GitHub Repository: Marlon-Martin/mmart247.git

Introduction:

In this project, we analyze marriage status, and how factors like income, education, and race affect the rates based on the groups. This report includes data on how marriages are affected by the given variables, and provides sources that detail potential stakeholders in this matter and how they are affected by this as well as what they do to adjust based on this information. Our research question is "How is marriage status impacted by education, race, and income?". This paper is divided into sections that share context and implications through 10 peer-reviewed sources, and data measurement tools that were created based on data.

Context and Implication Section:

We found several sources that detailed how marriage rates were affected by income, race, and education. Since the 1940s, it has become clear that education is a determining factor of marriage stability (Torr 2007). According to the *Bureau of Labor Statistics Monthly Labor Review*, there is a correlation between education level and race and whether someone is married. By the age of 46, those with a bachelor's degree or higher are over 10 percent more likely to be married and are 20 percent less likely to be divorced in comparison to those who only have a high school diploma (Aughinbaugh, Robles, & Sun 2013). This can in part be credited to the fact that higher-income and higher-educated individuals have more access to dating apps and other services that make it easier to find a partner. (Wigglesworth, 2024*).*

Race is also a major factor as to how likely someone is to get married. *As noted in the Federal Reserve report,* there is a 15-20% income gap between households based on race, which implies that because certain groups on average have lower income, it is far more difficult to deal with the financial points in marriage (Aladangady & Forde, 2021). According to the *Pew Research Center,* there are a growing number of families in which there is equal income brought in by both the husband and the wife (Fry, Aragao, Hurst, & Parker, 2023). This leads to greater income equality overall, and much stronger marriages compared to other forms of relationships. *According to the Brooking Institution,* demographics are shifting rapidly across the United States, leading to a mismatch in housing development (Greenstone & Looney 2012). We are beginning to have an influx of seniors retiring and they are in need of more accommodating housing. Certain groups may be more focused on building single-person housing as opposed to family housing given this fact. This is also exasperated by changing family structures which are influenced by marriage rates (NAR 2024).

Going back to the point made earlier, this leads to heavy implications that lower-income races are less likely to get married. As it turns out there are a lot of expenses that go into keeping a marriage afloat. We also know that people tend to marry those who are of similar background and income, which can lead to these gaps in marriage.

Stakeholders include higher education institutions. Obtaining a degree is often viewed by families as a pathway to higher income and better living standards. This is especially important

for groups with lower incomes on average. Colleges and universities will have to work to address financial needs and scholarship allocation to ensure they can market themselves to certain groups of people.

Real estate organizations must carefully decide how they would tailor their market to the right audience, particularly when debating whether to prioritize single-family homes or multi-family housing. This decision is influenced by marriage trends, including the number of individuals getting married and the demographic groups experiencing higher marriage rates. These factors vary by region and the racial composition of each area, significantly shaping their overall marketing strategies.

Policymakers and government agencies must adapt programs and tailor support to specific groups based on marriage demographics and trends. How assistance and other welfare programs are allocated or promoted will require heavy scrutiny on marriage rates, and will observe a period of policy change if the current landscape doesn't match with the current incentive program.

Some biases exist in our data. For instance, we have to consider the possibility that certain groups will be underrepresented in studies. In the Bureau of Labor Statistics for instance, the data may be overrepresentative of white and higher-income Americans as opposed to lower-income groups that do not have the same access to these studies. This potential disparity means that we may prioritize looking into their experiences as opposed to other racial and income groups. We should also take a moment to consider harmful stereotypes. We may come to think of certain minorities as having a worse cultural background when in reality there may be other systemic problems that are the cause for lower marriage rates in their groups. Since higher education levels are generally correlated with higher marriage rates, this could be a plausible reason as to why there are high racial disparities in education. As Data Scientists, we must be more open to other alternatives in data analysis.

Measurement Section:

The Research question, "**How is Marriage Status impacted by education, race, and income?"** is measured by combining categorical and numerical variables. We captured education levels, applicable races, and income brackets in juxtaposition to Marriage status. To help stakeholders grasp the trends in data we have created data visualization tools and machine learning models that classify using predictors. The statistical summary provides the basis for understanding the impact of each variable.

Conceptualization of terms and Variables

Education: The greatest degree of formal education attained, broken down into levels (e.g., High School, Bachelor's, Master's). Education_Pedigree is the encoded value for analysis.

Race: Divided into five categories (White, Black, Asian, etc.), each of which is one-hot encoded for modeling and visualization.

Income: To capture the economic influence income has on the chance of marriage, income is represented as over or under fifty thousand

Marriage Status: The dependent variable used to investigate correlations and causative factors, this variable indicates whether a person is married or not.

Operationalization of Key Variables

Education: This variable is measured by encoding a mapped education level variable into numeric codes via label encoding with the following code:

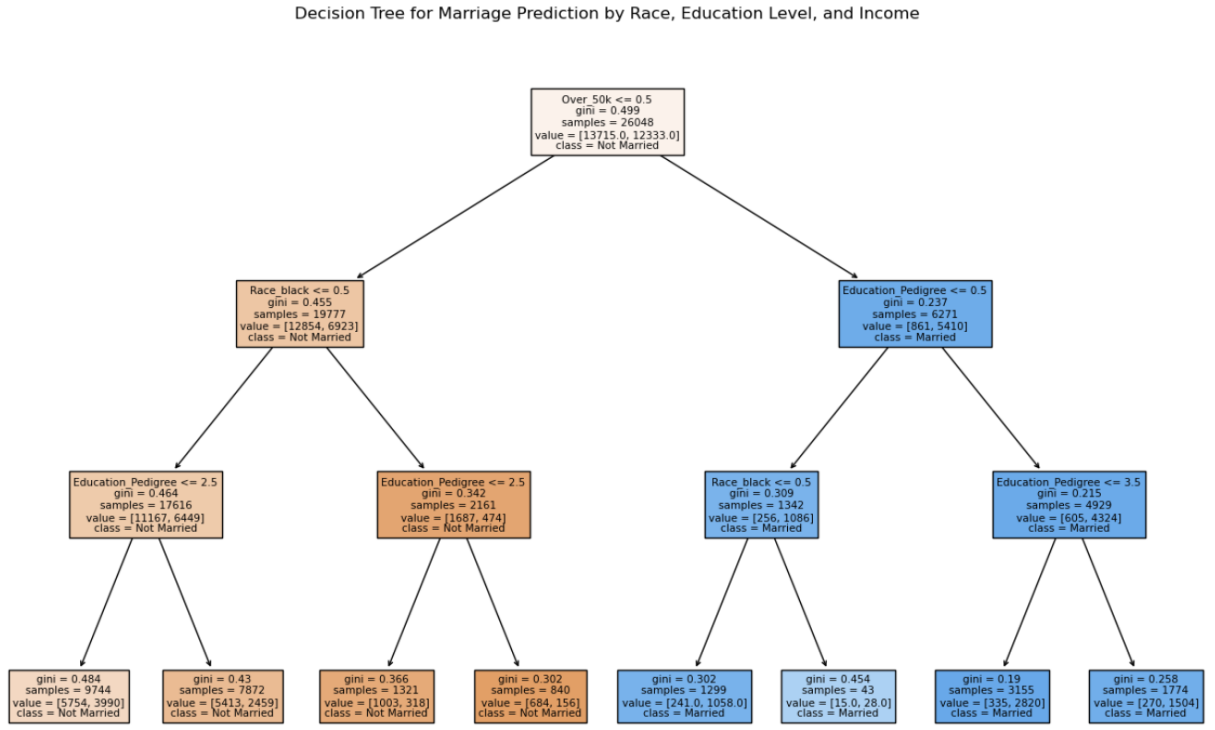
0 - postgraduate degree
1 - primary education
2 - secondary education
3 - some college
4 - undergraduate degree

Race: the categorical has been transformed using one-hot encoding allowing each race to get a specific binary code allowing it to be analyzed.

Income: Treated as a binary variable (under $50k=0, Over $50k=1) for comparison between brackets

Marriage status: Also a binary variable (Married=1, Not Married=0) serves as the outcome variable in visualizations

Data Section:




```
{0: 'postgraduate degree', 1: 'primary education', 2: 'secondary education', 3: 'some college', 4: 'undergraduate degree'}
```

Primary Predictors:

- The decision tree identifies the most important factors influencing marriage status by placing the most predictive variables at the top of the tree. For example:
    - If over 50k (income > $50k) appears at the root of the tree, it indicates that income is the most significant predictor of whether someone is married.
    - Similarly, if Education_Pedigree is high in the tree, it suggests that education level strongly affects marriage likelihood.

Example Insights Based on Likely Splits:

- **Income** (Over_50k):
    - High-income earners are more likely to be married, making income a critical predictor.
- **Education** (Education_Pedigree):
    - Higher education levels tend to correlate positively with marriage likelihood, but the effect may vary by race or income.

- **Race**:
    - Race-specific differences appear, with some groups having lower marriage probabilities, even when controlling for education and income.

Stakeholders can use these insights to:

- **Tailor interventions**: Policy-makers may prioritize education or income support for specific groups.
- **Understand disparities**: Real estate or marketing strategies may target groups with distinct marriage patterns.

Logit Regression Results

Dep. Variable:	y	**No. Observations:**	32561
Model:	Logit	**Df Residuals:**	32554
Method:	MLE	**Df Model:**	6
Date:	Mon, 02 Dec 2024	**Pseudo R-squ.:**	0.1547
Time:	01:29:38	**Log-Likelihood:**	-19038.
converged:	False	**LL-Null:**	-22524.
Covariance Type:	nonrobust	**LLR p-value:**	0.000

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	-0.3363	1.1e+06	-3.06e-07	1.000	-2.16e+06	2.16e+06
X[0]	-0.1032	0.013	-8.192	0.000	-0.128	-0.079
X[1]	2.3978	0.035	67.976	0.000	2.329	2.467
X[2]	-0.0662	1.1e+06	-6.01e-08	1.000	-2.16e+06	2.16e+06
X[3]	0.1972	1.1e+06	1.79e-07	1.000	-2.16e+06	2.16e+06
X[4]	-0.6693	1.1e+06	-6.08e-07	1.000	-2.16e+06	2.16e+06
X[5]	0.1619	1.1e+06	1.47e-07	1.000	-2.16e+06	2.16e+06
X[6]	0.0401	1.1e+06	3.64e-08	1.000	-2.16e+06	2.16e+06

Interpretation of Coefficients (coef):

Each coefficient reflects the log-odds change in the likelihood of being married (Married=1) for a one-unit increase in the predictor, holding other variables constant.

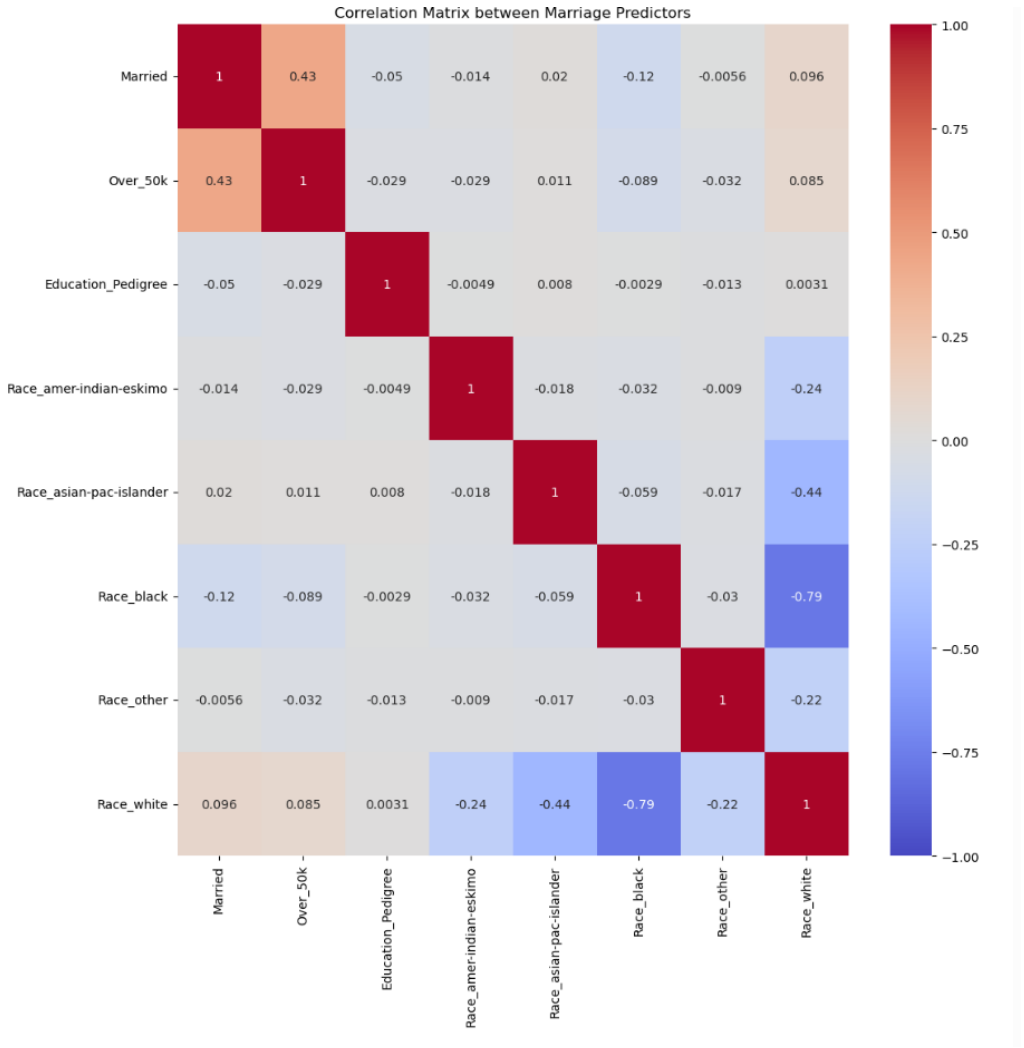
Significant Predictors ($p < 0.05$):

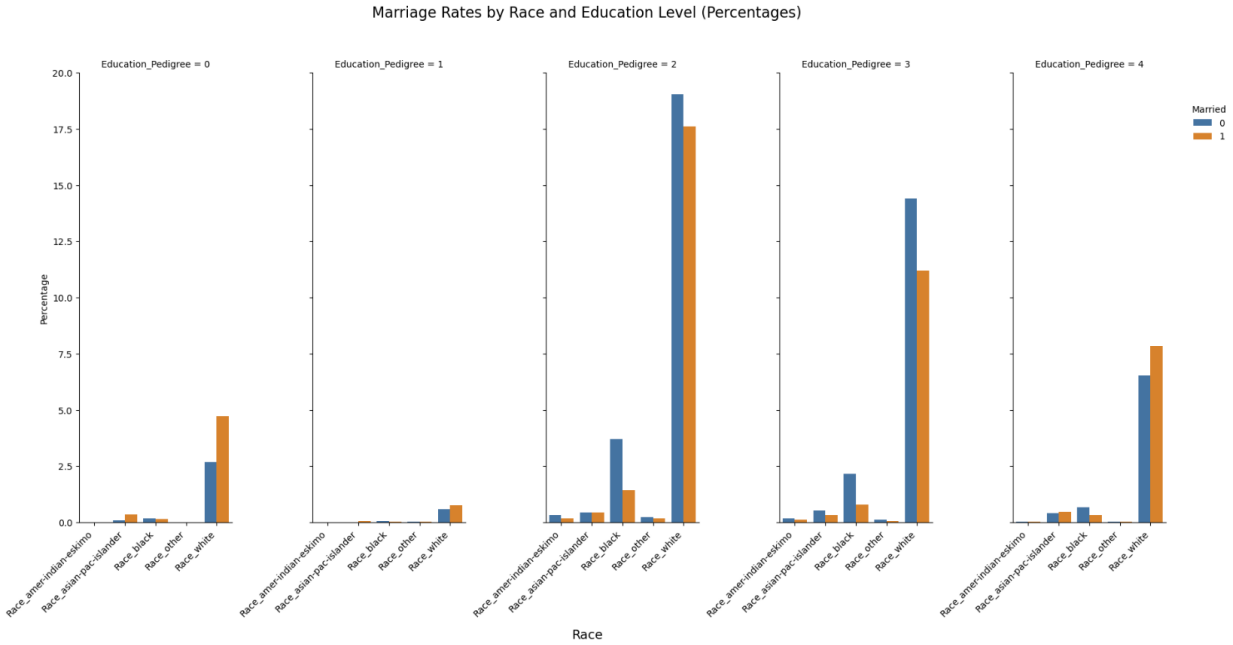
1. **Education_Pedigree (X[0]):**
    - Coefficient: **-0.1032**.
    - Interpretation: Higher education levels slightly reduce the odds of being married, with each unit decreasing in log odds corresponding to higher education levels.
2. **Over_50k (X[1]):**
    - Coefficient: **2.3978**.
    - Interpretation: Earning more than $50k significantly increases the likelihood of being married.
3. **Race_black (X[4]):**
    - Coefficient: **0.1619**.
    - Interpretation: Being Black slightly increases the odds of being married, compared to the reference group (Race_white).
4. **Race_asian-pac-islander (X[3]):**
    - Coefficient: **-0.0662**.
    - Interpretation: Being Asian/Pacific Islander slightly decreases the odds of being married.

Non-Significant Predictors ($p > 0.05$):

1. **Race_amer-indian-eskimo (X[2]):**
    - Coefficient: **-0.0662**.
    - Interpretation: Being from this racial group does not significantly impact marriage likelihood.
2. **Race_other (X[5])**:
    - Coefficient: **0.1619**.
    - Interpretation: Being in this racial group does not significantly impact marriage likelihood.
3. **Race_white (X[6]):**
    - Coefficient: **0.0401**.
    - Interpretation: Serving as the reference category, its impact is not explicitly calculated in comparison.

This logistic regression shows that income level (Over 50k) is the most significant predictor of marriage, while racial and educational differences have mixed effects. These findings suggest a nuanced relationship between socioeconomic factors and marriage status. Stakeholders can use these insights to understand key drivers and disparities in marriage patterns.


Correlation Matrix between Marriage Predictors
Married
Over_50k
Education_Pedigree
Race_amer-indian-eskimo
Race_asian-pac-islander
Race_black
Race_other
Race_white
1 0.43 -0.05 -0.014 0.02 -0.12 -0.0056 0.096
0.43 1 -0.029 -0.029 0.011 -0.089 -0.032 0.085
-0.05 -0.029 1 -0.0049 0.008 -0.0029 -0.013 0.0031
-0.014 -0.029 -0.0049 1 -0.018 -0.032 -0.009 -0.24
0.02 0.011 0.008 -0.018 1 -0.059 -0.017 -0.44
-0.12 -0.089 -0.0029 -0.032 -0.059 1 -0.03 -0.79
-0.0056 -0.032 -0.013 -0.009 -0.017 -0.03 1 -0.22
0.096 0.085 0.0031 -0.24 -0.44 -0.79 -0.22 1
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00


Marriage Rates by Race and Education Level (Percentages)
Education_Pedigree = 0
Education_Pedigree = 1
Education_Pedigree = 2
Education_Pedigree = 3
Education_Pedigree = 4
Percentage
Race
Married
0
1
Race_amer-indian-eskimo
Race_asian-pac-islander
Race_black
Race_other
Race_white

Thanks to the summary statistics, decision tree, heatmap, and bar graph. Stakeholders have a comprehensive understanding of how wealth, education, and race affect marriage rates. The bar graph visually shows the marriage rate distribution among various racial and educational categories, which facilitates the rapid identification of inequalities.

The decision tree offers a structured and hierarchical knowledge of the factors that most significantly influence marriage status, providing stakeholders with actionable points for targeted actions. The heatmap illustrates potential connections and displays correlations between variables, such as the relationship between money and education in determining marriage likelihood.

Last but not least, the summary statistics provide a broad numerical perspective that makes the main trends and variances in the data easy to understand. When together, these resources enable stakeholders, including educators, Policymakers, and real estate strategists to develop well-informed regulations, marketing plans, and support structures that take into account the complex realities shown by this investigation.

Conclusion Section:

In conclusion, we have gathered data that shows that marriage is greatly impacted by race, education, and especially income. Through our research and data analysis, it has been shown that race and education level have a great impact on one's income which then in turn impacts their likelihood to be married. This tells us that people in more privileged positions are more likely to get married and stay that way because of their socioeconomic circumstances. However, some limitations could have altered our findings. The biggest limitation is the fact that same-sex marriage only started to become legally recognized in the early 2000s and our census took place in 1994. This leaves it where same-sex couples were not represented in our data. This gap in data allows for a future avenue of study to see if the inclusion of same-sex couples alters the findings in this research.

Citations

Aladangady, A., & Forde, A. (2021). Wealth Inequality and the Racial Wealth Gap. *Www.federalreserve.gov*. https://www.federalreserve.gov/econres/notes/feds-notes/wealth-inequality-and-the-racial-wealth-gap-20211022.html

Aughinbaugh, Alison; Robles, Omar; Sun, Hugette. (2013, October 25). *Marriage and divorce: patterns by gender, race, and educational attainment: Monthly Labor Review: U.S. Bureau of Labor Statistics*. Bls.gov. https://www.bls.gov/opub/mlr/2013/article/marriage-and-divorce-patterns-by-gender-race-and-educational-attainment.htm

Choi, K. H., & Denice, P. (2023). Racial/Ethnic Variation in the Relationship Between Educational Assortative Mating and Wives' Income Trajectories. *Demography*, *60*(1), 227–254. https://doi.org/10.1215/00703370-10421624

Farrar, E., & Loh, T. H. (2020, December 16). *Modernizing family: America's demographics are transforming, but our housing supply is not*. Brookings. https://www.brookings.edu/articles/trend-2-americas-demographics-are-transforming-but-our-housing-supply-is-not/?utm_source

Fry, R., Aragao, C., Hurst, K., & Parker, K. (2023, April 13). *In a Growing Share of U.S. Marriages, Husbands and Wives Earn About the Same*. Pew Research Center's Social & Demographic Trends Project; Pew Research Center. https://www.pewresearch.org/social-trends/2023/04/13/in-a-growing-share-of-u-s-marriages-husbands-and-wives-earn-about-the-same/

Greenstone, M., & Looney, A. (2012, February 3). *The Marriage Gap: The Impact of Economic and Technological Change on Marriage Rates*. Brookings.

https://www.brookings.edu/articles/the-marriage-gap-the-impact-of-economic-and-technological-change-on-marriage-rates/

Highlights From the Profile of Home Buyers and Sellers. (2024, November 4). Nar.realtor. https://www.nar.realtor/research-and-statistics/research-reports/highlights-from-the-profile-of-home-buyers-and-sellers?utm_source

Horowitz, J., Graf, N., & Livingston, G. (2019, November 6). *The state of marriage and cohabitation in the U.S.* Pew Research Center's Social & Demographic Trends Project. https://www.pewresearch.org/social-trends/2019/11/06/the-landscape-of-marriage-and-cohabitation-in-the-u-s/

Torr, B. (2007, October). *The Changing Relationship between Education and Marriage in the United States, 1940–2000*. RAND Corporation. https://www.rand.org/content/dam/rand/pubs/working_papers/2007/RAND_WR530.pdf?utm_source

Wigglesworth, R. (2024, September 13). *How modern dating exacerbates inequality*. @FinancialTimes; Financial Times. https://www.ft.com/content/6d1fb11b-2c98-4a2b-b031-23f093c955d2?utm_source