Abrar, Sehr

Mentor Me Collective

Advanced Data Analytics

November 20, 2025

Data-Driven Solutions to Address Poverty in New York City

Introduction

New York City is a city of stark contrasts, where affluent neighborhoods exist alongside areas experiencing concentrated poverty. Our data analysis revealed that child poverty rates vary widely across boroughs, income distribution is highly uneven, and reliance on social assistance programs correlates strongly with specific socioeconomic and demographic factors. These findings underscore the urgent need for targeted, evidence-based interventions. By leveraging predictive modeling, exploratory data analysis (EDA), and historical datasets from sources such as the NYC Open Data Portal, we can identify high-risk neighborhoods and design solutions that efficiently reduce poverty while improving residents' quality of life.

1. Identifying High-Risk Neighborhoods

Using features like income brackets, racial and ethnic composition, employment types, and commute patterns, our predictive models—including Random Forest, Gradient Boosting, and HistGradientBoosting—were able to classify neighborhoods into high, medium, or low poverty risk.

For example, the Random Forest model achieved an R^2 of 0.90, demonstrating strong predictive accuracy in identifying areas with high poverty concentrations. EDA visualizations showed that neighborhoods with high proportions of residents earning less than $60,000 annually, combined with higher unemployment or reliance on public transportation, were consistently flagged as higher-risk.

By classifying neighborhoods, policymakers can allocate resources more efficiently: high-risk areas can receive targeted interventions such as funding, food assistance, and social services, while medium-risk neighborhoods can be monitored proactively. Historical trends also allow early identification of neighborhoods where poverty levels are rising, preventing crises before they worsen.

2. Targeted Food Assistance Programs

Analysis of poverty-related metrics alongside food assistance data revealed clear patterns: spikes in demand often occurred in neighborhoods with high child poverty and low median incomes. Predictive modeling can forecast future demand for food support programs, enabling more proactive and efficient allocation of resources.

For example, neighborhoods identified as high-demand by the model could benefit from mobile distribution units, extended pantry hours, or increased funding for soup kitchens. Partnerships with nonprofits and community organizations, informed by the EDA findings, can help reduce distribution bottlenecks, improve outreach, and ensure that aid reaches the residents who need it most.

3. Addressing Root Causes of Poverty

Poverty is multifaceted, requiring long-term solutions beyond immediate relief. Our data analysis highlighted several key drivers: For income support: neighborhoods with high proportions of residents in low-income brackets (<$60,000) corresponded strongly with higher overall poverty rates. Solutions such as wage subsidies, direct financial assistance, and job training programs can alleviate financial stress and improve economic stability. For education and youth support: data on child poverty revealed that certain boroughs have higher rates of youth at risk. Programs such as early childhood education, after-school initiatives, and mentorship can address intergenerational poverty. For employment Initiatives: employment patterns in our dataset show that neighborhoods with higher unemployment or reliance on low-wage service jobs are more prone to poverty. Targeted skill-building programs, professional development, and job placement services can increase earning potential and stability. These strategies, grounded in the data, aim to empower residents while systematically reducing neighborhood-level poverty.

4. Using Data to Inform Policy

EDA and predictive modeling results, visualized through choropleth maps, feature importance graphs, and predicted vs. actual poverty scatter plots, can guide policymakers and NGOs in making data-driven decisions. For example: Resource allocation for affordable housing and transportation subsidies can be prioritized in neighborhoods identified as high-risk. Feature importance rankings from models (e.g., income brackets, employment type, commute patterns) highlight which socioeconomic factors have the greatest influence on poverty, informing targeted policy interventions. Grounding decisions in quantitative analysis ensures that interventions are equitable, efficient, and effective.

5. Continuous Monitoring and Feedback

The research emphasizes the importance of maintaining an adaptive and iterative approach to poverty interventions. Predictive models should be regularly updated as new census data and food assistance records become available, ensuring that forecasts remain accurate and relevant. The effectiveness of implemented programs can be assessed using key metrics, such as reductions in poverty rates, decreases in spikes of food assistance demand, and improvements in neighborhood employment trends. In addition to quantitative monitoring, community feedback is critical: engaging local residents and organizations provides qualitative insights that can refine strategies and ensure interventions address real needs. By combining model-driven analysis with ongoing evaluation and stakeholder input, the city can continuously adjust and improve its efforts to combat poverty.

6. Future Opportunities

Looking ahead, data-driven poverty reduction efforts can be expanded by integrating additional datasets, including healthcare access, housing quality, and utilization of social services. These enriched models could uncover previously unrecognized risk factors and allow for more precise targeting of interventions. Real-time monitoring tools, such as public dashboards or mobile applications, could provide transparency and empower communities to track the impact of policies. Furthermore, collaborations with academic institutions and nonprofit organizations could facilitate longitudinal studies that evaluate the long-term effectiveness of interventions. Incorporating advanced machine learning techniques and broader datasets ensures that poverty-reduction strategies remain dynamic, adaptive, and grounded in evidence.

Conclusion

New York City's poverty crisis is complex, but it is not intractable. By combining data-driven insights, predictive modeling, and targeted interventions, policymakers and organizations can make smarter, faster, and more effective decisions. Equity should remain central to these efforts, ensuring that the neighborhoods and populations most affected by poverty receive the resources and support they need.

Citations

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. National Bureau of Economic Research. https://www.nber.org/papers/w25591

Kaggle. (n.d.). Exploratory Data Analysis techniques. Retrieved from https://www.kaggle.com/discussions/getting-started/42965

Machine Learning for Social Good. (n.d.). Retrieved from https://www.ml4good.org/

Nord, M., & Romig, K. (2006). Evaluating food assistance programs in urban communities. Journal of Policy Analysis and Management, 25(2), 275–294. https://www.jstor.org/stable/26347408

NYC Open Data Portal. (n.d.). Retrieved from https://opendata.cityofnewyork.us/

Scikit-learn developers. (n.d.). Scikit-learn: Machine learning in Python. Retrieved from https://scikit-learn.org/stable/

United Nations. (n.d.). Sustainable Development Goals: Goal 1 – No Poverty. Retrieved from https://sdgs.un.org/goals/goal1