2.⁠ ⁠Background Research

Research Papers Relevant to Topic/DS

Lemon et al. (2018): This work used machine learning methods like Naïve Bayes and Gradient Boosting Classifiers to examine the UCI Adult Census Dataset. The scientists found that income levels are highly influenced by educational attainment. This is consistent with our study's focus on the influence of education on income distribution. Their knowledge of preprocessing techniques and feature importance was crucial in helping us organise our strategy. (Chet Lemon, 2018)

Chakrabarty and Biswas (2018): A high validation accuracy of 88.16% was attained in this study by using Gradient Boosting Classifiers. The study underlined how important it is to deal with missing values and efficiently encode categorical categories. Education and work hours were found to be important predictors, confirming the importance of these factors in income analysis. (Navoneel Chakrabarty, 2018)

Ali (2020): This study showed how useful statistical techniques are for determining correlations between categorical variables, such as income and education. The study emphasised notable differences in income across educational levels by using chi-squared tests and visualisations. The statistical techniques and visualisation approaches used in our study are supported by these results. (Ali, April 2020)

Why RQ is of Interest

Economic outcomes are significantly shaped by education. "Is there a difference in proportions of income across different levels of education among adults in the USA?" is our research question. is essential to comprehending economic disparity. By focussing on education-driven interventions, the findings of this study can help guide strategies to lessen economic inequality.  
  
Our dataset's analysis demonstrates that greater income proportions are linked to higher education levels. Visualisations, for example, show that those with advanced degrees have a far better chance of earning more money than people with only a high school education. This finding is further supported by the chi-squared test, which shows statistically significant differences with a p-value < 0.05.

This study offers practical insights in addition to being scholarly. The results highlight how crucial it is for organisations and policymakers to invest in education in order to increase economic mobility and lower inequality. Future research should incorporate other socioeconomic factors, such geography or gender, to improve interventions and gain a deeper understanding.

# Bibliography

Ali, A., April 2020. *Adult Census Income Dataset Analysis,* s.l.: s.n.

Chet Lemon, C. Z. K. M., 2018. *Predicting if income exceeds $50,000 per year based on 1994 US Census Data with Simple Classification Techniques,* s.l.: s.n.

Navoneel Chakrabarty, S. B., 2018. *A Statistical Approach to Adult Census Income Level Prediction,* s.l.: s.n.