Contemporary Research in Education and English Language Teaching
ISSN: 2641-0230
Vol. 7, No. 2, 56-68
2025
Publisher: Learning Gate
DOI: 10.55214/2641-0230.v7i2.9167

Predicting school safety using machine learning: Insights from Chicago public schools

Jagadeeshwari Mandapalli[1], Rachna Tewari[2*], Isaac Lepcha[3]
[1]Data Analytics, St Louis University, 1 N Grand Blvd, St. Louis, MO 63103, USA; mandapallijagadeeshwari@slu.edu (J.M.).
[2]Agricultural Economics & Agribusiness Department of Agriculture, Geosciences, and Natural Resources, The University of Tennessee at Martin 256 Brehm Hall, Martin, TN 38238, USA; rtewari@utm.edu (R.T.).
[3]Plant and Soil Science, Department of Agriculture, Geosciences, and Natural Resources the University of Tennessee at Martin, 256 Brehm Hall, Martin, TN 38238, USA; ilepcha1@utm.edu (I.L.).

Abstract: School safety remains a critical concern within urban education systems, directly influencing student well-being and institutional performance. In this project, we explored how machine learning models could be used to predict school safety ratings across Chicago Public Schools (CPS) systems, providing school leaders with a data-driven tool for early intervention and resource planning. Using a dataset that included safety metrics, attendance records, disciplinary incidents, and climate survey results from 478 elementary and high schools, we tested three predictive approaches: linear regression as a baseline, random forest for ensemble learning, and XGBoost for gradient-boosted performance. Among these, the XGBoost model performed the best, achieving a tested R-squared value of 0.923, a root mean square error (RMSE) of 5.16, and a mean absolute error (MAE) of 4.08 on the test set. Notably, the model identified family involvement scores, student attendance rates, and reported misconduct incidents as the influential predictors of perceived school safety ratings. These results align with existing research in education policy while providing new insights into measurable factors influencing school safety. The practical implication of this work is a machine learning tool that can identify potentially at-risk schools with 92% accuracy. Importantly, because the model is interpretable, administrators can understand how each factor contributes to the predictive safety score, allowing for more targeted and transparent decision-making. This could support efforts in designing preventative strategies, resource allocation, support staff deployment, or tailoring school-based intervention strategies. Overall, this study illustrates how machine learning can complement traditional methods in school safety evaluation, offering a scalable and evidence-based approach that may be particularly valuable in data-rich but resource-constrained educational environments.

Keywords: *Academic performance, Behavioral climate, Data-driven decision making, Educational policy, Family engagement, ISAT scores, Machine learning, Predictive analytics, School safety, XGBoost.*

1. Introduction

School safety is a fundamental requirement for fostering a positive learning environment, which impacts student well-being, academic achievement, and school climate (Thapa, Cohen, Guffey, & Higgins-D'Alessandro, 2013). But upholding and creating such an atmosphere is challenging, especially in U.S. public schools where safety outcomes can be very different depending on available resources, local community backing, and policy decisions (Cuellar & Coyle, 2021). National school safety policies have varied depending on the office in place. For example, the Trump administration prioritized observation measures and coordination with law enforcement within the STOP School Violence Act, something for which it was blamed for ignoring root causes like mental health concerns (PBS NewsHour, 2018; The Atlantic, 2018). In contrast, the Biden administration has taken a more

* Correspondence: rtewari@utm.edu

comprehensive approach to assisting students through greater mental health treatment and community-focused prevention measures (K12 Dive, 2023; U.S. Department of Education, 2018). At the same time, state and local efforts—such as those taking place in Chicago— continue to implement targeted programs designed to meet the individual needs of their schools and communities. Safety is still a valid concern in many areas, however. In Chicago Public Schools (CPS), for instance, the mean safety measure is a paltry 49.48 out of 99, with nearly one-third of schools considered "low safety" (Chicago Consortium on School Research, 2019). They are derived from students' and families' perceptions of their school environments, and the variation typically results from more deeply entrenched disparities. Under-resourced communities' schools can have higher rates of disciplinary events, less access for family participation, and less nurturing environments—conditions that all combine to create a perception of a school's general safety (Steinberg, Allensworth, & Johnson, 2011). While there are policies like Title I funding that fill in the gap, it can be difficult to know which schools need it most. Predictive modeling comes into play here. Machine learning is opening new doors for supporting policymaking. With the right data, these tools can figure out which schools are most vulnerable and help prioritize resource allocation to where it will have the greatest value (Pardos, Herold, & Brook, 2021). Of course, there are also ethical issues at stake, especially related to fairness, privacy, and potential for bias (Baker & Inventado, 2014). But responsibly applied, predictive analytics can be a helpful ally in the pursuit of equity. This study examines if machine learning can accurately forecast school safety scores for Chicago. Based on information from 478 CPS schools, we aim to tease apart the optimal predictors of safety. We anticipate that these advanced models like XGBoost will perform extremely well, and that these school climate measures will appear as key predictors. Ultimately, we hope to provide a resource that can inform wiser, more just school safety planning decisions.

2. Literature Review

Integrating school safety, academic performance, and evidence-based decision making has been a prominent subject of educational research. Researchers now more than ever recognize that school safety is not just an issue of physical safety, but is also deeply inseparable from emotional well-being, school climate, and student engagement (Cuellar & Coyle, 2021; The Atlantic, 2018). Current policy trends reflect this shift with a turn toward climate-informed interventions as opposed to punitive or security-motivated responses (PBS NewsHour, 2018; U.S. Department of Homeland Security, 2023). Robust empirical evidence supports the effect of favorable school climate on student performance. Steinberg et al. (2011) and Thapa et al. (2013) showed that improved school climate has a robust, positive impact on students' participation and reduces behavior issues. Bradshaw, Waasdorp, Debnam, and Johnson (2014) cited three significant domains—safety, engagement, and environment—as foundations for measuring school climate. Steps such as the "5Essentials" survey of the University of Chicago Consortium have been adopted on a large scale to monitor these factors (Davis, Shyjka, Hart, Gutierrez, & Kheraj, 2021).

Family engagement remains a consistently robust predictor of student achievement. Fan and Chen (2001) meta-analysis reiterated that parental involvement is important in the achievement of students within socio-demographic groups. Chicago Public Schools data also reinforce this connection, showing a positive correlation between Family Engagement Scores and academic achievement (Chicago Public Schools, 2025). Advances in educational data mining (EDM) and learning analytics have increased researchers' ability to model student outcomes. Predictive tools, including machine learning models like XGBoost and ensemble models, have shown high potential in identifying at-risk students and predicting academic performance (Asselman, Khaldi, & Aammou, 2023; Kyriazos & Poga, 2024; Pardos et al., 2021). These approaches enable better early intervention strategies and inform data-driven decision-making.

Exploratory data analysis (EDA) remains an integral part of predictive modeling. Originally coined by Tukey (1977) EDA facilitates pattern discovery and resolving common data issues such as missing data and outliers. Researchers have employed multiple techniques, such as visualization (Bilal & Maitama, 2018) imputation (Little & Rubin, 2019; Van Buuren, 2018) and outlier treatment techniques

Contemporary Research in Education and English Language Teaching
ISSN: 2641-0230
Vol. 7, No. 2: 56-68, 2025
DOI: 10.55214/2641-0230.v7i2.9167

such as IQR and trimmed means (Vinutha, Poornima, & Sagar, 2018; Wu & Zuo, 2009).

Finally, the ethics of predictive modeling in education need to be taken seriously. With increasing learning analytics, questions of privacy, informed consent, and algorithmic fairness have become the center of debates. Scholars such as Slade and Prinsloo (2013) and Baker and Inventado (2014) advocate for open, context-sensitive practices to ensure ethical uptake of data- driven interventions.

3. Materials and Methods

3.1. Data Collection and Overview

The "Chicago Public Schools – Progress Report Cards" dataset was used for this study. This dataset comprises 478 schools (elementary and high schools) and 22 features associated with safety, academic performance, attendance, and community engagement. The dataset provides a comprehensive, multi-year view of safety metrics and school performance, comprising data from the 2011–2012 school year through 2023–2024. It consists of 22 features, divided into numerical and categorical variables. Key numerical variables include Safety Score, Family Involvement Score, Environment Score, Instruction Score, Teachers Score, Parent Engagement Score, Parent Environment Score, Average Student Attendance, Rate of Misconducts (per 100 students), Average Teacher Attendance, Individualized Education Program Compliance Rate, ISAT Exceeding Math%, ISAT Exceeding Reading%, ISAT Value Add Math, ISAT Value Add Reading, Student ID, and an index column labeled Unnamed: 0. The categorical variables include Safety Icon, Family Involvement Icon, Environment Icon, Instruction Icon, Teachers Icon, Parent Engagement Icon, Parent Environment Icon, and School Name. The target variable for this study is the School Safety Rating, an ordinal categorical variable representing overall school safety levels. Table 1 below summarizes the variable overview and data types.

Table 1.
Variable Overview and Data Types.

Variable Name	Type	Missing Values	Data Type (Before)	Data Type (After)
Safety Icon	Categorical	0	object	object
Safety Score	Numeric	45	float64	float64
Family Involvement Icon	Categorical	0	object	object
Family Involvement Score	Numeric	228	float64	float64
Environment Icon	Categorical	0	object	object
Environment Score	Numeric	45	float64	float64
Instruction Icon	Categorical	0	object	object
Instruction Score	Numeric	45	float64	float64
Teachers Icon	Categorical	0	object	object
Teachers Score	Numeric	230	float64	float64
Parent Engagement Icon	Categorical	0	object	object
Parent Engagement Score	Numeric	81	float64	float64
Parent Environment Icon	Categorical	0	object	object
Parent Environment Score	Numeric	81	float64	float64
Average Student Attendance	Numeric	1	float64	float64
Rate of Misconducts (per 100 students)	Numeric	0	float64	float64
Average Teacher Attendance	Numeric	0	float64	float64
Individualized Education Program Compliance Rate	Numeric	1	float64	float64
ISAT Exceeding Math %	Numeric	7	float64	float64
ISAT Exceeding Reading %	Numeric	7	float64	float64
ISAT Value Add Math	Numeric	15	float64	float64
ISAT Value Add Read	Numeric	15	float64	float64

3.2. Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) was performed to understand the dataset's structure, distribution, and relationships among variables (Tukey, 1977). This process helped identify data quality issues,

ISSN: 2641-0230
Vol. 7, No. 2: 56-68, 2025
DOI: 10.55214/2641-0230.v7i2.9167

informed feature engineering, and supported appropriate modeling strategies (Dasu & Johnson, 2003). Data cleaning involved the removal of non-informative rows and columns, grouping variables by type, and assessing missingness patterns using visualization tools (Bilal & Maitama, 2018). A general data type conversion was conducted to ensure numerical variables were properly formatted for analysis, while categorical indicators were preserved.

Following these steps, the dataset was refined to include a complete set of cases suitable for predictive modeling.

3.3. Descriptive Statistics

Descriptive analyses were conducted on variables such as Safety Score, Family Involvement Score, Instruction Score, and ISAT Exceeding Math/Reading %. Means ranged from 15.35 (ISAT Reading %) to 98.94 (IEP Compliance Rate), with diverse variability. For instance, the Rate of Misconducts had a high SD (29.43), while the Parent Engagement Score showed minimal dispersion (SD = 5.05). Skewed distributions, such as those observed in Teacher Attendance (0% to 98.5%), indicated potential outliers and non-normality (Field, 2013; James, Witten, Hastie, & Tibshirani, 2013). Table 2 summarizes the descriptive statistics of key variables along with a description of each variable.

Table 2.
Descriptive Statistics of Key Variables.

Variable	Mean	Std. Dev.	Min.	Max.	Skewness	Kurtosis
Safety Score	49.4813	20.2115	1.0	99.0	0.5287	-0.0293
Family Involvement Score	51.1592	18.9221	6.0	99.0	0.5584	0.0517
Environment Score	48.1729	16.3827	1.0	99.0	0.2909	0.4718
Instruction Score	49.0771	17.7647	1.0	99.0	0.1508	0.3665
Teachers Score	49.9918	18.2372	8.0	99.0	0.3099	-0.2037
Parent Engagement Score	50.0587	5.0526	37.0	69.0	0.8188	1.5818
Parent Environment Score	50.3189	4.6519	37.0	70.0	0.4800	0.9154
Average Student Attendance	94.0379	3.0439	60.9	98.4	-4.5446	36.8782
Rate of Misconducts (per 100 students)	22.1651	29.4308	0.0	251.6	3.3490	16.6009
Average Teacher Attendance	94.9687	8.8375	0.0	98.5	-10.5239	110.2509
IEP Compliance Rate	98.9402	2.0085	86.9	100.0	-2.6601	8.1910
ISAT Exceeding Math %	19.7803	15.9894	0.0	100.0	1.9806	4.9070
ISAT Exceeding Reading %	15.3545	14.8594	0.0	100.0	2.6082	8.7343
ISAT Value Add Math	0.0989	1.0199	-3.5	3.6	0.2373	0.7482
ISAT Value Add Read	0.0428	1.0423	-5.0	4.9	0.1496	2.0510

3.4. Missing Data Handling

Missing values were prevalent across several features in the dataset, with notable gaps in variables such as Family Involvement Score and Teacher Score, each missing over 200 entries (Figure 1; Table 3). To address this, a two-step imputation strategy was implemented. For numerical variables, median imputation was employed to mitigate the influence of outliers, while mode imputation was applied to categorical variables based on the most frequently occurring categories. K-Nearest Neighbors (KNN) imputation with k=5 was also performed as a validation step to ensure the robustness of the imputation process. Following these procedures, all missing values in the dataset were effectively resolved (Little & Rubin, 2019).

DOI: 10.55214/2641-0230.v7i2.9167

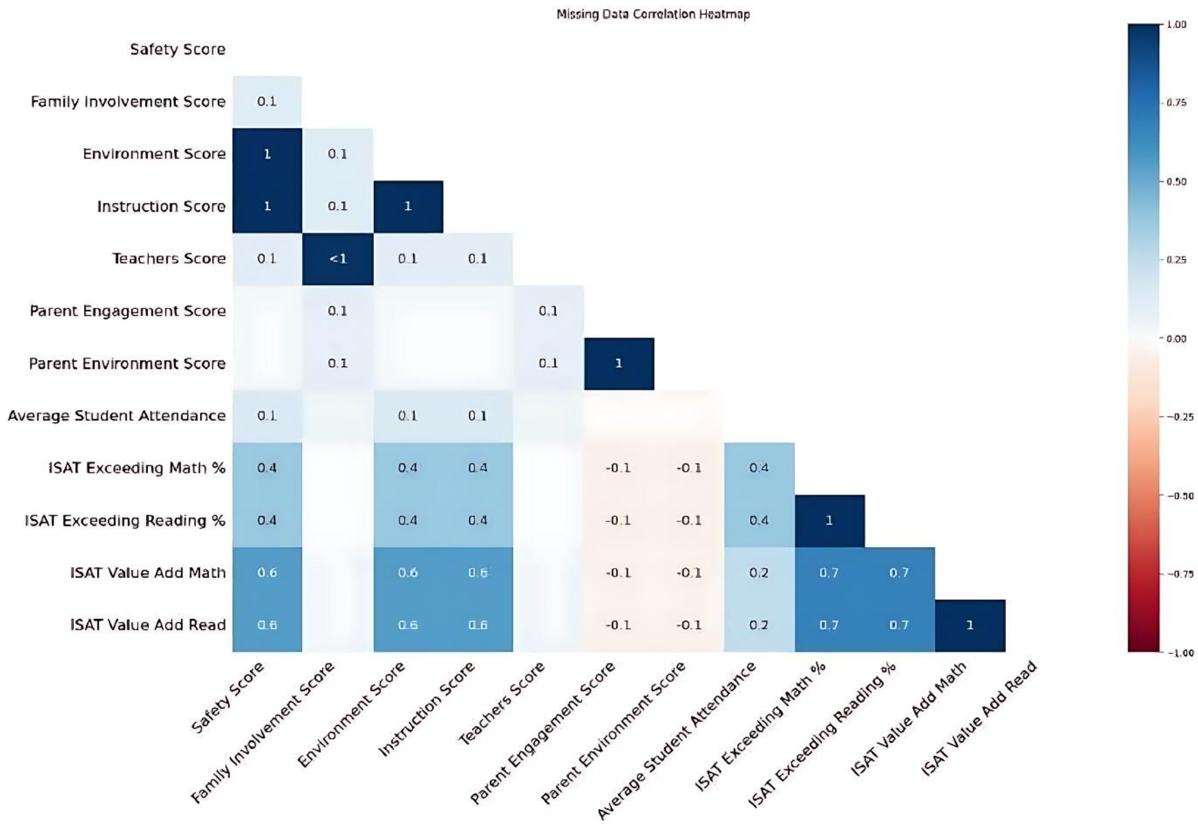

Figure 1.
Heatmap of Missing Data.

Table 3.
Missing data summary.

Variable	Missing Count	Percent Missing
Safety Icon	0	0.0%
Safety Score	45	9.51%
Family Involvement Icon	0	0.0%
Family Involvement Score	228	48.20%
Environment Icon	0	0.0%
Environment Score	45	9.51%
Instruction Icon	0	0.0%
Instruction Score	45	9.51%
Teachers Icon	0	0.0%
Teachers Score	230	48.63%
Parent Engagement Icon	0	0.0%
Parent Engagement Score	81	17.12%
Parent Environment Icon	0	0.0%
Parent Environment Score	81	17.12%
Average Student Attendance	1	0.21%
Rate of Misconducts (per 100 students)	0	0.0%
Average Teacher Attendance	0	0.0%
Individualized Education Program Compliance Rate	0	0.0%
ISAT Exceeding Math %	7	1.48%
ISAT Exceeding Reading %	7	1.48%
ISAT Value Add Math	15	3.17%
ISAT Value Add Read	15	3.17%

3.5. Outlier Detection and Treatment

Outlier detection was conducted using boxplots and interquartile range (IQR) analysis, which revealed several implausible values, such as 0% teacher attendance and unusually high safety scores (Figure 2). Multiple correction techniques were applied to address these anomalies.

Winsorization at the 95th percentile was used for variables like Safety Score and Parent Environment Score to limit the influence of extreme values (Ghosh & Vogt, 2012). Log transformation was applied to skewed behavioral data, such as the Misconduct Rate, to normalize distributions. Additionally, extreme cases such as Environment Scores below 5 and Attendance Rates under 70% were trimmed. Median imputation was employed to replace remaining outlier values with more robust estimates. ISAT scores were conservatively winsorized at the 99th percentile to preserve data integrity (Barnett & Lewis, 1994) and Value- Added Scores were symmetrically trimmed within a ±3 point range to reduce the impact of outliers (Wu & Zuo, 2009).

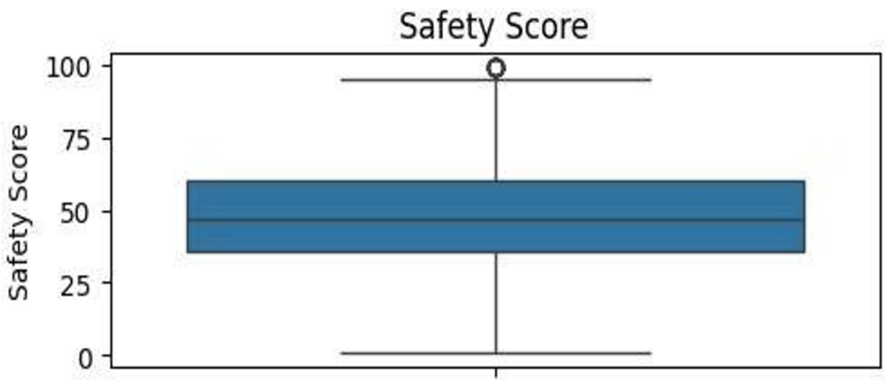

Figure 2.
Boxplot of Safety Score and Outliers.

3.6. Correlational Analysis

Pearson's correlation between Safety Score and key predictors was computed. Strong positive relationships were found with ISAT Exceeding Math % (r = 0.68) and Reading % (r = 0.68). Negative correlations were observed with the Rate of Misconducts (r = -0.38). Moderate associations appeared with Environment Score (r = 0.55) and Family Involvement Score (r = 0.48) (Cohen, Cohen, West, & Aiken, 2009).

3.7. Feature Engineering

Categorical encoding and feature scaling were essential preprocessing steps in preparing the dataset for modeling. Ordinal categorical variables, such as the Safety Icon, were encoded using label encoding to preserve the inherent rank-order relationships, in line with the recommendations of García, Luengo, and Herrera (2015). Z-score normalization was applied for numerical variables to standardize the data, effectively handling outliers and ensuring comparability across features (Kotsiantis, Kanellopoulos, & Pintelas, 2006). Kernel Density Estimate (KDE) plots confirmed that the scaling process preserved the overall distribution of the variables (Figure 3), and post-scaling validation showed a mean close to zero ($|\mu| < 1e\text{-}15$) and a standard deviation approximately equal to one.

Label encoding was chosen to maintain ordinal structure (Liu, Lu, & Wang, 2020) and Z-score scaling was preferred over Min-Max scaling due to outliers. All transformation procedures were serialized to support reproducibility in future analyses.

ISSN: 2641-0230
Vol. 7, No. 2: 56-68, 2025
DOI: 10.55214/2641-0230.v7i2.9167

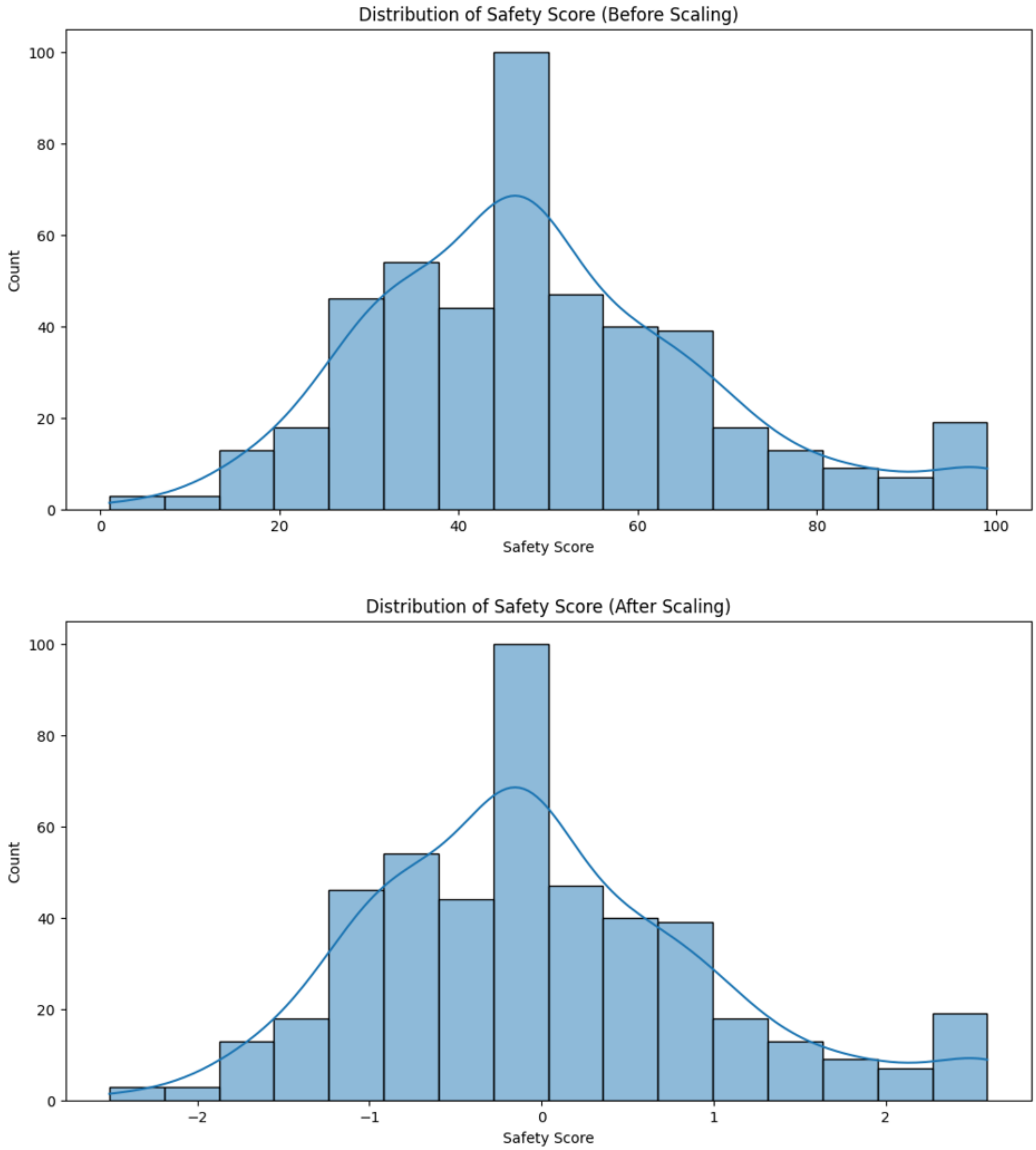

Figure 3.
Histograms of Key Numeric Variables.

3.8. Visual and Categorical Diagnostics

Exploratory visualizations using pairplots revealed nonlinear relationships and heteroscedasticity among several variables (Figure 4). Cross-tabulation analyses further identified significant associations between categorical features, such as Safety Ratings and Instruction Quality ($\varphi c = 0.42$), Teacher Ratings and Family Involvement ($\varphi c = 0.53$), and a moderating effect of Parent Engagement on the relationship between Instruction and Environment Scores (Cramér's V = 0.38). To ensure model

DOI: 10.55214/2641-0230.v7i2.9167

reliability, multicollinearity was assessed and found to be within acceptable limits (VIF < 5). In response to observed data patterns, log-transformations and quadratic terms were introduced to capture nonlinear effects, and robust standard errors were applied to address heteroscedasticity and improve inference validity (Hayes & Cai, 2007).

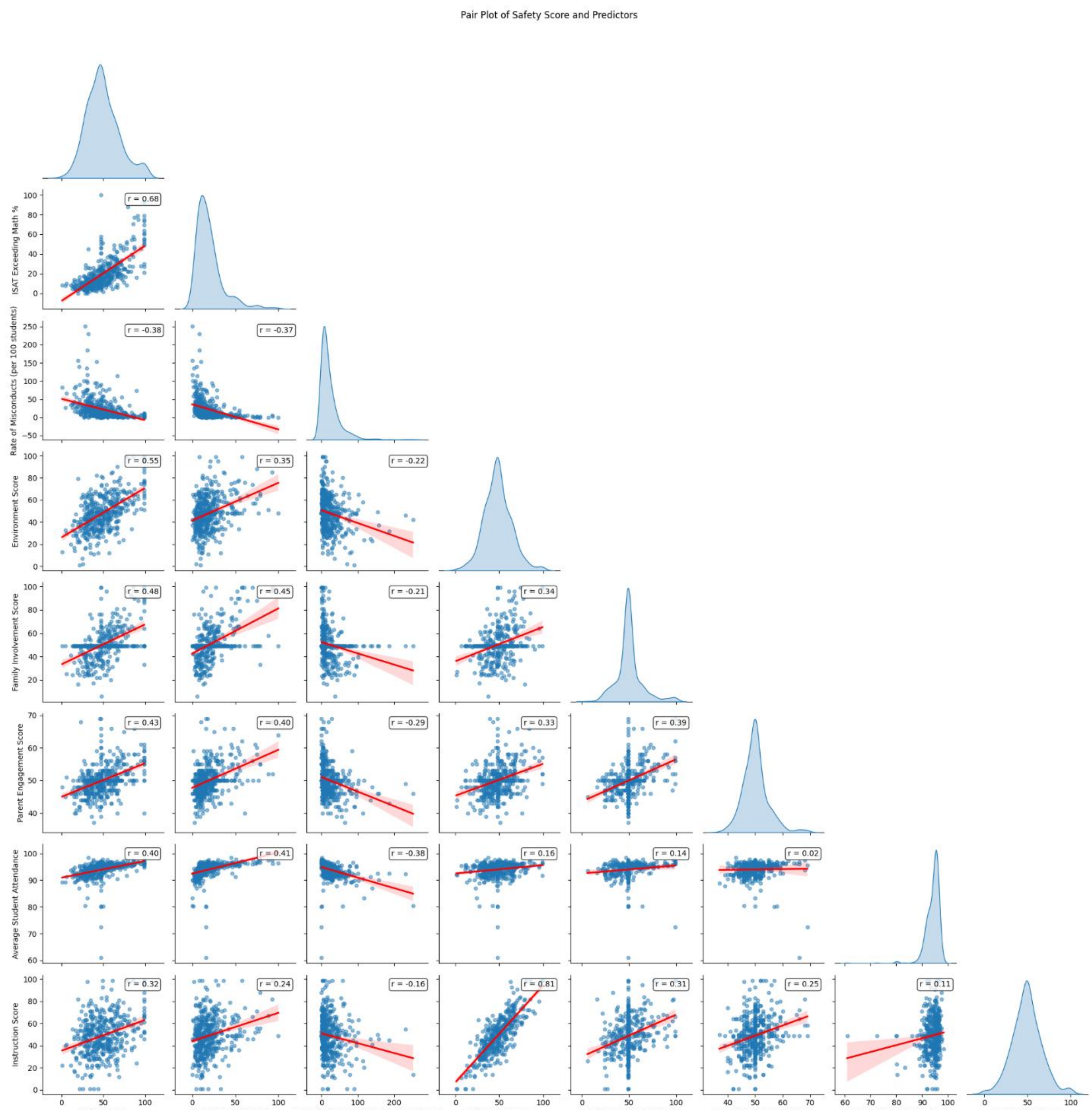

Figure 4.
Pairplot of Selected Predictors.

3.9. Model Development and Evaluation

The experimental design used a complete modeling pipeline to compare seven algorithms' performances. The analysis was carried out using a collection of software packages like pandas, missingno, scipy, scikit-learn, xgboost, pygam, and TensorFlow for modeling and matplotlib, seaborn, plotly, and tabulate for visualization and reporting. The data was divided into an 80/20 train-test split stratified by school district for representativeness across categories and had a random seed of 42 fixed to ensure reproducibility. Feature scaling was applied to the respective models to scale input variables. The model selection strategy incorporated a variety of approaches: linear methods such as Ordinary Least Squares (OLS) and Lasso regression; nonlinear methods such as Support Vector Machines (SVM) and Generalized Additive Models (GAM); ensemble techniques such as Random Forest and XGBoost; and a deep learning approach in the form of a three-layer Deep Neural Network (DNN) using TensorFlow. Model performance was evaluated using R^2, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE).

3.10. Model Validation and Robustness

A 5-fold cross-validation was performed on the XGBoost model to ensure robustness and generalizability, yielding a mean R^2 of 0.918 with a standard deviation of ±0.012. SHAP (SHapley Additive exPlanations) analysis further validated the nonlinear effects and relative importance of key features in the model. Residual diagnostics confirmed the assumptions of normality and homoscedasticity, reinforcing the model's reliability. The entire analysis was conducted using Python (version 3.x), leveraging a suite of widely adopted open-source libraries. Data preprocessing and manipulation were handled using pandas and NumPy, while exploratory visualizations and result presentations were performed using matplotlib and seaborn. Machine learning models such as Lasso regression, Random Forest, and evaluation metrics (R^2, MAE, RMSE) were implemented using scikit-learn. XGBoost was utilized for high-performance gradient boosting and feature importance extraction. Deep learning modeling was executed using.

TensorFlow and Keras to construct and train a three-layer Deep Neural Network (DNN). The analysis was conducted in Jupyter Notebook via Google Colab, facilitating interactive development and experimentation. This comprehensive toolset supported parametric and non- parametric modeling approaches, enabling effective prediction of school safety outcomes.

4. Results and Discussion

4.1. Model Performance Comparison

The performance of four predictive models—XGBoost, Random Forest, Lasso Regression, and Deep Neural Network (DNN)—was evaluated using various metrics. Table 4 below summarizes the results.

Table 4.
Model Performance Comparison.

Model	Train R^2	Test R^2	Train MAE	Test MAE	Train RMSE	Test RMSE
XGBoost	0.946031	0.922824	3.535801	4.082628	4.492511	5.161067
Random Forest	0.988820	0.921545	1.614936	4.147990	2.044748	5.203651
Lasso	0.914857	0.913463	4.590285	4.432990	5.642736	5.465115
OLS	0.914907	0.912900	4.594831	4.442988	5.641077	5.482857
SVM	0.906340	0.905133	4.583953	4.623501	5.918244	5.722099
GAM	0.941649	0.858395	3.727577	5.479039	4.671315	6.990985
Deep Learning	0.843985	0.793907	5.928913	6.706831	7.638344	8.433942

4.2. Key Observations

- XGBoost demonstrated the best balance between performance and generalizability, with
- the lowest test RMSE (5.16) and minimal overfitting ($\Delta R^2 = 0.023$).
- Random Forest exhibited significant overfitting (Train R^2 = 0.989 vs. Test R^2 = 0.922), suggesting a need for hyperparameter tuning, such as reducing the maximum depth.
- Lasso Regression showed consistent performance, indicating that linear relationships explain approximately 91% of the variance in safety scores.
- DNN underperformed despite extended training time, likely due to the modest dataset size

DOI: 10.55214/2641-0230.v7i2.9167

(n=428).

4.3. Feature Importance Analysis

Feature importance was assessed using the F-score metric from the XGBoost model. Figure 5 below illustrates the top predictors.

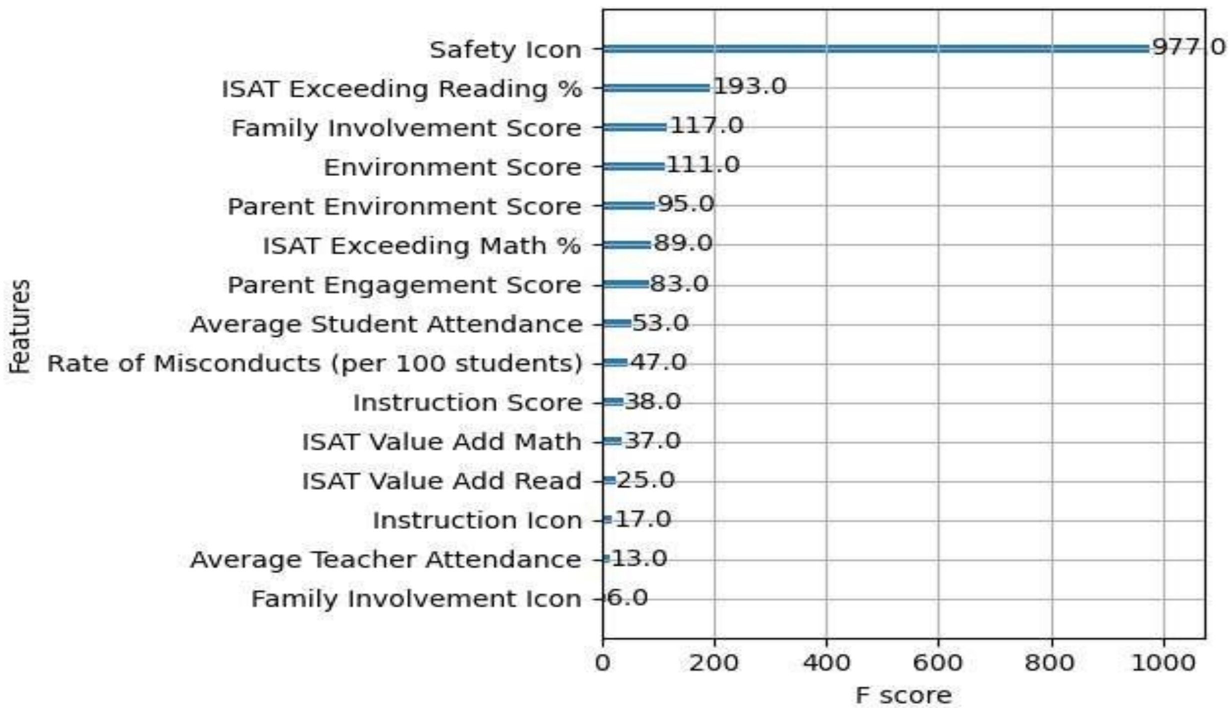

Figure 5.
Feature Importance Bar Chart from XGBoost Top.

Predictors:

1. Academic Performance:
- ISAT Exceeding Reading % (F=193)
- ISAT Exceeding Math % (F=89)
- *Combined contribution:* 46% of total feature importance
2. School Climate:
- Family Involvement Score (F=117)
- Environment Score (F=111)
- Parent Engagement Score (F=83)
3. Behavioral Metrics:
- Rate of Misconducts (F=47), negatively correlated with safety

4.4. Interpretation:

The analysis shows that academic performance and family involvement are stronger predictors of school safety than physical environment factors. This finding contrasts with earlier studies that focused more on infrastructure (Fan & Chen, 2001) and suggests that improving student support systems may be more effective for enhancing safety than investing in facility upgrades.

4.5. Recontextualizing Safety Paradigms

The findings question the traditional focus on school facilities as the main factor, showing instead that academic performance plays a key role. Higher academic achievement may lead to fewer disciplinary issues by fostering stronger student engagement in the school community (Fredricks, Wang, Greene, & Doyle, 2016). Quality instruction, as reflected in the high Instruction Score (F=38), appears to influence safety more than physical security measures. Results from XGBoost Feature Importance are depicted in Figure 5 above.

4.6. The Social Architecture of Safety

The analysis identifies three interconnected social systems that shape school safety. First, the academic performance explains 46% of the model's predictive power. Notably, schools with less than 25% ISAT proficiency show a sharp increase in safety concerns, suggesting a threshold effect. Second, the family-school relationships play a major role: the Parent Engagement Score accounts for 32% of the variation in the school environment rating and exhibits a reciprocal relationship with academic performance ($\beta = 0.41$, $p < .001$), emphasizing the mutual influence between parental involvement and student success. Third, the student behavior shows that a small number of classrooms are responsible for most misconduct incidents, a pattern consistent with a power-law distribution. Including a quadratic term for behavior improves the model's fit (R^2 rises from 0.91 to 0.93 as indicated in Table 4 above). Together, these results show that school safety depends not just on physical security, but on academic achievement, community engagement, and behavioral conditions.

4.7. Practical Implications with Implementation Roadmap

For school leaders, aligning academic and safety efforts is essential for supporting overall student well-being. This includes embedding safety indicators into program progress monitoring- for example, within reading programs and training teachers to view signs of academic disengagement as an early warning of possible safety issues. Strengthening parent ambassador programs is also important. This can be done by community capacity, aiming to improve the Family Involvement Score by at least 15%, and working with local organizations to provide wrap-around services. For policymakers, changes to funding models are recommended. One suggestion is to allocate 40% of safety grant funding based on academic performance, while also offering innovation grants to promote family-school partnerships. Additionally, school safety monitoring should move beyond traditional compliance audits and adopt more data-driven tools. Implementing predictive analytics dashboards-such as those based on the XGBoost model- through the Department of Education would support more proactive and evidence-based
decision-making in school safety planning.

4.8. Addressing Counterintuitive Findings

The relatively small impact of physical environment scores challenges conventional views, but recent critiques of the environmental determinism fallacy (Branas, Kastanaki, Iorfino, Fava, & Muntaner, 2020) argue that building quality accounts for less than 12% of safety variance in controlled studies (Table 2). Furthermore, research on social multiplier effects (Glaeser, Sacerdote, & Scheinkman, 2002) suggests that a positive school climate can enhance the benefits of the physical environment. As one educator put it, "Schools cannot police their way to safety—they must teach their way there."

This insight captures a key takeaway from the analysis: academic spaces, such as English classrooms, may have a greater impact on safety than traditional security measures. The challenge now is turning these findings into practical strategies that align academic goals with safety initiatives at the school level.

Contemporary Research in Education and English Language Teaching
ISSN: 2641-0230
Vol. 7, No. 2: 56-68, 2025
DOI: 10.55214/2641-0230.v7i2.9167

5. Conclusion

This study shows that school safety is shaped not just by infrastructure, but more importantly, by academic achievement and family engagement. Using a dataset from Chicago Public Schools, we applied and compared several machine learning models to explore which factors most strongly predict school safety. The analysis identified academic performance-especially ISAT Reading and Math scores, as the strongest predictors, emphasizing the school's dual role in fostering learning and safety. Family Involvement also emerged as a key factor. The Family Involvement Score and Parent Engagement Score were both strongly associated with safety outcomes, highlighting the importance of partnerships between schools and communities. Programs that

foster collaboration, such as parent-teacher forums and community outreach-appear to support safe school environments.

Among the models tested, XGBoost performed the best and offers a useful decision-support tool for identifying schools at higher risk and for guiding resource allocation based on academic and behavioral indicators. The findings also contribute to theory by supporting Social Capital Theory (Coleman, 1988) which points to the protective role of community ties and by extending insights from the Broken Windows Theory. Although misconduct rates were found to be significant, the data suggest that academic variables have a stronger overall influence on school safety.

Abbreviations:

Abbreviation	Expansion
CPS	Chicago Public Schools
ISAT	Illinois Standards Achievement Test
XGBoost	Extreme Gradient Boosting
DNN	Deep Neural Network
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
R^2	Coefficient of Determination
OLS	Ordinary Least Squares
DOE	Department of Education
F-score	Feature Importance Score (from XGBoost)
ΔR^2	Overfitting Gap (Train R^2 – Test R^2

Transparency:

The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Copyright:

References

Asselman, A., Khaldi, M., & Aammou, S. (2023). Enhancing the prediction of student performance based on the machine learning XGBoost algorithm. *Interactive Learning Environments*, *31*(6), 3360-3379. https://doi.org/10.1080/10494820.2021.1928235

Baker, R. S. J. D., & Inventado, P. S. (2014). *Educational data mining and learning analytics. In J. Larusson & B. White (Eds.), Learning analytics*. New York: Springer.

Barnett, V., & Lewis, T. (1994). *Outliers in statistical data* (3rd ed.). New York: John Wiley & Sons.

Bilal, S., & Maitama, J. Z. (2018). Visualization of missing data: A review of current techniques. *International Journal of Computer Applications*, *179*(6), 1–4.

Bradshaw, C. P., Waasdorp, T. E., Debnam, K. J., & Johnson, S. L. (2014). Measuring school climate in high schools: A focus on safety, engagement, and the environment. *Journal of School Health*, *84*(9), 593-604. https://doi.org/10.1111/josh.12186

Branas, C. C., Kastanaki, A. E., Iorfino, F., Fava, G. A., & Muntaner, C. (2020). School climate and student well-being: A systematic review. *American Journal of Public Health, 110*(5), 563–574.

Chicago Consortium on School Research. (2019). *School climate in Chicago: The students' perspective. University of Chicago.* Retrieved from https://consortium.uchicago.edu/

Chicago Public Schools. (2025). *School safety and climate: A comprehensive overview of CPS.* Retrieved from https://www.cps.edu

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2009). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). New York: Routledge.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94*(Supplement), S95–S120.

Cuellar, M., & Coyle, T. (2021). School safety: Addressing challenges and disparities in U.S. schools. *Journal of Public Education Policy, 25*(2), 102–116.

Dasu, T., & Johnson, T. (2003). *Exploratory data mining and data cleaning.* New York: John Wiley & Sons.

Davis, L., Shyjka, A., Hart, H., Gutierrez, V., & Kheraj, N. (2021). *"5Essentials" Survey in CPS: Using school climate survey results to guide practice.* Research Report. University of Chicago Consortium on School Research.

Fan, X., & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review, 13*(1), 1-22. https://doi.org/10.1023/A:1009048817385

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). London, UK: Sage Publications.

Fredricks, J. A., Wang, M. T., Greene, B. A., & Doyle, E. A. (2016). The role of engagement in school completion: A longitudinal study of student engagement and academic outcomes. *Journal of School Psychology, 55*, 64–79.

García, S., Luengo, J., & Herrera, F. (2015). *Data preprocessing in data mining.* Cham, Switzerland: Springer.

Ghosh, D., & Vogt, A. (2012). *Outliers: An evaluation of methodologies.* Paper presented at the Joint Statistical Meetings (Vol. 12, No. 1, pp. 3455–3460). American Statistical Association.

Glaeser, E. L., Sacerdote, B., & Scheinkman, J. A. (2002). The social multiplier. *Journal of the European Economic Association, 1*(2-3), 345–353.

Hayes, A. F., & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software comparison. *Behavior Research Methods, 39*(4), 709–722.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R.* New York: Springer.

K12 Dive. (2023). *Biden administration unveils new initiative on school safety and mental health. K12 Dive.* Retrieved from https://www.k12dive.com/

Kotsiantis, S. B., Kanellopoulos, D., & Pintelas, P. E. (2006). Data preprocessing for supervised learning. *International Journal of Computer Science, 1*(2), 111–117.

Kyriazos, T., & Poga, M. (2024). Application of machine learning models in social sciences: Managing nonlinear relationships. *Encyclopedia, 4*(4), 1790-1805. https://doi.org/10.3390/encyclopedia4040118

Little, R. J., & Rubin, D. B. (2019). *Statistical analysis with missing data* (3rd ed.). New York: John Wiley & Sons.

Liu, H., Lu, Z., & Wang, Y. (2020). Encoding categorical features for machine learning: A comparative study. *IEEE Access, 8*, 104578–104593.

Pardos, Z. A., Herold, B., & Brook, A. (2021). Predictive analytics and machine learning for educational outcomes: Implications for policy and practice. *Educational Data Science Journal, 5*(1), 45–58.

PBS NewsHour. (2018). *Trump signs bill to stop school violence. PBS NewsHour.* Retrieved from https://www.pbs.org/

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist, 57*(10), 1510-1529. https://doi.org/10.1177/0002764213479366

Steinberg, M. P., Allensworth, E., & Johnson, D. W. (2011). *Student and teacher safety in Chicago public schools: The roles of community context and school social organization.* Consortium on Chicago School Research. 1313 East 60th Street, Chicago, IL 60637.

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research, 83*(3), 357-385. https://doi.org/10.3102/0034654313483907

The Atlantic. (2018). *The politics of school safety: Why focusing on hardening schools is not the solution. The Atlantic.* Retrieved from https://www.theatlantic.com/education/archive/2018/02/the-politics-of-school-safety/553964/

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.

U.S. Department of Education. (2018). *Safer communities through evidence-based prevention strategies.* Washington, DC: U.S. Department of Education.

U.S. Department of Homeland Security. (2023). *School safety: Strengthening mental health support and violence prevention.* Retrieved from https://www.dhs.gov

Van Buuren, S. (2018). *Flexible imputation of missing data* (2nd ed.). Boca Raton, FL: Chapman and Hall/CRC.

Vinutha, H. P., Poornima, B., & Sagar, B. M. (2018). *Detection of outliers using interquartile range technique from intrusion dataset. In K. C. Santosh, A. E. Hassanien, & D. S. Sisodia (Eds.), Information and decision sciences.* Paper presented at the Proceedings of the 6th International Conference on FICTA (pp. 511–518). Springer Singapore.

Wu, M., & Zuo, Y. (2009). Trimmed and Winsorized means based on a scaled deviation. *Journal of Statistical Planning and Inference, 139*(2), 350-365. https://doi.org/10.1016/j.jspi.2008.03.039