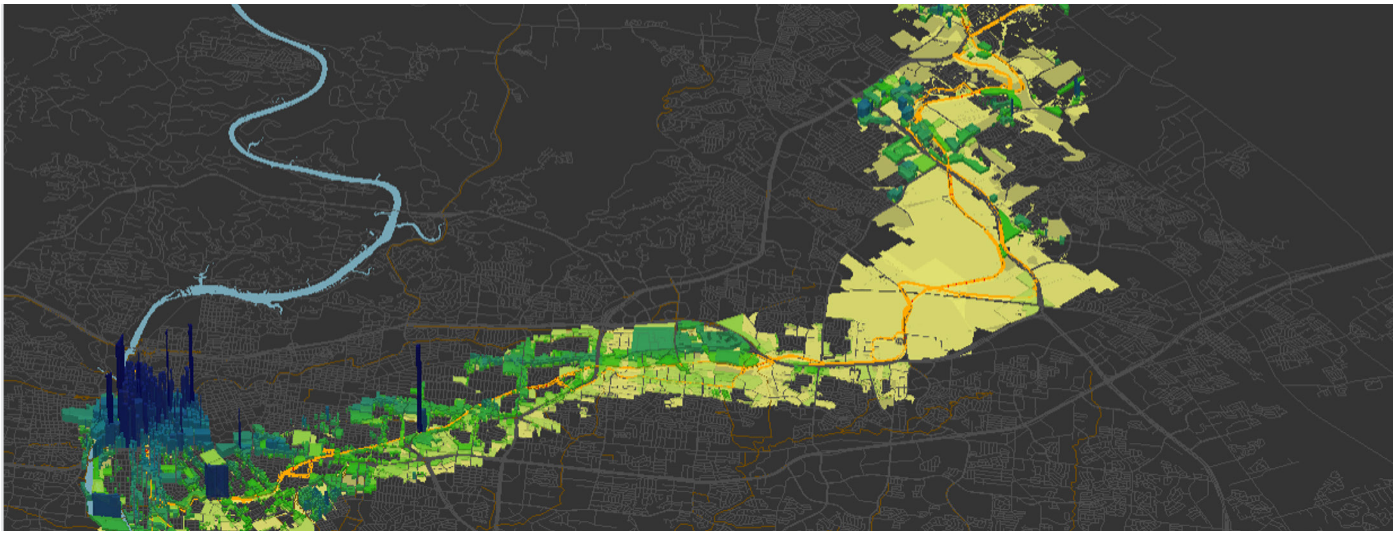
Parkhill

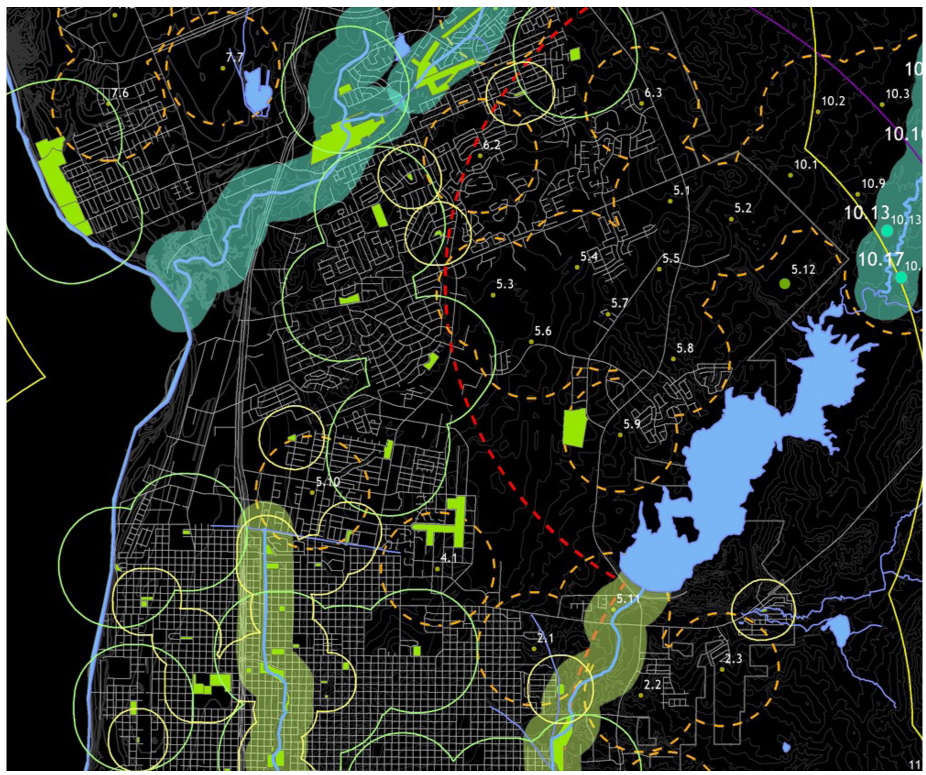
School Performance Improvement: Planning with Data & Algorithms– A Case Study

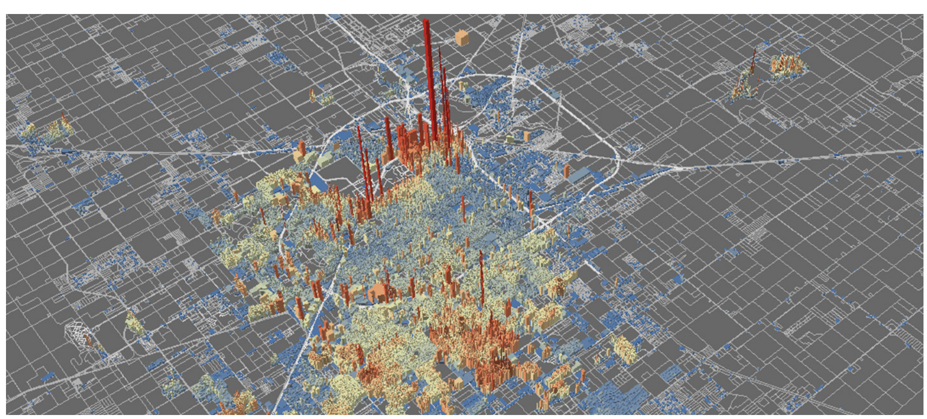
JULIO CARRILLO - OCTOBER 2022



Artificial Intelligence (AI), Algorithms, and Geospatial Analytics are terms that sound technologically charged, with a great sense of complexity behind them. Introducing successful technologies into the world creates rapid development, with an organic interest of corporations adopting them in search of their benefits, resulting in driving the evolution and path of future applications of these technologies. Rightfully so, applications of these technologies are adopted at a slower pace in the public sector and more often in small towns and communities with a higher rate of economically disadvantaged populations.

In the spirit of Building Community, this article explores data from school districts where academic success is challenging, shining lights on solutions that touch beyond the educational focus. The analytic approach of this study exposes the immense opportunity for public sector applications of these technologies, but also the integration of School Districts as key stakeholders in Community Planning.





SCHOOL DISTRICTS | GEOSPATIAL ASSESMENTS

In Texas, the Texas Education Agency (TEA) defines academic success. School districts and campuses get a score for academic success and performance based on various factors such as test results, graduation rates, etc. All these factors are currently accounted for while defining success, but often relegated, there is the vital link to the built environment and socioeconomic aspects that could impact these success scores.

School districts can have a portion of their population living where critical systems, such as internet connectivity, water/ sewage access, or other aspects linked to infrastructure service zones, have been unintentionally neglected, leading to a burden on students. Due to the increased number of barriers, such as access to internet connectivity, school districts will find it difficult to perform successfully by TEA standards, even when those factors are not in the scoring matrix.
Depending on the urban dissemination of the student population across the city or community, socially and economically disadvantaged districts are more burdened to succeed in academic performance.

A geospatial equity assessment is an intriguing procedure that can help visualize where a school district stands in the performance of socio-economic indicators based on place/location indicators. It can potentially open a uniform standard for comparison with nearby peers. Still, ultimately, it is a way to observe how the district's population falls in terms of disadvantage burdens (either economically or physically).

These assessments might not create solutions or reaffirm the conditions already known by the district. But, if aspects link to the physical environment, the evaluation might highlight potential interventions that could improve academic success and reduce the implicit burdens.

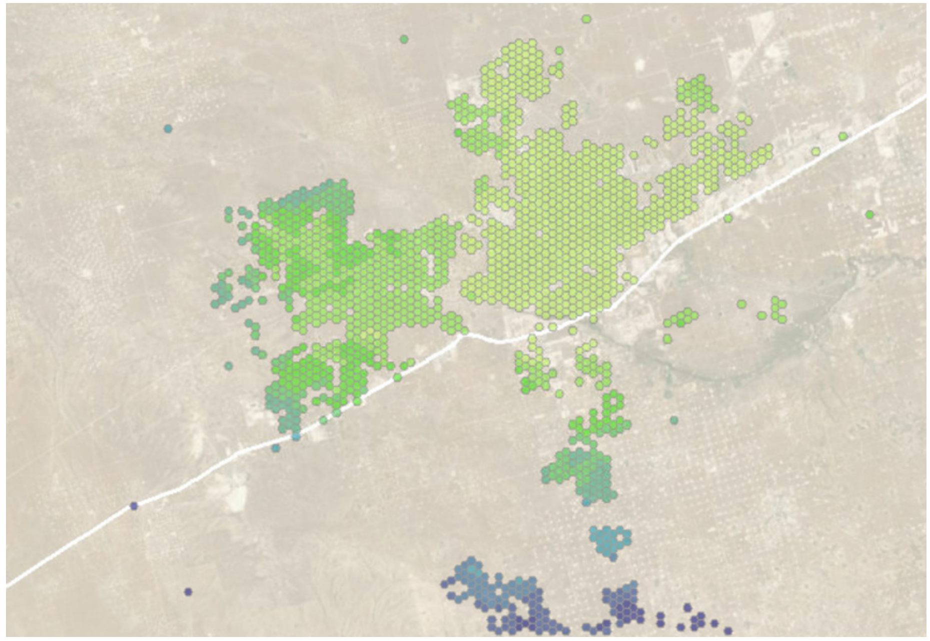


FIGURE 1 | ASSESSMENT OF SPATIAL-BURDENED POPULATION FOR ECTOR COUNTY SCHOOL DISTRICT. HEXAGONS IN DARK BLUE ARE MORE IMPACTED IN A COMBINATION OF SERVICE ACCESS TO POTABLE WATER AND/OR SEWAGE; BUT ALSO FACTORING PROXIMITY TO SCHOOLS, PARKS, NOISE, URBAN HEAT, AND FLOODPLAINS.

Incorporating spatial analysis into a process that does not include these in a traditional approach helps define what School Districts can advocate for in their communities for a successful future.

BUILDING COMMUNITY WITH TECHNOLOGY: ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

In a technologically advanced world, education and community planning should skillfully connect to achieve common goals and effectively improve performance. The most efficient way to combine both "sectors" is through clever use of data.

The example showcased in this article brings to light the connection between the physical environment and the socio-economic performance of a school district's population. In this example, Parkhill's Planning+Analytics team explored algorithms such as random forests and k-means clustering to reaffirm that conditions in the physical environment in a community link to the socio-economic burden of students in the district. This exploration also highlights that geospatial factors, i.e., the physical environment, are a valuable way to predict the socio-economic limitations of the district population.

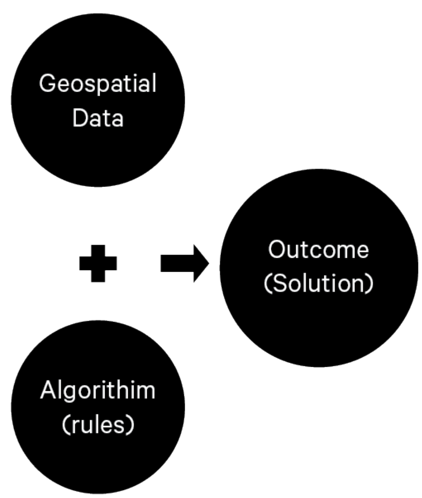


FIGURE 2 | DIAGRAM OF A "TRAINING MODEL" IN THE ARTIFICIAL INTELLIGENCE / MACHINE LEARNING REALM

ALGORITHMS FOR BUILDING COMMUNITY

In Artificial Intelligence, algorithms create predictive models to help you make decisions once inputting new data. In Community Planning, this process is very similar to suitability analyses or assessments.

It is typical in geospatial analysis to define layers for suitability or an assessment process. These processes in geospatial analysis help define areas most "suitable" for specific solutions or help assess regions with high concentrations of particular aspects.

We use "supervised algorithms" to define the outcome zones in the geospatial analysis. As an example, the supervised algorithm used to generate the heatmap of Figure 1 (assessment) exemplifies the "sum of burdens" of the geospatial (physical environment) layers, such as proximity to flood plains, parks, schools, urban heat zones, or noise zones. This algorithm generated an outcome that outlined that the greater the combination of burden layers, the darker the blue for each hexagon.

One can get the same output for assessing burdens by using an "unsupervised algorithm". A clustering algorithm (k-means) or a classification algorithm (random forests) can also help define these outcome zones.

During testing, both supervised and unsupervised algorithms showed that socio-economic and location-based factors have similar high burdens.

In this specific study (figures 1 and 3), Ector County's student population areas were assigned a cluster id (label) or an impact level (number). From one edge of the hexagon to the other represents about a 5-min walk. Each hexagon received these identifiers as an indicator of a prediction.

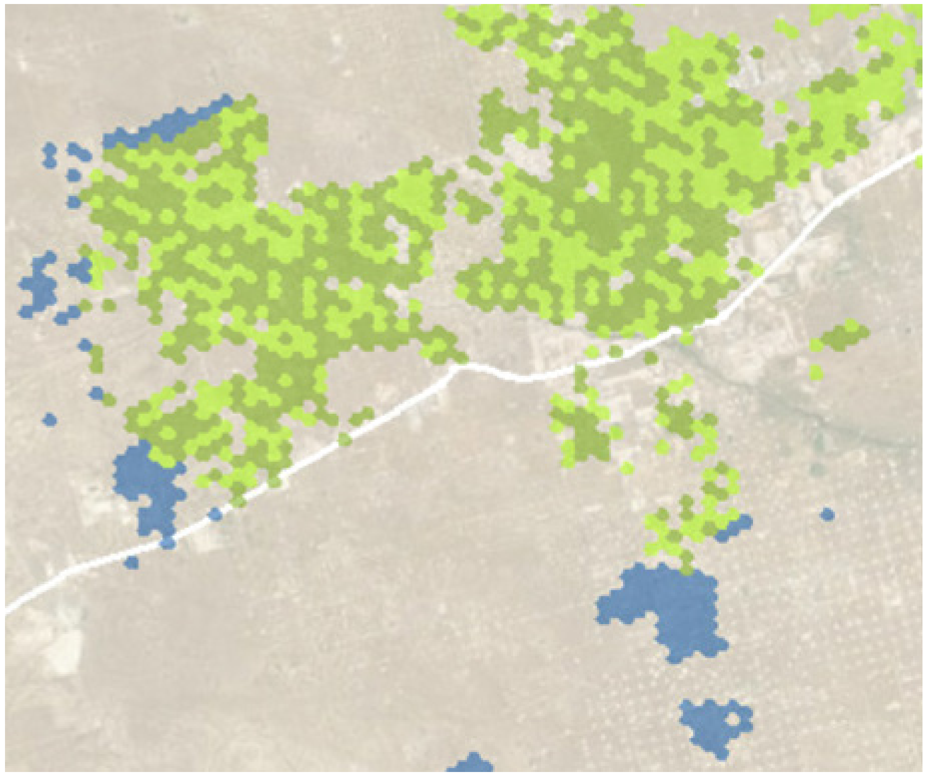


FIGURE 3 | ASSESSMENT OF SPATIAL BURDENED POPULATION IN ECTOR COUNTY SCHOOL DISTRICT, DEFINED IN 3 CLUSTERS BY THE UNSUPERVISED K-MEANS ALGORITHM.

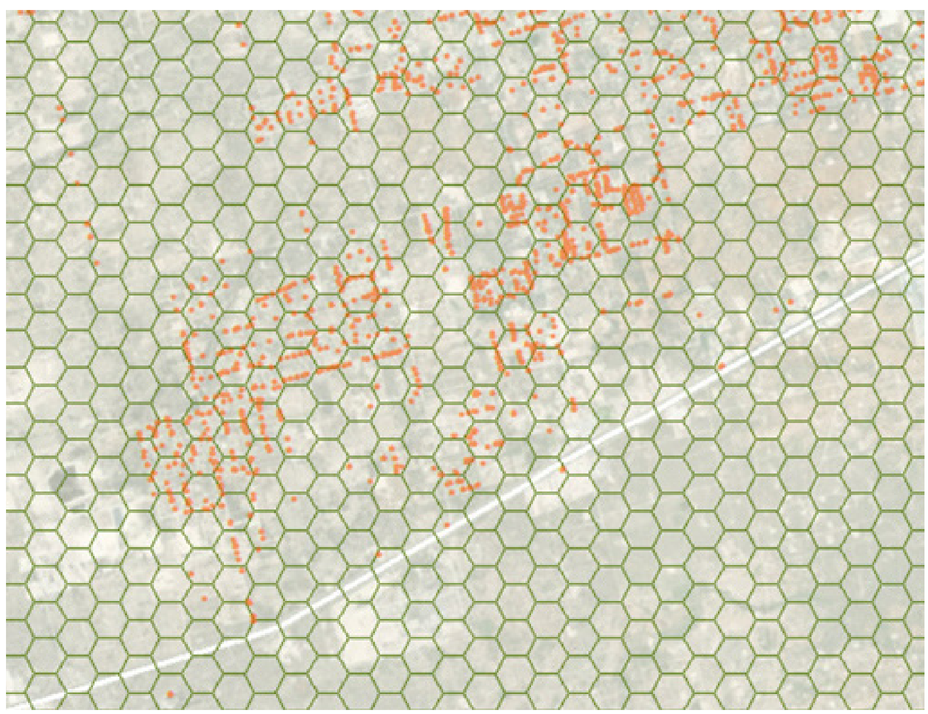


FIGURE 4 | STUDENT POPULATION (ORANGE DOTS) WITHIN HEXAGON AREAS USED FOR GEOSPATIAL ANALYSIS.

With these classifications assigned, each hexagon area helps predict the socio-economic burden of the student population. A random forest algorithm processed statistics based on geospatial data to classify the student population into ECODIS (economically disadvantaged) or Not-ECODIS, based on their physical geolocation.

> *"The resulting prediction matched the School District's classification with an accuracy of 73.5%, meaning that the physical environment factors alone could define an ECODIS characteristic with almost 75% accuracy."*

It is common practice within school districts to seek self-reporting of socio-economic status by the families. Self-reporting is very helpful in defining the ECODIS categories of the student population, but in later years, this has been increasingly challenging due to the reduction of reporting ratios. This algorithm can help school districts target areas where they prioritize self-reporting and enrolling.

DATA-DRIVEN VS DATA-INFORMED SOLUTIONS

When building a community, data analytics and geospatial analysis are vital for processing the essential pieces of information within various systems in a specific area. Artificial intelligence and machine learning methods help make future predictions in similar studies.

In this specific community, the clustering exercise helped define which portion of the student population should receive relief or targeted support to eliminate the socio-economic burdens and barriers.

In the Artificial Intelligence realm, most solutions amplify systemic biases and barriers. Still, the approach aims to target solutions for this specific example's burdened sectors or clusters.

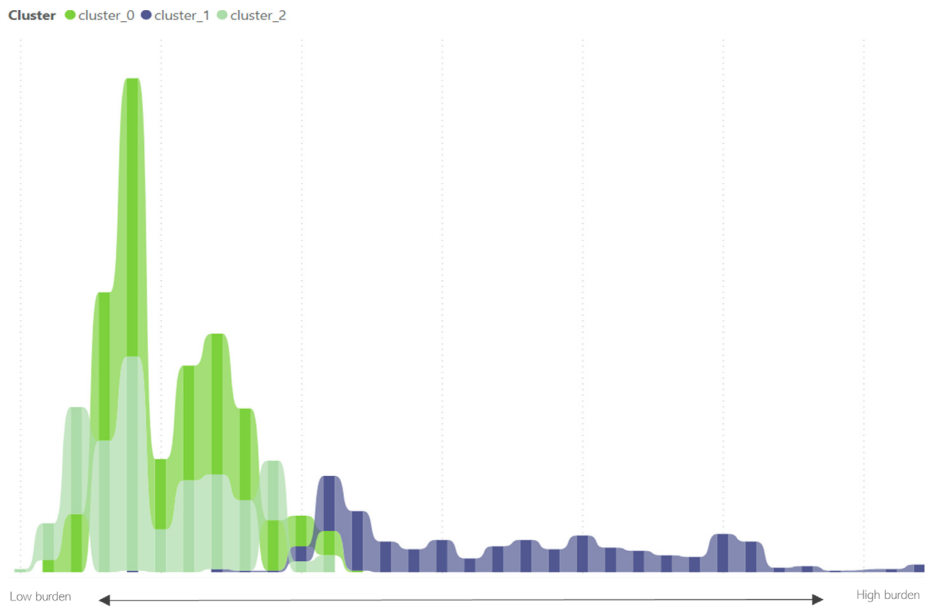


FIGURE 5 | DISTRIBUTION OF STUDENT POPULATION IN A GEOSPATIAL BURDENED SCALE AND CLASSIFIED BY K-MEANS CLUSTERS. THIS CHART SHOWS THE SAME CLUSTERS HIGHLIGHTED IN FIG3 (MAP) BUT CLARIFIES THE AMOUNT OF THE POPULATION THAT IS CURRENTLY BURDENED.

In our study, Cluster_1 (blue) represents only 7.29% of the total student population, equaling 2,389 students. Around 60% (1,416) of this cluster are classified as ECODIS.

Although the most burdened population (blue cluster) is not the majority of the total ECODIS population, this geospatial assessment with a clustering algorithm allowed us to identify those areas within the community that needs a higher level of support. These areas are also highlighted in blue color in Figure 3.

FUTURE STEPS & RECOMMENDATIONS | PLANNING2050 INITIATIVE

Geospatial Assessments might only reaffirm pre-known conditions by districts, cities, and communities but, at the same time, bring to light the opportunity to implement targeted strategies for specific community sectors when the support needed is beyond an academic focus.

These assessments also allow the definition of a metric that needs to be improved. Sometimes this metric results from a mix of actions performed by the school district and the city/community in those specific jurisdictions.
After an assessment, it is time for school districts (in collaboration with the city/community) to perform two additional steps which can be beneficial in a search for long-term success:

- Commitments – This step will help define a threshold (metric) of improvement with a specific target timeline. Example: "By 2050, Cluster 1 population shall ___ to improve its access to ___"
- Database –Ongoing monitoring needs reports from actions performed to achieve the commitment goals. This continuous monitoring will help track progress and confirm activities are on track to achieve the commitment goals.

The framework described above is currently under development as a national initiative called "PLANNING 2050", with a focus on a data-driven

& data-informed search for solutions and is presently supported by many professional institutions representing professions transforming the natural and built environments.

With Planning 2050, Cities and Communities will ensure critical systems are equitable, resilient, and carbon neutral by 2050.

CONCLUSIONS & OPPORTUNITIES

Socio-economically disadvantaged communities are less likely to utilize AI technologies to improve their performance. Knowledge and Capabilities for using this type of technology are often not within reach of this type of community.

Drafted by a clear goal and commitment, small towns and historically underserved communities can apply for funding to build capacity within these communities. A city or community auditioning for PLANNING2050 will open the opportunity to intentionally create a long-term goal by implementing strategies funded by the Infrastructure Investment and Jobs Act (2021) and the Inflation Reduction Act (2022).

Geospatial Assessments with the support of Artificial Intelligence algorithms are a clever approach to predicting academic barriers and socio-economic burdens in a school district population. The geospatial aspects of a community have a solid link to their outcome in academic performance.

Even if technologies like AI can be powerful and criticized for helping increase systematic inequalities, there is hope in using these technologies to improve communities. We truly embrace a culture of purposeful improvement for economically burdened sectors of our cities and communities when technology enhances a community's outputs or performance.

Building | Community

Parkhill

Parkhill.com