SAN MATEO COUNTY OFFICE OF EDUCATION

A Place-Based Approach to Chronic Absenteeism Risk

Yung-Hui Chien, Lihua Yang, Anya Macomber,
Jeremy Del Carpio, Leena Bhai, Jared Prolo

San Mateo County Office of Education
Center for Learning Analytics

Contact:
Yung-Hui Chien
yhchien@smcoe.org

Suggested Citation:

Chien, Y.-H., Yang, L., Macomber, A., Del Carpio, J., Bhai, L., & Prolo, J. (2026). *A Place-Based Approach to Chronic Absenteeism Risk*. San Mateo County Office of Education.

A Place-Based Approach to Chronic Absenteeism Risk

Yung-Hui Chien, Lihua Yang, Anya Macomber, Jeremy Del Carpio, Leena Bhai, Jared Prolo

San Mateo County Office of Education, Center for Learning Analytics

Abstract

Chronic absenteeism reflects overlapping school, family, and community conditions that district-level averages may obscure. This study developed a place-based risk profile for 30 ZIP codes in San Mateo County using publicly available educational and community indicators. District outcomes were allocated to ZIP codes through a population-informed crosswalk. Correlation screening, principal component analysis, and regression identified eight indicators associated with absenteeism and produced relative local risk scores. Mapping revealed variation within and across district boundaries, while an indicator-contribution heat map showed that similar scores reflected different local conditions. The approach supports targeted inquiry, cross-sector coordination, and equitable, context-responsive attendance improvement planning.

Purpose

Chronic absenteeism has become an important measure of school performance and accountability across the United States (U.S. Department of Education, 2025). In California, it was adopted in 2016 as a state indicator for grades Kindergarten to eighth grade and incorporated into the California School Dashboard's accountability framework (Gee & Kim, 2019). Its inclusion reflects a basic but important principle: students must be present to benefit from high-quality instruction, supportive relationships, and the programs schools provide. Even well-designed services have limited impact when students are not consistently present to access them.

Chronic absenteeism is also used to identify local educational agencies for additional assistance through California's system of Differentiated Assistance. Districts may qualify for support when student groups demonstrate low performance across multiple state or local indicators (California Department of Education, 2025). This accountability structure has increased attention to attendance and encouraged districts to monitor disparities across student groups. However, after several years of implementation, many districts continue to struggle with identifying and addressing the conditions that contribute to chronic absenteeism.

Part of the challenge is that chronic absenteeism is influenced by multiple and often overlapping factors. Families may be managing unstable housing, health needs, transportation problems, or work schedules that make attendance routines difficult to sustain. Schools may be managing staffing instability, gaps in communication, or weak student connection. These conditions can overlap and reinforce one another. Common strategies, such as attendance letters or general awareness campaigns, may address the visible outcome without addressing the conditions behind it. Districts may see improvement for one student group or school, followed by declining attendance elsewhere. This can create a recurring cycle of response without a clear understanding of why attendance patterns are changing.

California's Differentiated Assistance process encourages districts to use Improvement Science to investigate root causes and test potential changes in their local contexts (California County Superintendents, 2025). In practice, however, districts and schools may have limited time, analytic capacity, or access to data that describe the broader environments in which students live. Attendance analyses often focus on individual students, student groups, schools, or districts. While each level is important, none fully captures variation among neighborhoods

within the same district. District-level averages may obscure variation in attendance-related conditions among communities within the same district.

This study addresses that gap by examining factors associated with chronic absenteeism and estimating how risk varies by ZIP code. The purpose is not to replace student-level attendance monitoring or school-level accountability data, but to add a place-based layer that shows where risk is concentrated and which local conditions are most strongly connected to it. By centering geographic context, the study seeks to make inequities more visible and shift attention away from explanations that place responsibility primarily on students and families. The approach is intended to help leaders identify areas for further inquiry and support more targeted, equitable action.

The study focuses on transitional kindergarten through eighth grade (TK-8) because early and middle childhood are critical periods for developing academic foundations, social-emotional skills, school relationships, and attendance routines. Absence during these years has been associated with academic underachievement, grade retention, and increased risk of dropping out later in a student's educational experience (Calabro-Pepin, 2020). It can also affect classmates through disruptions to instruction and classroom continuity (Gottfried, 2019). Earlier identification of attendance risk may therefore allow districts and community partners to intervene before patterns become more difficult to change.

Conceptual Framework

This study is grounded primarily in Bronfenbrenner's ecological systems theory, which frames student outcomes as the result of interactions across multiple overlapping systems, including families, schools, neighborhoods, institutions, and broader social and economic

structures (Bronfenbrenner, 1979). From this perspective, chronic absenteeism is not simply a matter of student motivation, family behavior, or school practice. It can reflect pressures and constraints across systems, such as unstable housing, limited access to care, or difficulty remaining connected to school (Gottfried & Gee, 2017). This framing shifts the focus from asking why individual students are not attending school to asking what conditions make regular attendance more difficult in certain communities.

The study also draws on Vygotsky's sociocultural theory to explain how these broader conditions influence daily attendance practices. From a sociocultural perspective, participation is supported through relationships, routines, and tools that help people function within social institutions (Vygotsky, 1978). Consistent attendance depends partly on whether families can communicate with schools, understand expectations, maintain stable schedules, and navigate education systems. When families experience overcrowded housing or limited language access, the issue is not necessarily that school is less valued. Rather, the support that helps sustain attendance routines may be weak, difficult to reach, or hard to maintain.

A geospatial equity lens extends these theories by recognizing that opportunity and risk are unevenly distributed across places. Attendance is associated with neighborhood and broader structural conditions, suggesting that where students live may affect their exposure to attendance barriers (Singer et al., 2021). For example, housing instability in San Mateo County was associated with higher chronic absenteeism and disparities in other educational outcomes (John W. Gardner Center for Youth and Their Communities, 2022). ZIP code-level analysis can provide a way to examine where housing-related and other barriers may be concentrated.

Together, these perspectives frame chronic absenteeism as structural rather than primarily individual, socially mediated rather than simply behavioral, and geographically patterned rather

than evenly distributed. Ecological systems theory identifies the multiple settings and systems that may shape attendance, sociocultural theory explains how these conditions are experienced through everyday relationships, routines, communication, and access to institutional supports, and the geospatial equity lens draws attention to how these conditions are distributed across communities.

This combined framework supports the study's examination of school and community indicators within the same model and guides the interpretation of ZIP code differences as patterns of unequal exposure to supportive resources and attendance barriers, rather than as inherent characteristics of students or neighborhoods. The resulting risk profiles are therefore interpreted as contextual signals for systems-level inquiry.

Research Questions

This study is guided by three research questions:

(1) What educational and socioeconomic factors are most strongly associated with TK-8 chronic absenteeism across ZIP codes?

(2) How can these factors be integrated into a regression-based model for estimating chronic absenteeism risk across local contexts?

(3) How does chronic absenteeism risk vary across ZIP codes?

Methods

The study integrates publicly available educational and community indicators from 2020 through 2025. When possible, indicators were selected from comparable reporting periods. The resulting dataset was analyzed cross-sectionally to examine variation across districts and ZIP

codes rather than change over time. Educational measures, including chronic absenteeism, student subgroup enrollment, and staffing characteristics, were drawn from California Department of Education (CDE) sources (California Department of Education, n.d.). Student wellness and school climate measures were incorporated from the California Healthy Kids Survey (CHKS) where available (WestEd, n.d.). Community-level measures came from the American Community Survey (ACS) (U.S. Census Bureau, n.d.) and California Department of Public Health (CDPH) public data portal (California Department of Public Health, n.d.). The resulting dataset represents six domains: demographics, student subgroups, health and safety, economic and housing conditions, school structure, and staff characteristics.

The main analytic decision was to examine absenteeism risk at the ZIP code level. ZIP codes were selected because several public datasets can be linked or approximated at that geographic unit. Although ZIP codes do not perfectly represent communities, they provide a practical way to connect district outcomes with local social and economic conditions.

A district-to-ZIP code crosswalk was developed to connect district outcomes with community measures while preserving local variation. The crosswalk was created using the National Center for Education Statistics (NCES) School District Geographic Relationship Files, which link school districts to ZIP Code Tabulation Areas (ZCTAs) (National Center for Education Statistics, n.d.). District-level values were allocated to ZCTAs through a population-informed crosswalk. For each district-ZCTA intersection, the proportion of the ZCTA's land area within the district was multiplied by the ZCTA population to estimate the intersection population. These estimates were then converted into weights representing each district's share of the estimated population within the ZCTA. The ZIP code chronic absenteeism estimate was calculated as the weighted average of the chronic absenteeism rates of the overlapping districts.

The district-level source data included 20 school districts, and the crosswalk produced estimates for 30 residential ZIP codes in San Mateo County.

Because residential ZIP codes in the county correspond to their associated ZCTAs, the terms “ZIP code” and “ZIP” are used for both geographic units throughout the paper. Using ZIP codes as the unit of analysis involves several methodological considerations. District and ZIP code boundaries do not align perfectly, and students may attend school outside the ZIP code where they live. The crosswalk therefore provides an analytic approximation rather than a direct assignment of individual students to neighborhoods. This approach was considered appropriate because the purpose was to estimate community context, not to reproduce student-level attendance records.

Pearson correlations were first used to identify variables associated with chronic absenteeism. Variables with correlations of $|r| \geq .50$ were retained for further analysis, and descriptive statistics were examined for the selected measures. When conceptually overlapping variables met the threshold, such as English learner and long-term English learner concentration, the variable with the stronger correlation was retained. The selected variables were standardized using z scores so that variables on different scales could be compared. Pairwise correlations among the retained variables were then reviewed for multicollinearity.

Principal component analysis (PCA) was used to reduce overlap among the retained variables. PCA summarizes shared variation among related indicators in a smaller set of uncorrelated components, which reduces redundancy before regression modeling. A scree plot was examined to identify plausible component solutions. The two- and three-component regression models were then compared using an incremental F-test to determine whether adding a third component significantly improved model fit.

The retained components were entered as predictors in a linear regression model of chronic absenteeism. The fitted regression equation was then used to estimate chronic absenteeism risk for the 30 ZIP codes. The model was designed to estimate a current place-based risk profile using the most recently available indicators, rather than forecast future chronic absenteeism. Coefficients for the original indicators were then derived by combining the regression coefficients for the components with the corresponding component loadings.

The resulting regression-based estimates were then converted into risk scores using min-max scaling. Scores ranged from 0 to 100, with 100 assigned to the highest estimated value in the county sample and 0 assigned to the lowest. These scores represent relative positions within this analysis and should not be interpreted as probabilities or percentages of chronic absenteeism. These risk scores were mapped to examine their geographic distribution across the county and in relation to school district boundaries.

Finally, to identify the indicators most relevant to each ZIP code score, each standardized ZIP code value was multiplied by its coefficient from the PCA-based regression equation. The resulting products were treated as model-based indicator contributions and displayed in a heat map table to show their relative magnitude and direction within each ZIP code. These contributions describe how each indicator enters the estimated score and should not be interpreted as causal effects.

Results

The initial correlation analysis identified several strong relationships with chronic absenteeism (see Table 1). Hispanic student concentration and English learner concentration showed the strongest positive correlations. Student-reported sadness, homelessness, and

overcrowded housing were also positively correlated with absenteeism, while the percentage of fully credentialed teachers showed a negative relationship. These associations describe how the variables varied with chronic absenteeism in the sample and do not establish causality.

Eight variables met the model inclusion criteria of having a Pearson correlation coefficient of $|r| \geq .50$ and acceptable data quality. Descriptive statistics for the retained variables are presented in Table 2. The variables varied considerably across ZIP codes, providing local variation for the subsequent analyses. After the retained variables were standardized, pairwise correlations were used to assess the relationships among the variables (Table 3). Several variables were strongly related, such as Hispanic student and English learner concentration. Because these variables were correlated, their contributions should not be interpreted as independent effects. The broader pattern of intercorrelations supported the use of PCA to reduce overlap among variables.

The PCA scree plot indicated that either a two- or three-component solution was plausible. The first two components accounted for 71% of the total variance, while the first three accounted for 82% (see Figure 1). An incremental F test comparing a regression model containing the first two principal components, PC1 and PC2, with a model that also included the third principal component, PC3, showed that adding PC3 did not significantly improve model fit ($F(1, 26) = 0.08$, $p = .78$). The two-component regression model was therefore retained.

The component loadings show how each retained variable contributed to each of the principal components (see Table 4). PC1 was characterized most strongly by Hispanic student and English learner concentration in one direction and by the percentage of students identifying with two or more races in the other. PC2 was characterized most strongly by homelessness and

Asian student concentration. Each component combines demographic, community, and school variables and is treated as a statistical dimension rather than a single underlying cause.

The two-component regression model explained approximately 86% of the observed variation in chronic absenteeism (R^2 = .86, adjusted R^2 =.84, $F(2,27) = 79.82$, $p<.001$) (see Table 5). PC1 was statistically significant, whereas PC2 did not reach the conventional .05 significance level. PC2 was retained to preserve the two-dimensional PCA representation used to construct the current risk profiles, although its coefficient should be interpreted cautiously.

The coefficients derived from the PCA-based regression equation were positive for English learner concentration, Hispanic student concentration, homelessness, overcrowded housing, and student-reported sadness, and negative for Asian student concentration, the percentage of fully credentialed teachers, and the percentage of students identifying with two or more races (see Table 6). The resulting chronic absenteeism risk scores varied substantially across the county (see Table 7). ZIP code 94303 received the highest relative risk score, followed by 94063 and 94060. ZIP codes 94065 and 94070 received the lowest relative risk scores.

Table 8 provides additional detail about the indicator contributions underlying each ZIP code's estimated risk score. The magnitude and direction of the contributions varied across ZIP codes, showing that similar scores can reflect different combinations of community and school conditions. Positive values raised the estimated score, whereas negative values lowered it.

The countywide map shows that estimated risk was unevenly distributed across ZIP codes (see Figure 2). Some high-risk areas did not stand out in district-level averages, and neighboring ZIP codes sometimes showed different profiles even when they shared schools or services. When school district boundaries were added, estimated risk did not align neatly with

district geography (see Figure 3). Together, the maps make variation within and across district boundaries more visible.

The chronic absenteeism risk scores describe community-level patterns, not outcomes for individual students. A high-risk score indicates that a ZIP code has a combination of conditions associated with higher absenteeism. It does not mean that every student in that area will be chronically absent. The scores should therefore be interpreted as contextual estimates that guide further inquiry.

Discussion

The findings support the central argument of this study: chronic absenteeism should be understood as a place-based and systems-level indicator. Students miss school for immediate reasons, but those reasons are shaped by broader conditions.

This matters for accountability. Accountability systems can identify where rates are high, but they do not always help leaders understand why rates are high. A school or district can be identified for chronic absenteeism without enough information about whether the underlying barriers are related to housing, health, transportation, school climate, staffing, family engagement, or community stress. Furthermore, when these place-based conditions span district boundaries, the accountability systems can misidentify, or fail to identify areas of need. The place-based model adds an interpretive layer to accountability data by shifting the question from who has a high rate to what local conditions may be shaping that rate.

A comparison of the two highest-risk ZIP codes shows how different local conditions can produce distinct risk profiles. ZIP code 94303 received the highest chronic absenteeism risk score in the model, followed by 94063. The three largest indicator contributions to the 94303

score were associated with its teacher credentialing rate, English learner concentration, and Hispanic student concentration. This profile suggests that local inquiry in 94303 should examine staffing stability and school capacity, as well as the accessibility of support for multilingual students and families. Local evidence could then determine whether responses related to credentialing pathways, staffing stability, or linguistically accessible family communication are warranted.

The 94063 profile differs in that overcrowded housing, along with Hispanic student and English learner concentration, makes one of the largest indicator contributions to its score. Language access and responsive family engagement still warrant examination, but the prominence of overcrowded housing suggests that local inquiry should also assess housing-related attendance barriers and the possible need for coordination with housing agencies and family resource providers. Schools cannot resolve overcrowded housing, but they can identify related barriers, connect families with available support, and bring education data into broader community planning.

The demographic indicators require especially careful interpretation. Hispanic student and English learner concentration should not be treated as explanations for absenteeism. Instead, these indicators should prompt examination of the institutional and community-based conditions surrounding students, including whether services, communication, and engagement practices are accessible and responsive.

The housing, well-being, and staffing results also point to different areas for inquiry. Housing instability can make attendance routines harder to sustain by affecting transportation, sleep, and communication with schools. Where student-reported sadness contributes strongly to a risk profile, districts may need to examine school connectedness and access to support. The

negative association with fully credentialed teachers may reflect a protective dimension of staffing stability, although this study did not test that mechanism directly. Local evidence could then determine whether the relevant staffing issue involves credentialing, recruitment, retention, assignment patterns, or another dimension of school capacity.

In practice, the model can serve as the first stage of a structured inquiry process. Leaders can use the countywide map to identify locations for closer examination and the contribution heat map to develop questions about the conditions associated with each score. Real-time attendance data can then show which students are beginning to miss school, while school records, service data, and input from students and families can help explain why. When multiple sources identify a common barrier, districts and partners can select a focused strategy, establish targeted measures, and monitor whether the response improves attendance. This sequence connects place-based analysis with Improvement Science by using the model to guide inquiry rather than determine services or interventions.

This process may also reveal attendance barriers that require cross-sector or regional coordination. Schools cannot address housing, health, and other community conditions on their own, but partnerships with community organizations and public agencies can help them respond more effectively (Germain et al., 2024). For behavioral-health-related barriers, California's Children and Youth Behavioral Health Initiative (CYBHI) offers one example of a statewide framework connecting schools, behavioral health systems, community partners, youth, and families to strengthen prevention, early intervention, and access to services (California Health and Human Services Agency, n.d.). County-level mapping can complement such efforts by showing where agency responsibilities and services overlap, whether support reaches higher-risk communities, and when a challenge extends across district boundaries. Used alongside local

knowledge and student-level data, the community profile can reveal when what appears to be an isolated family challenge may reflect a broader pattern affecting multiple families in the same area, supporting a more preventive and coordinated system of support.

Several limitations of this study should be considered. The analysis is correlational and does not establish causality. ZIP codes may hide finer variation within communities, and public datasets cannot capture barriers that are visible only in student-level data. The population-informed crosswalk assumes that population is distributed proportionally across the land area of each ZCTA, which may introduce error where residential development is unevenly distributed. The model reflects relationships observed in the most recently available data and should not be assumed to remain stable over time. Because community and school conditions may change, the model should be updated as new data become available. The model should therefore be used as a source of current contextual evidence, and indicators with large contributions should not be interpreted as direct causes.

Future research should update the model as new data become available, examine the stability of its risk profiles over time, and assess whether the observed relationships hold across other settings. Additional work could examine smaller geographic units where data availability and privacy protections allow and include qualitative evidence from families and educators. Studying positive outliers, including communities where estimated risk is high but observed absenteeism is lower than expected, may reveal protective practices that can inform attendance improvement elsewhere.

Conclusion

Overall, this study shows that chronic absenteeism is not only a school outcome. It is also a signal of community conditions and educational opportunity. By linking education data with community indicators at the ZIP code level, the analysis reveals local patterns that district averages may obscure. Instead of treating chronic absenteeism only as a lagging accountability measure, districts and community partners can use the model to identify underlying conditions and coordinate support earlier. In this way, attendance data become a tool for understanding housing instability, student well-being, language access, belonging, and school capacity. This shift is essential to ensuring that every learner, regardless of where they live, has a fair opportunity to participate and succeed in school.

References

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.

Calabro-Pepin, K. L. (2020). *A collective actions approach to reducing chronic absenteeism: An analysis of student responses, family court perspectives, & key characteristics of successful and sustainable interventions* [Doctoral dissertation, Johnson & Wales University]. ProQuest Dissertations Publishing.

California County Superintendents. (2025). *Differentiated assistance for local educational agencies facilitation guide: 2025-26 (Version 5.0).* https://www.cacountysupts.org/_files/ugd/63f997_76e5b3ba8c074e72aa58bcbeccda2984.pdf

California Department of Education. (2025). *System of support*. https://www.cde.ca.gov/ta/ac/cm/sysofsupport.asp

California Department of Education. (n.d.). *DataQuest.* https://dq.cde.ca.gov/dataquest/

California Department of Public Health. (n.d.). *Public health data portal.* https://data.chhs.ca.gov/organization/california-department-of-public-health

California Health and Human Services Agency. (n.d.). *About the Children and Youth Behavioral Health Initiative.* https://cybhi.chhs.ca.gov/about/

Germain, E., Hernández, L. E., Klevan, S., Levine, R. S., & Maier, A. (2024). *Reducing chronic absenteeism: Lessons from community schools.* Learning Policy Institute. https://doi.org/10.54300/510.597

Gee, K. A., & Kim, C. (2019). *Chronic absence in California: What new dashboard data reveals about school performance*. Policy Analysis for California Education. https://edpolicyinca.org/publications/chronic-absence-dashboard

Gottfried, M. A. (2019). Chronic absenteeism in the classroom context: Effects on achievement. *Urban Education*, 54(1), 3-34. https://doi.org/10.1177/0042085915618709

Gottfried, M. A., & Gee, K. A. (2017). Identifying the determinants of chronic absenteeism: A bioecological systems approach. *Teachers College Record*, 119(7), 1-37. https://doi.org/10.1177/016146811711900704

John W. Gardner Center for Youth and Their Communities. (2022). *Housing instability and educational outcomes of San Mateo County youth.* Stanford University. https://eric.ed.gov/?id=ED624057

National Center for Education Statistics. (n.d.). *School district geographic relationship files.* https://nces.ed.gov/programs/edge/Geographic/DistrictBoundaries

Singer, J., Pogodzinski, B., Lenhoff, S. W., & Cook, W. (2021). Advancing an ecological approach to chronic absenteeism: Evidence from Detroit. *Teachers College Record*, 123(4), 1-36. https://doi.org/10.1177/016146812112300406

U.S. Census Bureau. (n.d.). *American Community Survey (ACS).* https://www.census.gov/programs-surveys/acs

U.S. Department of Education. (2025). *Chronic absenteeism*. https://www.ed.gov/teaching-and-administration/supporting-students/chronic-absenteeism

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

WestEd. (n.d.). *California Healthy Kids Survey.* https://calschls.org/

Tables and Figures

Table 1

Correlations Between Collected Measures and ZIP Code Chronic Absenteeism Rate

Demographics		**Student Subgroup**		**Health & Safety**	
Variable	*r*	Variable	*r*	Variable	*r*
% African American	.312	% English Only	-.738***	% Population Fully Vaccinated	-.360
% American Indian or Alaska Native	-.053	% Initially Fluent English Proficient	-.382*	% Feeling Safe at School	-.411*
% Asian	-.515**	% English Learner	.855***	% Feeling Frequently Sad	.549**
% Filipino	-.151	% Reclassified Fluent English Proficient	.261	% Feeling Connected to School	-.450*
% Hispanic or Latino	.883***	% Long-term English Learner	.676***	% Ever Used Alcohol/Drugs	-.406*
% Pacific Islander	.436*	% Foster Youth	.423*		
% White	-.322	% Homeless	.523**		
% Two or More Races	-.800***				
Staff Characteristics		**Economic & Housing**		**School Structure**	
% Associate Degree	-.074	Day Care Center/Preschool Spending-to-Income Ratio	.107	Small/Medium Classes Ratio	-.331
% Baccalaureate Degree	.209	Home Child Care Spending-to-Income Ratio	.293	Student/Teacher Ratio	-.262
% Baccalaureate plus Extra Coursework	.111	Overcrowded Households Rate	.505**	Student/Administrator Ratio	-.056

Staff Characteristics		Economic & Housing		School Structure	
Variable	*r*	Variable	*r*	Variable	*r*
% Master's Degree	-.287	Poverty Rate	.267	Student/Pupil Services Ratio	-.028
% Master's Plus Extra Coursework	-.303	% Population Taking Public Transportation to Work	-.273	Student/Non-Instructional Support Staff Ratio	-.294
% Doctorate Degree	.160	% of Household Spending on Healthcare	-.049		
% No Degree	.134	% Households with Children Receiving SNAP	.090		
% Experienced Staff (> 2 Years Experience)	-.356	% Households without Internet Access	.148		
% of Not Fully Credentialed Teachers	.474**	% Single-Parent Households	-.011		
% of Fully Credentialed Teachers	-.592***	% Households Burdened by Housing Costs	.043		
		% Households without a Car	-.095		
		% Children Living Below Poverty Line	.208		
		% Children in Single-Parent Households	.033		

Note. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 2

Descriptive Statistics for Retained Variables Across ZIP Codes

Variables	*n*	*M*	*Median*	*SD*
Chronic Absenteeism Rate	30	.143	.140	.055
% English Learner	30	.213	.189	.101
% Homeless	30	.027	.014	.031
% Feeling Frequently Sad	30	.164	.157	.053
Overcrowded Households Rate	30	.070	.055	.061
% of Fully Credentialed Teachers	30	.926	.963	.135
% Asian	30	.151	.161	.131
% Hispanic or Latino	30	.414	.410	.170
% Two or More Races	30	.088	.089	.041

Table 3

Pearson Correlations Among Retained Variables (Pairwise)

Variable	1	2	3	4	5	6	7	8
% English Learner	—							
% Homeless	.71***	—						
% Feeling Frequently Sad	.48**	.02	—					
Overcrowded Households Rate	.52**	.14	.25	—				
% of Fully Credentialed Teachers	-.56**	-.39*	-.30	-.19	—			
% Asian	-.31	.08	-.51**	-.21	.26	—		
% Hispanic or Latino	.85***	.44*	.59***	.46*	-.45*	-.72***	—	
% Two or More Races	-.75***	-.14	-.63***	-.55**	.35	.60***	-.86***	—

Note. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 4

Principal Component Loadings for Retained Variables

Variable	PC1	PC2
% English Learner	.438	-.308
% Homeless	.217	-.680
% Feeling Frequently Sad	.330	.328
Overcrowded Households Rate	.279	.003
% of Fully Credentialed Teachers	-.282	.279
% Asian	-.310	-.460
% Hispanic or Latino	.463	.043
% Two or More Races	-.430	-.208

Note. Loadings are from the unrotated two-component solution.

Table 5

Linear Regression Predicting Chronic Absenteeism from Principal Components

Predictor	*b*	*SE*	*t*	*p*
(Intercept)	0.000	0.072	0.000	1.000
PC1	0.443	0.035	12.503	< .001
PC2	-0.111	0.060	-1.825	.079

Note. b = unstandardized coefficient; *SE* = standard error. Model summary: $F(2, 27) = 79.82$, $p < .001$, $R^2 = .86$, adjusted $R^2 = .84$.

Table 6

Indicator Coefficients Derived From the PCA-Based Regression

Variable	Coefficient
% English Learner	0.228
% Hispanic or Latino	0.201
% Homeless	0.172
Overcrowded Households Rate	0.123
% Feeling Frequently Sad	0.110
% Asian	-0.086
% of Fully Credentialed Teachers	-0.156
% Two or More Races	-0.168

Note. Coefficients were derived by combining the regression coefficients for PC1 and PC2 with the corresponding component loadings.

Table 7

Chronic Absenteeism Risk Scores by ZIP Code

ZIP Code	Chronic Absenteeism Risk Score	ZIP Code	Chronic Absenteeism Risk Score
94303	100	94019	30
94063	53	94074	30
94060	52	94018	28
94021	49	94037	28
94061	49	94062	23
94015	42	94025	20
94020	42	94005	19
94014	38	94030	10
94401	36	94044	10
94066	33	94027	8
94080	33	94010	6
94038	32	94002	4
94404	32	94028	3
94402	31	94065	0
94403	31	94070	0

Note. Scores are min-max scaled within the county sample and represent relative risk, not probabilities or percentages.

Table 8

Model-Based Indicator Contributions to Chronic Absenteeism Risk Scores by ZIP Code

	Indicator Contributions							
ZIP Code	% English Learner	% Hispanic or Latino	% Homeless	Overcrowded Households Rate	% Feeling Frequently Sad	% Asian	% Fully Credentialed Teachers	% Two or More Races
94002	-0.23	-0.30	-0.06	-0.02	-0.11	-0.14	-0.06	-0.21
94005	-0.07	-0.07	-0.13	-0.05	0.02	-0.04	-0.03	-0.10
94010	-0.26	-0.29	-0.09	-0.02	-0.10	-0.12	-0.04	-0.13
94014	0.15	-0.01	-0.08	0.09	0.20	-0.02	-0.02	0.08
94015	0.19	0.02	-0.04	0.11	0.22	-0.01	-0.02	0.12
94018	-0.09	0.10	-0.12	-0.12	0.10	0.09	-0.08	0.09
94019	-0.09	0.10	-0.12	-0.03	0.09	0.09	-0.08	0.08
94020	0.12	0.18	-0.15	-0.03	0.03	0.09	0.08	0.27
94021	0.13	0.18	-0.15	0.24	0.03	0.09	0.09	0.27
94025	-0.05	-0.07	0.00	-0.05	-0.14	0.00	-0.03	-0.07
94027	-0.18	-0.15	-0.07	-0.14	-0.18	-0.01	-0.07	-0.13
94028	-0.39	-0.30	-0.15	-0.10	-0.15	0.03	0.14	-0.25
94030	-0.08	-0.23	-0.11	-0.05	-0.11	-0.21	-0.02	-0.06
94037	-0.09	0.10	-0.12	-0.14	0.09	0.09	-0.08	0.08
94038	-0.09	0.10	-0.12	0.04	0.09	0.09	-0.08	0.08
94044	-0.31	-0.17	-0.08	-0.09	0.06	0.05	0.01	-0.33
94060	0.13	0.18	-0.15	0.39	0.03	0.09	0.09	0.27
94061	0.30	0.30	0.22	0.03	-0.05	0.06	-0.08	0.08
94062	-0.07	-0.03	-0.01	-0.09	-0.02	0.05	-0.06	-0.03
94063	0.30	0.30	0.22	0.25	-0.05	0.06	-0.08	0.08
94065	-0.28	-0.33	-0.10	-0.05	-0.11	-0.15	-0.06	-0.23
94066	0.01	0.01	0.05	0.01	0.13	-0.03	-0.04	0.05
94070	-0.29	-0.26	-0.07	-0.11	-0.14	-0.06	-0.03	-0.32
94074	-0.05	0.11	-0.12	-0.14	0.08	0.09	-0.05	0.12
94080	0.07	0.04	0.01	0.04	-0.01	-0.01	0.00	0.02
94303	0.74	0.53	0.49	0.05	0.20	0.10	0.77	0.29
94401	0.14	0.00	0.27	0.09	-0.05	-0.07	-0.05	-0.03
94402	0.12	-0.01	0.27	-0.08	-0.05	-0.07	-0.05	-0.03
94403	0.10	-0.03	0.24	0.01	-0.06	-0.07	-0.05	-0.05
94404	0.13	-0.01	0.27	-0.05	-0.05	-0.07	-0.05	-0.03

Note. Contributions are reported in standardized units. Color indicates the direction of each contribution, and darker shading indicates a larger absolute contribution.

Figure 1

Variance Explained by Principal Components

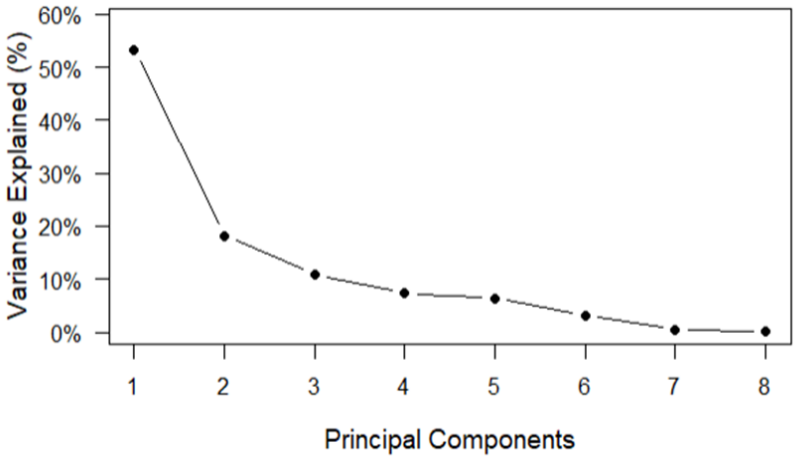


Figure 2

County Map of Chronic Absenteeism Risk Scores by ZIP Code

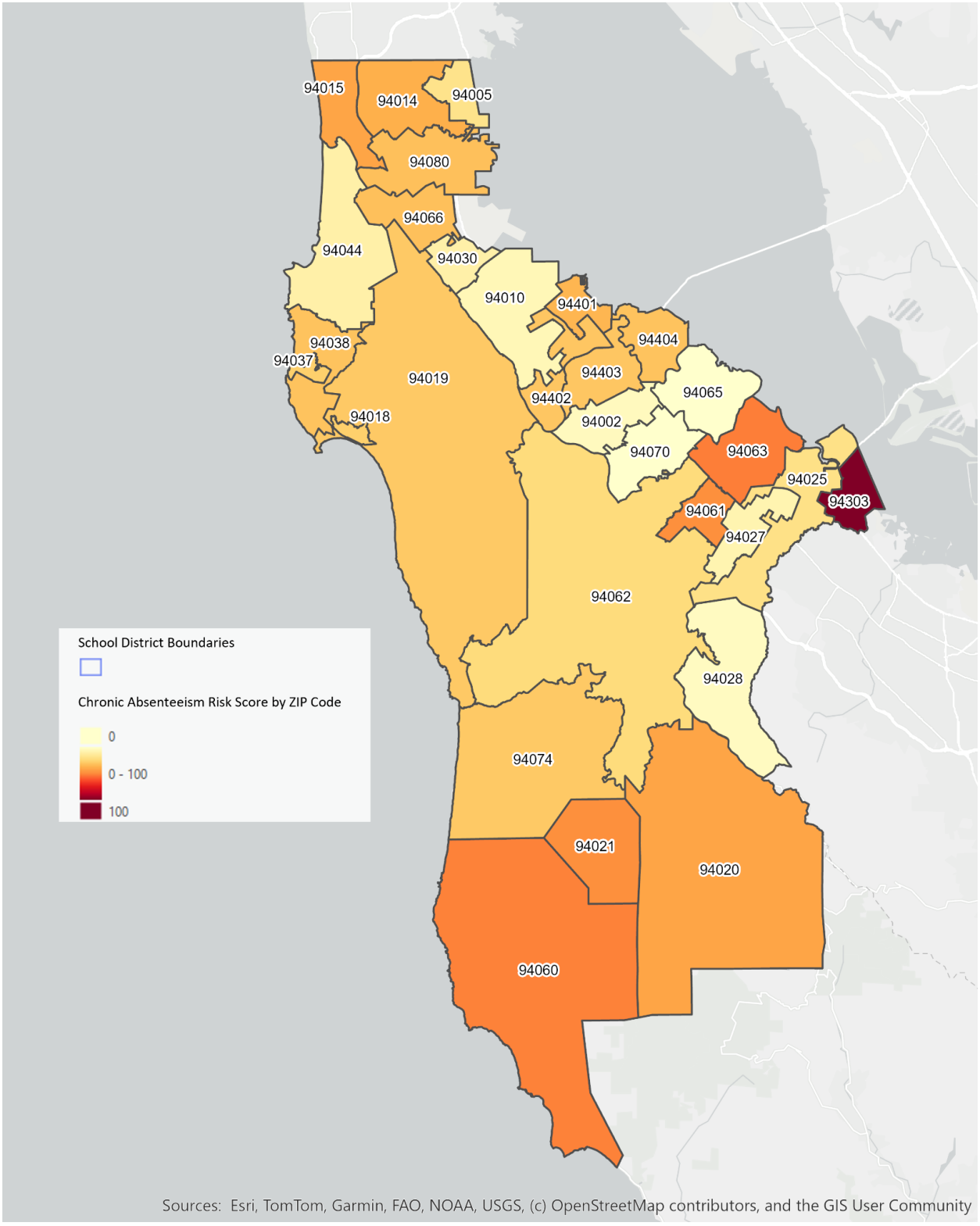


Figure 3

County Map of Chronic Absenteeism Risk Scores by ZIP Code and School District Boundaries

