

Storytelling Insights: Household Demographic Data Analysis

This project explores how household income is influenced by age and education level, transforming raw demographic data into meaningful socioeconomic insights.

The analysis reveals that household income generally increases with age as individuals gain experience and career stability. Income reaches its peak during the mid-career stage, particularly between the ages of 40 and 45. Beyond this point, income begins to decline, reflecting reduced work participation, career plateauing, or preparation for retirement. This pattern highlights the natural income lifecycle associated with professional growth and aging.

Education level emerges as the strongest determinant of income. Households led by individuals with post-graduate education consistently show higher and more stable income levels compared to other groups. Graduates display moderate income with greater variability, while households with primary education earn the lowest income overall. Secondary education falls between these extremes, occasionally producing higher-income outliers. This indicates that higher education not only increases earning potential but also reduces income volatility.

Income distribution analysis shows a right-skewed pattern, where the majority of households fall within a low-to-middle income range, and only a small proportion earn significantly higher incomes. This uneven distribution reflects real-world income inequality and suggests that most economic activity is driven by middle-income households rather than high-income earners.

Further statistical analysis highlights substantial income variability even among individuals with similar educational backgrounds. This suggests that while education provides access to better opportunities, income outcomes are strongly influenced by additional factors such as skills, industry, experience, and geographic location.

An important technical insight from the project is the handling of real-world data inconsistencies. The dynamic column normalization approach ensures the script remains robust against formatting variations in datasets, demonstrating practical data engineering skills beyond basic analysis.

Overall, the project presents a realistic picture of how income evolves across life stages and education levels, offering valuable insights for business strategy, public policy formulation, workforce planning, and career decision-making.