Introduction:

Understanding the factors that influence higher income is essential for examining social mobility, economic opportunities, and inequality. Income-related disparities have been a key focus in socio-economic research. Kochhar and Moslimani (2021) highlight the persistent wealth inequalities in the United States, reporting that Asian and White households hold significantly more wealth than Black and Hispanic households, primarily due to differences in educational attainment and systemic inequities. Similarly, Ahrenfeldt et al. (2020) underscore the importance of income stability, demonstrating that lower and declining incomes are strongly linked to increased health risks and higher mortality rates, especially among men. Notably, income levels during an individual’s late 50s have been identified as a critical predictor of health outcomes later in life.

This project investigates the research question: What are the key factors that best predict whether an individual’s annual income exceeds $50,000, and how effectively can these variables be used in a predictive model? To address this question, we use the "Adult Income" dataset from the UCI Machine Learning Repository. Rather than relying on manually selected variables, we employ LASSO regression to systematically identify the most relevant predictors from a diverse set of features, including age, education level, work hours, and occupation. This data-driven approach allows us to construct a predictive model that not only answers our research question but also provides insights into the complex relationships between these variables and income levels.

Reference:

Ahrenfeldt, L. J., Pedersen, J. K., Thinggaard, M., Christensen, K., & Lindahl-Jacobsen, R. (2019). Sex differences in health and mortality by income and income changes. *Journal of Epidemiology and Community Health*, *74*(3), 225–231. <https://doi.org/10.1136/jech-2019-213096>

Kochhar, R., & Moslimani, M. (2023, December 4). *Wealth gaps across racial and ethnic groups*. Pew Research Center. <https://www.pewresearch.org/2023/12/04/wealth-gaps-across-racial-and-ethnic-groups/>