Visualization Analysis of Income Prediction for UVW College Marketing Campaigns

Ravi Tej Chaparala

Abstract

This report leverages United States Census Bureau data to develop targeted marketing strategies for UVW College, focusing on individuals around the $50,000 income threshold. By analyzing demographic factors like education, occupation, and marital status, we aim to identify key demographics for enrollment strategies.

Goals and Business Objectives

- *Goal***:** Utilize data visualization to identify demographic predictors of income levels.

- *Objective***:** Increase UVW College's enrollment by tailoring marketing to demographics most likely to benefit from advanced education.

Assumptions

- The dataset from the United States Census Bureau accurately represents the broader population relevant to UVW College.

- The $50,000 income threshold significantly influences educational needs and enrollment decisions.

- Demographic factors (age, education, marital status, occupation) remain stable predictors of income.

User Stories

Here are the user stories that were pivotal to this project:

1. *Educational Influence on Income*:

   - As a member of the UVW marketing team, I need to understand how different levels of educational attainment relate to income levels, so we can tailor our marketing approach to emphasize the financial benefits of higher education.

2. *Racial Income Distribution*:

   - The marketing director seeks to identify how income is distributed across races to develop campaigns that are inclusive and resonate with diverse population segments, enhancing the college's appeal across all demographics.

3. *Age, Education, and Income Correlation*:

   - As a data analyst, I am examining the correlation between age and education on income to help the marketing team craft messages that resonate with professionals seeking career advancement through education.

4. *Occupational Impact on Earnings*:

- The marketing team requires insights into how occupation and educational levels combined affect income, which will assist in designing specialized campaigns for various professional sectors.

5. *Capital Gains and Age as Income Indicators*:

    - As part of the analytical team, I want to investigate the impact of age and capital gains on income levels to determine if these factors can predict the likelihood of financial success for potential students interested in financial planning or investment courses.

6. *Marital Status and Gender Income Dynamics*:

    - As a UVW College enrollment officer, it is essential to understand how marital status and gender together influence income so that we can identify demographic segments with distinct educational needs and financial capabilities.

Visualizations and Interpretations
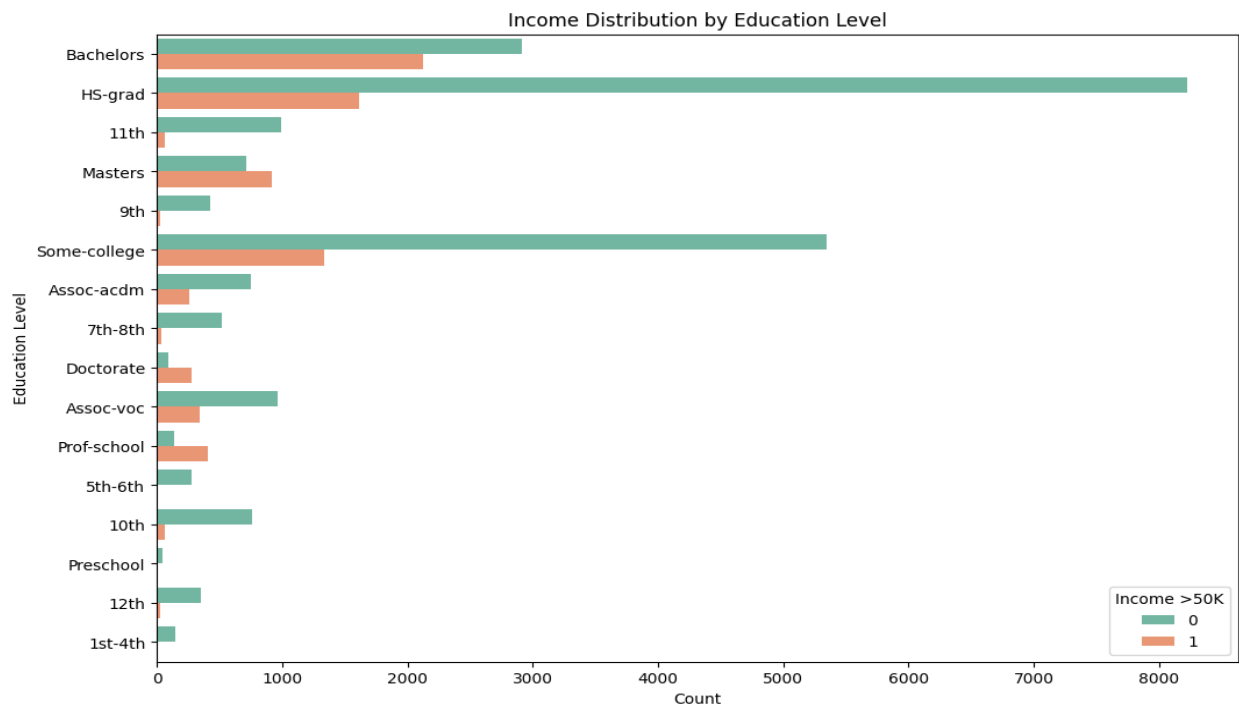
Visualization 1: Income Distribution by Education Level

Fig. 1. Bar chart showing the number of individuals by education level, segmented by income above and below $50K.

- *Demonstration*: This visualization shows the count of individuals by education level, categorized by income greater than or less than $50K.

- *Reason for Creation*: Created to elucidate the correlation between educational attainment and income, it serves as a basis for advocating the value of higher education.

- *Design Process*: We chose a horizontal count plot for clarity and ease of comparison among education levels. The two-tone color palette effectively differentiates income categories.

- *Conclusion*: Higher educational attainment often aligns with higher income levels, emphasizing the importance of UVW College's programs in elevating economic prospects.

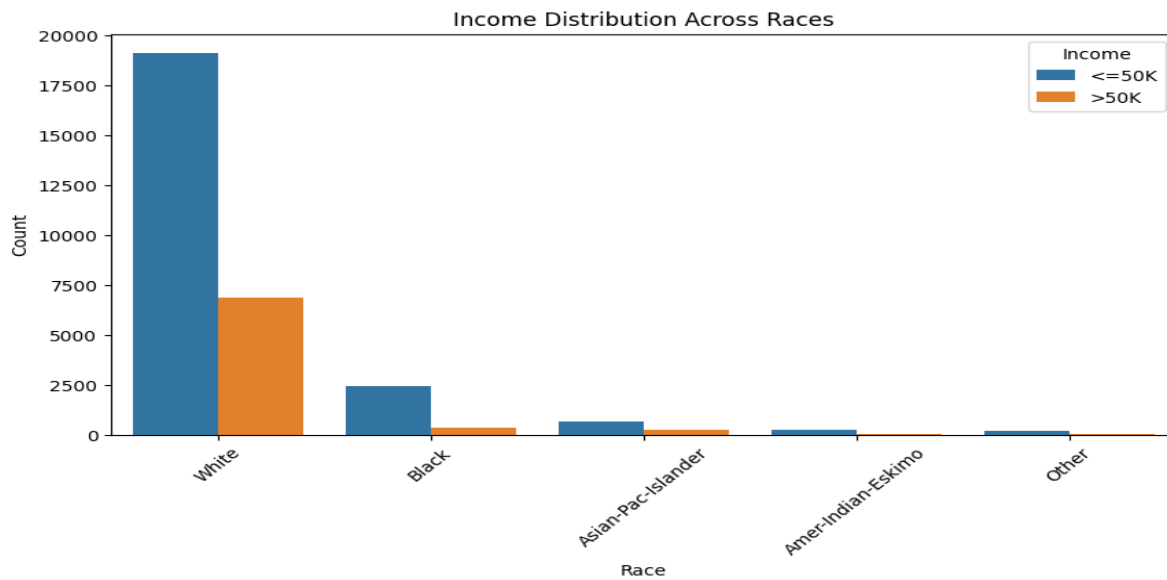*Visualization 2: Income Distribution Across Races*

Fig. 2. Count plot of income distribution across different racial groups, categorized by income threshold.

- *Demonstration*: This count plot compares income distribution across different racial groups.

- *Reason for Creation*: It was designed to highlight racial disparities in income which can inform inclusive marketing.

- *Design Process*: The x-axis represents races, ensuring each category is distinguishable. The count plot allows for a direct visualization of the income disparity across races.

- *Conclusion*: Notable income disparities exist among racial groups, pointing towards a need for marketing strategies that resonate with and empower underrepresented groups.

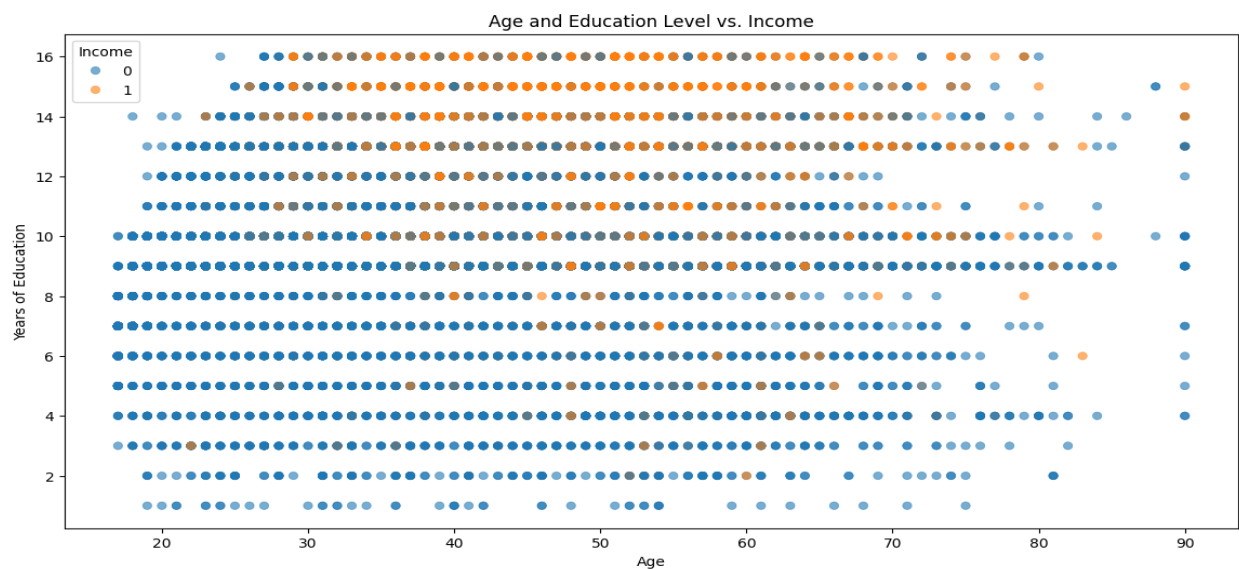*Visualization 3: Age and Education Level vs. Income*

Fig. 3. Scatter plot illustrating the relationship between age, years of education, and income categories.

- *Demonstration*: A scatter plot that visualizes the interaction between age, years of education, and income.

- *Reason for Creation*: To illustrate how age combined with educational investment impacts earning potential.

- *Design Process*: Age and education-num were plotted along the axes with income as the hue. The scatter plot was chosen for its ability to show the distribution and clustering of data points.

- *Conclusion*: There's a positive trend between age plus education and income, supporting the promotion of UVW College's programs to mid-career individuals.

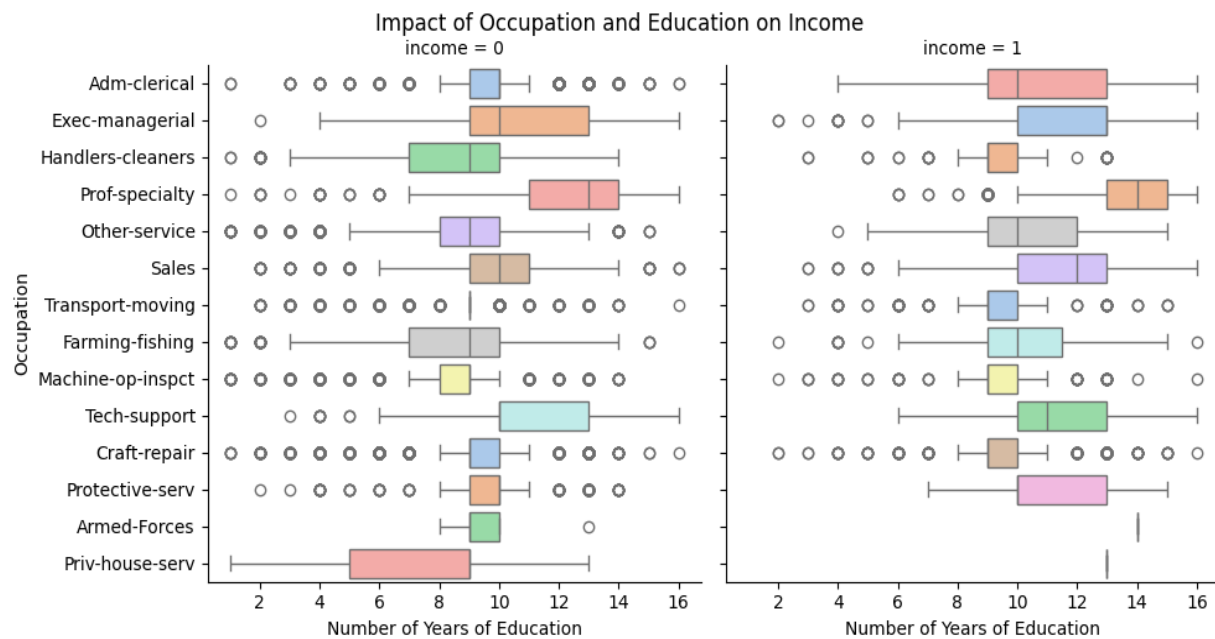*Visualization 4: Impact of Occupation and Education on Income*

Fig. 4. Facet grid of box plots depicting the interaction between occupational fields, education level, and income.

- *Demonstration*: Facet grids show box plots of years of education against occupation, split by income category.

- *Reason for Creation*: To observe the combined effect of occupation and educational attainment on income.

- *Design Process*: Box plots within a facet grid allow for a side-by-side comparison between income categories while detailing the variance within each occupation.

- *Conclusion*: Certain occupations yield higher incomes with higher education levels, indicating the value of targeted educational advancement programs.

Visualization 5: 3D Scatter Plot of Age, Capital Gain, and Income
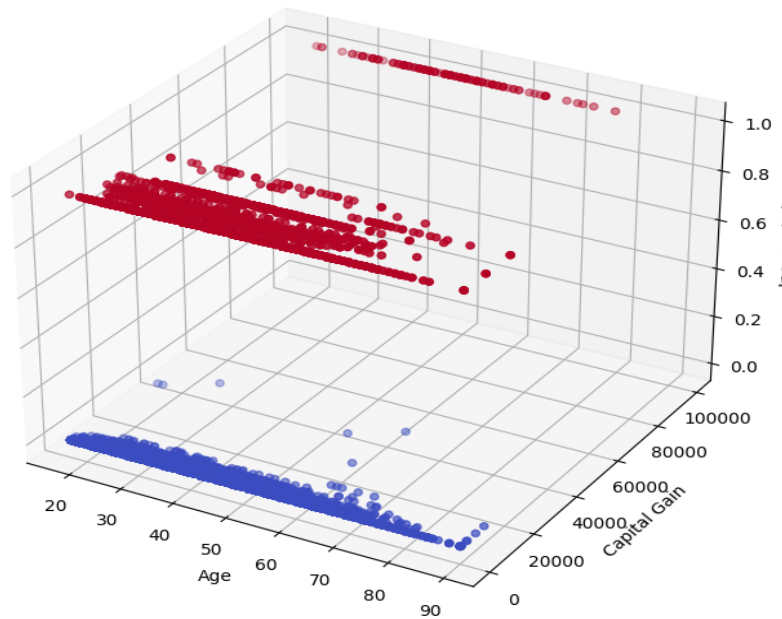
3D Scatter Plot of Age, Capital Gain, and Income

Fig. 5. Three-dimensional scatter plot displaying the correlations among age, capital gains, and income.

- *Demonstration*: A 3D scatter plot displaying the relationship between age, capital gain, and income level.

- *Reason for Creation*: To analyze the multidimensional relationship between age, investment success, and income.

- *Design Process*: A 3D scatter plot was chosen to simultaneously display three continuous variables and their interactions.

- *Conclusion*: Older age and higher capital gains tend to associate with higher income levels, implying an older demographic may benefit from financial advancement courses.

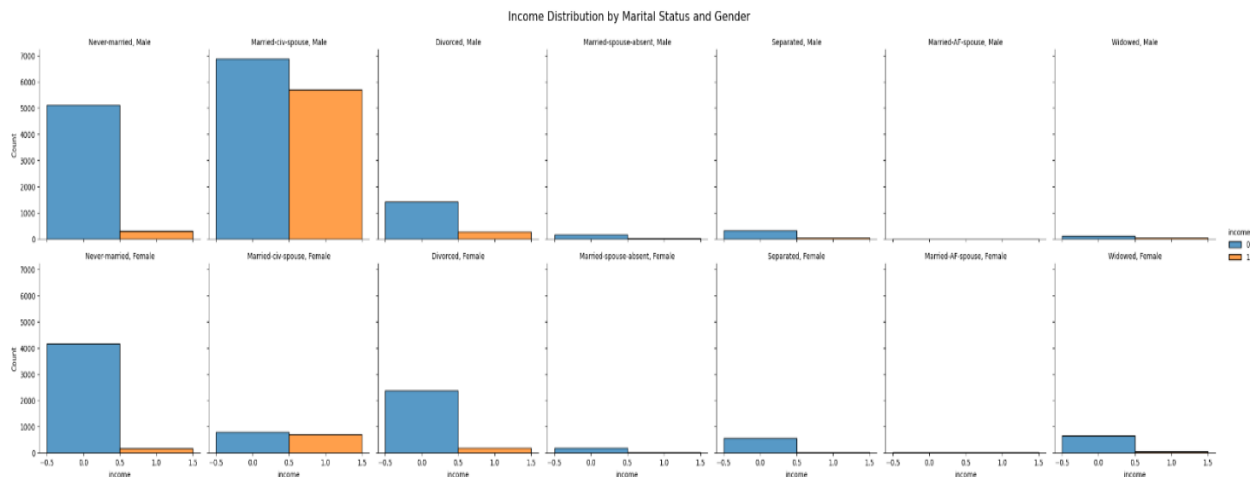*Additional Visualization: Marital Status and Gender vs. Income*

Fig. 6. Histograms within a FacetGrid showing income distribution by marital status and gender.

- *Demonstration*: A series of histograms within a FacetGrid showcasing the distribution of income across different marital statuses and genders.

- *Reason for Creation*: To discern the complex interplay of marital status and gender on income levels.

- *Design Process*: Histograms were plotted within a FacetGrid to provide a multi-faceted view that allows for comparison between the different subgroups.

- *Conclusion*: Marital status and gender significantly affect income distribution, pointing to the importance of personalized marketing approaches for various demographic segments.

Questions and Solutions

- *Data Completeness Concern*: Ensured that missing data handling did not skew results by using robust statistical methods.

- *Visualization Clarity*: Enhanced visualizations with user-friendly designs and annotations to ensure clarity and accessibility.

- *Complex Data Relationships*: Applied multivariate analysis techniques to capture complex interactions between variables.

Future Work

- *Integration of Real-time Data*: Plan to include more dynamic, up-to-date datasets to reflect current economic conditions.

- *Predictive Modeling Enhancement*: Future models will incorporate machine learning to predict income levels more accurately.

- *Development of Interactive Tools*: Aim to create interactive visualizations that allow stakeholders to explore data scenarios in real-time.

Conclusion

This project successfully applied data visualization to analyze and interpret key demographic factors affecting income levels, providing UVW College with actionable insights for targeted marketing strategies. The visualizations developed have laid a foundational understanding, enabling UVW College to tailor its outreach and program offerings more effectively to prospective students' financial and educational backgrounds.

References

1. United States Census Bureau. Adult Income Dataset. Retrieved from https://archive.ics.uci.edu/ml/machine-learning-databases/adult/adult.data.

2. Seaborn and Matplotlib documentation for data visualization.

3. Pandas documentation for data manipulation.