design

This current study applies mathematical modeling through regression and classification algorithms to cross-sectional PISA 2022 data (Organisation for Economic Co-operation and Development (OECD), 2023). The purpose was to explore how emotional, instructional, familial, and technological disruptions influenced students' mathematics performance during the pandemic, using both linear regression and CART modeling.

### 2.2. Data Source and participants

Data were drawn from the publicly available PISA 2022 international dataset, specifically the subset of countries and regions that administered the COVID-19 disruption survey items. The final analytic sample comprised 109,097 secondary students aged approximately 15 years from diverse national and socioeconomic backgrounds. Countries included in the sample were those that provided complete responses to the core COVID-related items used in this analysis.

## 2.3. Measures

### 2.3.1. Outcome variable

Mathematics Achievement: To simplify the analysis, students' mathematics performance was measured using the simple average score (PVMEAN_MATH) derived from ten PISA plausible values (PV1MATH–PV10MATH). For the CART analysis, a binary variable named MATH_ACHIEVEMENT was created to classify students as high achievers (1) or low achievers (0) based on their relative standing in the sample.

$$Y_i = \frac{1}{10} \sum_{j=1}^{10} PV_{ij}$$

Where $PV_{ij}$ are $j$-th plausible value for student $i$.

For CART:
MATH_ACHIEVEMENT$_i$ = 1 if the student's score $Y_i$ is greater than or equal to the median.
MATH_ACHIEVEMENT$_i$ = 0 otherwise.

### 2.3.2. Predictor variables

The COVID-19 disruption-related variables used in this study were constructed using item blocks from the official PISA 2022 COVID-19 disruption module (OECD, 2023a, 2023b, 2024), developed by the OECD and administered in participating countries. Although the individual items (e.g., FL166Q01HA to FL174Q04JA) are standardized and internationally validated, PISA 2022 does not provide composite indices for broader constructs such as emotional impact, learning loss, family support, or access to digital resources. Accordingly, this study grouped conceptually related items into exploratory indices based on face validity and thematic coherence, following the structure of COVID-19 disruption dimensions discussed in previous literature (Engzell et al., 2021; König et al., 2020). Five major predictors were used in the analysis, each representing a dimension of COVID-19 disruption variables (see Table 1 for the specific items used):

- ESCS (Economic, Social, and Cultural Status): A standardized PISA index based on parents' education, occupation, home possessions, and cultural capital.
- COVID_EMO_IMPACT: Students' self-reported emotional strain due to the pandemic (e.g., stress, worry, anxiety). Using the mean value from 3 items (FL166Q01HA – FL166Q03HA).
- COVID_LEARNING_LOSS: Students' perceived loss in learning progress due to school closures. Using the mean value from 5 items (FL166Q05HA – FL166Q07HA, FL170Q01JA–FL170Q02JA).
- FAMILY_SUPPORT: Students' perceived emotional and academic support from their family during the pandemic. Using the mean value from 5 items (FL171Q01JA – FL171Q05JA).
- COVID_ACCESS: A composite index reflecting students' access to devices, internet, and learning platforms at home during remote learning periods. Using the mean value from 4 items (FL174Q01JA – FL174Q04JA).

Reliability and validity were assessed using confirmatory factor analysis (CFA) with the Diagonally Weighted Least Squares (DWLS) (Li, 2016) estimator and McDonald's (1999) omega. Emotional Impact showed high reliability ($\omega$ = .86), Learning Loss moderate (ω = .58), Family Support acceptable (ω = .71), and Digital Access lower, but interpretable ($\omega$ = .57). CFA also supported a four-factor structure with acceptable fit (Comparative Fit Index; CFI = .95; Tucker–Lewis Index; TLI = .94; Standardized Root Mean Square Residual; SRMR = .071; Root Mean Square Error of Approximation; RMSEA = .085) following established thresholds (CFI/TLI ≥ .90, SRMR ≤ .08, RMSEA ≤ .08) (Hu & Bentler, 1999; Schreiber et al., 2006). Together, these results support using the indices as exploratory but valid dimensions of pandemic disruption.

### 2.3.3. Demographic variables

- Age: Measured in years (M = 15.81, SD = 0.29).
- Gender: Coded as male or female.
- Country: A nominal variable indicating the student's country of participation.

**Table 1:** COVID-19 Disruption Variables Items Codebook

| Item Code | Item Meaning |
|---|---|
| Emotional Impact | |
| FL166Q01HA | Felt anxious or worried during school closures |
| FL166Q02HA | Felt overwhelmed during school closures |
| FL166Q03HA | Felt isolated or lonely during school closures |
| Learning Loss | |
| FL166Q05HA | Had difficulty concentrating while learning at home |
| FL166Q06HA | Fell behind in learning during school closures |
| FL166Q07HA | Worried about learning progress due to school closures |
| Family Support | |
| FL171Q01JA | Felt emotionally supported by family during closures |
| FL171Q02JA | Family helped organize learning time at home |
| FL171Q03JA | Family encouraged learning during closures |
| FL171Q04JA | Family helped with school assignments |
| FL171Q05JA | The family discussed the learning progress with the student |
| Access to Digital Resources | |
| FL174Q01JA | Had access to a quiet place to study |
| FL174Q02JA | Had access to a computer or digital device |
| FL174Q03JA | Had an internet connection at home |
| FL174Q04JA | Had access to digital learning platforms |

### 2.5. Data analysis procedures

All analyses were conducted using SPSS for descriptive statistics, correlations, and multiple linear regression, and R through the rpart package (Therneau & Atkinson, 2025) for CART analysis. More specific procedures are as follows:

- Descriptive statistics were first calculated for all key variables to understand their distribution. For each variable, the mean and standard deviation (SD) were computed:
- Pearson correlations were used to examine bivariate relationships between mathematics achievement and the predictors.
- Multiple linear regression was performed to assess the strength and direction of each predictor's contribution to PVMEAN_MATH.
- CART analysis was used to model non-linear and interaction effects among predictors. The resulting tree was pruned using the default complexity parameter to prevent overfitting. Lastly, the CART model is evaluated based on interpretability and predictive utility. To explore non-linear interactions, a CART was fitted. At each node $t$, the Gini impurity was computed:

$$G(t) = 1 - \sum_{c=1}^{C} p(c|t)^2$$

Where $p(c|t)$ is the proportion of class $c$ at node $t$. The optimal split $s$ was chosen to maximize information gain:

$$IG(s,t) = G(t) - \sum_{k\in\{L,R\}} \frac{N_k}{N_t} G(t_k)$$

Where $N_k$ is the number of observations in child node $k$. The recursive partitioning continued until stopping criteria were reached, and the tree was pruned using the default complexity parameter to prevent overfitting.

## 3. Results

### 3.1. Patterns of COVID-19-related disruptions across countries

Table 2 presents the demographic and COVID-19-related disruption characteristics of the 109,097 secondary school students who participated in the PISA 2022 assessment across 17 countries. The average age of students across the sample was 15.81 years (SD ≈ 0.29), with country-specific means ranging from 15.72 (Hungary) to 15.88 years (Brazil). The overall mean score for mathematics achievement (PVMEAN_MATH) was 480.68, though considerable variation was observed between countries. Students from the Netherlands (Mean or M = 518.38), Belgium (M = 527.15), and the Czech Republic (M = 517.90) reported the highest average mathematics scores, while students from Brazil (M = 404.68), Peru (M = 404.92), and Malaysia (M = 418.08) reported the lowest on average.
In terms of socioeconomic status (based from ESCS), the sample-wide mean was slightly above zero (M = 0.031), with the highest average ESCS observed in Canada (M = 0.403), Denmark (M = 0.413), and Belgium (M = 0.299), while the lowest was in Peru (M = –1.022), Brazil (M = –0.798), and Malaysia (M = –0.606). These disparities highlight the socioeconomic diversity of the sampled countries. With regards to COVID-19-related disruptions, the mean emotional impact score was 1.77 (on a 1 to 3 scale), with slightly higher levels of emotional strain reported in the UAE (M = 1.96), Malaysia (M = 1.94), and Canada (M = 1.91). Students from Italy (M = 1.55) and Belgium (M = 1.59) reported lower emotional impact. While the average perceived learning loss was 1.94 (on a 1 to 5 scale), with the highest reported in Bulgaria (M = 2.40) and the Netherlands (M = 2.17), reported higher-than-average instructional disruption, while Portugal (M = 1.78) and Italy (M = 1.85) reported relatively lower levels.
Family support during the pandemic also showed moderate cross-national consistency. Denmark (M = 2.98), the Netherlands (M = 2.97), and Poland (M = 2.95) had the highest reported levels of family support, while Peru (M = 2.55) and Spain (M = 2.52) had the lowest. Finally, access to digital learning resources (measured on a 1 to 4 scale) had an overall mean of 2.24. Specifically, Belgium (M = 2.51), Hungary (M = 2.46), and Portugal (M = 2.45) reported the highest levels of digital access, while students from Peru (M = 1.82) and Malaysia (M = 2.11) reported the least access. These results highlight the presence of a persistent digital divide that could impact student learning outcomes, particularly in times of remote instruction. Collectively, the data illustrate considerable variability in students' backgrounds and pandemic-related experiences, highlighting the importance of incorporating these contextual factors into predictive and policy-oriented models.

**Table 2:** Demographics of the Participants

| Country | *n* | Mean Age | PVMEAN Math | ESCS | Emotion | Learning Loss | Family Support | IT Access |
|---|---|---|---|---|---|---|---|---|
| Austria | 4170 | 15.83 | 512.48 | 0.175 | 1.72 | 1.90 | 2.74 | 2.10 |
| Belgium | 3159 | 15.87 | 527.15 | 0.299 | 1.59 | 2.00 | 2.81 | 2.51 |
| Brazil | 4834 | 15.88 | 404.68 | -0.798 | 1.77 | 1.85 | 2.88 | 2.11 |
| Bulgaria | 2574 | 15.74 | 455.80 | -0.072 | 1.69 | 2.40 | 2.85 | 2.19 |
| Canada | 9667 | 15.85 | 501.56 | 0.403 | 1.91 | 1.96 | 2.81 | 2.22 |
| Czech | 5898 | 15.79 | 517.90 | 0.079 | 1.71 | 1.96 | 2.91 | 2.17 |
| Denmark | 3833 | 15.74 | 494.27 | 0.413 | 1.83 | 1.96 | 2.98 | 2.14 |
| Hungary | 4617 | 15.72 | 495.91 | 0.149 | 1.78 | 2.03 | 2.69 | 2.46 |
| Italy | 7391 | 15.76 | 488.73 | -0.023 | 1.55 | 1.85 | 2.63 | 2.18 |
| Malaysia | 5319 | 15.84 | 418.08 | -0.606 | 1.94 | 1.94 | 2.75 | 2.11 |
| Netherlands | 3511 | 15.79 | 518.38 | 0.339 | 1.84 | 2.17 | 2.97 | 2.34 |
| Peru | 5141 | 15.87 | 404.92 | -1.022 | 1.92 | 1.96 | 2.55 | 1.82 |
| Poland | 4063 | 15.74 | 512.86 | -0.008 | 1.61 | 2.08 | 2.95 | 2.32 |
| Portugal | 5278 | 15.74 | 486.58 | -0.168 | 1.66 | 1.78 | 2.63 | 2.45 |
| Spain | 21779 | 15.81 | 495.72 | 0.084 | 1.65 | 1.79 | 2.52 | 2.42 |
| UAE | 14712 | 15.85 | 457.75 | 0.367 | 1.96 | 2.03 | 2.85 | 2.15 |
| USA | 3151 | 15.84 | 476.42 | 0.093 | 1.71 | 1.94 | 2.85 | 2.15 |
| TOTAL | 109097 | 15.81 | 480.68 | 0.031 | 1.77 | 1.94 | 2.74 | 2.24 |

### 3.2. Relationships of mathematics achievement and disruption variables

Pearson correlation analysis was conducted to examine the relationships between students' mathematics performance (PVMEAN_MATH) and key predictors, including socioeconomic status (ESCS), emotional impact of the pandemic, perceived learning loss, family support, and access to digital resources. The correlation matrix is presented in Table 3. As expected, mathematics achievement was positively correlated with ESCS ($r = .430$, $p < .01$), indicating that students from higher socioeconomic backgrounds tended to perform better in mathematics. Learning loss (or COVID_LEARNING_LOSS) ($r = –.207$, $p < .01$) and emotional impact (or COVID_EMO_IMPACT) ($r = –.140$, $p < .01$) were both negatively correlated with mathematics performance, suggesting that students who experienced greater emotional disruption or perceived higher learning loss reported lower mathematics scores. A weaker, though still significant, negative correlation was observed between family support (or FAMILY_SUPPORT) and mathematics achievement ($r = –.062$, $p < .01$), a somewhat unexpected finding that may reflect indirect effects mediated by other factors. In contrast, access to digital resources (or COVID_ACCESS) was positively correlated with mathematics performance (or PVMEAN_MATH) ($r = .122$, $p < .01$), supporting the idea that better access to learning technology during the pandemic was associated with improved academic outcomes. Intercorrelations among the predictors also revealed several significant associations. Notably, emotional impact and learning loss were moderately correlated ($r = .248$, $p < .01$), as were learning loss and family support ($r = .413$, $p < .01$). These relationships suggest interconnected experiences of emotional strain, learning disruption, and perceived support during the pandemic.

**Table 3:** Correlation Matrix of Math Achievement and COVID-19 Related Disruption Variables

| Variables | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| (1) PVMEAN_MATH | 1 | 0.430 | -0.140 | -0.207 | -0.062 | 0.122 |
| (2) ESCS | | 1 | -0.008 | 0.017 | 0.099 | 0.051 |
| (3) COVID_EMO_IMPACT | | | 1 | 0.248 | 0.233 | -0.204 |
| (4) COVID_LEARNING_LOSS | | | | 1 | 0.413 | -0.062 |
| (5) FAMILY_SUPPORT | | | | | 1 | -0.075 |
| (6) COVID_ACCESS | | | | | | 1 |

*Notes*. All correlations are significant at the .01 level (2-tailed). $N = 109{,}097$.

### 3.3. Predicting mathematics achievement

To assess the relative contribution of students' socioeconomic background and COVID-19-related experiences to their mathematics performance, a multiple linear regression analysis was conducted. The dependent variable was mathematics achievement, and the predictors included ESCS, emotional impact, learning loss, family support, and access to digital resources. All variables were entered simultaneously using the enter method. The results are summarized in Table 4. The overall regression model was statistically significant, $F(5, 109091) = 7004.73$, $p < .001$, and explained approximately 24.30% of the variance in students' mathematics achievement (Adjusted $R^2 = .243$). Among the predictors, ESCS emerged as the strongest and most significant positive predictor of mathematics performance ($\beta = .429$, $p < .001$), reinforcing the well-established relationship between socioeconomic advantage and academic success.

**Table 4:** Multiple Linear Regression Predicting Mathematics Achievement

| Predictor | $B$ | SE $B$ | $\beta$ | $t$ | $p$ |
|---|---|---|---|---|---|
| (Constant) | 519.539 | 1.451 | – | 358.119 | < .001 |
| ESCS | 38.853 | 0.240 | 0.429 | 161.647 | < .001 |
| COVID_EMO_IMPACT | -11.082 | 0.424 | –0.073 | –26.106 | < .001 |
| COVID_LEARNING_LOSS | -22.199 | 0.343 | –0.190 | –64.782 | < .001 |
| FAMILY_SUPPORT | -0.357 | 0.334 | –0.003 | –1.068 | .286 |
| COVID_ACCESS | 10.455 | 0.383 | 0.074 | 27.313 | < .001 |

*Note*. $N = 109{,}097$.

In contrast, learning loss was the strongest negative predictor ($\beta = –.190$, $p < .001$), followed by emotional impact ($\beta = –.073$, $p < .001$), indicating that both perceived instructional disruption and emotional strain during the pandemic adversely affected mathematics achievement. Furthermore, access to digital resources was also a statistically significant positive predictor ($\beta = .074$, $p < .001$), suggesting that better access to digital resources during the pandemic contributed positively to learning outcomes. Interestingly, family support did not significantly predict mathematics performance in the model ($\beta = –.003$, $p = .286$), possibly due to overlapping variance with other predictors or non-linear effects not captured in the linear model. These results highlight the central role of structural inequality and subjective learning disruptions in shaping students' academic trajectories during the pandemic.

### 3.4. Modeling non-linear effects using CART analysis

To further explore complex and non-linear interactions among the predictors of mathematics achievement, a CART analysis was conducted. The outcome variable was the binary-coded MATH_ACHIEVEMENT (1 = high, 0 = low), and the predictors included ESCS, learning loss, emotional impact, family support, and access to digital resources. The CART algorithm identified ESCS as the most important predictor (as the primary split) and used it as the root node, splitting at a threshold of 0.31955 or 0.32 (see Figure 1). Students above this threshold (high-ESCS) and below it (low-ESCS) followed sharply divergent trajectories.

Within the High-ESCS branch (left side of the figure): Among students with ESCS ≥ 0.32, the next split was perceived learning loss (orange box; threshold = 2.64 on a 1–5 scale). Which means that students with lower learning loss (< 2.64) had a 63.0% probability of being high achievers (above-median math scores). In contrast, students with higher learning loss (≥ 2.64), despite advantaged backgrounds, dropped to only 29.8% high achievers, illustrating that even privilege could not fully buffer against significant instructional disruption.

For the Low-ESCS branch (right side of the figure): Students with ESCS < 0.32, the model classified 71.4 or 71% as low achievers, regardless of their emotional strain, family support, or digital access. This branch was the largest terminal node, underscoring that socioeconomic disadvantage alone was enough to predict underperformance, and that other supports offered limited compensatory power within this subgroup.

Overall, the CART analysis tells a clear story, wherein students from wealthier families who did not fall far behind during the pandemic had the best chances of doing well in math. While students from poorer families were at high risk of lower performance, no matter how much support or technology they had at home. Interestingly, factors like family support and access to computers or the internet did not play a big role in the decision tree. This suggests that their influence was smaller compared to family background and learning loss, or that these supports were already tied up with socioeconomic status.

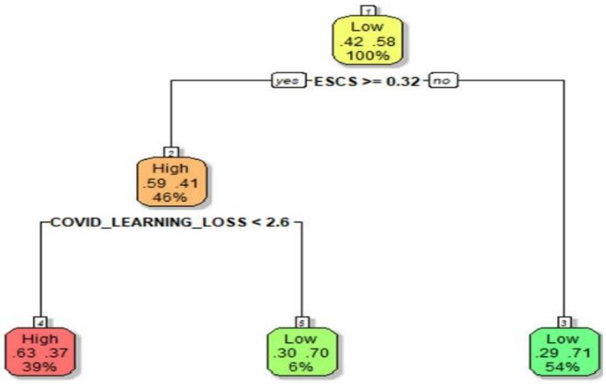


**Fig. 1:** Classification and Regression Tree (CART) Predicting Mathematics Achievement.

*Note*: The root split is ESCS (0.31955). High-ESCS students were further split by learning loss (2.64), while low-ESCS students formed the largest terminal node with 71.4% low achievers.

Overall, the CART model revealed interpretable, non-linear interactions among variables that were not captured in the linear regression model. More specifically, it illustrated how the combination of high socioeconomic status and low learning disruption defined the most successful student profile, whereas low ESCS alone was enough to predict underperformance. Family support and digital access, although included in the model, did not appear as splitting variables, possibly due to their weaker predictive power or limited internal consistency.

### 3.5. Comparison of regression and CART results

The linear regression and CART models each provided complementary insights into the factors affecting students' mathematics achievement during the COVID-19 pandemic. While the regression model quantified the independent effects of each predictor variable, the CART model revealed interaction effects and decision-rule pathways that segment students into performance profiles. In the regression analysis, ESCS was the strongest positive predictor of math achievement ($\beta$ = .429, $p$ < .001), followed by negative effects of COVID-related learning loss ($\beta$ = –.190) and emotional impact ($\beta$ = –.073). Digital access ($\beta$ = .074) also emerged as a modest positive predictor, while family support did not significantly contribute when controlling for other variables.

The CART analysis echoed the central role of ESCS, which formed the root node of the decision tree. However, unlike regression, CART illustrated how specific combinations of predictors produced different risk profiles. For instance, students with high ESCS and low learning loss had a 63.0% likelihood of being high achievers, while students with low ESCS were overwhelmingly predicted to be low achievers, regardless of other variables. This suggests that the impact of learning loss is conditional on socioeconomic advantage, and that low-ESCS students may require more intensive, structural interventions. Importantly, some variables that were statistically significant in the regression model, such as emotional impact and digital access, did not appear in the CART tree. This difference reflects how CART prioritizes variables that maximize information gain at each split, meaning that once a powerful predictor like ESCS is considered, other variables may no longer contribute to further splits if their incremental value is low.

Taken together, these findings highlight both the individual and interaction-based effects of COVID-19 disruptions on mathematics achievement. While regression emphasizes linear contributions, CART illustrates the threshold effects and student segmentation patterns that are crucial for designing targeted interventions. For policymakers and educators, this suggests that recovery efforts should not only be guided by average effects but also by student profiles that reflect the complex interplay of socioeconomic status, learning loss, and access disparities.

## 4. Discussions

The findings of this study extend the growing body of research on the educational impacts of the COVID-19 pandemic by identifying specific factors that influenced mathematics achievement across a large, diverse international sample. Guided by Bronfenbrenner's (2000) Ecological Systems Theory, the results support the notion that student learning outcomes are not shaped by isolated variables, but by the interaction of emotional, cognitive, familial, and contextual systems, all of which were disrupted during the pandemic. Consistent with prior research, socioeconomic status emerged as the most powerful predictor of math achievement (Easterbrook et al., 2023; Xian et al., 2023). Both the linear regression and CART analyses revealed that students from higher-ESCS backgrounds consistently outperformed their peers, particularly when paired with lower perceived learning loss. From a methodological perspective, the regression model formalizes the linear relationship between mathematics performance and its predictors, while the CART model operationalizes decision rules by minimizing impurity functions, thereby providing a complementary mathematical framework for analyzing complex, heterogeneous populations. This aligns with previous studies showing that advantaged families were better able to buffer the effects of school closures by providing structured support, resources, and educational continuity (Gaxiola Romero et al., 2022; Scrimin et al., 2022).

Importantly, the study also confirms the detrimental role of perceived learning loss on mathematics achievement, in line with findings that mathematics learning, more than other subjects, suffered from disrupted instructional time and lack of scaffolding (Donnelly & Patrinos, 2022; Hammerstein et al., 2021; Nesher, 1986). Students who reported higher levels of learning disruption performed significantly worse, even when other factors like ESCS or emotional well-being were considered. The CART model underscored this by showing that learning loss was the key determinant separating high and low-achieving students within the higher-ESCS group. Interestingly, while emotional impact was a statistically significant predictor in the regression model, it did not appear as a major splitting variable in the CART tree. This suggests that while emotional strain negatively affects achievement (Bubb & Jones, 2020; Camacho et al., 2021), its predictive power may be more diffuse or indirect when considered alongside stronger contextual variables. Nonetheless, it reinforces prior literature emphasizing the importance of emotional well-being as a prerequisite for engagement and cognitive function (Zaccoletti et al., 2020).

Conversely, the findings on family support and access to digital resources were more inconsistent. Despite theoretical support for their relevance (Ramsetty & Adams, 2020; Sosa Díaz, 2021)These constructs did not emerge as strong predictors in either model. This finding aligns with prior evidence that connectivity gaps directly exclude students from learning opportunities (Ramsetty & Adams, 2020), but contrasts with studies such as Mathrani et al. (2022), which argues that access alone is insufficient without pedagogical support and structured home routines. Taken together, these inconsistencies suggest that digital access may only enhance learning when embedded within broader instructional and family systems, highlighting the need for more integrative research designs.

In addition, while family support was theoretically expected to buffer against disruption, it did not emerge as a significant predictor in the regression model and did not appear in the CART tree. One possible explanation is mediation, wherein family support may work indirectly through reducing emotional strain or perceived learning loss, which were already strong predictors in the models. Another explanation is measurement limitations, wherein the support index had only moderate reliability, and self-reports may underestimate informal or non-academic help. Finally, family support may be context-dependent, such as its protective role could be stronger in early grades or in single-country studies, but diluted in large cross-national data where cultural norms of parental involvement vary widely.

Overall, the comparison between regression and CART analyses highlights the added value of modeling non-linear, interaction-based patterns. While regression helps isolate the individual contributions of predictors, CART demonstrates how certain combinations of conditions, especially when low ESCS is paired with any degree of disruption, can sharply distinguish student performance outcomes. This complementarity illustrates the utility of applying both parametric and non-parametric models: regression quantifies global parameter estimates, whereas CART generates piecewise-defined decision functions that reveal threshold effects. These findings have practical value for educational policy and recovery planning. Interventions must not only address average trends, but also target specific student profiles most vulnerable to pandemic-induced learning gaps. Finally, the study's ecological framing emphasizes that COVID-19 amplified existing inequities by simultaneously disrupting the nested systems that support student development. In recovery planning, education systems should consider holistic interventions that span school, home, and digital environments, ensuring that support for learning includes not only cognitive resources but also emotional stability and equitable access. Although applied here to pandemic-related disruptions, the mathematical modeling framework demonstrated in this study can be generalized to other systemic shocks in education, such as technological change or climate-related disruptions, underlining the interdisciplinary role of mathematics in policy-relevant social research.

## 5. Conclusions

The current study examined the influence of COVID-19-related disruptions, such as emotional impact, perceived learning loss, family support, and access to digital resources, on secondary students' mathematics achievement using PISA 2022 data. Grounded in Bronfenbrenner's Ecological Systems Theory, the findings reinforce that academic outcomes are shaped by the dynamic interplay between individual, familial, and contextual systems, all of which were impacted by the pandemic. Regression analysis highlighted the independent contributions of socioeconomic status, learning loss, emotional disruption, and digital access. CART analysis further revealed that socioeconomic background served as the most critical stratifier of academic performance, particularly when combined with students' perceived learning loss. The comparative modeling approach demonstrated that while each variable plays a role, the interaction of multiple disruptions, especially among disadvantaged students, can significantly magnify academic vulnerability. Beyond substantive findings, this study highlights how applied mathematics, in the form of regression equations and decision-tree classifiers, can rigorously model both linear and interaction effects in international education data. As education systems continue to recover, these findings highlight the importance of targeted, data-informed interventions. Policies should not only address average learning outcomes, but also prioritize students who face compounding disadvantages across multiple ecological domains. Ensuring equitable recovery from COVID-19 requires a comprehensive focus on academic, emotional, and infrastructural support. The work, therefore, contributes not only to educational policy debates but also to the broader methodological role of mathematics as a tool for understanding human development across nested ecological systems.

While this study provides valuable insights, several limitations should be acknowledged. First, the analysis is based on self-reported survey data, which may be subject to response biases, especially in items relating to emotions or perceived learning loss. Second, while the disruption indices were constructed using conceptually aligned PISA items, not all showed strong internal consistency, particularly the digital access and family support scales. As such, results involving these variables should be interpreted with caution. Future studies may benefit from refining these measures or incorporating additional qualitative or observational data. Third, although the PISA dataset provides a robust international sample, the study was limited to 17 countries that administered the COVID-19 module. This may affect the generalizability of results across all PISA-participating nations. Cultural and systemic differences, such as parental involvement norms, duration of school closures, recovery programs, and digital infrastructure, likely moderated how disruptions translated into achievement outcomes. These contextual factors could not be fully modeled here, meaning that the observed patterns reflect cross-national trends within this subset of countries rather than universally generalizable findings.

Anchored in Bronfenbrenner's Ecological Systems Theory, future research should further investigate how disruptions operate across nested levels of influence. More specific research questions may include: How do teacher-level factors, such as instructional quality or emotional support, mediate the relationship between perceived learning loss and student outcomes? To what extent do school-level recovery initiatives moderate the association between socioeconomic background and mathematics achievement? Such questions would allow a more precise understanding of the mechanisms through which ecological systems interact in shaping resilience and vulnerability during crises. Methodologically, subsequent studies could adopt mixed-methods designs that combine the breadth of large-scale assessments with the depth of qualitative inquiry, thereby capturing both statistical patterns and lived experiences of students and families. Extending CART and related decision-tree approaches with additional predictors, such as teacher support, institutional recovery programs, or national policy responses, would also provide greater insight into how multiple layers of support interact with structural disadvantage. Finally, although this study focused on COVID-19, the analytic framework can be applied to other systemic disruptions (e.g., climate-related school closures, political instability, or rapid technological change), enabling cross-contextual comparisons of how disadvantage and disruption jointly shape learning trajectories. By pursuing these directions, future research can both deepen the ecological interpretation of educational disruptions and enhance the policy relevance of data-driven decision models.

**Declaration of AI Use in the Preparation of the Manuscript**: The authors used Wordtune to do a language check and improve the grammar and clarity of the manuscript. All content was carefully reviewed and edited by the authors, who take full responsibility for the final version.

# References


[1] Bronfenbrenner, U. (2000). Ecological systems theory. In A. E. Kazdin (Ed.), *Encyclopedia of psychology* (Vol. 3, pp. 129-133). Oxford University Press. https://doi.org/10.1037/10518-046.

[2] Bubb, S., & Jones, M.-A. (2020). Learning from the COVID-19 home-schooling experience: Listening to pupils, parents/carers and teachers. *Improving Schools, 23*(3), 209-222. https://doi.org/10.1177/1365480220958797.

[3] Camacho, A., Correia, N., Zaccoletti, S., & Daniel, J. R. (2021). Anxiety and social support as predictors of student academic motivation during the COVID-19. *Frontiers in Psychology, 12*, 644338. https://doi.org/10.3389/fpsyg.2021.644338.

[4] Chin, J. M.-C., Ching, G. S., del Castillo, F., Wen, T.-H., Huang, Y.-C., del Castillo, C. D., Gungon, J. L., & Trajera, S. M. (2022). Perspectives on the barriers to and needs of teachers' professional development in the Philippines during COVID-19. *Sustainability, 14*(1), 470. https://doi.org/10.3390/su14010470.

[5] Contini, D., Di Tommaso, M. L., Muratori, C., Piazzalunga, D., & Schiavon, L. (2022). Who lost the most? Mathematics achievement during the COVID-19 pandemic. *The B.E. Journal of Economic Analysis & Policy, 22*(2), 399-408. https://doi.org/10.1515/bejeap-2021-0447.

[6] d'Orville, H. (2020). COVID-19 causes unprecedented educational disruption: Is there a road towards a new normal? *Prospects, 49*, 11-15. https://doi.org/10.1007/s11125-020-09475-0.

[7] Di Pietro, G. (2021). Changes in Italy's education-related digital divide. *Economic Affairs, 41*(2), 252-270. https://doi.org/10.1111/ecaf.12471.

[8] Di Pietro, G., Biagi, F., Costa, P., Karpiński, Z., & Mazza, J. (2020). *The likely impact of COVID-19 on education: Reflections based on the existing literature and recent international datasets*. EU Science Hub.

[9] Donnelly, R., & Patrinos, H. A. (2022). Learning loss during Covid-19: An early systematic review. *Prospects, 51*, 601–609. https://doi.org/10.1007/s11125-021-09582-6.

[10] Easterbrook, M. J., Doyle, L., Grozev, V. H., Kosakowska-Berezecka, N., Harris, P. R., & Phalet, K. (2023). Socioeconomic and gender inequalities in home learning during the COVID-19 pandemic: Examining the roles of the home environment, parent supervision, and educational provisions. *Educational and Developmental Psychologist, 40*(1), 27-39. https://doi.org/10.1080/20590776.2021.2014281.

[11] Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences, 118*(17), e2022376118. https://doi.org/10.1073/pnas.2022376118.

[12] Gaxiola Romero, J. C., Pineda Domínguez, A., Gaxiola Villa, E., & González Lugo, S. (2022). Positive family environment, general distress, subjective well-being, and academic engagement among high school students before and during the COVID-19 outbreak. *School Psychology International, 43*(2), 111-134. https://doi.org/10.1177/01430343211066461.

[13] Hammerstein, S., König, C., Dreisörner, T., & Frey, A. (2021). Effects of COVID-19 related school closures on student achievement: A systematic review. *Frontiers in Psychology, 12*, 746289. https://doi.org/10.3389/fpsyg.2021.746289.

[14] Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118.

[15] König, J., Jäger-Biela, D. J., & Glutsch, N. (2020). Adapting to online teaching during COVID-19 school closure: Teacher education and teacher competence effects among early career teachers in Germany. *European Journal of Teacher Education, 43*(4), 608-622. https://doi.org/10.1080/02619768.2020.1809650.

[16] Kania, N., & Juandi, D. (2023). Does self-concept affect mathematics learning achievement? *Journal of Education and Learning, 17*(3), 455-461. https://doi.org/10.11591/edulearn.v17i3.20554.

[17] Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher, 49*(8), 549-565. https://doi.org/10.3102/0013189X20965918.

[18] Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods, 48*, 936–949. https://doi.org/10.3758/s13428-015-0619-7.

[19] Mathrani, A., Sarvesh, T., & Umer, R. (2022). Digital divide framework: online learning in developing countries during the COVID-19 lockdown. *Globalisation, Societies and Education, 20*(5), 625-640. https://doi.org/10.1080/14767724.2021.1981253.

[20] McDonald, R. P. (1999). *Test theory: A unified treatment*. Psychology Press. https://doi.org/10.4324/9781410601087.

[21] Mehenni, T., & Moussaoui, A. (2012). Data mining from multiple heterogeneous relational databases using decision tree classification. *Pattern Recognition Letters, 33*(13), 1768-1775. https://doi.org/10.1016/j.patrec.2012.05.014.

[22] Nesher, P. (1986). Learning mathematics: A cognitive perspective. *American Psychologist, 41*(10), 1114-1122. https://doi.org/10.1037/0003-066X.41.10.1114.

[23] OECD. (2023a). *PISA 2022 results (Volume 1): The state of learning and equity in education*. PISA, OECD Publishing. https://doi.org/10.1787/53f23881-en.

[24] OECD. (2023b). *PISA 2022 results (Volume 2): Learning during- and from - disruption*. PISA, OECD Publishing. https://doi.org/10.1787/a97db61c-en.

[25] OECD. (2024). *PISA 2022 technical report*. PISA, OECD Publishing. https://doi.org/10.1787/01820d6d-en.

[26] Organisation for Economic Co-operation and Development (OECD). (2023). *PISA 2022 database* OECD. https://www.oecd.org/en/data/datasets/pisa-2022-database.html.

[27] Ramsetty, A., & Adams, C. (2020). Impact of the digital divide in the age of COVID-19. *Journal of the American Medical Informatics Association, 27*(7), 1147-1148. https://doi.org/10.1093/jamia/ocaa078.

[28] Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research, 99*(6), 323-338. https://doi.org/10.3200/JOER.99.6.323-338.

[29] Schult, J., Mahler, N., Fauth, B., & Lindner, M. A. (2022). Long-term consequences of repeated school closures during the COVID-19 pandemic for reading and mathematics competencies. *Frontiers in Education, 13*, 867316. https://doi.org/10.3389/feduc.2022.867316.

[30] Scrimin, S., Mastromatteo, L. Y., Hovnanyan, A., Zagni, B., Rubaltelli, E., & Pozzoli, T. (2022). Effects of socioeconomic status, parental stress, and family support on children's physical and emotional health during the COVID-19 pandemic. *Journal of Child and Family Studies, 31*, 2215–2228. https://doi.org/10.1007/s10826-022-02339-5.

[31] Sosa Díaz, M. J. (2021). Emergency remote education, family support and the digital divide in the context of the COVID-19 lockdown. *International Journal of Environmental Research and Public Health, 18*(15), 7956. https://doi.org/10.3390/ijerph18157956.

[32] Tashtoush, M. A., Wardat, Y. A., & Elsayed, A. M. (2023). Mathematics distance learning and learning loss during COVID-19 pandemic: Teachers' perspectives. *Journal of Higher Education Theory and Practice, 23*(5), 162-174. https://doi.org/10.33423/jhetp.v23i5.5933.

[33] Therneau, T., & Atkinson, B. (2025). *rpart: Recursive Partitioning and Regression Trees.* In (Version R package version 4.1.24) https://cran.r-project.org/web/packages/rpart/index.html.

[34] Xian, X., Zhang, Y., Bai, A., Zhai, X., Hu, H., Zhang, J., & Ye, M. (2023). Association between family support, stress, and sleep quality among college students during the COVID-19 online learning period. *International Journal of Environmental Research and Public Health, 20*(1), 248. https://doi.org/10.3390/ijerph20010248.

[35] Zaccoletti, S., Camacho, A., Correia, N., Aguiar, C., Mason, L., Alves, R. A., & Daniel, J. R. (2020). Parents' perceptions of student academic motivation during the COVID-19 lockdown: A cross-country comparison. *Frontiers in Psychology, 11*, 592670. https://doi.org/10.3389/fpsyg.2020.592670.