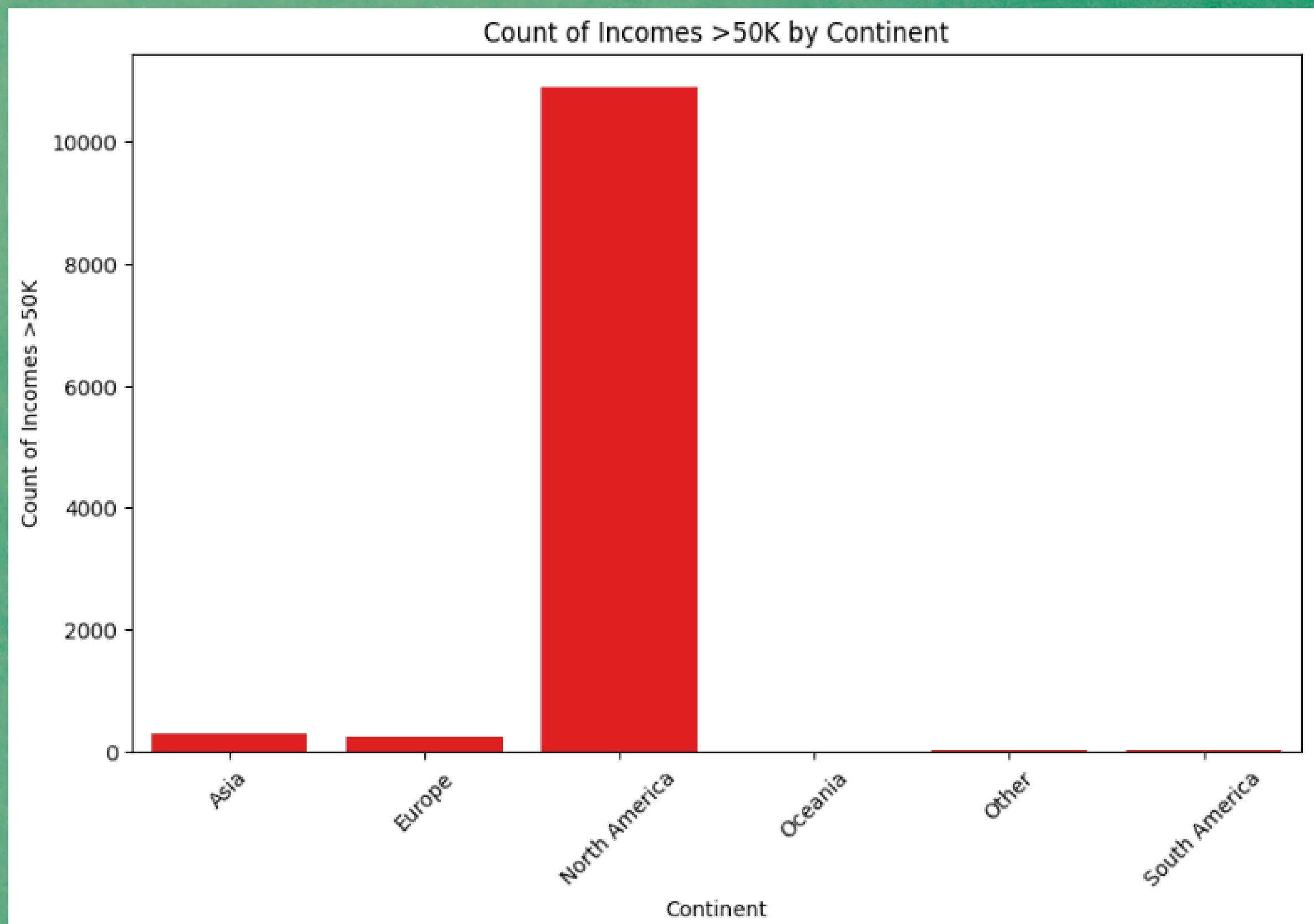


Midterm Skills Exam: Data Wrangling and Analysis

Census Income Dataset

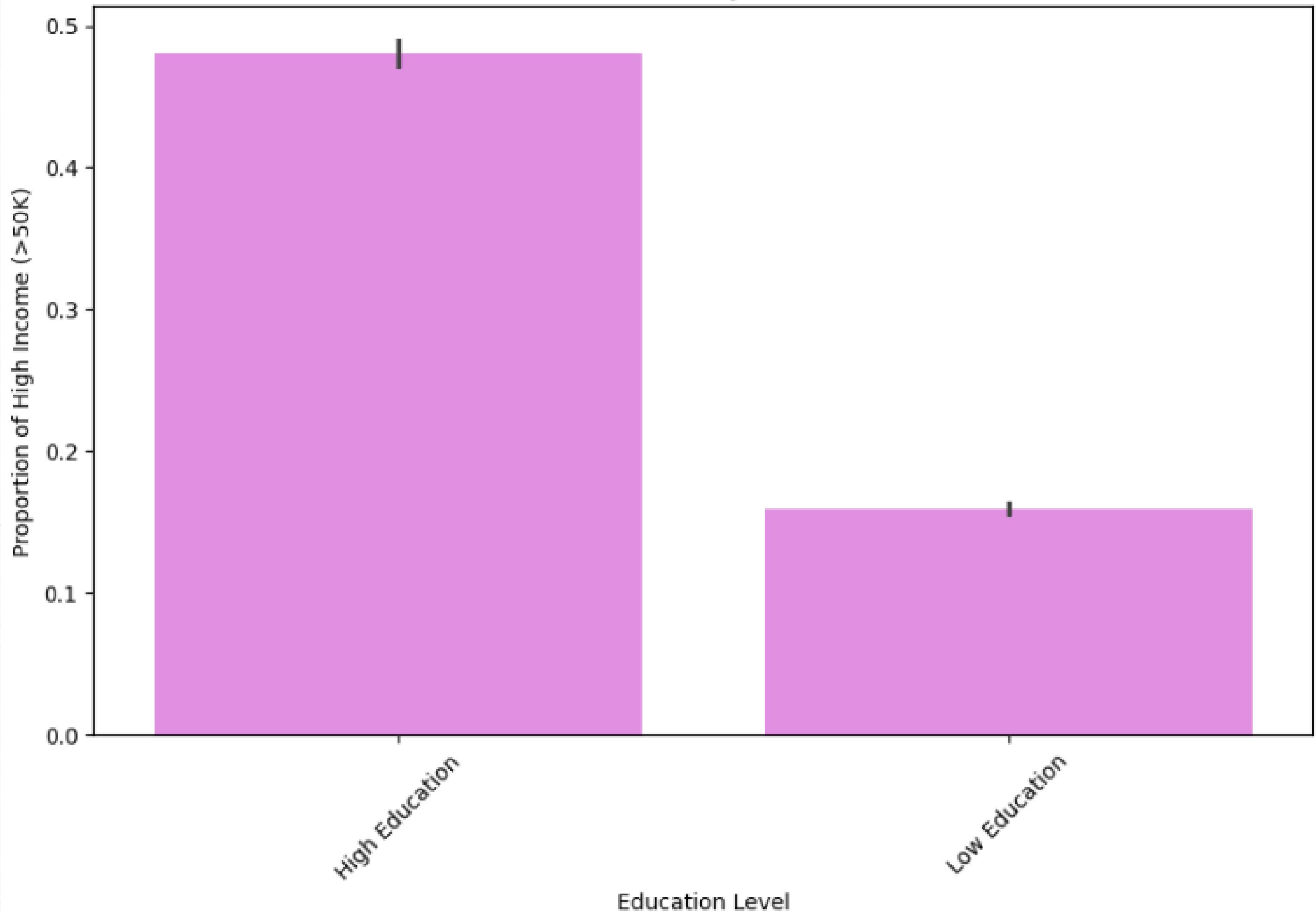
By: Keith Leigh Zhen Marquez

This shows
the count
of incomes
greater
than 50K
for each
continent



North America has the most people with the highest income this means have stronger economies compared to other continents, leading to higher average incomes and more people earning incomes greater than 50K

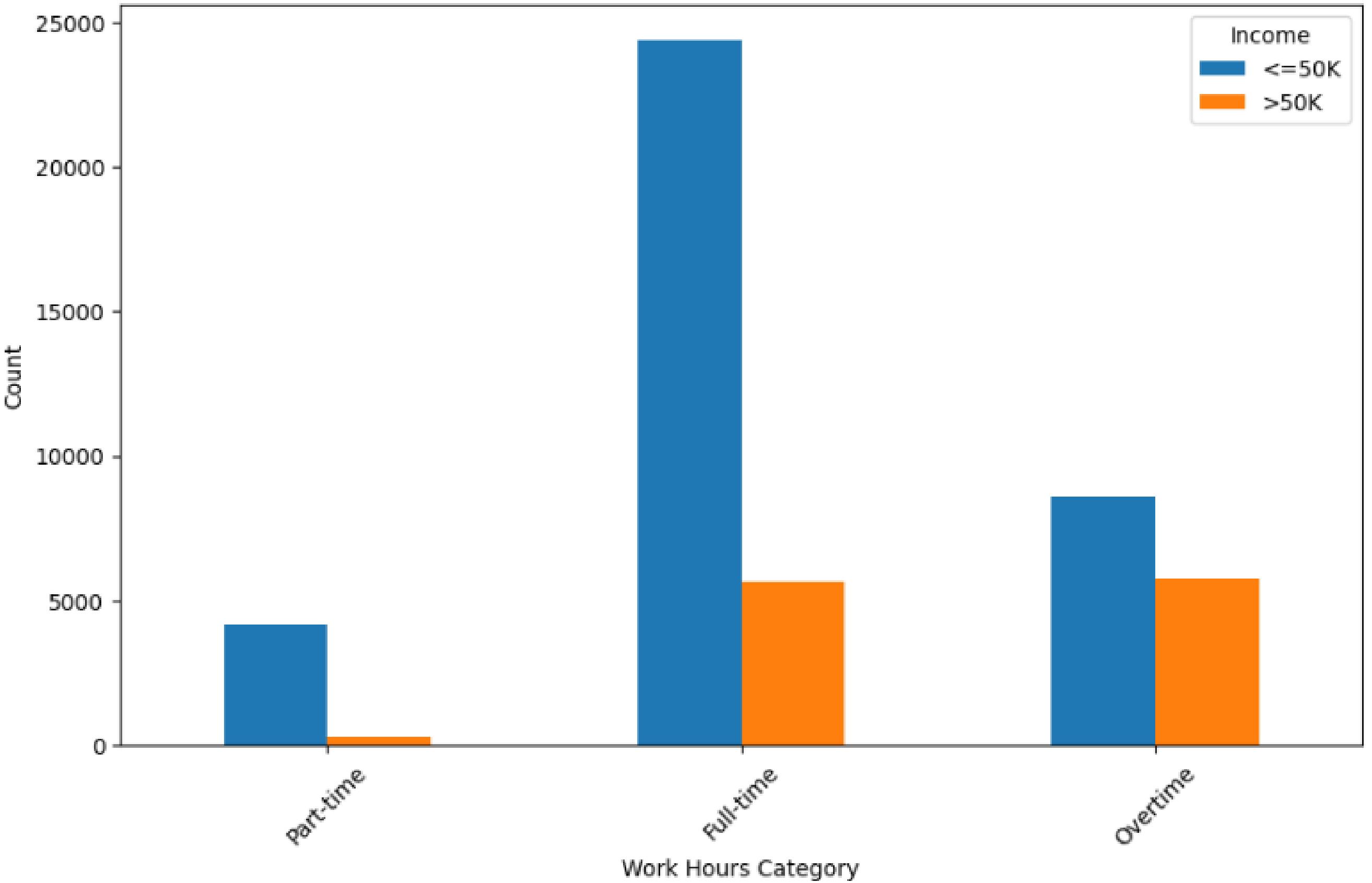
Income Distribution by Education Level



The Higher
the
Education
Level the
Higher the
Income



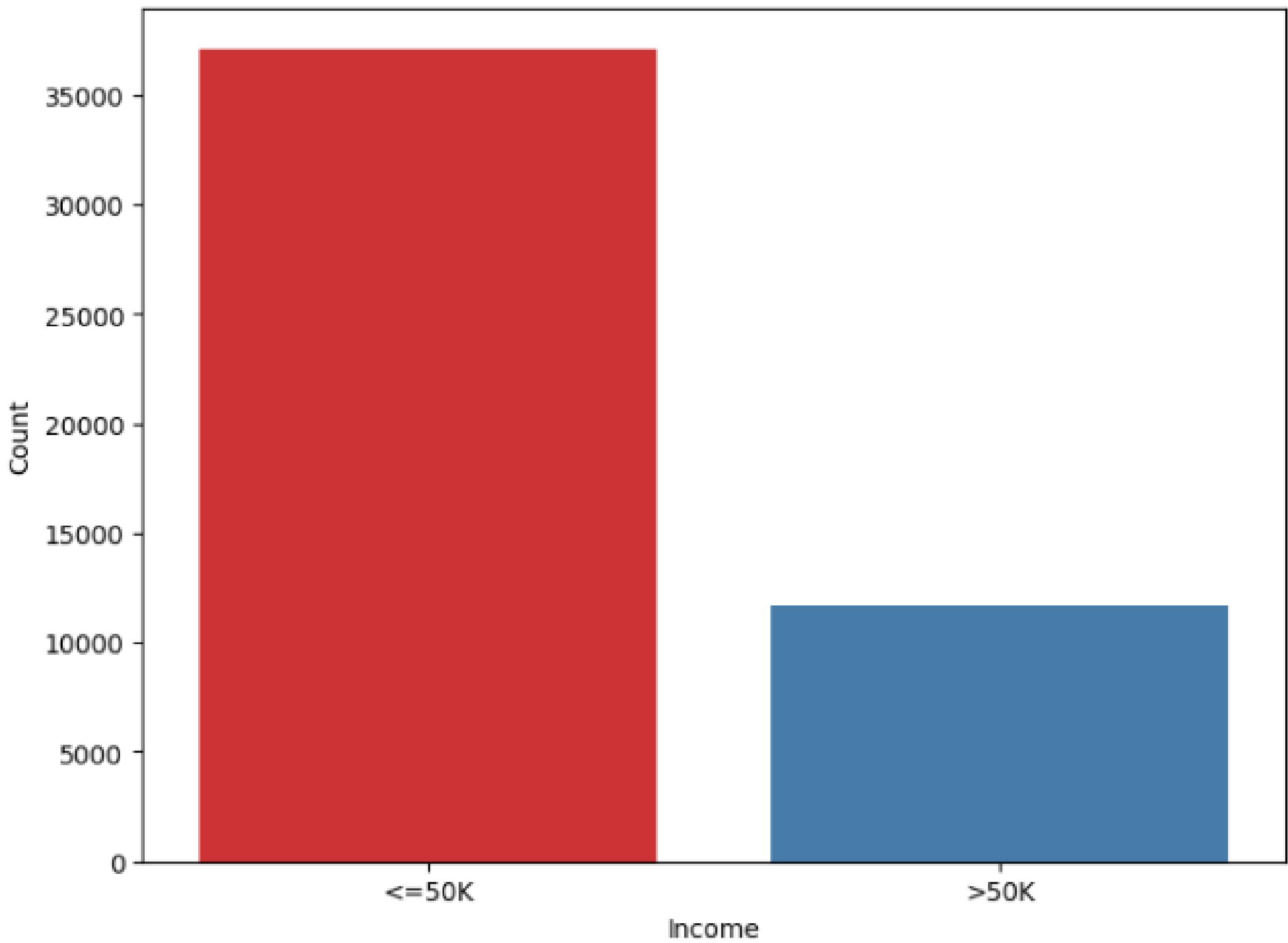
Counts of Individuals by Work Hours Category and Income



The Full-time category has the highest count of individuals for both income levels.

In all categories, there are more individuals earning $\leq 50K$ than those earning $>50K$.

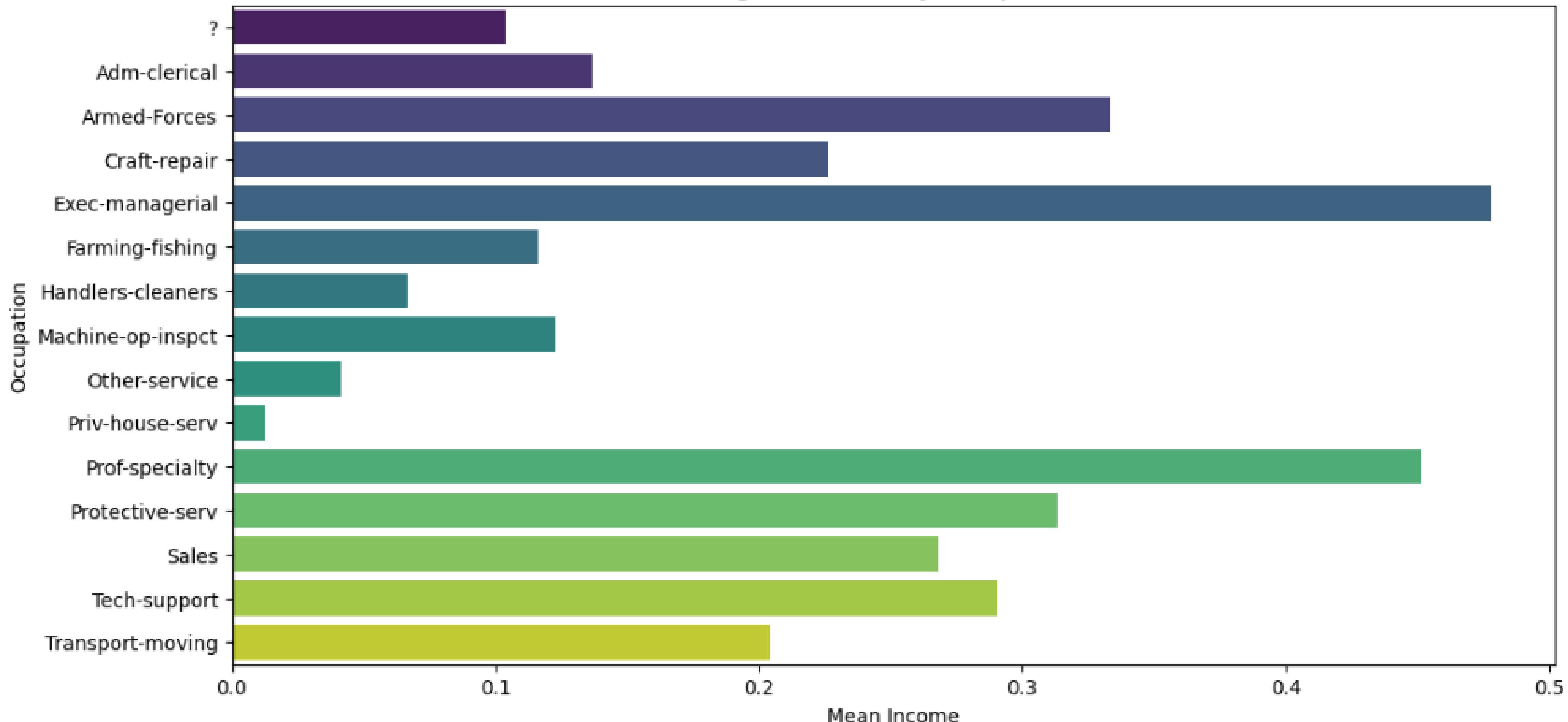
Income Distribution



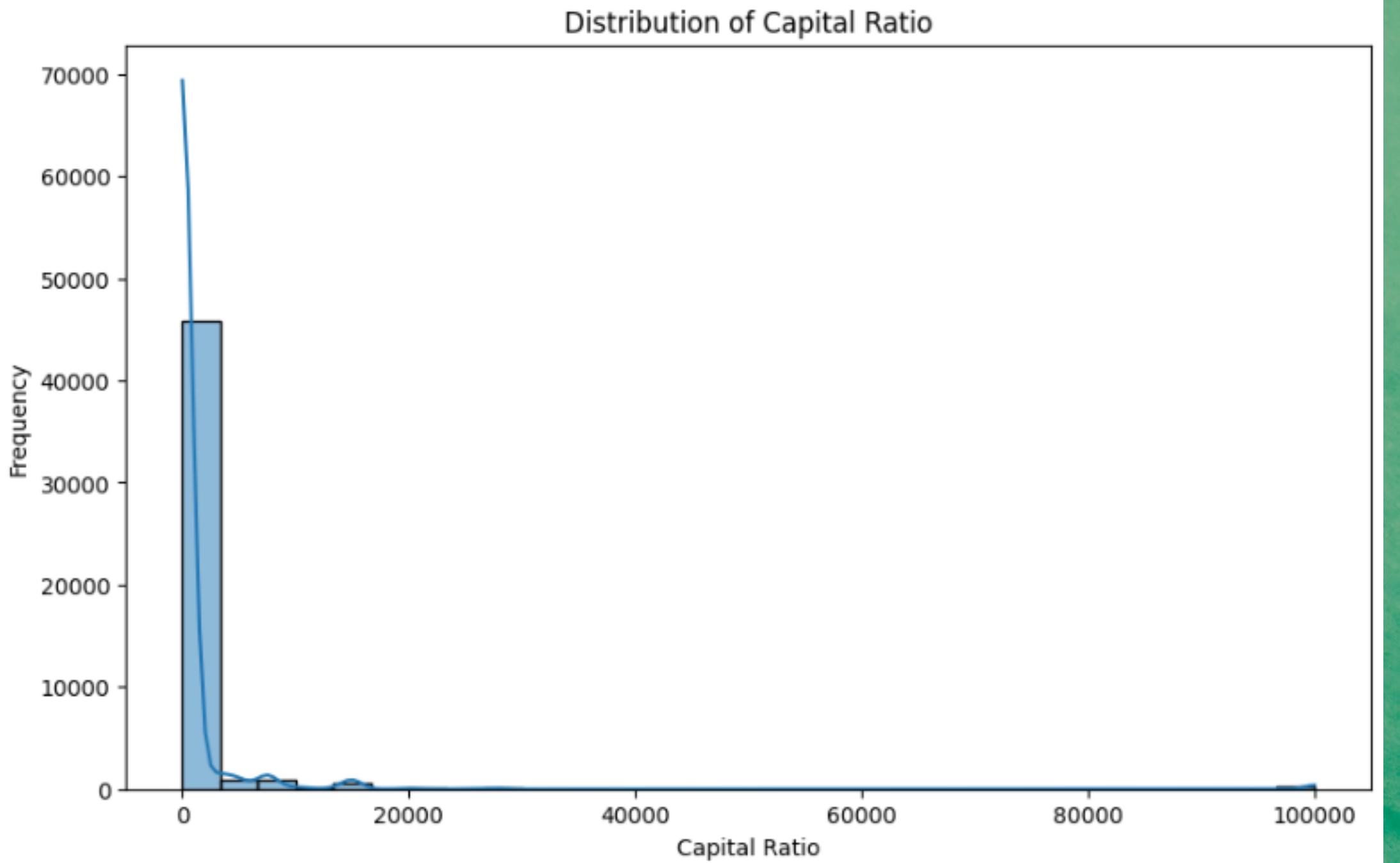
(37,128 out of 48,813) earn less than or equal to 50K

the dataset contains a large proportion of individuals with lower income

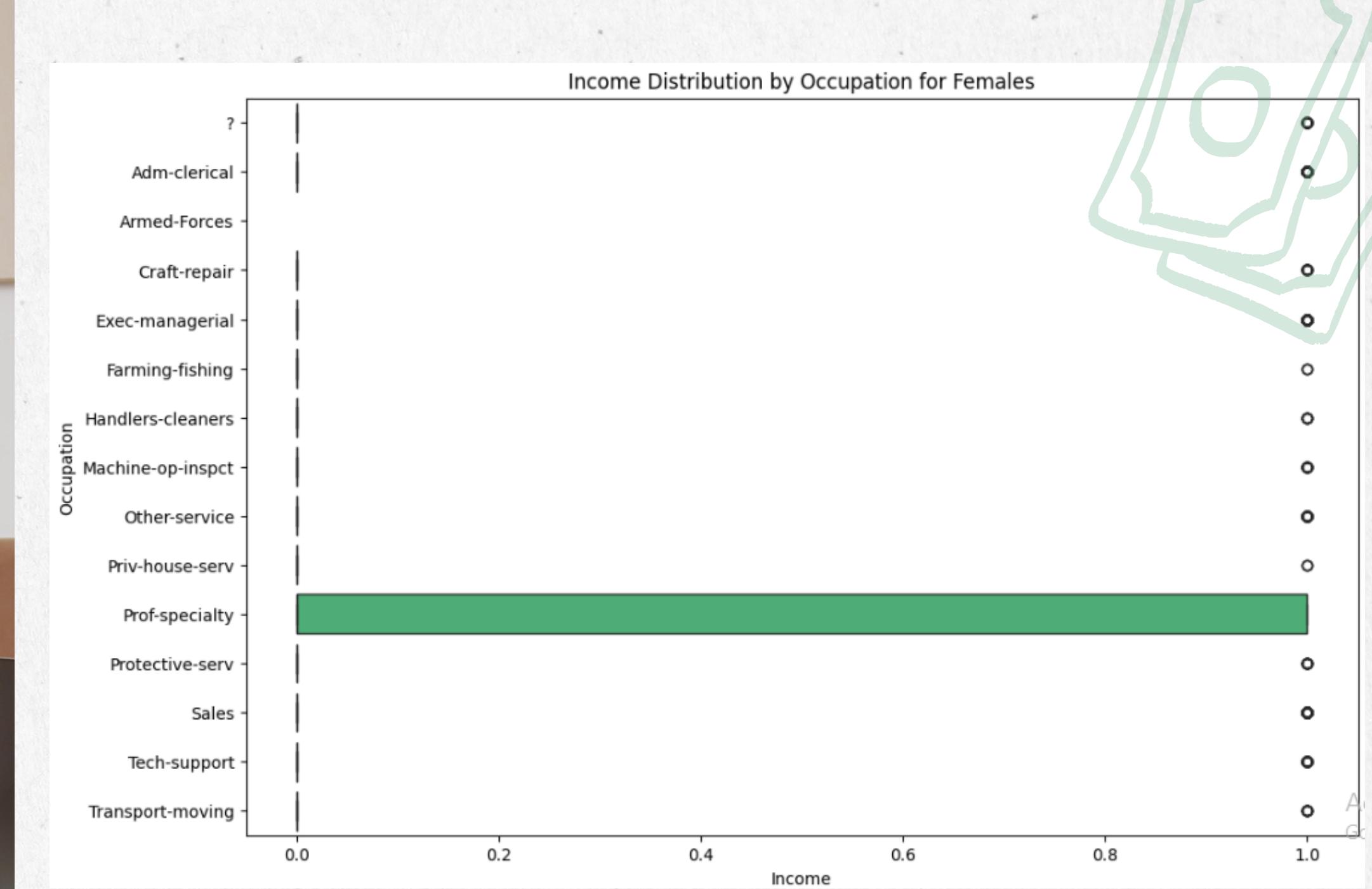
Highest Income by Occupation



The 'Exec-managerial' occupation has the highest mean income among all the occupations in the dataset.



The 'capital_ratio' in the dataset presents insights into the relative balance of capital gains and losses among sampled individuals. Analysis for the first row illustrates a substantial capital gain, indicative of potentially financial activities. For the row that have 0.0, suggesting financial inactivity or unreported information.

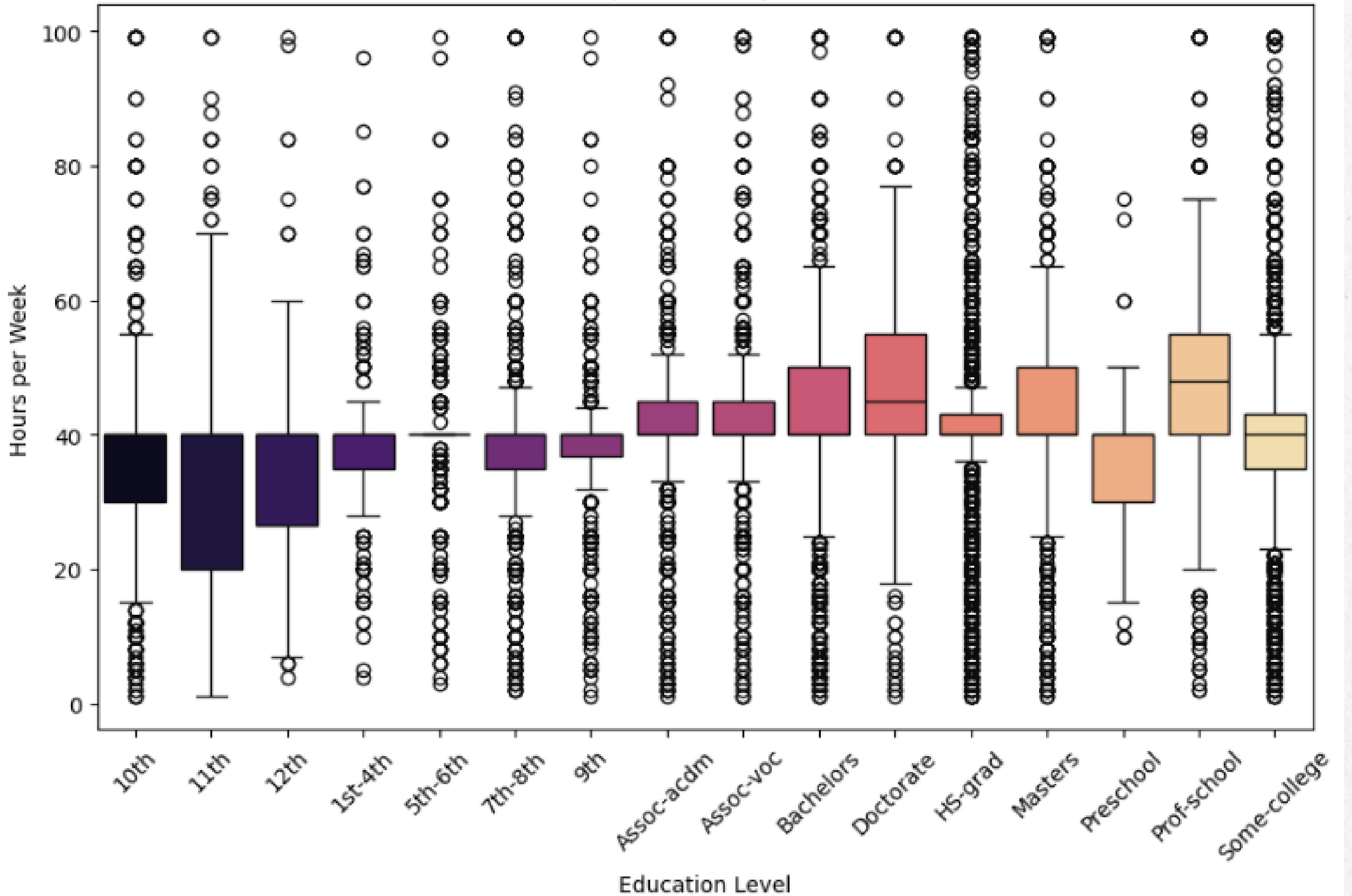


Prof-specialty have the highest income in the occupation for females.



It also implies that there is a correlation between high education level and higher income, as professions requiring advanced education ('Prof-specialty') have a higher income

Hours Worked per Week by Education Level



Work hours seem to increase with higher education levels, suggesting that individuals with more education may be engaged in jobs that require, or allow, more hours of work per week.

Key findings include insights into the distribution of features such as age, education, and work hours, as well as their relationships with income. For instance, box plots and histograms revealed income distributions across different age groups, education levels, and work hour categories. Moreover, summary statistics highlighted differences in income between various groups, providing valuable insights for further analysis.