The Impacts of Sociodemographic Factors and Economic Development on Marriage: A Multivariate Regression Analysis

Bashiru Mukaila

2025-11-18

1 Introduction

Marriage is regarded to be an important social institution in many cultures across different countries in the world, and its rate has been a topic of interest for researchers for many decades. In recent years, there has been a growing interest in understanding the factors that influence marriage rates across different countries. This research aims to examine the relationship between marriage age, schooling year, GDP per capita growth, GDP per capita, and mental disorder. And the impact of these variables on marriage rates and marriage age. The motivation behind these variables is driven by the fact that economic conditions, education, and mental health are all factors that have been shown to affect people's decisions regarding marriage. Marriage age and School year are expected to have a negative relationship with marriage rates, as higher levels of education and older age at marriage have been associated with delayed marriage. Also GDP growth and GDP per capita are expected to have a positive relationship with marriage rates, as economic growth and higher income levels are often linked with increased marriage rates. Lastly, mental disorder is expected to have a negative relationship with marriage rates, as individuals with mental health issues may experience difficulties in forming and maintaining relationships.

All data for this research are collected from the 'Our World in Data' website. The data been merged in each column is the average number of all raw data after the year 2000. Section 2 consists of an exploratory data analysis of marriage data and explores the potential relationship between marriage rate and marriage age, education level, GDP per capita, speed of GDP growth, rate of mental disorders and the age structure.

2 Exploratory data analysis

Figure 1 shows the marriage rate and marriage age in some selected countries. The average values of the data after year 2000 represent the situation of each country, supported by three

reasons as follow: Firstly, the slope of the distribution curve was relatively stable. Secondly, the stability of using average values provided a reliable measure. and thirdly, missing data presented a challenge to drawing conclusions. Additionally, the use of average values after 2000 ensured the data was more recent and relevant, while minimizing the impact of data fluctuations. We can see that there is generally a decreasing trend in marriage rate overtime while there is an increasing trend in the first marriage age. This implies that on average people wait longer to venture into marriage in recent times.

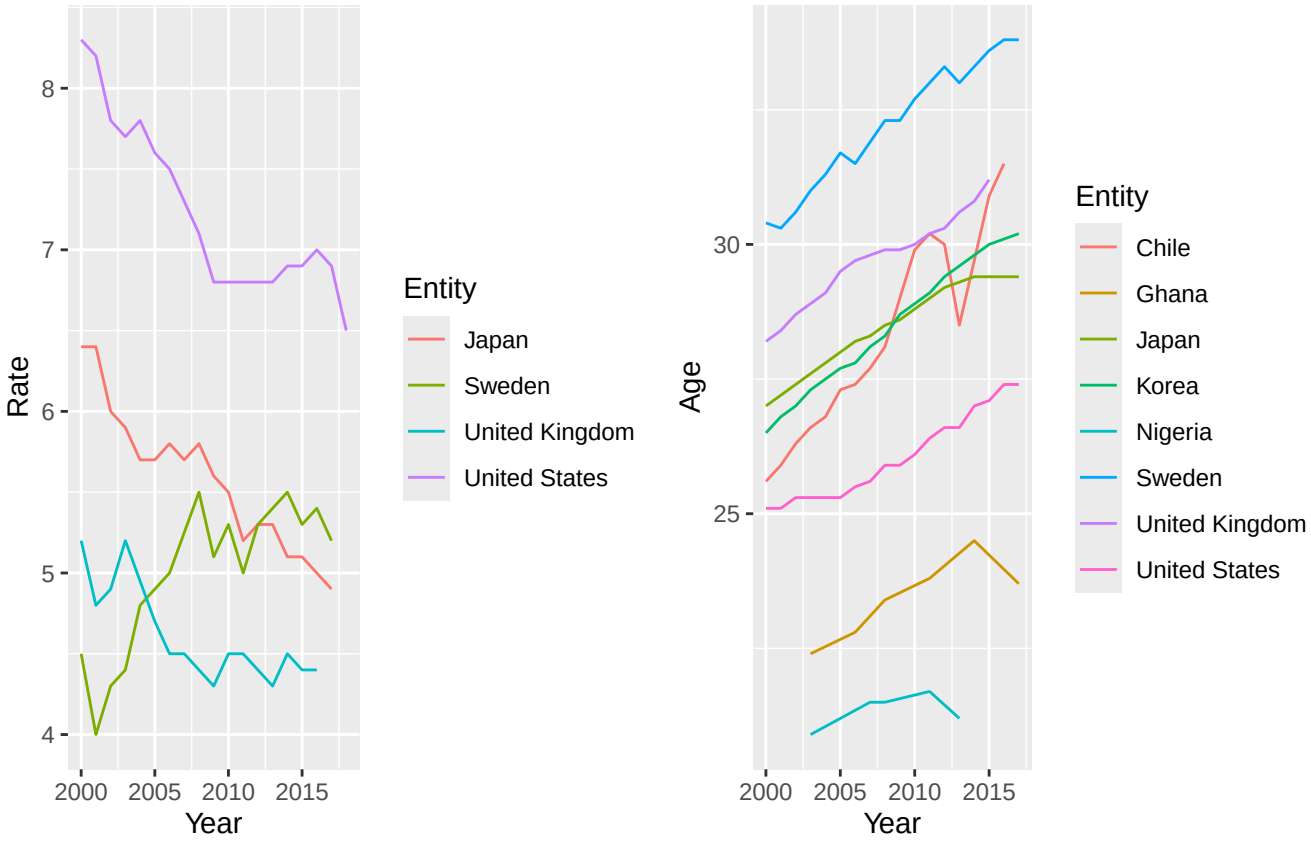

Figure 1: Change of marriage rate (left) and the marriage age (right) of some countries after 2000 .

Table 1 below displays the first five rows of the data. The marriage rate is the number of marriages per 1,000 people in the population. The first marriage age refers to the average age of women at their first marriage. The median age divides the population into two parts of equal size. The schooling year is a common measure to show the population's education level in a country. The unit of the GDP per capita in dollar, and the unit for the growth speed of GDP is percentage. The mental disorders refers to share of population with any mental health such as depression and anxiety.

Table 1: The first 5 rows of the data

Entity	MarriageRate	FirstMarriageAge	Schoolingyear	Medianage	GDPpercapita	GDPgrowth	Mental_Disorders
Albania	7.792857	24.04000	9.327778	32.14091	10137.72	4.3145433	11.6135
Algeria	9.050000	29.13333	7.105556	24.85000	10724.76	1.1450594	14.7905
Argentina	3.033333	24.60000	9.427778	28.95455	21392.59	0.3674278	13.6800
Armenia	5.490909	23.47500	11.127778	31.39545	9326.25	6.1320920	11.4490
Australia	5.311111	27.80000	12.266667	35.87727	44672.38	1.2917527	19.0775

Table 2 contains summarized statistical data on the marriage rate and marriage age of 105 countries all over the world. First, it is noticeable that the maximum marriage rate is nearly 7 times of the minimum marriage rate. Secondly, the middle 50% of the median age lies between 25.23 and 38.34, with an average age of 31.36. The 3rd quantile of the two GDP data are both around quadruple of the 1st quantile. And also, it is remarkable that the GDP per capita shows a huge range between each countries. The proportions of population suffering from mental health disorder ranging between 9.6 to 19.1.

Table 2: The summary statistics of marriage rate, Marriage age, schooling year, median age, GDP per capita, GDP Growth and mental disoder

MarriageRate	FirstMarriageAge	Schoolingyear	Medianage	GDPpercapita	GDPgrowth	Mental_Disorders
Min. : 2.300	Min. :18.73	Min. : 4.578	Min. :17.14	Min. : 2460	Min. :-1.8764	Min. : 9.625
1st Qu.: 4.533	1st Qu.:23.32	1st Qu.: 8.072	1st Qu.:25.23	1st Qu.: 10294	1st Qu.: 0.8235	1st Qu.:11.575
Median : 5.482	Median :25.50	Median : 9.878	Median :31.40	Median : 18353	Median : 1.5846	Median :13.552
Mean : 6.154	Mean :25.69	Mean : 9.613	Mean :31.36	Mean : 25558	Mean : 2.0744	Mean :13.683
3rd Qu.: 7.364	3rd Qu.:28.12	3rd Qu.:11.328	3rd Qu.:38.34	3rd Qu.: 37727	3rd Qu.: 3.3541	3rd Qu.:15.075
Max. :14.200	Max. :32.23	Max. :13.361	Max. :44.40	Max. :111599	Max. : 8.0821	Max. :19.078

Figure 2 illustrates the spread of average marriage rate and marriage age of the 105 countries in a map. As we can see in the plot on the left, darker blue means the country has a lower marriage rate and lighter blue means a higher first marriage rate. It is significant that, the countries in Asia have higher marriage rates and the countries in other continents like South America and North America have relative lower marriage rates.

The figure on the right hand side illustrates the spread of average marriage age. It is noticeable that the countries that have a relative higher marriage rate usually have lower marriage age.

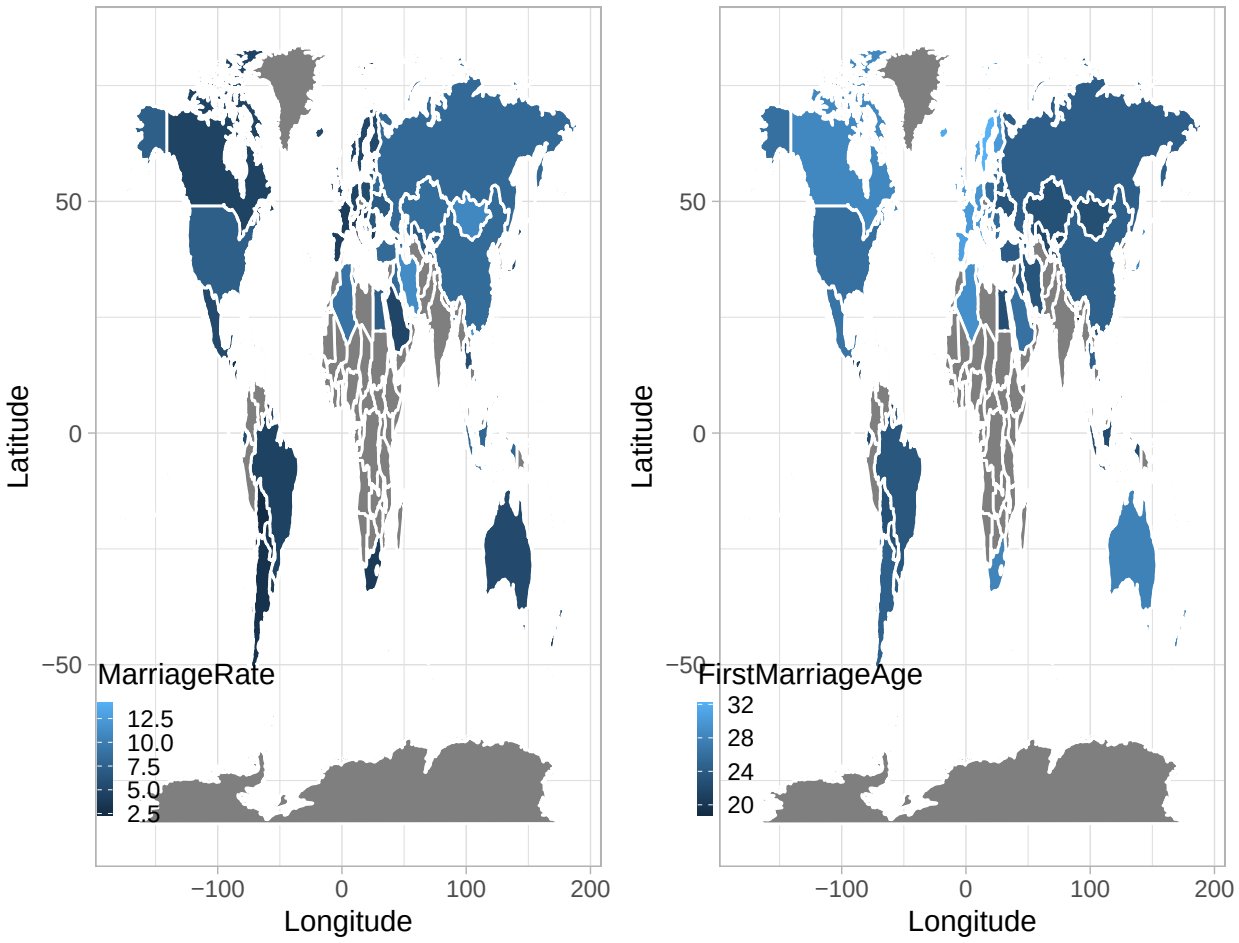

Figure 2: Marriage Rate(left) and Marriage Age(right) Distribution in World map.

Figure 3 displays correlations (as scatterplots), distribution and also prints correlation coefficient between 2 response variables and 5 explanation variables.

It is recognizable that each explanatory variable has a moderate or even strong correlation relation with first marriage age (as we can see from the significant coefficient in the 2nd line). Thus, we can construct a multiple linear regression model with these variables. AS for marriage age and other potential explanatory variables, GDP per capita has the most significant correlation coefficient. It is notable that marriage age has significant relationships with all other values, which means a single linear model is the most effective method to fit the model. GDP per capita is highly correlated with all other explanatory variables, then multicollinearity could be a problem when constructing multiple regression model. Hence, we may consider to remove GDP per capita when fitting the data with multivariate linear model.

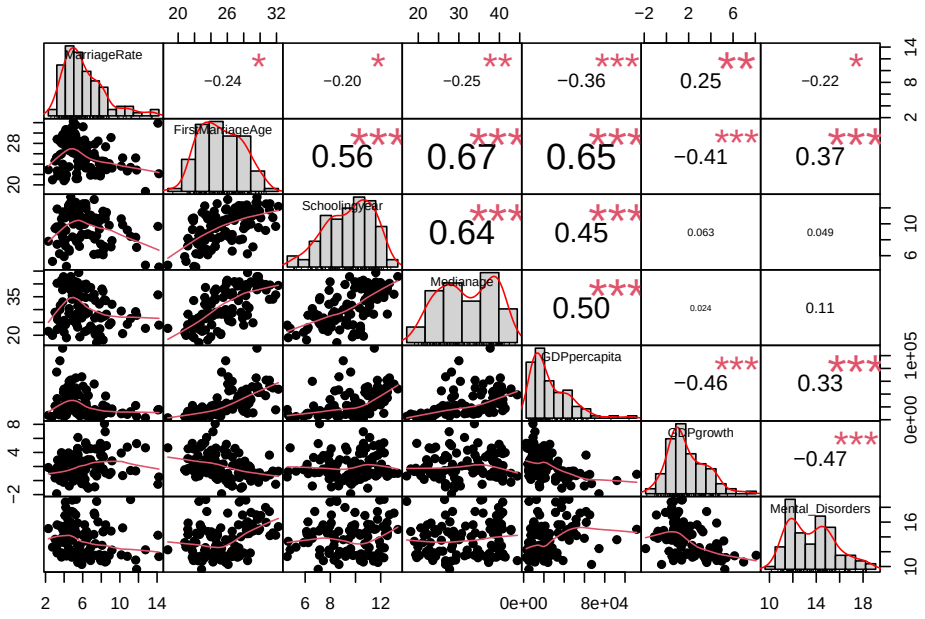

Figure 3: Scatterplot matrix and Correlation.