Team DA54

Table of Contents:

Abstract

In a landscape where healthcare extends beyond traditional medical interventions, Social Determinants of Health (SDOH) play a critical role in shaping individual and community health. This report highlights the significant impact of factors like income, education, and ethnicity on health disparities and the effectiveness of healthcare. We examine the role of SDOH in driving health outcomes, underlining their importance in advancing health equity and preventive healthcare.

We explore the challenges and opportunities in integrating SDOH into healthcare practices, focusing on the work of Patient Navigators (PNs) and physicians in addressing the complex social contexts of patients. The report presents scenarios demonstrating the practical benefits of effectively managing SDOH, emphasising patient-focused care and efficient use of healthcare resources.

Addressing the challenges in SDOH assessment, we call for sophisticated, integrated approaches to understand and act on these determinants. The development of a sophisticated dashboard exemplifies our commitment to leveraging SDOH insights for proactive healthcare interventions, advocating for a healthcare system that prioritises empathy and effectiveness.

Ultimately, this report urges the healthcare community to fully integrate SDOH into their strategies, envisioning a future where healthcare equality is achieved through informed, compassionate care.

Executive Summary

Context

The project zeroes in on the critical examination of Social Determinants of Health (SDoH) and their influence on health outcomes. SDoH represents the array of conditions in which individuals are born, grow, live, work, and age, encompassing various factors that influence health status, functioning, and quality of life. To delve deeper and craft a more nuanced analysis, the team has elected to explore an expanded set of ten pivotal SDoH variable groups that are considered most pertinent for thorough examination. These groups are:

1. Housing Stability
2. Economic Instability
3. Health Care Access
4. Education Access
5. Technology Access
6. Social Support and Security
7. Transportation
8. Environmental Conditions
9. Food Security
10. Child Care Access

This augmented focus aims to encompass a broader spectrum of factors influencing health outcomes, reflecting a comprehensive approach to understanding and addressing the multifaceted nature of health determinants. By emphasising these ten variable groups, the project intends to shed light on the interconnected aspects of SDoH and inform more effective strategies in healthcare and public policy.

The selection of these ten major Social Determinants of Health (SDoH) variable groups for analysis is grounded in a strategic desire to holistically understand the multifactorial influences on health outcomes. Here's a rationale for each category along with recommendations for supporting literature:

Rationale

1. **Housing Stability**: Essential for overall health, unstable housing correlates with various adverse health conditions. Housing and Health: An Overview of the Literature by Health Affairs for insights into this link.

2. **Economic Instability**: Economic hardship directly influences health disparities. The Relationship Between Income Inequality, Poverty, and Globalization in the International Journal of Health Services for Economic Health Impacts.

3. **Health Care Access**: Essential for disease management and prevention, unequal access impacts health equity. Access to Health Care and Population Health: A Literature Review in Milbank Quarterly.

4. **Education Access**: Directly tied to health behaviors and outcomes, higher education fosters better health. Education and Health: Evaluating Theories and Evidence by Cutler and Lleras-Muney.

5. **Technology Access**: Affects health through information access and telehealth. Digital Health: A Path to Validation in NPJ Digital Medicine for its significance.

6. **Social Support and Security**: Influences health via stress reduction and behavior norms. Social Relationships and Health: A Flashpoint for Health Policy in the Journal of Health and Social Behavior.

7. **Transportation**: Impacts access to care and social determinants. Transportation Barriers to Accessing Health Care for Urban Children in the Journal of Health Care for the Poor and Underserved.

8. **Environmental Conditions**: Affects health through pollution and green space access. Environmental Conditions and Health Outcomes: A Comprehensive Review in Lancet Planetary Health.

9. **Food Security**: Critical for nutritional status and health. Food Insecurity and Health Outcomes in Health Affairs for its health implications.

10. **Child Care Access**: Influences early development and long-term health. Early Childhood Care and Development: A Lifecourse Approach to Health in Developmental Origins of Health and Disease.

By focusing on these ten SDoH variable groups, the research aims to address the comprehensive array of factors that impact health, advocating for a multi-sectoral approach to health promotion and disease prevention.

Data Resources

The Social Care Scorecard project utilizes key data sources to inform its predictive models:
1. **AHRQ Social Determinants of Health (SDOH) Data**: Central to analyzing community-level social determinants impacting health.
2. **Health-Related Census Tract Data**: Supplements the SDOH data with health outcomes, insurance, and disability statistics.
3. **U.S. Small-area Life Expectancy Estimates Project (USALEEP)**: Provides life expectancy data by census tract, aiding in refining the predictive model.

For the Texas state model, the project employs several data sources to construct a detailed support database across key categories:

1. **Healthcare Facilities and Providers**:
- Extract data from CMS on hospitals, home health, and long-term care facilities, plus healthcare provider information including specialities and Medicare status.

2. **Living Facilities and Care Services**

- Gather data on Assisted Living, Independent Living, and Memory Care facilities using Caring.com scripts.

3. **Community and Support Services**:

- Compile information on services like food delivery, transportation, and elder care from MealsonWheelsAmerica.org and CommunityResourceFinder.org.

4. **Data Integration and Mapping**:

- Employ the HUD 2020 Zip Code - Census Tract Crosswalk for precise service mapping in Texas, aligning resources with the scorecard's geographic specificity.

This approach ensures a targeted and relevant compilation of resources, facilitating the connection between identified social needs and available community support within the Texas region, and enhancing the practical utility of the Social Care Scorecard.

Data Scraping:

This task outlines the methodology used to traverse a sample of the zip codes in Texas, focusing on those with the best RES_RATIO, BUS_RATIO, OTH_RATIO, and TOT_RATIO. The extraction process involved gathering essential data such as establishment names, addresses, distances, contact numbers, and ratings. Additionally, recommendations were generated based on either the shortest distance or the highest ratings available on the website.

Methodology:

1. **Sample Selection:** A sample of the Texas zip codes was selected based on criteria such as RES_RATIO, BUS_RATIO, OTH_RATIO, and TOT_RATIO as provided.
2. **Data Traversal:** The project involved traversing through zip codes of Texas on the designated websites using a combination of Selenium and Beautiful Soup.
3. **Data Extraction:** Relevant data, including establishment names, addresses, distances, contact numbers, and ratings, were extracted from the HTML source code of the web pages.
4. **Recommendation Process:** The top three recommendations were made based on either the shortest distance or the highest ratings retrieved from the website.

Data Scraping from Given Websites:

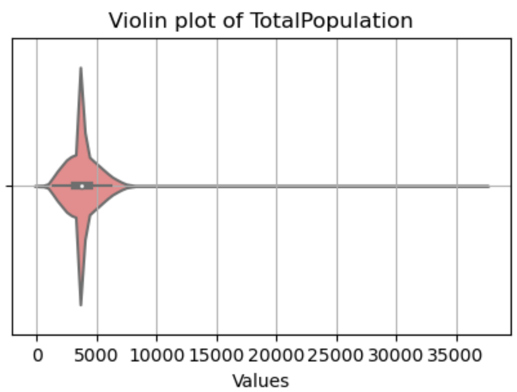
The primary task involved scraping data from the designated websites provided in the problem statement. Using a combination of Selenium and Beautiful Soup, the following steps were executed:

1. **Navigating to Web Pages:** Selenium WebDriver was used to navigate to the specified URLs of the target websites.
2. **Scanning HTML Structure:** Beautiful Soup parse the HTML structure of the web pages, allowing the identification of relevant data elements such as tags, classes, and attributes.
3. **Extracting Data:** Specific data points such as text, links, and other relevant content were extracted from the HTML source using Beautiful Soup's functions.
4. **Data Processing:** Extracted data were processed, cleaned, and stored in appropriate formats for recommendations.

Recommendations:

Based on the extracted data, the top three recommendations were made considering either the smallest distance or the highest ratings available on the website. This recommendation process ensured providing users with the best possible options according to their preferences.

Exploratory Data Analysis:

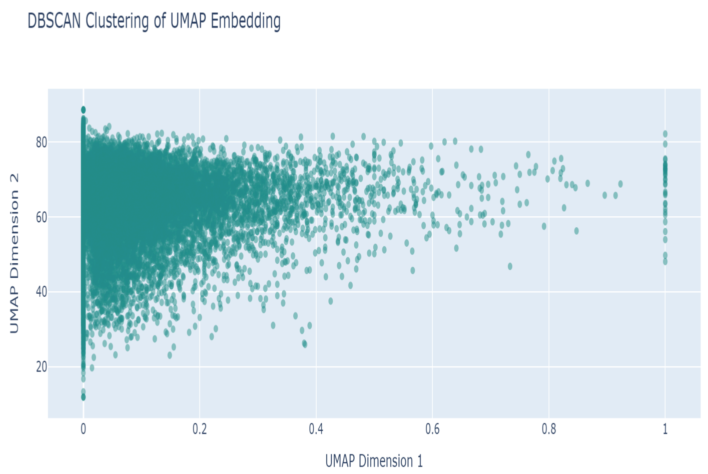


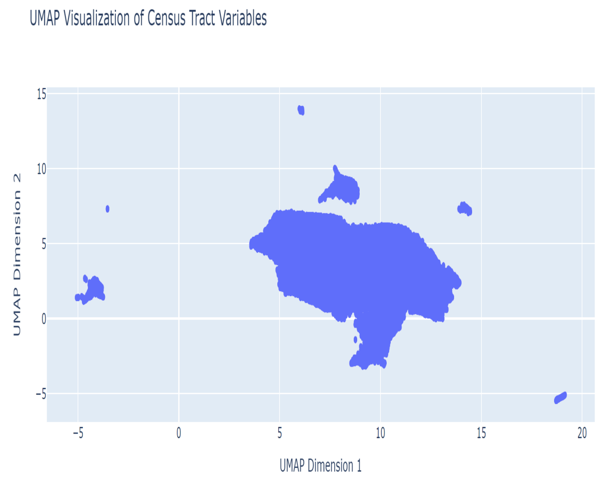
Data Normalization

In our data normalization process, we employed two prominent scaling techniques: Standard Scaler and MinMax Scaler, with a preference for the Standard Scaler approach due to its efficacy in our analysis.
Standard Scaler: It is effective in handling outliers and ensuring a consistent scale across all variables. Its effectiveness in our context can be attributed to its ability to facilitate faster convergence of algorithms and maintain the relative distances between feature values.
Isolation Forest for Null Value Imputation: To address the issue of missing values in our dataset, we utilized the Isolation Forest method. The choice of Isolation Forest was based on its efficiency in handling large datasets and its non-parametric nature, allowing it to deal with the complexities and nuances of our comprehensive data.

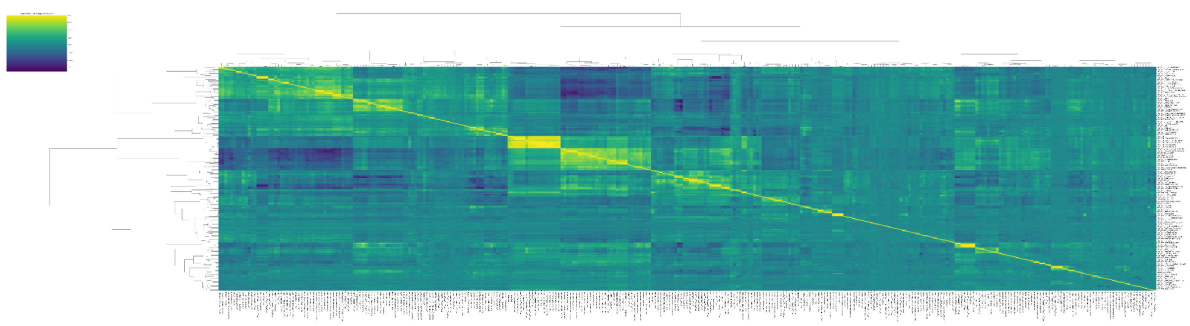
Feature Reduction

In the feature selection phase, our team explored several advanced techniques to manage and interpret the high-dimensional data inherent in our dataset. Initially, we experimented with dimensionality reduction and visualization methods like **DBSCAN, UMAP, and t-SNE**. However, these approaches did not yield significant insights, potentially due to the complex and **intertwined nature of the health-related variables** within our dataset.

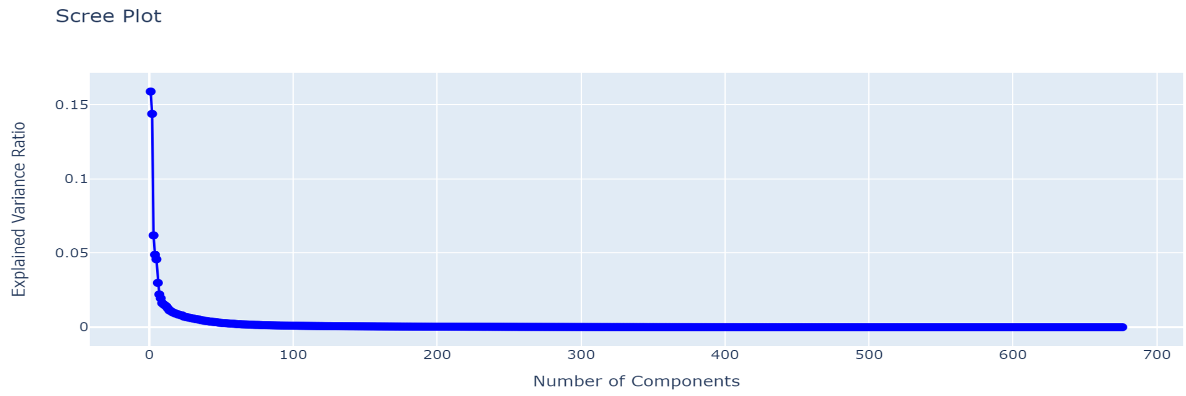




Subsequently, we shifted our focus to **Hierarchical Clustering** and utilized **dendrogram** plots to assess the potential clustering of variables. Despite this sophisticated approach, the clusters derived did not reveal actionable or **coherent patterns**. This outcome suggests the complexity of the health data may surpass the **discriminatory power of these clustering methods**, or the inherent **inter-variable relationships are too subtle** for such distinct segmentation.

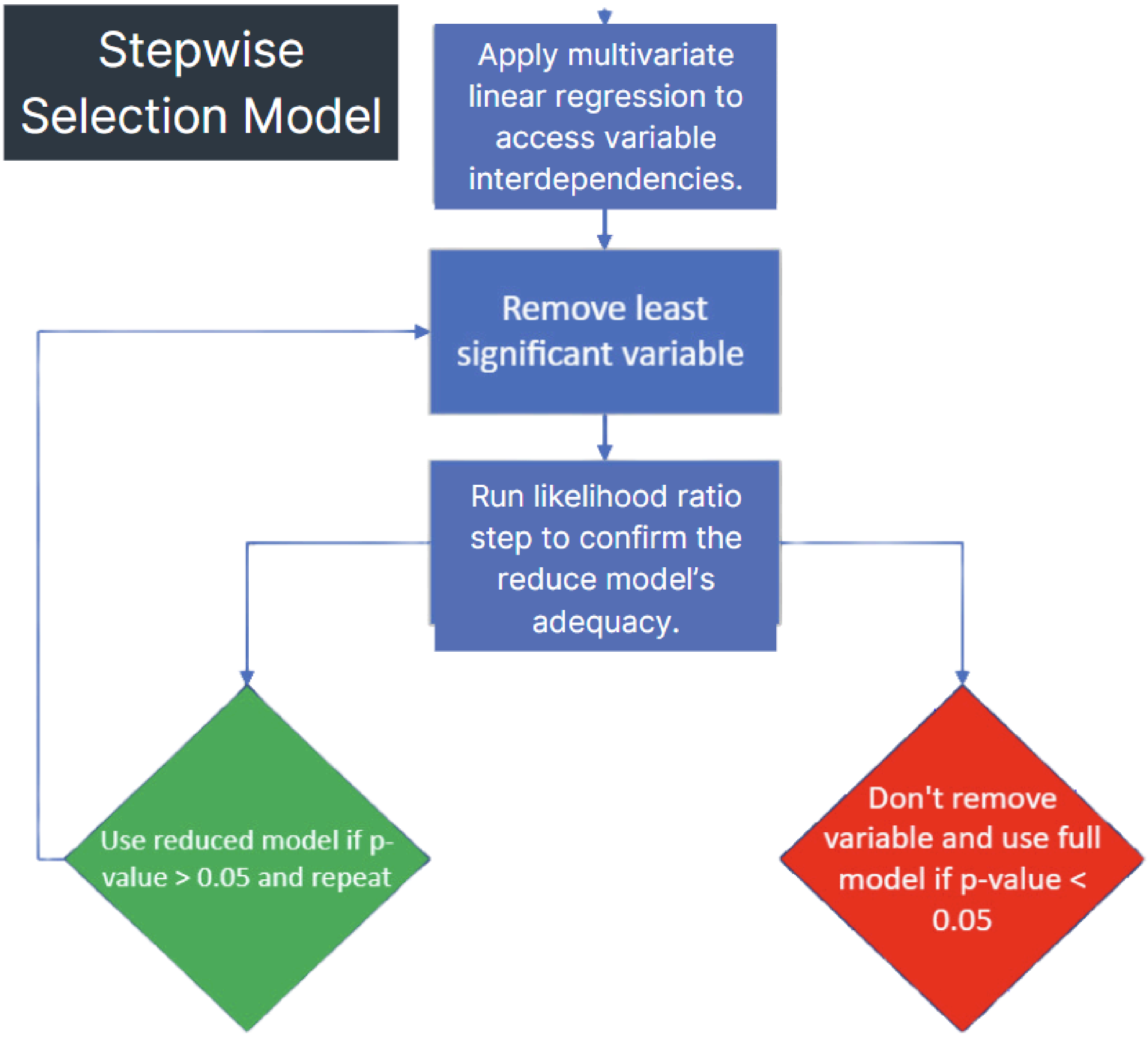


The rationale behind using these advanced techniques was to uncover inherent data structures that could inform more targeted feature selection. Despite the lack of clear results from these methods, the process was valuable for confirming the complexity and nuanced interrelations within our dataset, guiding us to adopt alternative strategies for feature selection that better accommodate the dataset's intricate nature.



For feature reduction, we applied **Principal Component Analysis (PCA) and Factor Analysis** to our dataset. Based on the scree plot analysis, we reduced the dimensions to 11 and 49, which represented the most significant variance in the data. After transforming the dataset, we examined the new dimensions and their constituent variable groups. Although this analysis identified distinct dimensions, the variable groups lacked clear interpretability, indicating a need for further refinement to achieve actionable insights.

Our Approach



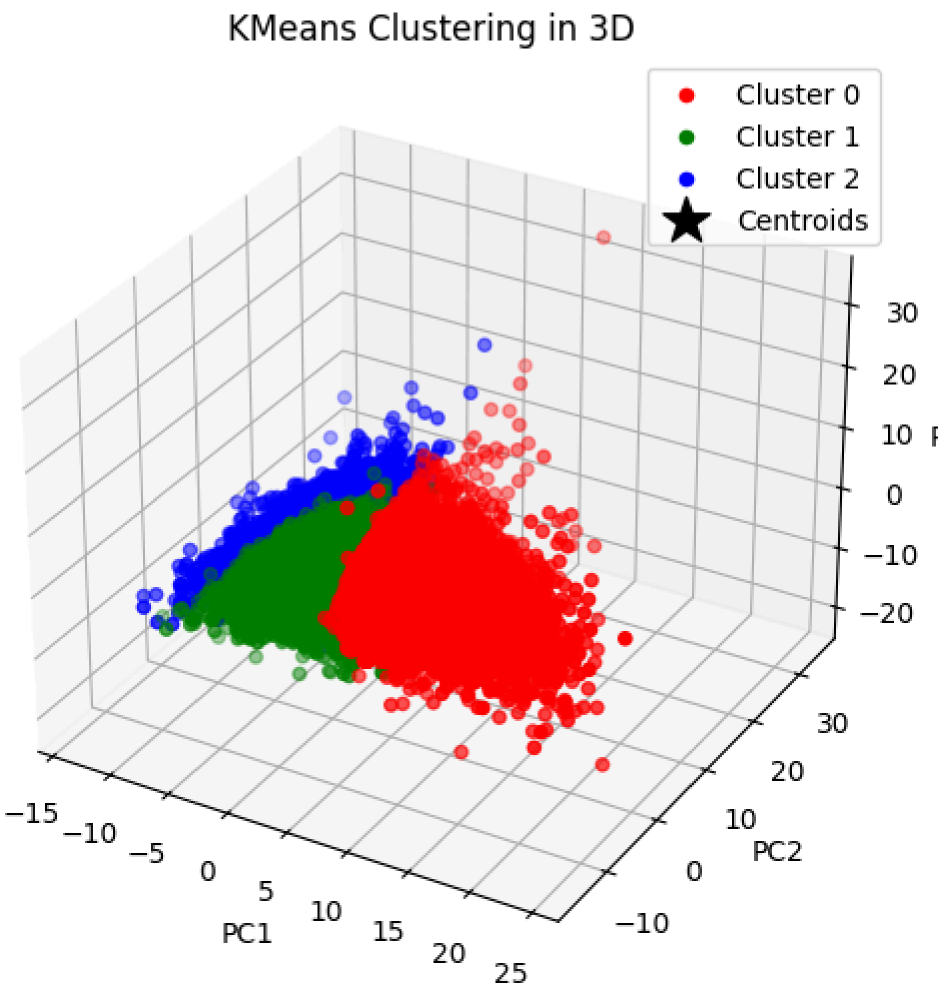
Constructed a correlation matrix to detect multicollinearity, eliminating variables with near-perfect correlations to preserve model integrity.

- Applied **Multivariate Linear Regression (MLR)** to assess variable interdependencies and predict health outcomes.
- Compared model scores to select the most predictive MLR model, incorporating both main effects and interaction terms.
- Utilized stepwise selection based on model score to pare down to significant variables, enhancing model simplicity and interpretability.
- Employed Likelihood-Ratio tests to confirm the reduced model's adequacy, with a p-value threshold of 0.05 as the cutoff for model acceptance.
- Choose the final model for each health statistic after ensuring reduced models pass the Likelihood-Ratio test, focusing on the most impactful variables.

Finally, we dropped the redundant, unnecessary variables and also performed some effective feature engineering to condense the features. This resulted in the reduced dimension of **153 variables.**

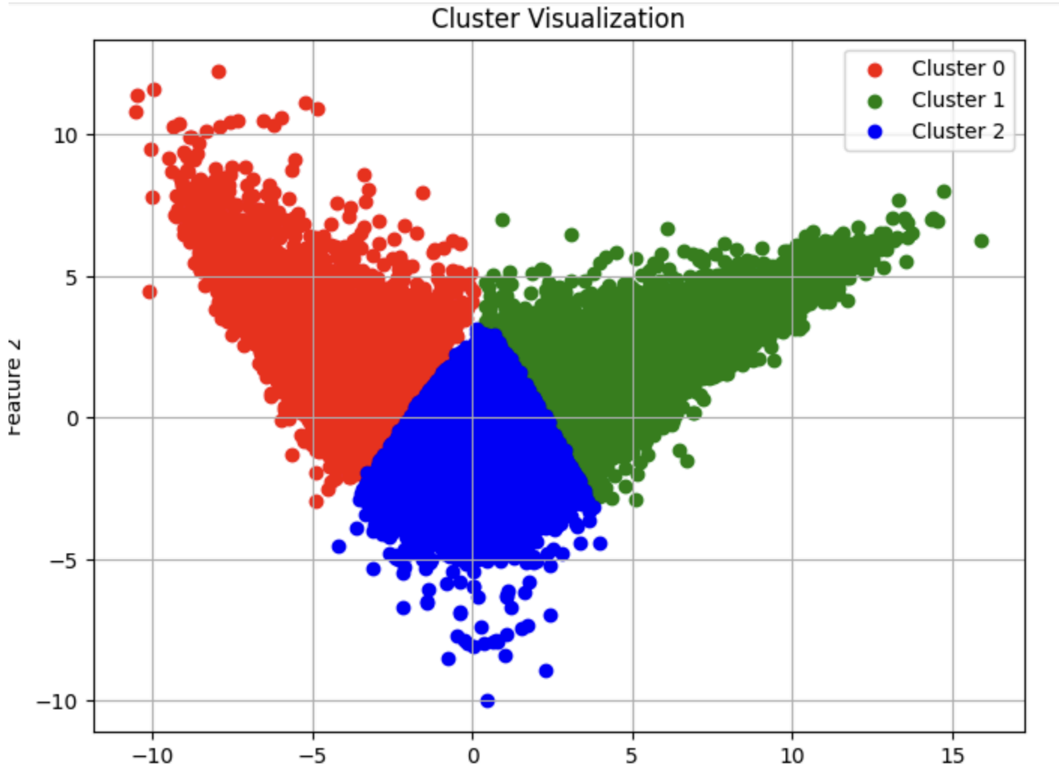
Methodology:

Approach 1:



Data Segmentation: Partitioned the dataset according to the 10 pre-defined Social Determinants of Health to facilitate focused analysis.
Clustering Analysis: Executed clustering within each determinant category using K-Means, identifying natural groupings within the data.
Risk Labeling: Assigned risk labels ('high', 'medium', 'low') based on the characteristics of each cluster's centroid.
Risk Scoring: Calculated inverse linear distance scores for each data point relative to cluster centroids to quantify risk levels.

Optimal Clusters Determination: Applied the elbow method to determine the appropriate number of centroids for the K-Means algorithm.



Clustering Validation: Validated the K-Means clustering with several metrics, including the Silhouette Score, Davies-Bouldin Index, and Calinski-Harabasz Index, to ensure the robustness and relevance of the clusters formed.

Approach 2

- Utilized a suite of machine learning models, including **Linear Regression, Lasso Regression, Elastic Net, Decision Tree, Random Forest, Gradient Boosting, SVR, and KNR**, to analyze the relationship between SDOH census tract data and outcomes such as **life expectancy and health status**.

Models	**Linear Regression**	**Lasso Regression**	**Elastic Net**	**XGB**	**SVR**	**KNN Regressor**	**Ridge Regressor**
MSE	0.57	0.65	0.65	0.57	0.58	0.66	0.57
MAE				0.56			
R2	0.11			0.13			
MAPE				392.33			

- Validated the models against a range of metrics: **RMSE, MAPE, MSE, R² score, and Mean Absolute Error**. Linear Regression and XGBoost emerged as the top performers.

- Determined variable significance through permutation importance, aiming to distill the variables' aggregated impact on the training dataset, inclusive of all census tracts.

- To tailor the risk scores to individual census tracts, normalized the feature importance and multiplied by the variable values, subsequently summing these products to derive category-specific risk scores.

- Implemented two methods for calculating risk scores:

1. Direct multiplication of risk scores with variable values.
2. Multiplication of normalized feature importance with variable values.
- This process was replicated for both the XGBoost and MLR models to ensure robustness.
- Normalized the final risk scores to a 0-1 scale, providing a standardized metric for assessing and comparing the impact of SDOH on health outcomes and life expectancy across different census tracts.

Resource Recommendation

For resource recommendation within the report, include the following points:

- Implemented a targeted recommendation strategy, prioritizing the top three resources for individuals categorized within the highest risk group.
- Utilized geospatial mapping techniques to relate users' zip codes to census tracts, enhancing the accuracy of resource allocation.
- Employed web scraping tools and APIs, including those from the US Census Bureau and various healthcare databases, to compile a comprehensive list of community resources and services.
- Recommended resources were carefully chosen based on key parameters such as service quality, accessibility, and relevance to the user's specific needs.
- For Texas users, the top three resources were identified and recommended based on the most critical support requirements determined by the risk assessment model.

Appointment But No Shows

In addressing the decision-making needs of physicians, our methodology pivots from broad evaluations to an issue-centric approach, enabling a more practical application of the Social Determinants of Health. This refined strategy is a departure from standard tools like PRAPARE, which often fall short in providing an integrated, issue-driven perspective that directly informs clinical decisions.

We introduce an enhanced scorecard featuring critical categories such as "Appointment No Shows," which are not merely indicators of patient engagement but also critical factors influencing healthcare delivery efficiency and financial sustainability. Recognizing the significant impact of missed appointments, we have developed a composite risk score encompassing underlying factors such as transportation difficulties, economic pressures, and educational barriers.

This holistic scoring model does not simply enumerate risks; it elucidates the interconnectedness of social determinants and their direct bearing on patient behavior. It empowers providers with actionable insights, translating complex social data into tailored recommendations that can proactively mitigate risks of non-attendance and enhance patient-provider interactions.

As a result, physicians are equipped with a nuanced understanding of each patient's unique challenges, fostering a more empathetic and informed approach to care that extends beyond the clinic and into the socio-economic realities of the patients they serve.

Education	• Explore Adult Education: Look into local adult education programs for getting your GED or improving literacy. • Learn about Health Matters: Access simplified health info for better understanding of preventive care and managing chronic conditions.

	• Discover Community Resources: Find nearby centers for vocational training and literacy programs.
Environment	• Limited Green Spaces: Explore community garden initiatives for outdoor activities. Advocate for more parks and indoor exercise options. • Water Contamination Risks: Look into installing water filtration systems and testing water regularly for safety. Support efforts to improve water infrastructure.
Economic Instability	• Healthcare Access: Explore affordable healthcare options like Medicaid or community clinics for medical needs. • Employment Help: Consider job training and resume support for stable employment opportunities. • Food and Housing Assistance: Look into programs providing food and housing support for those in need.
Access To Technology	• Internet Assistance: Ask about initiatives to improve internet access in your area. • Device Distribution Services: Learn about programs providing devices for connecting with healthcare and support services. • Device Borrowing Programs: Inquire about borrowing devices for telehealth appointments and health education.
Childcare excess	• Flexible Work Policies: Inquire about flexible work arrangements to balance work and childcare responsibilities. • Subsidized Childcare Programs: Explore options for affordable childcare programs available for families in need. •
Housing Stability	• Join Support Groups: Consider joining counseling or support groups for help during challenging times. • Caregiver Assistance: Seek support and education for caregivers to ensure proper care for your well-being.

Additional Inputs Adjustment

To refine the model's sensitivity to individual circumstances, we incorporated a dynamic weighting system based on feature importance, enhancing the personalization of risk scores for each determinant.

- Model Calibration: Adjusted risk scores for each Social Determinant of Health using a weighted approach, grounding the assessment in the feature importance derived from the variables most pertinent to the user.

- Importance-Driven Adjustment: Leveraged a Multivariate Linear Regression model, training it with a range of SDOH variables alongside user-specific inputs as independent variables, with the dependent variables being the 10 risk scores developed by our team.

- Targeted Risk Assessment: Upon determining feature importance, we applied these weights to the user-input variables, thereby tailoring the risk scores in direct response to the individual's unique profile.

Rationale:

- Enhanced Personalization: This method allows for a model that not only assesses risk based on general data but also adjusts its predictions in light of the specific attributes of an individual, leading to a more nuanced and actionable assessment.

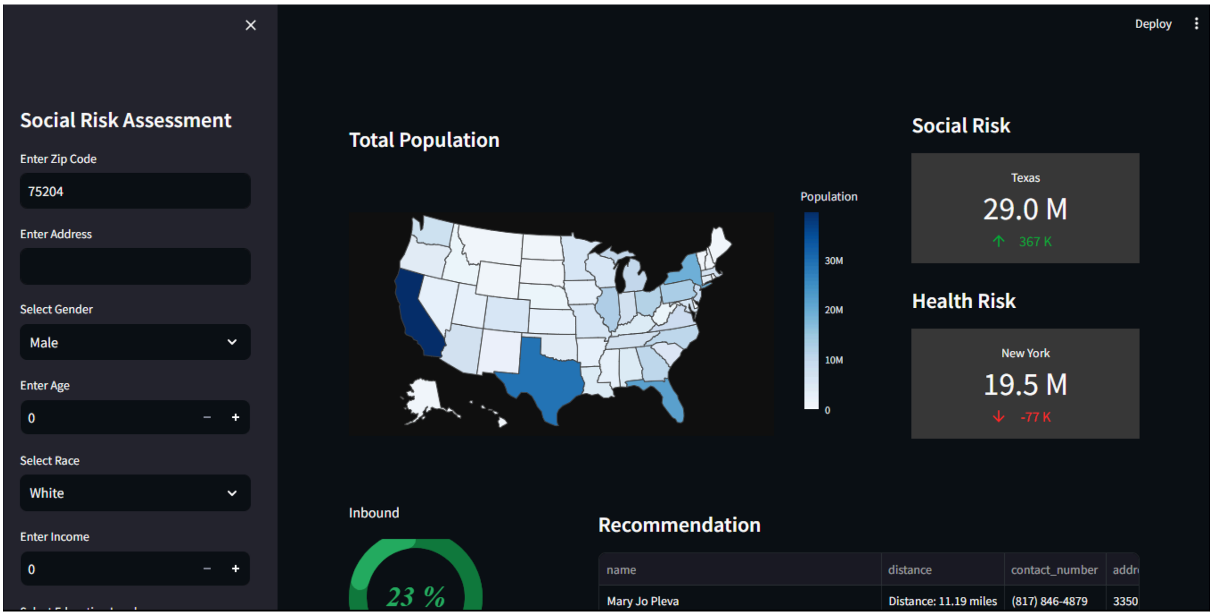
- Improved Predictive Accuracy: By focusing on the variables with the greatest impact on the calculated risk scores, the model prioritizes factors most likely to affect the outcome, thereby increasing its predictive precision.

- Adaptive Framework: This strategy ensures that the model remains flexible and responsive to new information, reflecting the evolving nature of healthcare needs and interventions.

In essence, this tailored approach ensures that each individual's risk assessment is a true reflection of their personal SDOH landscape, directly informing healthcare providers and enabling informed decision-making.

Score-Card Dashboard:

Social-Risk:



Health-Risk:

Social Determinants of Health (SDOH) in India

Introduction:

India, with its vast population and diverse social landscape, presents unique challenges in understanding and addressing Social Determinants of Health (SDOH). This report explores the key SDOH factors impacting health outcomes in India, highlighting relevant data sources and limitations compared to the US.

SDOH Factors in India:

1. **Housing Stability:** The Ministry of Housing and Urban Affairs (MoHUA) offers data on housing conditions (https://mohua.gov.in/), but pincode-level data is unavailable. Rural-urban disparities exist, with slums and inadequate sanitation impacting health in urban areas.

2. **Economic Instability:** Data on income distribution is available from sources like income tax returns (https://www.incometax.gov.in/iec/foportal/statistics-data) but lacks pincode-level granularity. Caste and religion can significantly influence economic opportunities, requiring further research.

3. **Health Care Access:** Public health data is available district-wise through the Census (https://censusindia.gov.in/), but pincode-level details are absent. Rural areas often have limited access to quality healthcare facilities.

4. **Education Access:** The Unified District Information System for Education Plus (UDISE+) provides educational statistics by state and district (https://dashboard.udiseplus.gov.in/). Literacy rates vary across regions and castes, impacting health literacy.

5. **Technology Access:** Internet penetration data is available from Statista (https://www.statista.com/chart/30029/internet-penetration-rate-in-india/), but pincode-level details are missing. The digital divide between rural and urban areas creates disparities in accessing health information and services.

6. **Social Support and Security:** The National Family Health Survey (NFHS) provides district-wise data on family health (https://github.com/topics/nfhs). Caste-based discrimination and limited social safety nets can negatively impact mental and physical health.

7. **Transportation:** Public transportation data might be available through specific government agencies, but pincode-level details are likely absent. Limited access to reliable transportation can hinder access to

healthcare and healthy food options.

8. **Environmental Conditions:** Data on air and water quality is likely available from government agencies, but pincode-level details are scarce. Pollution levels are a major health concern in India, particularly impacting urban populations.

9. **Food Security:** While some data might be available through government channels, pincode-level details are likely missing. Rural poverty and lack of access to fresh produce contribute to malnutrition, especially among children.

10. **Childcare Access:** Data on childcare facilities might be available through Women and Child Development ministry websites, but pincode-level details are likely absent. Lack of affordable childcare options can limit women's employment opportunities and impact child development.

11. **Veteran Prosperity:** Data specific to veterans' well-being might be available from the Ministry of Defence, but pincode-level details are unlikely. Social support programs tailored to veterans' needs may be lacking.

Data Limitations Compared to US:

- **Granularity:** Indian data is primarily available at the district level, while the US has data readily available at county and even zip code levels. This limits our understanding of health disparities within districts, especially in geographically vast states.
- **Caste & Religion:** Caste and religion play a significant role in health outcomes in India, but data collection on these factors is often limited due to sensitivities. Understanding the health impact of these factors requires further research and sensitive data collection methods.

Possible Solutions:

- **Leveraging Technology:** Invest in technology solutions for data collection and analysis, potentially using mobile phone surveys and anonymized data to capture pincode-level insights.
- **Community-based Data Collection:** Partner with local NGOs and community workers to collect data that reflects specific needs and challenges.
- **Standardization and Collaboration:** Encourage collaboration between government agencies and research institutions to standardize data collection methods and facilitate data sharing.
- **Data Privacy:** Develop robust data privacy frameworks to ensure data security while enabling research and policy development.

Conclusion

Understanding SDOH is crucial for improving health outcomes in India. While data limitations exist compared to the US, ongoing efforts to improve data collection, analysis, and collaboration can bridge these gaps. By addressing the social determinants of health, India can create a more equitable and healthy future for all its citizens.

Appendix

- *Anderson, L. M., Scrimshaw, S. C., Fullilove, M. T., Fielding, J. E., & Normand, J. (2003). Culturally competent healthcare systems. American Journal of Preventive Medicine, 24(3), 68–79. https://doi.org/10.1016/s0749-3797(02)00657-8*
- *APHA. (2018, November 13). Violence is a public health issue: Public Health is essential to understanding and treating violence in the U.S. American Public Health Association. Retrieved September 10, 2022, from https://apha.org/policies-and-advocacy/public-health-policy-statements/policy-database/2019/01/28/violence-is-a-public-health-issue*
- *Artiga, S., Hill, L., & Haldar, S. (2022, April 19). Key facts on health and health care by race and ethnicity - social determinants of health. KFF. Retrieved December 10, 2022, from https://www.kff.org/report-section/key-facts-on-health-and-health-care-by-race-and-ethnicity-social-determinants-of-health/*
- *Brou, L. (2019, June 7). Immigration as a Social Determinant of Health. University of Colorado School of Medicine. Retrieved December 9, 2022, from https://medschool.cuanschutz.edu/family-medicine/about/news/disrupting-the-status-quo-blog/disrupting-the-status-quo/immigration-as-a-social-determinant-of-health*
- *Bureau, U. S. C. (2022, August 3). What We Do. Census.gov. Retrieved October 25, 2022, from https://www.census.gov/about/what.html*
- *Catalyst, N. E. J. M. (2017, December 1). Social Determinants of Health (SDOH). NEJM Catalyst. Retrieved September 10, 2022, from https://catalyst.nejm.org/doi/full/10.1056/CAT.17.0312*
- *Centers for Disease Control and Prevention. (2022, December 8). Social Determinants of Health at CDC. Centers for Disease Control and Prevention. Retrieved September 10, 2022, from https://www.cdc.gov/about/sdoh/index.html*
- *Correlation and Regression with R. Simple linear regression. (2016). Retrieved October 25, 2022, from https://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/R/R5_Correlation-Regression/R5_Correlation-Regression4.html*
- *Dan Turtureanu, Demetri Pananos. (1968, October 1). How the poisson distribution is used in regression? Cross Validated. Retrieved October 25, 2022, from https://stats.stackexchange.com/questions/542155/how-the-poisson-distribution-is-used-in-regression*
- *Douthit, N., Kiv, S., Dwolatzky, T., & Biswas, S. (2015). Exposing some important barriers to healthcare access in the Rural USA. Public Health, 129(6), 611–620. https://doi.org/10.1016/j.puhe.2015.04.001*
- *Egerter, S., Dekker, M., An, J., Grossman-Kahn, R., & Braveman, P. (2008, December). The commission to build a Healthier America. The Commission to Build a Healthier America. Retrieved December 14, 2022, from http://www.commissiononhealth.org/*
- *Foundation, the A. E. C. (2010, January 1). Reading by third grade. The Annie E. Casey Foundation. Retrieved December 14, 2022, from https://www.aecf.org/resources/early-warning-why-reading-by-the-end-of-third-grade-matters*
- *Frederick, T. J., Chwalek, M., Hughes, J., Karabanow, J., & Kidd, S. (2014). How stable is stable? defining and measuring housing stability. Journal of Community Psychology, 42(8), 964–979. https://doi.org/10.1002/jcop.21665*

- *Gulati, R. K., & Hur, K. (2021). Association between limited English proficiency and healthcare access and utilization in California. Journal of Immigrant and Minority Health, 24(1), 95–101. https://doi.org/10.1007/s10903-021-01224-5*

- *Han, B., Cohen, D., Derose, K., Li, J., & Williamson, S. (2018, January 12). Violent Crime and Park Use in Low-Income Urban Neighborhoods. American Journal of Preventive Medicine. Retrieved September 10, 2022, from https://doi.org/10.1016/j.amepre.2017.10.025*

- *Hill, J., Rodriguez, D. X., & McDaniel, P. (2021, October 29). Immigration status as a social determinant of health during covid-19. Public Health Post. Retrieved December 9, 2022, from https://www.publichealthpost.org/research/immigration-status-covid-19/*

- *Kenton, W. (2022). What is social capital? Investopedia. Retrieved September 10, 2022, from https://www.investopedia.com/terms/s/socialcapital.asp*

- *Kushel, M. B., Gupta, R., Gee, L., & Haas, J. S. (2006). Housing instability and food insecurity as barriers to health care among low-income Americans. Journal of General Internal Medicine, 21(1), 71–77. https://doi.org/10.1111/j.1525-1497.2005.00278.x*

- *LexisNexis Risk Solutions, L. N. R. (2017, October 24). Milliman MedInsight to use LexisNexis risk solutions socioeconomic health attributes to help enhance healthcare intelligence. Milliman MedInsight to Use LexisNexis Risk Solutions Socioeconomic Health Attributes to Help Enhance Healthcare Intelligence. Retrieved October 25, 2022, from https://www.prnewswire.com/news-releases/milliman-medinsight-to-use-lexisnexis-risk-*

-

- *solutions-socioeconomic-health-attributes-to-help-enhance-healthcare-intelligence-300541930.html*

- *Medicare Signed into Law. U.S. Senate: Medicare Signed into Law. (2019, July 12). Retrieved September 15, 2022, from https://www.senate.gov/artandhistory/history/minute/Medicare_Signed_Into_Law.htm#:~:text=On%20July%2030%2C%201965%2C%20President,Harry%20Truman%20had%20proposed%20it*

- *Michele Ver Ploeg, V. B. (2009, June 25). Access to affordable and nutritious food-measuring and understanding food deserts and their consequences: Report to Congress. USDA ERS. Retrieved October 12, 2022, from http://www.ers.usda.gov/publications/pub-details/?pubid=42729*

- *Milliman (2022). The Work Matters. Retrieved October 25, 2022, from https://www.milliman.com/en*

- *Published by Statista Research Department, & 20, J. (2022, June 20). Distribution of Medicare beneficiaries by age 2020. Statista. Retrieved September 15, 2022, from https://www.statista.com/statistics/248035/distribution-of-medicare-beneficiaries-by-age/*

- *Published: Feb 13. (2019, February 13). An Overview of Medicare. KFF. Retrieved December 14, 2022, from https://www.kff.org/medicare/issue-brief/an-overview-of-medicare/#:~:text=%23Medicare%20plays%20a%20key%20role,physician%20services%2C%20and%20prescription%20drugs*

- *Singh, G. K., & Siahpush, M. (2006). Widening socioeconomic inequalities in US life expectancy, 1980–2000. International Journal of Epidemiology, 35(4), 969–979. https://doi.org/10.1093/ije/dyl083*
- *The college payoff: More education doesn't always mean more earnings. CEW Georgetown. (2022, May 16). Retrieved October 10, 2022, from https://cew.georgetown.edu/cew-reports/collegepayoff2021/*
- *United States Government. (2022, November 29). Rural Poverty & Well-being. USDA ERS - Rural Poverty & Well-Being. Retrieved September 16, 2022, from https://www.ers.usda.gov/topics/rural-economy-population/rural-poverty-well-being/*
- *United States Government. (2022, October 17). Key Statistics & Graphics. USDA ERS - Key Statistics & Graphics. Retrieved November 3, 2022, from https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/key-statistics-graphics.aspx*
- *US Department of Health and Human Services. (n.d.). Access to health services. Access to Health Services - Healthy People 2030. Retrieved September 10, 2022, from https://health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/access-health-services#cit4*
- *US Department of Health and Human Services. (n.d.). Access to primary care. Access to Primary Care - Healthy People 2030. Retrieved September 10, 2022, from https://health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/access-primary-care*
- *US Department of Health and Human Services. (n.d.). Crime and Violence. Crime and Violence - Healthy People 2030. (2022). Retrieved September 10, 2022, from*
- *https://health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/crime-and-violence*
- *US Department of Health and Human Services. (n.d.). Employment. Employment - Healthy People 2030. Retrieved December 14, 2022, from https://health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/employment#cit1*
- *US Department of Health and Human Services. (n.d.). Enrollment in higher education. Enrollment in Higher Education - Healthy People 2030. Retrieved September 10, 2022, from https://health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/enrollment-higher-education*
- *US Department of Health and Human Services. (n.d.). Environmental Conditions. Environmental Conditions - Healthy People 2030. (2022). Retrieved December 14, 2022, from https://health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/environmental-conditions*
- *US Department of Health and Human Services. (n.d.). High school graduation. High School Graduation - Healthy People 2030. Retrieved September 10, 2022, from https://health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/high-school-graduation*
- *US Department of Health and Human Services. (n.d.). Language and literacy. Language and Literacy - Healthy People 2030. Retrieved September 10, 2022, from https://health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/language-and-literacy#cit5*

- *US Department of Health and Human Services. (n.d.). Social cohesion. Social Cohesion - Healthy People 2030. Retrieved September 10, 2022, from https://health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/social-cohesion*
- *US Department of Health and Human Services. (n.d.). Social Determinants of Health. Social Determinants of Health - Healthy People 2030. Retrieved September 10, 2022, from https://health.gov/healthypeople/objectives-and-data/social-determinants-health*
- *Vanhove, J. (2019, April 11). Before worrying about model assumptions, think about model relevance. Jan Vanhove :: Before worrying about model assumptions, think about model relevance. Retrieved October 25, 2022, from https://janhove.github.io/analysis/2019/04/11/assumptions-relevance#:~:text=The%20general%20linear%20model's%20assumptions,)%2C%20normality%2C%20and%20independence*
- *Ministry of Housing and Urban Affairs (MoHUA): https://mohua.gov.in/*
- *Census of India: https://censusindia.gov.in/*
- *National Family Health Survey (NFHS): https://github.com/topics/nfhs*
- *Unified District Information System for Education Plus (UDISEP)*