

Equity in Education: Tackling Disparities in California's Schools

Michael Marutyan

ID: 8844052

Introduction

The state of California represents a microcosm of educational complexity and potential in the United States, embodying unprecedented demographic diversity, significant socioeconomic variations, and complex educational challenges. With over 6 million students spanning a vast geographical and cultural landscape, California's educational system serves as a critical case study in understanding the intricate relationships between student characteristics, systemic resources, and academic outcomes. This analysis seeks to unravel the multifaceted dimensions of California's educational ecosystem by comprehensively examining student demographics, socioeconomic conditions, institutional capacities, and academic performance metrics. By employing a holistic approach that integrates quantitative data and nuanced contextual understanding, this study aims to provide policymakers, educators, and stakeholders with a granular view of the current educational landscape—highlighting both the remarkable strengths and significant challenges that define California's commitment to equitable and high-quality public education. The research draws upon extensive state-level data, longitudinal studies, and comprehensive demographic surveys to construct a detailed portrait of how factors such as racial composition, economic background, and staff salaries intersect to shape students' educational experiences and trajectories in one of the nation's most demographically complex and economically significant states.

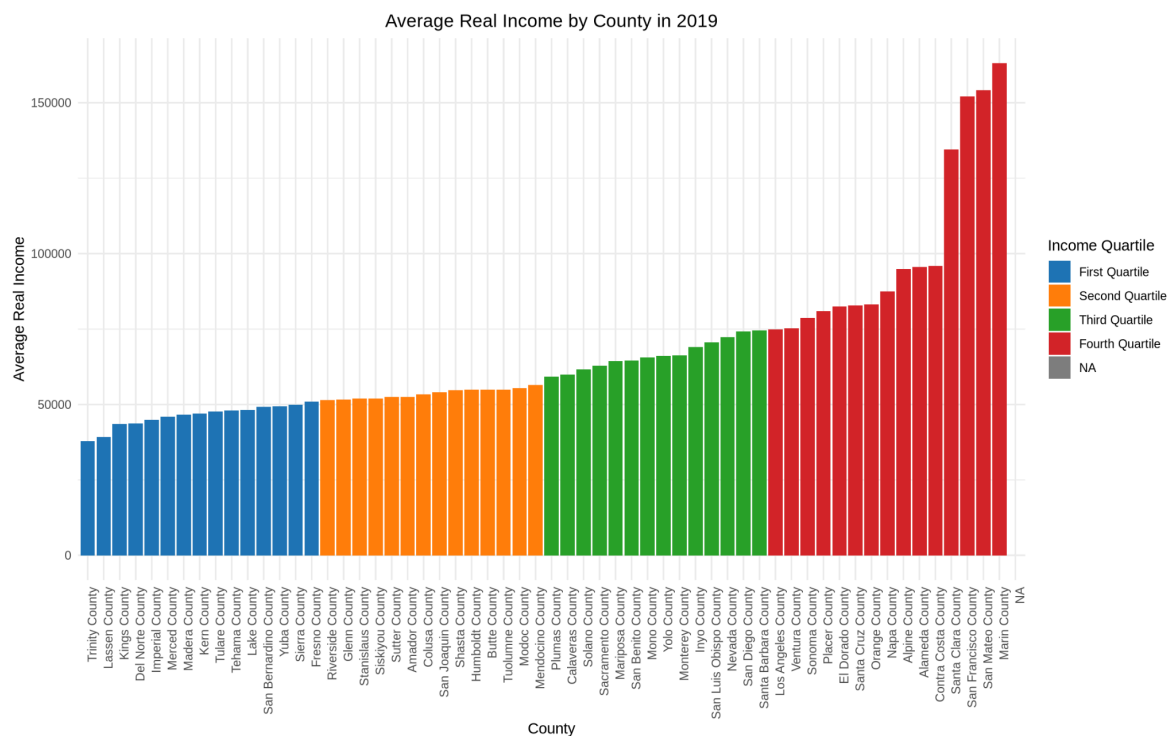
Data Description

The original datasets include school, district, county income, teacher salary, and consumer price index (CPI) data, along with an additional dataset on graduation rates. The school dataset provides details such as total student enrollment, eligibility for free or reduced lunch, and demographic breakdowns by ethnicity, while the district data includes information about the number of teachers, total students, and individualized education program (IEP) enrollment. The county income data captures income levels, population, and geographic identifiers specific to California counties, while the teacher salary dataset outlines employer details, annual salaries, and related variables. The CPI data standardizes economic measurements across years by scaling inflation-adjusted figures to 2023. Graduation data, from `graduationData.csv`, includes high school completion rates and other relevant metrics at the school and district levels, aggregated by year and location. To be able to combine all of the data, the function “mutate” was used to combine several columns to create school IDs to match with the same ID numbers within the `schoolData` table. To ensure data quality, several cleaning functions were applied throughout the process. The “`clean_names`” function standardized column names into “snake_case” for consistency and easier handling. Custom functions like “`cleanValues`” and “`name_cleaner`” were

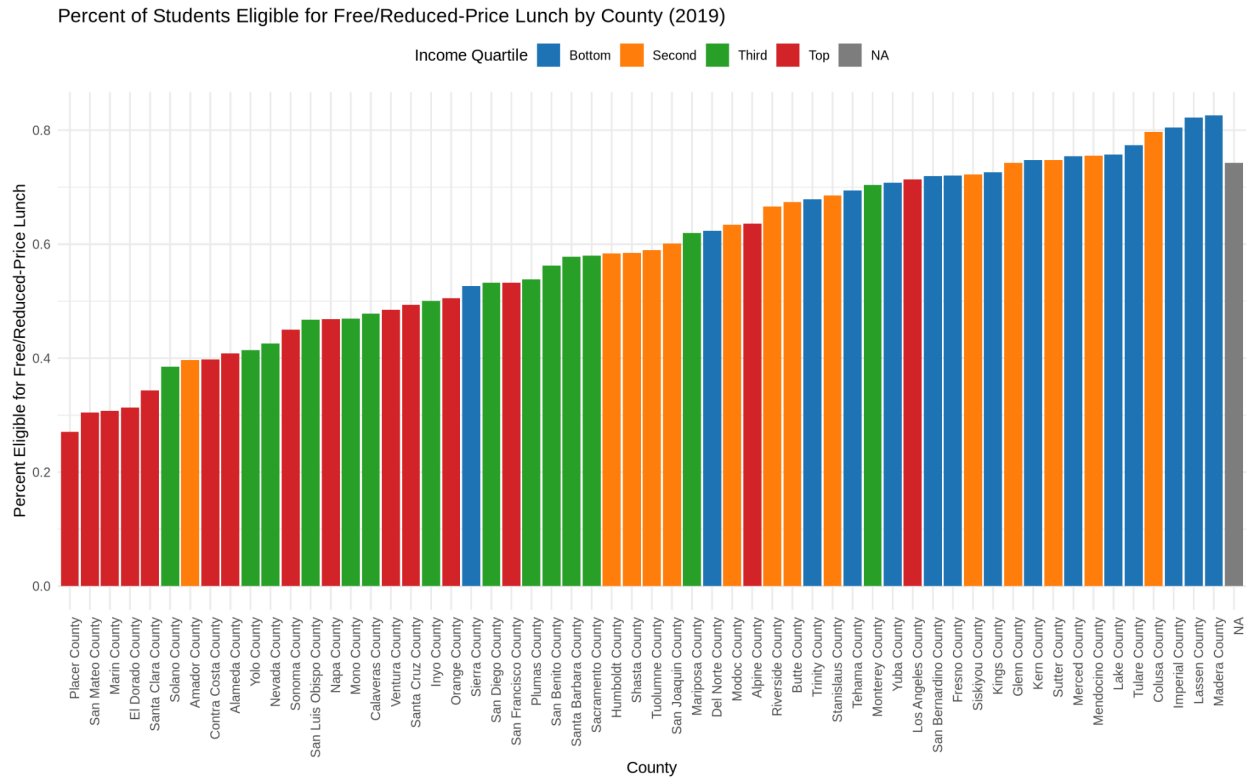
used to remove punctuation, convert numeric data, and normalize text fields, such as school and district names. Missing or invalid key identifiers were filtered out using functions like “filter” and “mutate” to avoid incomplete data impacting the analysis. Data was summarized at higher levels of aggregation, such as districts or counties, using the weighted mean formula to represent key metrics accurately. The weighted mean formula calculates the average by multiplying each value by its corresponding weight (e.g., total students in a school) and dividing by the sum of the weights, ensuring larger populations contribute proportionally more to the final result. This approach was particularly important for percentages of student demographics and eligibility for free or reduced lunch. Missing values were addressed by imputing adjacent years' data or using comparable records where applicable, and completely missing records were removed when critical variables were unavailable. These steps ensured a reliable, consistent dataset for analyzing trends in school demographics, teacher salaries, economic data, and graduation outcomes.

Free and Reduced-Price Lunch Analysis

The first figure, "Average Real Income by County in 2019", highlights the significant disparities in income levels across counties, grouped by income quartiles. Counties such as Marin and Santa Clara, in the fourth quartile, demonstrate considerably higher real incomes compared to counties like Trinity and Lassen in the first quartile. This data is crucial as income levels often correlate with access to resources and educational opportunities. Lower-income counties are likely to have higher eligibility for free and reduced-price lunch programs, making them key areas for intervention.

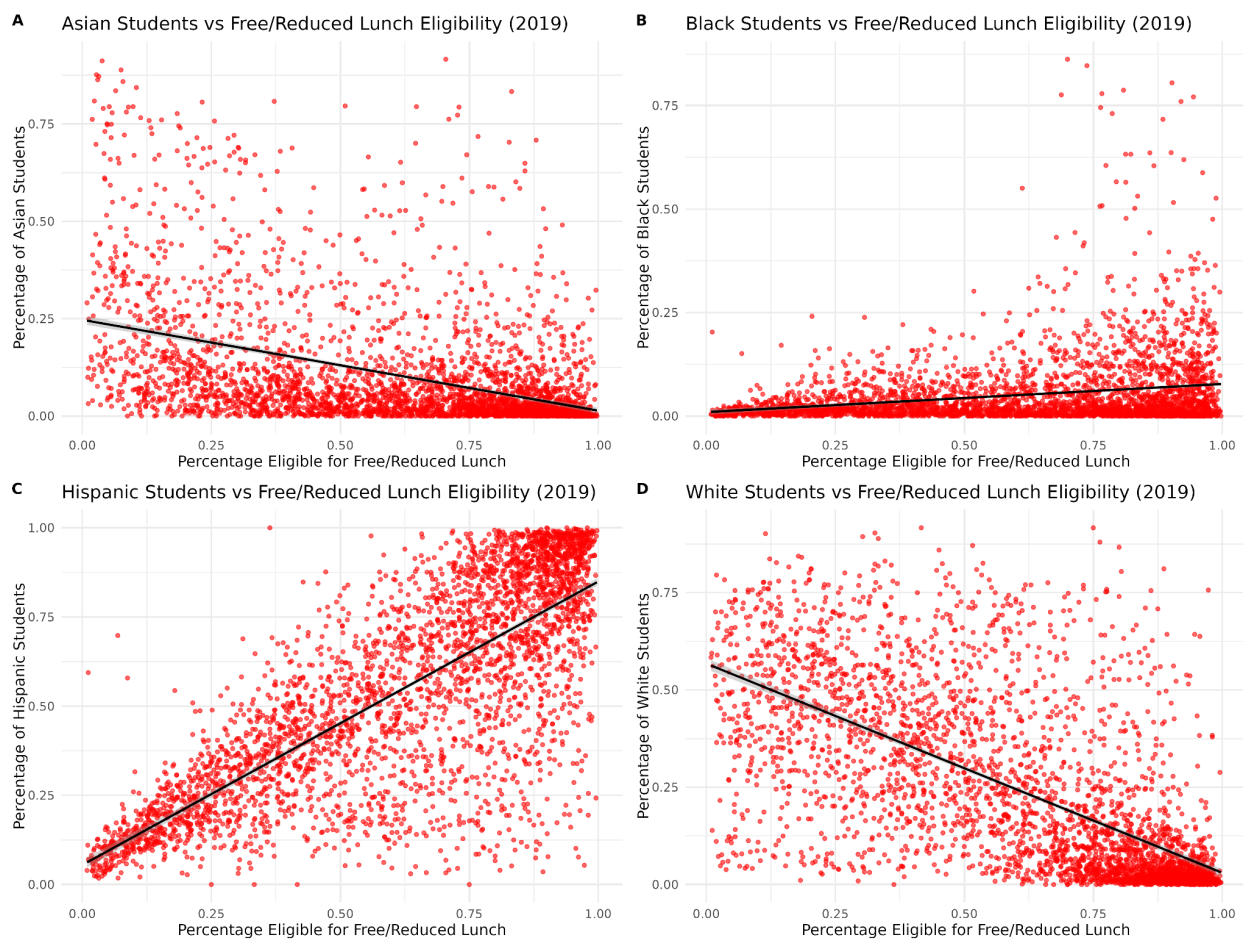


The second figure, "Percent of Students Eligible for Free/Reduced-Price Lunch by County in 2019", further emphasizes this disparity. Counties in the bottom income quartile, such as Madera and Lassen, show significantly higher percentages of students eligible for these programs, often exceeding 70%. These counties are critical targets for the Department of Education to allocate additional funding and support, as the high eligibility rates suggest a concentration of low-income families, which can negatively impact student outcomes, such as graduation rates and standardized test performance.



To deepen the analysis, a scatter plot was created visualizing the relationship between eligibility for free/reduced-price lunch and demographic variables (e.g., Hispanic and Black student percentages). The scatter plot shows a clear positive trend between the percentage of students eligible for free/reduced lunch and the representation of historically underserved student populations. This highlights the overlap between low-income status and minority representation,

suggesting that targeted programs should account for these intersecting vulnerabilities.



Finally, a trend analysis over time, visualized through the evolution of free and reduced-price lunch eligibility by locale (e.g., city vs. rural areas), reveals the impact of programs like the Community Eligibility Option (CEO). Starting in 2014, the CEO enabled schools with high percentages of low-income students to provide free meals universally, which significantly reduced administrative barriers. Rural areas saw the most pronounced uptake of this program by 2022, with over 40% of schools offering free meals under the CEO, indicating a broader reliance on such initiatives in regions with fewer resources.

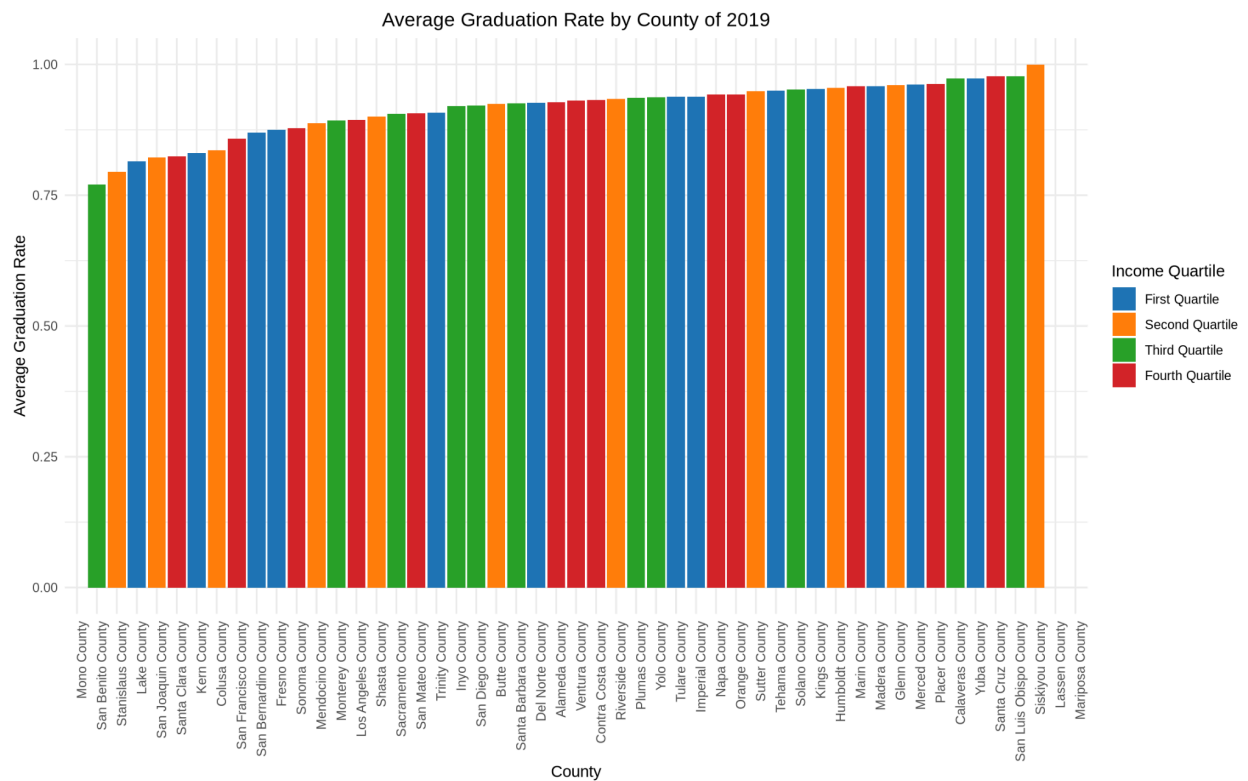
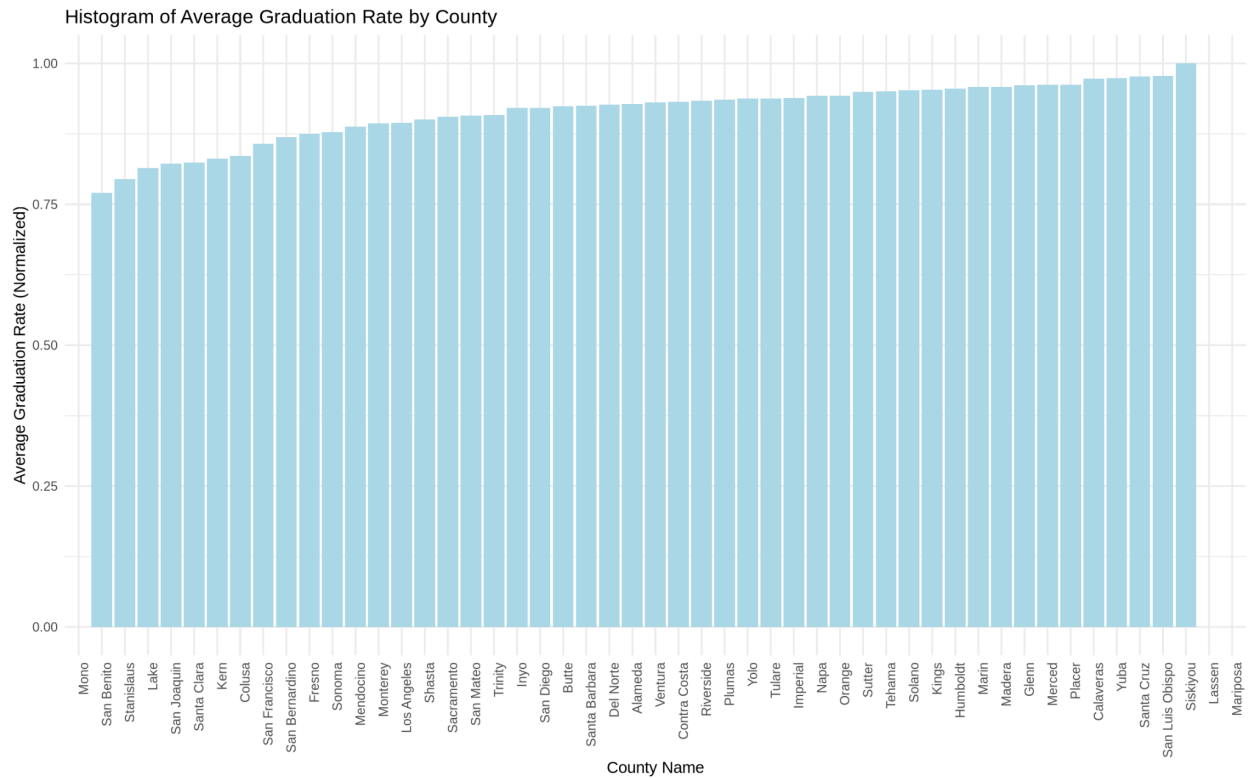
City				
year	No	Yes with CEO	Yes without CEO	
2013	72.9%	0.0%	27.1%	
2014	1.7%	4.1%	94.2%	
2015	73.2%	14.2%	12.6%	
2016	69.8%	20.6%	9.6%	
2017	67.7%	25.8%	6.6%	
2018	63.9%	31.6%	4.6%	
2019	60.2%	36.7%	3.0%	
2022	0.6%	64.1%	35.4%	

Rural				
year	No	Yes with CEO	Yes without CEO	
2013	87.2%	0.0%	12.8%	
2014	9.0%	2.0%	89.0%	
2015	85.4%	4.8%	9.9%	
2016	82.3%	7.5%	10.2%	
2017	84.8%	9.9%	5.3%	
2018	71.8%	23.4%	4.7%	
2019	66.6%	28.4%	5.0%	
2022	0.3%	42.4%	57.4%	

It is recommended that the counties within the first quartile should be targeted with more free and reduced lunch policies, such as Trinity County and Del Norte County, as their income is within the lowest 25% and include the highest number of underserved minority groups and rural communities as they have lower percentages of CEOs being present compared to urban city environments.

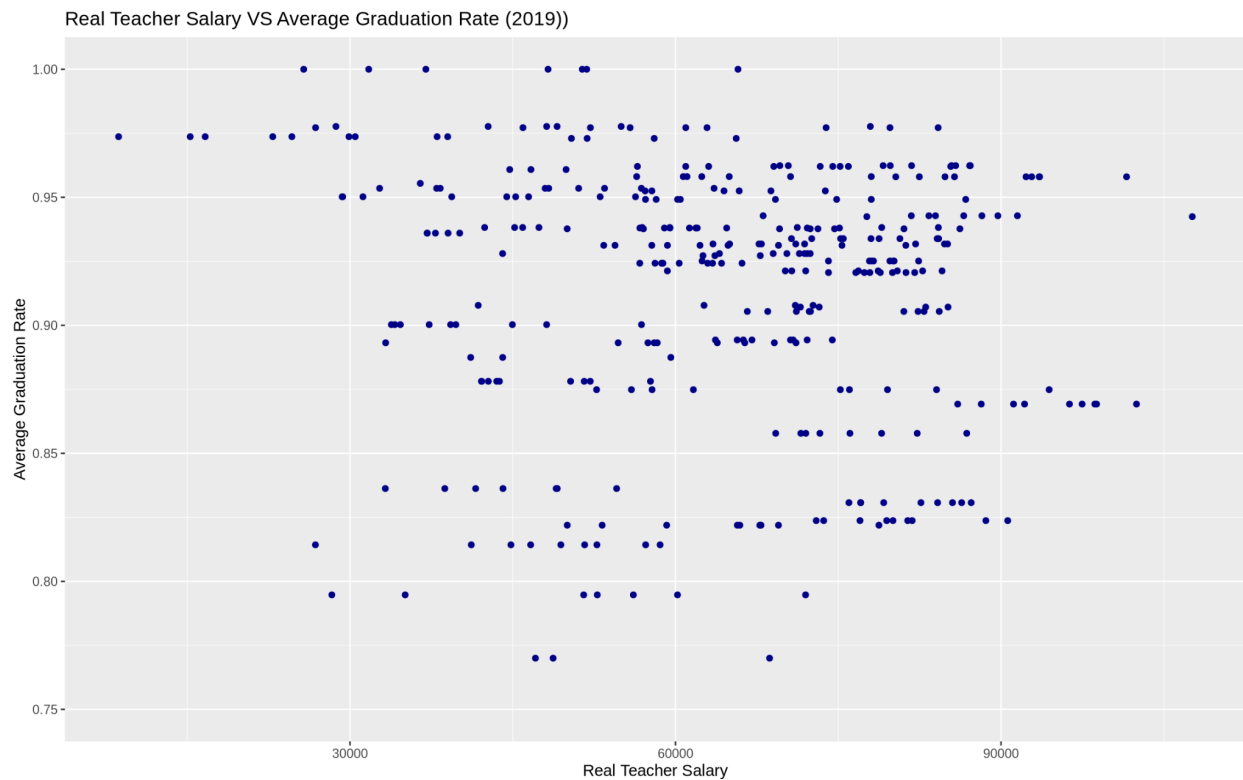
Student Outcomes Analysis

Using the graduationData.csv, which provides detailed high school graduation rates at the school level, we analyzed how graduation outcomes vary across counties, income levels, and other influencing factors. Our findings reveal significant disparities in graduation rates across counties, ranging from above 95% in wealthier counties like Santa Clara and Marin to closer to 80% in economically disadvantaged areas like Trinity and Lassen. The data indicates a strong association between socioeconomic status and educational outcomes, with counties in the bottom income quartiles exhibiting lower average graduation rates. The second histogram shows that the spread of counties is relatively spread out, but the majority of counties with their income within the first and second quartile on the lower rate of graduation while the counties within the third and fourth quartiles are on the higher rate of graduation, though it isn't a very strong relationship between county income averages and graduation rates.



Additionally, a positive correlation between teacher salaries and graduation rates emerged, as demonstrated by a scatter plot showing that counties with higher real teacher salaries tend to achieve better graduation outcomes. For instance, regions with higher compensation for teachers,

often found in the fourth income quartile, consistently performed better in terms of student graduation. Although the scatterplot's points are highly varied, there is a concentrated area along the top right corner. This suggests that adequate teacher compensation is a key factor in enhancing educational success. Over time, the disparities between wealthier and lower-income counties have persisted, underscoring the systemic challenges faced by underfunded regions in addressing educational inequities.



To address these issues, it is recommended increasing funding for teacher salaries in counties within the first and second income quartiles, as this could help attract and retain qualified educators who are critical to student success. Additionally, expanding programs such as free and reduced-price lunch initiatives and after-school tutoring could mitigate socioeconomic barriers that impede academic achievement. By targeting funding and resources to counties with both low incomes and low graduation rates, the Department of Education can improve equity and overall outcomes for students in these areas. Systematic tracking of graduation rates and evaluating policy impacts will also be critical in ensuring sustained improvements in educational equity. With an addition to increases in teacher salaries, schools should open more after school academic opportunities and tutoring services to help increase students' grades and increase overall graduation rates with their new increased funding.

Conclusion

California's education system serves as a vivid example of the profound intersection between socioeconomic factors, resource allocation, and student outcomes. The analysis highlights

significant disparities in access to resources and educational achievements across the state's diverse counties. Wealthier counties like Santa Clara and Marin consistently exhibit higher graduation rates and lower eligibility for free or reduced-price lunch programs, reflecting their access to greater economic and institutional resources. Conversely, economically disadvantaged counties such as Trinity and Lassen face systemic barriers that result in lower graduation rates and higher rates of lunch program eligibility, underscoring the critical need for targeted intervention. The correlation between higher teacher salaries and improved graduation rates further emphasizes the importance of adequate compensation in fostering student success. To combat these inequities, it is essential to prioritize funding for low-income counties by increasing teacher salaries, expanding access to universal meal programs, and investing in after-school and tutoring initiatives. These measures can address socioeconomic disparities and provide underserved communities with the tools they need to succeed. Furthermore, tracking the long-term impact of these policies will ensure accountability and sustained progress toward educational equity. California's commitment to addressing these challenges will not only improve student outcomes but also serve as a blueprint for addressing educational inequities nationwide.