Analysis Report

Impact of Mental Health on Marriage in the USA (2021/2022)

1 Introduction

Marriage is a critical component of society institutions, serving as the foundation for family systems and social stability. However, mental health illnesses, which impact around one in every five persons in the United States each year, have emerged as key elements that may influence marriage dynamics. Understanding the link between mental health and marriage rates is critical for identifying hurdles that may prevent people from developing long-term marriages.

This report investigates the association between state-level mental health metrics and average marriage rates across the United States. By focusing on mental health conditions such as Serious Mental Illness (SMI) and suicidality factors hence in the end of this report a conclusion will be drawn to answer the main question

- Does mental health, specifically mental illness and suicidality, impact marriage rates in different USA states?

2 Used Data and Pipeline output

As a reminder in table 1 the original data as shown in table 1

Dataset 1: Marriage Rates by State (2019-2022)	Mental Illness and Suicidality (2021-2022)
National Center for Health Statistics This dataset shows the number of marriages per 1,000 people in each U.S. state for 2021/2022	National Survey on Drug Use and Health with - Confidence Intervals 95% Age Groups: Includes specific data for 12-17, 18-25, 26+, and 18+. 7 factor scope "Any Mental Illness," "Serious Mental Illness," "Received Mental Health Treatment," "Major Depressive Episode," "Thoughts of Suicide," "Suicide Plans," and "Attempted Suicide."

The ETL pipeline produces a structured SQLite database containing one table *mental_Marrige_Data* . The tabular data has 24 columns and 51 rows. As shown in figure 1 .

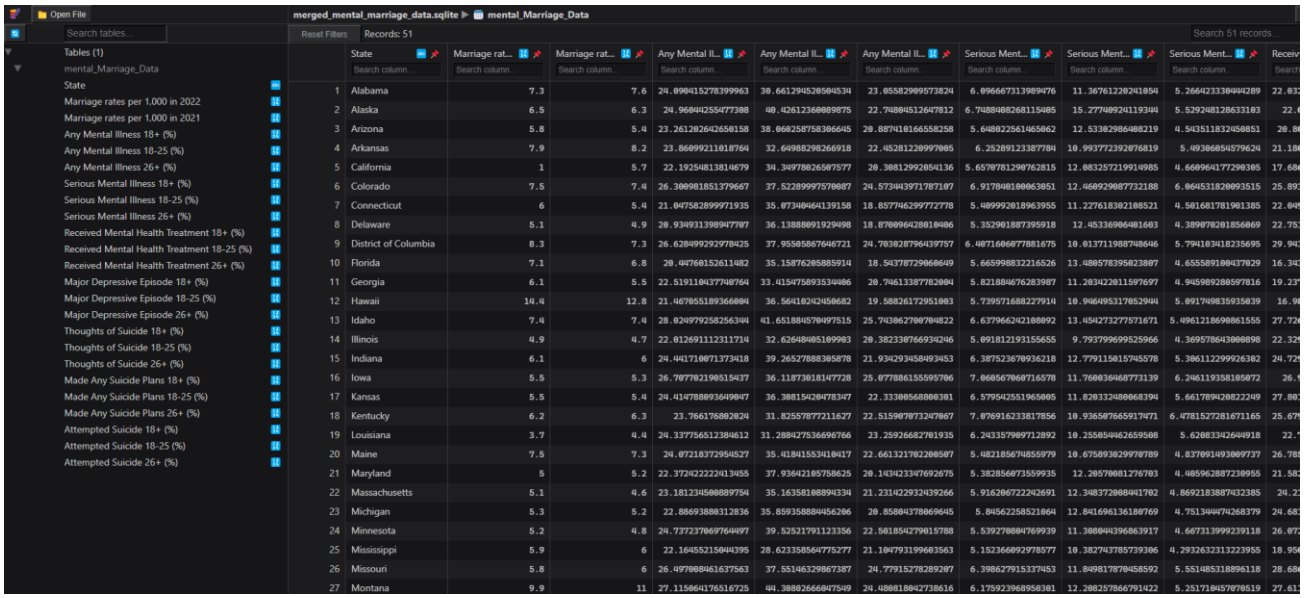


Figure 1 the output from the pipeline (cleaned data)

The Table will show the datatype and what is the expectation for the analysis

The "State" column identifies each state by name. The "Marriage rates per 1,000 in 2022" and "Marriage rates per 1,000 in 2021" columns, expressed as rates per 1,000 individuals, measure the frequency of marriages relative to the population size. While factors like Mental health-related columns include "Any Mental Illness 18+ (%)" and its age-specific counterparts, "Any Mental Illness 18-25 (%)" and "Any Mental Illness 26+ (%)," which represent the percentage of individuals within these age groups who reported experiencing any mental health conditions. Additional columns provide details on more severe conditions, such as "Serious Mental Illness 18+ (%)" and its age-specific breakdowns, indicating the proportion of individuals with significant mental health impairments. Furthermore, "Major Depressive Episode 18+ (%)" and its age-specific versions. Suicide-related metrics are covered in columns such as "Thoughts of Suicide 18+ (%)" and its age-specific versions, as well as "Made Any Suicide Plans 18+ (%)" and "Attempted Suicide 18+ (%)," with similar age-based details. Hence the output is a combination of percentages for factors and rates per 1,000 people for marriage data

Data license and rules: The datasets are open licensed as discussed deeply in the previous report with privacy protection rules regarding the inclusion of people identity one more condition to comply with is not to use the data in any commercial use.

3 Analysis

3.1 Some Conditions for Accurate Analysis

While determining the best approach to answer the question, several challenges were faced:

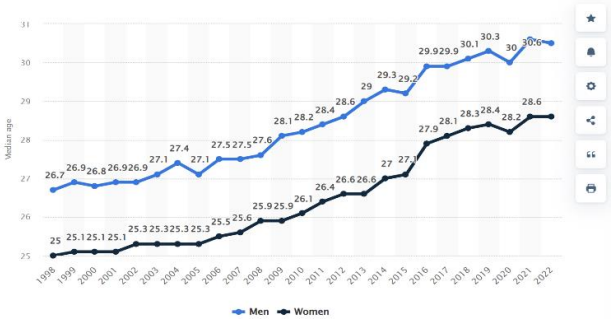


Fig2: Estimated median age of Americans at their first wedding in the US from 1998 to 2022, by sex

1. **Age Group Selection**: The original dataset included four age groups. However, the median age for marriage in the U.S. is 30 for men and 28 for women,
2. making the age group 26+ most relevant for analysis.
3. **Limited Temporal Scope**: Only two years of data (2021 and 2022) were available. To mitigate temporal noise, the average marriage rate was calculated for these years.
4. **Influence of External Factors**: Marriage rates are affected by socioeconomic conditions, cultural norms, and religious diversity, especially in a migrant-heavy country like the U.S.

To solve these problems

1. **Age Group Focus**: Restrict the analysis to adults aged 26+ to align with marriage demographics.
2. **Noise Reduction**: Analyzed based on state-level data instead of yearly trends, accounting for higher sample size based on the state (52 state is better than 2 year).
3. **Simplification**: Computed the average marriage rate for 2021 and 2022.

3.2 Methodology

To answer the main question of the project these steps are followed in sequence to draw an accurate conclusion:

1. **Set Hypothesis**:

($H0$): There is no significant correlation between marriage rates and mental health metrics or Suicidality Factors . Any observed correlation is due to random chance.
($H1$): There is correlation between marriage rates and mental health metrics or Suicidality Factors. this relationship is unlikely to be due to random chance.

2. **Heatmap Visualization**: Generating a heat map to display the relationships between Average Marriage Rate and mental health metrics for (adults aged 26+) as shown in figure 3.

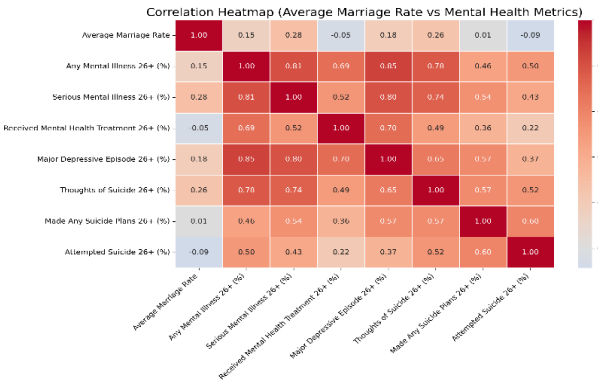


Fig 3: Correlation matrix (Heatmap correlation)

3. Geographic Analysis:

Choropleth maps were created to visualize the geographic distribution of both marriage rates and mental health metrics and suicidal metrics.

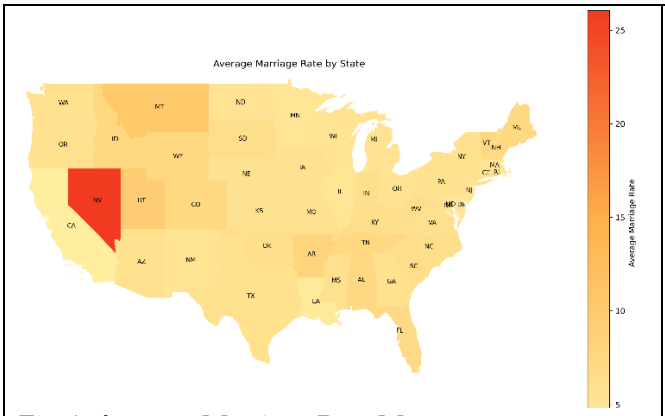


Fig 4: Average Marriage Rate Map

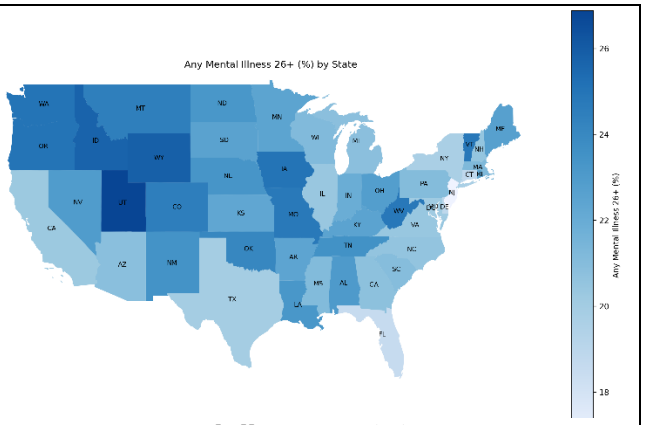


Fig 5: Any Mental Illness 26+ (%) Map

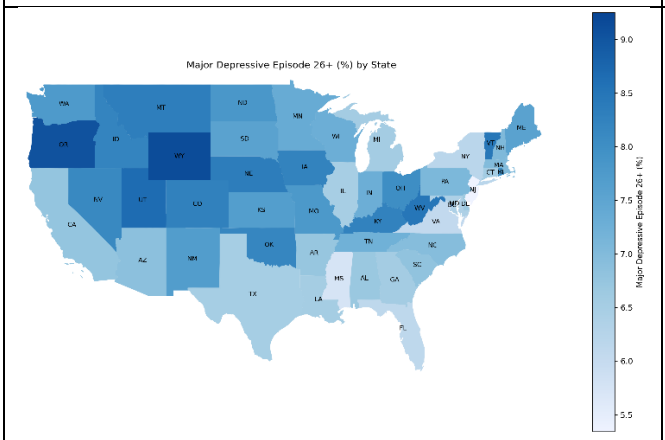


Fig 6: Major Depressive Episode 26+ (%) Map

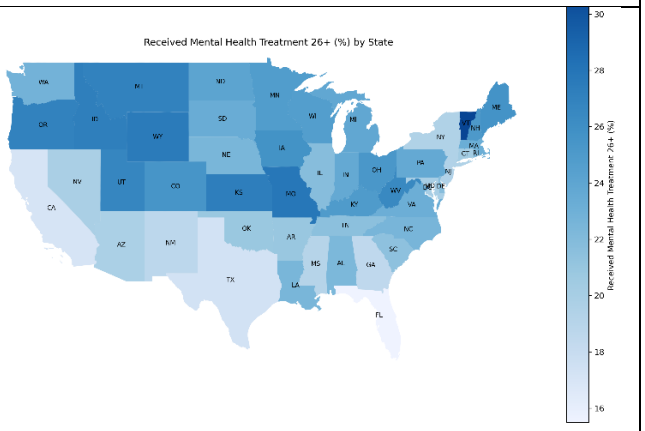


Fig 7: Received Mental Health Treatment Map

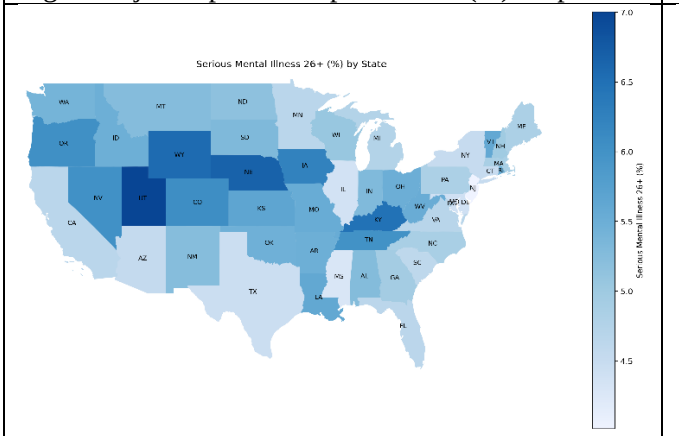


Fig 8: Serious Mental Illness 26+ (%) Map

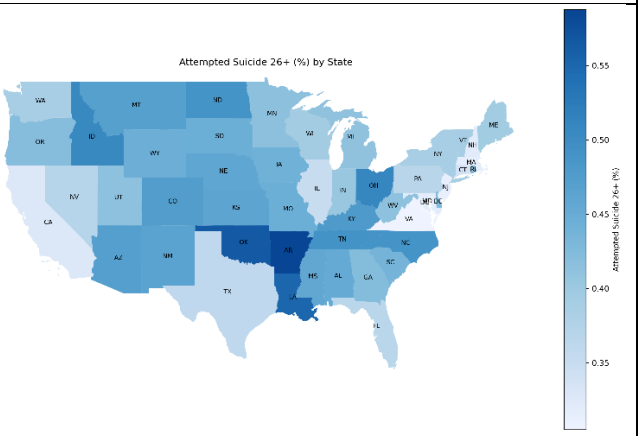


Fig 9: Attempted_Suicide_26+_ (%)_Map

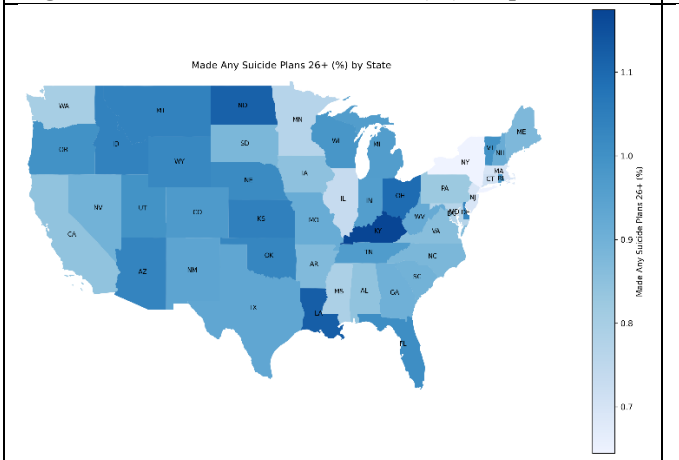


Fig 10: Made Any Suicide Plans 26+ (%) Map

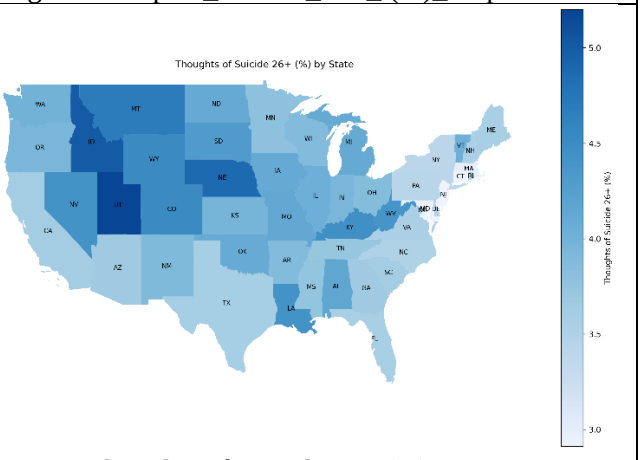


Fig 11: Thoughts of Suicide 26+ (%) Map

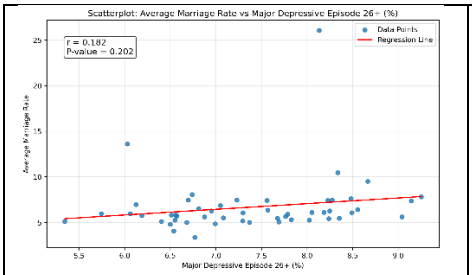
4. **Correlation Analysis**:
    a. Pearson correlation coefficients were computed to measure the linear relationship between **Average Marriage Rate** and mental health metrics.
    b. P-values were calculated to assess the statistical significance of these correlations.

    Formulas Used:
    c. Correlation Coefficient (r): Measures the linear relationship and t-statistic for correlation.
    d. Then **Degrees of Freedom (df)** selected as df =n−2 = 50-2 =48**:** 50 is the number of states
    e. **P-value:** Computed from the t-statistic using the t-distribution with degrees of freedom.

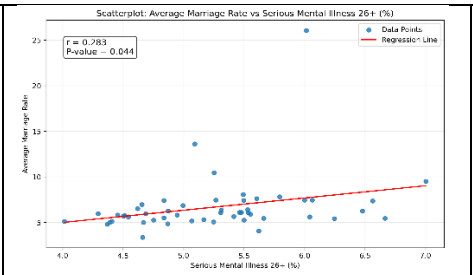
    Result: Top 3 Strongest Correlations:

Factor	Correlation	P-value
Mental Health Metric: Serious Mental Illness 26+ (%)	0.283	0.044
Thoughts of Suicide 26+ (%)	0.256	0.070
Mental Health Metric: Major Depressive Episode 26+ (%)	0.182	0.202

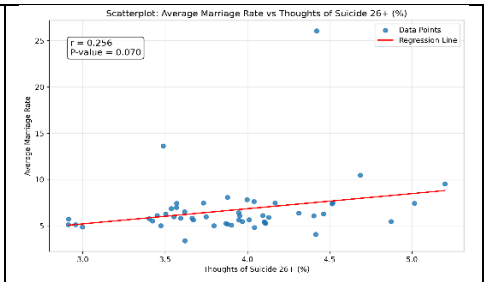
5. **Regression Analysis**:
    a. with **Average Marriage Rate** as the dependent variable and individual mental health and suicidal metrics as independent variables to relate the r square factor.
    b. Scatterplots were generated with regression lines to illustrate these top 3 relationships.



Thoughts of Suicide 26+ (%)
R-squared:0.065
P-value:0.07
Rejected



Serious Mental Illness 26+ (%)
R-squared:0.08
P-value:0.0445



Major Depressive Episode 26+ (%)
R-squared:0.03
P-value:0.2
Rejected

4 Result summary & Key Findings

- geo-visualizations provided additional insights by illustrating the spatial distribution of marriage rates and mental health metrics across the U.S. states. Notably, there is no similarity or contradictory map which is hard to see a relationship between each factor and marriage rate
- There is no conclusive evidence to reject the null hypothesis for most metrics, except Serious Mental Illness 26+ (%), which showed a moderate positive correlation.
- Mental health metrics have limited explanatory power for marriage rates, as evidenced by low R-squared values.
- Correlation does not imply causation; the observed relationships could be influenced by unmeasured variables such as socioeconomic factors or cultural norms.

5 Adjusted Hypothesis:

There is no significant correlation between marriage rates and mental health metrics or suicidality factors. Any observed correlation is likely due to random chance, except for Serious Mental Illness 26+ (%), which warrants further study.

6 Conclusion and discussion

As Mental health and suicidal metrics and marriage rates in the U.S., revealing a surprising moderate positive correlation between Serious Mental Illness 26+ (%) and marriage rates. This result challenges conventional expectations of a negative relationship that comes from logical perspective if there are people who have serious mental health it is usually hard for them to find partners etc., That suggests potential nuances that merit further exploration. However, low R-squared values indicate that mental health metrics are weak predictors and are not explanatory of marriage rates overall.

In future work while expanding the temporal scope for more accuracy for e.g. from 1990 to 2024. That will be huge work in term of data collection, but it will help in drawing more accurate conclusion also to clarify the unexpected positive correlation and enhance the understanding of mental health's role in marital trends.