

Factors That Impact Marriage Duration

Sungju Kim
Barry Lin
Ashley Nguyen
April 4, 2022

1. Introduction & Research Question

What are the factors that impact the marriage duration of divorced couples in Mexico?

We aim to investigate the correlation between the marriage duration of divorced couples in Mexico as the dependent variable and the demographic factors such as partners' age, partner's monthly income, partner's education level, type of divorce, number of children, child custody as independent variables. This research question is important because marriage and divorce are two huge milestones in anyone's life and they also have a significant impact (both positive and negative) on society. Therefore, studying relevant factors that affect the marriage duration of divorce couples will potentially help us create better support systems and socioeconomic policies to facilitate the divorce process.

2. Data:

The dataset, found on Kaggle, is the Mexican government official dataset on divorces from 01-01-2000 to 31-03-2015 in Xalapa, Veracruz, Mexico. The government collected this dataset for the purpose of archiving, since we did not find any government-affiliated studies that used this dataset. We also expect to encounter missing values in some of the variables, but we can safely assume that the missing values are random and will not cause bias for our regression results. This dataset is a good representative of Mexico's population because Xalapa is the capital city of Veracruz where there are not many foreigners and its GDP per capita is relatively similar to Mexico's GDP per capita. There are 3,845 observations in total and the data points are in Spanish, but the names of variables are translated to English. The independent and dependent variables are defined as below:

- Marriage duration: The legal marriage period from the date of marriage registration to date of divorce in years. This is calculated from the raw data in the original dataset.
- Husband's age: Age of husband at the date of divorce in years. This is obtained from their date of birth in the dataset.

- Wife's age: Age of wife at the date of divorce in years. This is obtained from their date of birth in the dataset.
- Husband's monthly income: Monthly income of husband in pesos.
- Wife's monthly income: Monthly income of wife in pesos.
- Husband's education level: A qualitative categorical variable measures the level of education completed by the husband. There are 5 categories: (0) Sin Escolaridad: Without any schooling (1) Primaria: completed primary school (2) Secundaria: completed secondary school (3) Preparatoria: completed high school (4) Professional: completed any higher education level, having either an undergraduate or graduate degree
- Wife's education level: A qualitative categorical variable measures the level of education completed by the wife. There are 5 categories: (0) Sin Escolaridad: Without any schooling (1) Primaria: completed primary school (2) Secundaria: completed secondary school (3) Preparatoria: completed high school (4) Professional: completed any higher education level, having either an undergraduate or graduate degree
- Type of divorce: A qualitative categorical variable indicates the two types of divorce: (0) Necessary - Necessary divorce means there is disagreement regarding the divorce between the spouses or when one of the parties does not accept the divorce, (1) Voluntary - There is mutual consent or both spouses agree to the divorce (Immigration and Refugee Board of Canada).
- Number of children: Number of children that they have together.
- Child custody: A qualitative categorical variable indicates the person who gets the legal right and obligation to decide on the child's upbringing. There are 3 categories: (0) Mother, (1) Father, (2) Other.
- Unilateral: Mexico implemented no-fault unilateral divorce law in 2008 (Hoehn-Velasco & Penglase, 2021), so we generated a dummy categorical variable to control for this change in the legal system. There are two categories: (0) Date of divorce before 2008, (1) Date of divorce after 2008.

Figure 1. Descriptive Statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
Husband's age	4816	39.444	10.397	19	91
Husband's monthly income	3504	10920.113	70569.764	2.4	3150242
Wife's age	4772	36.958	9.935	17	84
Wife's monthly income	2804	7374.254	16337.049	3.5	708652
Marriage duration	4688	11.724	9.299	1	61
Children	3011	1.815	.931	1	10
Divorce year	4923	7.063	4.278	0	15
Unilateral	4923	.485	.5	0	1
Husband edu	4605	3.236	.996	0	4
Wife edu	4530	3.275	.969	0	4
Divorce type	4923	.486	.5	0	1
Custody	2072	.035	.212	0	2

The dummy variable 'unilateral' was later generated to take into account the effect of unilateral divorce being introduced in Mexico in 2008.

3. Literature Review

Since divorce has legally existed in Mexico, the divorce rates have exhibited an upward trend in past decades, and many believe that the introduction of unilateral divorce, along with the mutual-agreement divorce provision, caused the remarkable growth in divorce rate. Previous studies (Aguirre, 2019) (Lew et al., 2008) researched the effect of adopting unilateral divorce provision on the divorce rate in the Latin American country Mexico. Besides the effect of unilateral divorce policy, other factors that may influence divorce rate in Mexico have been rarely discussed. Thus, it is essential to take the influence of age, income, education, number of children, and child custody into consideration, all of which represents the quality of marriage. Our research would provide insights on to what degree the marriage duration of divorced couples is influenced by these factors in Mexico.

To start with, we believe that the husband's age & wife's age would affect the duration of marriage. According to a research of BGSU national center for family and marriage research (Allred, 2019), the rate of divorce declined with the age of marriage in the US. Thus, we deduce that a similar trend would occur across different age groups in our dataset.

Furthermore, the husband's monthly income and the wife's husband income could potentially affect the duration of marriage as well. A research about divorce rate in Sweden (Liu & Vikat,

2008) sheds light on the effect of income on the divorce rate. They found out that the higher the wife's share in the couple's income the higher the divorce risk. However, the research has also found that when the combined income of the household is higher then there is generally a lower risk of divorce.

In addition, the husband's education level and wife's education level would also play a role. A research by Pew Research Center (Wang, 2015) indicates that college-educated women are more likely to have a long-lasting relationship compared with that of the women who have a high school degree or less. Thus, we expect a positive correlation between education level and divorce rate.

Regarding the family, the number of children might affect the duration of marriage. A 2015's research on the effect of children on the divorce rate in China (Xu et al., 2015) concluded that the higher the number of children, the more stable the marriage. We expect a similar relationship found out in our research.

Child custody could also be associated with the duration of the marriage. Current research mainly focuses on the emotional effect of divorce on children and their future development, but rarely discusses the effect of child custody on the divorce. We hope we could find some insights from the dataset to further discuss the relationship.

Researching these factors that influence the marriage duration of couples will allow us to assess what are the variables that escalate or de-escalate the process of a married couple arriving to the stage of divorce. Such understanding can allow policy makers to alleviate social problems or factors that cause divorce. Similarly, policy makers can provide assistance so that the marriage duration is shortened for a divorcing couple in the case where a longer duration of maintaining the marriage causes suffering of any of the family members. For example, let's say a mother who is suffering from domestic violence wants a divorce but is unable to do so because of her financial instability. If the government can provide some sort of financial support or job security, the mother could quickly get a divorce without having to endure the suffering and choosing divorce as a last resort.

4. Methodology:

$$\begin{aligned} \text{marrieduration} = & \beta_0 + \beta_1 \text{husbandage} + \beta_2 \text{husbandmincome} + \beta_3 \text{wifeage} \\ & + \beta_4 \text{wifemincome} + \beta_5 \text{husbandedu} + \beta_6 \text{wifeedu} + \beta_7 \text{divorcetype} \\ & + \beta_8 \text{children} + \beta_9 \text{custody} + \beta_{10} \text{unilateral} + \epsilon \end{aligned}$$

We start our analysis with a simple regression model (stated above) and take out the independent variables that are not statistically significant. Such variables are husband's monthly income, wife's monthly income, divorce type, and custody. We will also check for multicollinearity and homoskedasticity for the preliminary model (model (1)) to make sure that our regression models don't violate any Gauss-Markov assumptions.

Firstly, we plotted the graph matrix to see how the observations are distributed according to the independent variables. We can observe an overview and the relationship between the dependent with each independent variable.

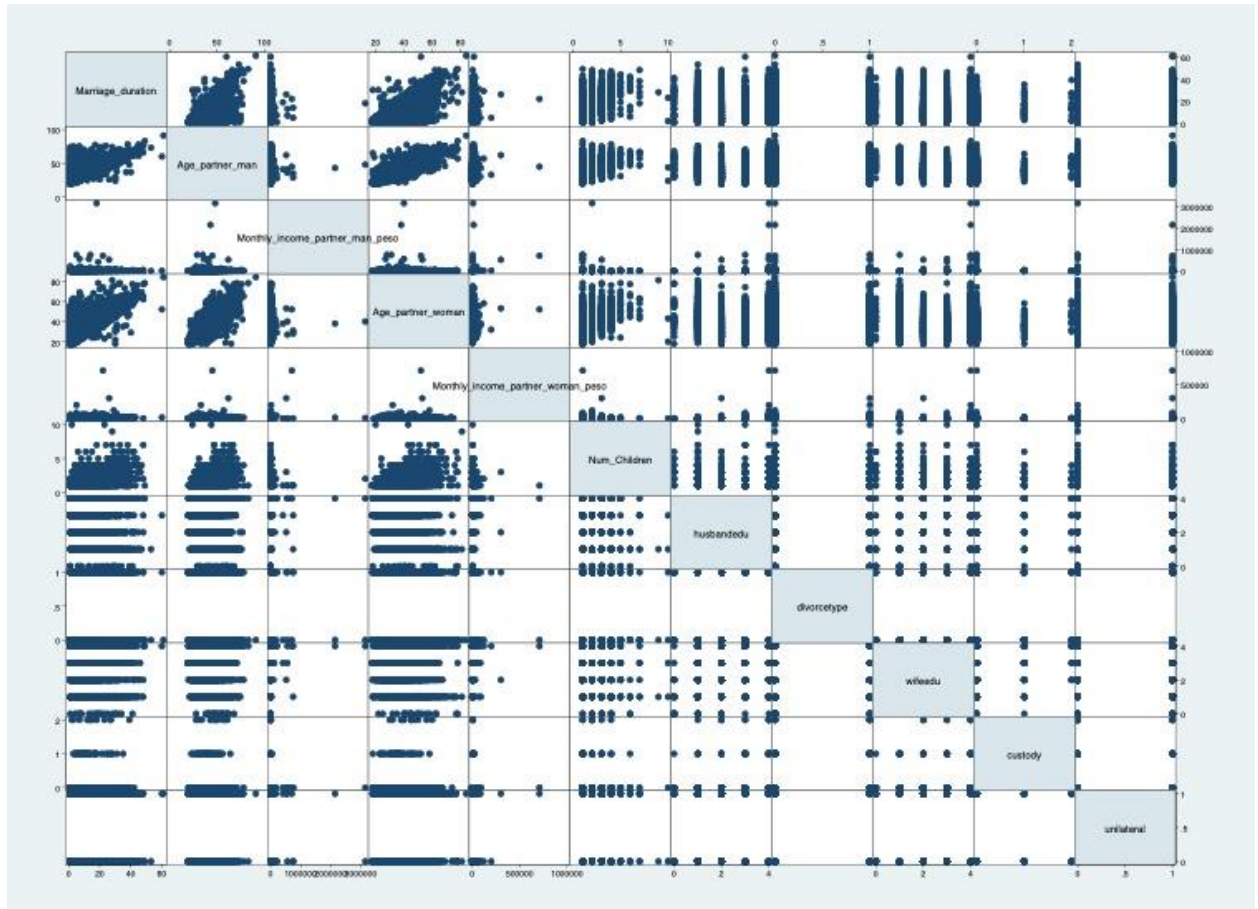


Figure 2. Graph matrix of all variables

a. Multicollinearity:

We also checked the multicollinearity between variables. The table shows that there are no multicollinearity we should be concerned about.

	VIF	1/VIF
Wife's age	3.189	.314
Husband's age	3.119	.321
Divorce type	2.317	.432
Unilateral	2.313	.432
Wife edu	1.419	.705
Husband edu	1.41	.709
children	1.198	.835
Wife's monthly income	1.066	.938
Husband's monthly income	1.05	.953
Custody	1.012	.988
Mean VIF	1.809	.

Figure 3. Variance inflation factor

b. Homoskedasticity:

White's test		
H0: Homoskedasticity		
Ha: Unrestricted		
heteroskedasticity		
chi2(63) = 124.19		
Prob > chi2 = 0.0000		
Cameron & Trivedi's		
decomposition of IM-test		
chi2	df	p
124.190	63	0.000
8.300	10	0.599
3.820	1	0.051
136.310	74	0.000

Figure 4. White test

We checked for heteroskedasticity using white test. The p-value from the chi-square test (0.000) is smaller than 0.05, so our model violated the homoskedasticity assumption.

c. Ordinary Linear Regression

Variables	(1) Marriage duration
Husband's age	.224*** (.028)
Husband's monthly income	0 (0)
Wife's age	.352*** (.03)
Wife's monthly income	0 (0)
children	2.189*** (.186)
Husband edu	-.439*** (.166)
Divorce type	.046 (.388)
Wife edu	-.523*** (.182)
custody	.542 (.751)
unilateral	-.105 (.375)
constant	-10.174*** (.893)
Observations	1078
R-squared	.602

Standard errors are in parentheses

*** $p < .01$, ** $p < .05$, * $p < .1$

Figure 5. OLS Regression of the initial model

The OLS of our initial models shows insignificance with the husband and wife's monthly income, divorce type, custody and unilateral variable. Moreover, the numbers listed for husband and wife's monthly income is strange.

5. Model Improvement:

After running the preliminary regression with all the variables included (Regression model (1) in Figure 5), we have come up with the following improvements for the variables to improve the regression models.

a. Scaling and adjusting the outliers in the income variables

Because the beta coefficients of both income variables are close to 0 (Regression model (1) in Figure 5), we decide to scale both husband and wife monthly income variables by 100,000 pesos. By summarizing the income variables in detail and showing the distribution on the first histogram, we propose truncating the 1st and 99th percentile of the income variables to better represent the general population. The 1st and 99th percentile income range might have different factors that affect their marriage duration that are not common in the general population, so truncating the income variable better serves our research question which is to investigate the factors that affect the marriage duration among the Mexican population. The effect of the truncation is shown in the second histogram which shows a better distribution and is a better representative of the population.

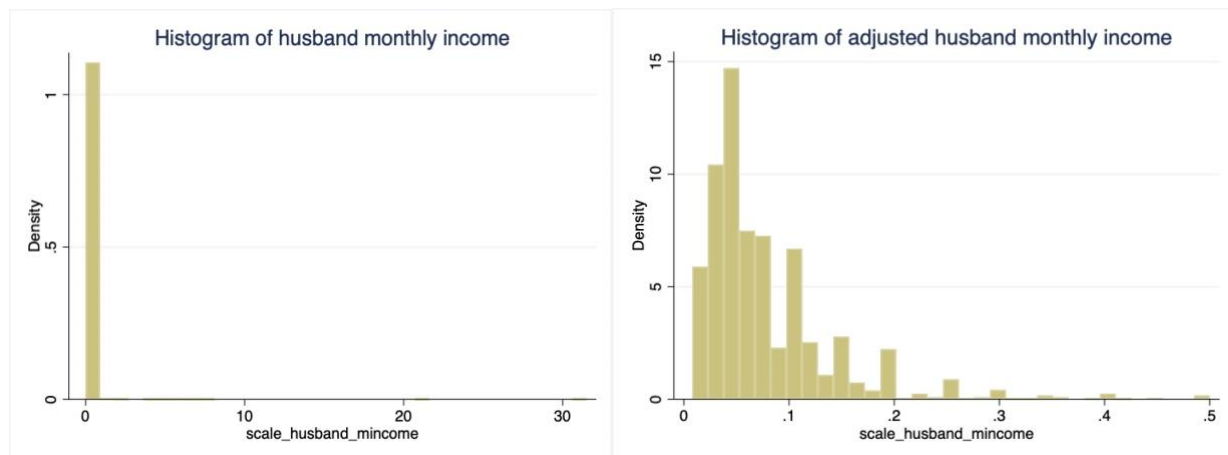


Figure 6. Two histograms showing the density of the scaled husband monthly income and scaled, truncated husband monthly income.

b. Combine income variables

From our literature review, Liu & Vikat, 2008 explores the effect of total income on the duration of marriage, so generating a combined income variable might help to improve the fit of regression models. We combined wife and husband monthly income after scaling and truncating them to create a family monthly income which is included in the regression model (9).

6. Results:

From the number of observations listed at the bottom of the regression table, it shows that the number of observations are different for all seven regressions. However, the difference is not significant to the extent that we have to disregard the adjusted R-squared values when comparing for model fit.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Marriage duration	Marriage duration	Marriage duration	Marriage duration	Marriage duration	Marriage duration	Marriage duration
Husband age	.231*** (.029)	.229*** (.029)	.231*** (.029)	.252*** (.029)	.231*** (.029)	.255*** (.029)	.257*** (.029)
Wife age	.345*** (.031)	.343*** (.031)	.344*** (.031)	.456*** (.03)	.344*** (.031)	.454*** (.03)	.452*** (.03)
Scale and adjusted husband monthly income	4.433** (2.157)	3.754* (2.074)	4.423** (2.152)	4.874** (2.116)	4.444** (2.155)	3.919* (2.014)	
Scaled and adjusted wife monthly income	-3.628 (3.17)		-3.64 (3.166)	-2.225 (3.1)	-3.693 (3.163)		
children	2.189*** (.19)	2.196*** (.19)	2.189*** (.19)	1.76*** (.182)	2.194*** (.189)	1.755*** (.181)	1.767*** (.181)
Husband edu	-.499*** (.174)	-.499*** (.174)	-.5*** (.173)	-.496*** (.168)	-.497*** (.173)	-.534*** (.168)	-.507*** (.167)
Wife edu	-.502** (.199)	-.558*** (.193)	-.501** (.199)	-.352* (.19)	-.504** (.199)	-.388** (.185)	-.394** (.189)
Divorce type	.034 (.399)	.053 (.398)		1.162*** (.329)	.13 (.268)		
Custody	.477 (.764)	.481 (.764)	.481 (.762)		.47 (.763)		
Unilateral	-.125 (.383)	-.152 (.382)	-.149 (.258)	.354 (.325)			
Scaled and adjusted family monthly income							1.828 (1.458)
constant	- 10.081*** (.931)	-9.917*** (.92)	- 10.054*** (.877)	- 15.079*** (.856)	- 10.171*** (.889)	- 14.297*** (.819)	- 14.365*** (.831)
Observations	1039	1039	1039	1425	1039	1425	1425
R-squared	.603	.602	.603	.679	.603	.675	.675
Adj R ²	.599	.599	.599	.677	.599	.674	.673

Standard errors are in parentheses

*** $p < .01$, ** $p < .05$, * $p < .1$

Figure 7. All Regression models

From figure 7, we see that the new combined income variable of scaled and adjusted family monthly income turned out to be insignificant.

For our final model we chose model 4 because as the adjusted R-square value was the highest compared to the other models despite the slight difference in the number of observations. The difference in the number of observations per model is because our data set of age, income children, and all other variable information come from a single couple. Stata will omit this observation if it has certain information missing that is required for our model.

Therefore, the less information required for the model, the greater the number of observations.

Since we know that our initial model's residuals were heteroscedastic, we will check the white's test result for our final model as well.

White's test		
H0: Homoskedasticity		
Ha: Unrestricted		
heteroskedasticity		
chi2(52) = 139.58		
Prob > chi2 = 0.0000		
Cameron & Trivedi's		
decomposition of IM-test	df	p
chi2		
139.580	52	0.000
31.130	9	0.000
12.670	1	0.000
183.370	62	0.000

Figure 8. White's Test for model 7

Given that our final model is also heteroscedastic, we will run a robust standard error test.

	(7) Marriage duration
Husband age	.252*** (.035)
Wife age	.456*** (.034)
Scaled and adjusted monthly husband income	4.874*** (1.876)
Scaled and adjusted monthly husband income	-2.225 (3.401)
children	1.76*** (.192)
Husband edu	-.496*** (.176)
Wife edu	-.352* (.213)
Divorce type	1.162*** (.368)
unilateral	.354 (.369)
constant	-15.079*** (.904)
Observations	1425
R-squared	.679
<i>Robust standard errors are in parentheses</i>	
*** $p < .01$, ** $p < .05$, * $p < .1$	

Figure 9. Robust standard error regression for model 7

Given that our final model is also heteroscedastic, we will run the final regression model with robust standard errors. We can observe that the model with robust standard errors have higher values of standard errors than the model without. However, the coefficients and the significance level of all the independent variables do not change much, so we decided to use the model without robust standard errors for interpretation.

Final Model (Model 4)

$$\begin{aligned} \widehat{\text{marriage duration}} &= 0.252 \text{ husbandage} + 0.456 \text{ wifeage} + 4.874 \text{ scaledhusbandmincome} \\ &\quad - 2.225 \text{ scalewifemincome} + 1.76 \text{ children} - 0.496 \text{ husbandedu} \\ &\quad - 0.352 \text{ wifeedu} + 1.162 \text{ divorcetype} + 0.354 \text{ unilateral} \\ &\quad + 15.079 \end{aligned}$$

We choose model 4 since it has the highest adjusted R-square among all models. Although the scaled wife monthly income and unilateral variable are not statistically significant, we will keep it in the model for the goodness-of-fit.

With the assumption of *ceteris paribus*, on average and keeping all the other independent variables constant, we can infer these relationships from the final model (model 4). Both husband and wife's age are associated with an increase in years of marriage duration, specifically: one year increase in the husband's age increases the marriage duration by 0.252 years and one year increase in the wife's age, the marriage duration increases by 0.456 years. In contrast, the education level of both husband and wife is associated with a decrease in years of marriage duration. Since the education level variables are categorical and defined from having no education to finishing some forms of postgraduate degree, an increase in education level of husband decreases marriage duration by 0.496 years, and an increase in wife's education level decreases marriage duration by 0.352. If the divorce type is voluntary, the marriage duration increases by 1.162 years. In addition, two independent variables that is associated with an increase in marriage duration are number of children and the unilateral divorce law. An increase in number of children in the household is associated with an increase of 1.76 years in marriage, while the legalization of unilateral divorce law increases the marriage duration by 0.354 years.

We truncate the 1st and 99th percentile of both income variables to erase the extreme value as the extreme value is not representative to the whole dataset. After truncation and scaling, the husband's income is statistically significant while the wife's income is not statistically significant. On average, with an increase of 100,000 pesos in the husband's income, the marriage duration increases by 4.874 years with all other variables constant and all else equal. On average, with an increase of 100,000 pesos in the wife's income, the marriage duration decreases by 2.225 years with all other variables constant and all else equal.

In terms of the age of husband and wife, the increase in age is associated with the increase in the marriage duration, overall. We believe that, as people become more mature, they would cherish marriage more and believe their partner is the right one. Otherwise, people would seek divorce at a young age rather than later. This result is consistent with previous research conducted in the US by BGSU national center for family and marriage research (Allred, 2019).

For education, the regression indicates that both the husband and wife's education level is negatively associated with marriage duration. In terms of the wife's education level and marriage duration, previous research by Pew Research Center (Wang, 2015) has different findings. We believe that the difference is because higher education people have better legal knowledge so they will not stay in an unwanted marriage.

Regarding the divorce type, couples choosing voluntary divorce tend to have longer marriage duration. We believe the reason behind this phenomenon is that couples choosing voluntary divorce have a more equal status between wife and husband rather than abusive. In this situation, they can solve issues through communications. That might be the reason why the marriage duration is longer than that of couples choosing necessary divorce.

Number of children also play a role for the marriage duration: they tend to maintain the marriage longer. Indeed, many parents see children as a bond between husband and wife. Worrying about the childhood of the children, couples may perform more cooperative and lower the divorce wills. This result is consistent with the research conducted in China (Xu et al., 2015).

In regards to the relationship between income and marriage duration, we adopt three ways to explore such associations: husband's income, wife's income, and combined income. For the husband's income, it is positively correlated with the marriage duration. This might be because of the fact that women prefer to have a high-income partner who can financially support the family, according to a report by the Washington Post (Cunha, 2015). Furthermore, the relationship between wife's income and marriage duration as well as the relationship between combined family income and marriage are not statistically significant. One potential explanation of this phenomenon is that, due to cultural influences, Mexican men play the breadwinner role, while women are not expected to work outside of the family. In this way, women's monthly income are not statistically significant in correlation to the duration of the marriage, so as to the combined family income.

7. Limitations

Because we obtained this dataset from Kaggle, an open-source website, the quality of data is relatively good, except that it has quite a lot of missing values. We use Stata to report all the missing values out of 3845 observations in Figure 10. This could be one of the reasons that the coefficients for wife monthly income and custody are not significant since we don't have enough data in the dataset to show evidence. Since we assume that the missing values from the dataset were random from the data collection process, it is safe to say that our regression results and discussion are still valid. Moreover, since this dataset is about divorce and contains a lot of private information such as income, occupation, and date of birth, it is possible to identify who these divorced individuals are. Therefore, there could be some concern related to the privacy or consent of individuals for this dataset to be used publicly.

Figure 10. All missing values in the dataset

Variable	#
Husband age	107
Husband monthly income	1419
Wife age	151
Wife monthly income	2119
Marriage duration	235
children	1912
Husband edu	318
Wife edu	393
custody	2851

Our methodology only uses one sample set to infer the result for the whole population of Mexico, so the limitation could be that the sample size is not big enough to come to a valid conclusion. We could improve this by finding similar datasets from different states in Mexico to merge them and analyze all together which could result in a fairer sample.

8. Conclusions

We conclude that the factors that are associated with an increase in marriage duration among divorced couples in Mexico are husband & wife's age, number of children, voluntary divorce, and husband's monthly income, while the factors that are associated with a decrease in marriage duration are husband and wife's level of education and necessary divorce. We also found that the person getting custody rights of the children has no significant effect on the marriage duration.

9. References (APA)

- Allred, C. (2019). Age variation in the divorce rate, 1990 & 2017. *Family Profiles*, FP-19-13. Bowling Green, OH: National Center for Family & Marriage Research. <https://doi.org/10.25035/ncfmr/fp-19-13>.
- Aguirre, E. (2019). Do changes in divorce legislation have an impact on divorce rates? The case of unilateral divorce in Mexico. *Latin American Economic Review*, 1. <https://doi.org/10.1186/s40503-019-0071-7>
- Cunha, D. (2015, May 4). *Men are to blame for the high divorce rate among America's poor*. The Washington Post. Retrieved April 28, 2022, from <https://www.washingtonpost.com/posteverything/wp/2015/05/04/men-are-to-blame-for-the-high-divorce-rate-among-americas-poor/>
- Immigration and Refugee Board of Canada. (2013, November 7). *Mexico: Divorce law and practice, including separation agreements, custody, property dispositions, consent requirements, related*. Department of Justice. Retrieved March 2, 2022, from <https://www.justice.gov/sites/default/files/eoir/legacy/2013/11/07/MEX43266.E.pdf>
- Lew, N., & Beleche, T., & Agüero, J., & Ashley, E., & Bedard, K., & Fairris, D., & Sorensen, T, (2008). 'Til Laws Do Us Part? The Impact of Changing Divorce Laws on Divorce Rates in Mexico.
- Liu, G., & Vikat, A. (2004). *Does divorce risk depend on spouses' relative income? A register-based study of first marriages in Sweden in 1981–1998* (No. WP-2004-010). Max Planck Institute for Demographic Research, Rostock, Germany.
- Wang, W. (2015). *The link between a college education and a lasting marriage* | Pew Research Center. Pew Research Center; <https://www.facebook.com/pewresearch>. Retrieved March 3, 2022, from <https://www.pewresearch.org/fact-tank/2015/12/04/education-and-marriage/>
- Xalapa City Hall. (2017, May 3). *Datos Abiertos de México - Registro Civil*. Datos Abiertos de México - Datos.Gob.Mx. <https://datos.gob.mx/busca/dataset/registro-civil>
- Xu, Q., Yu, J., & Qiu, Z. (2015). The impact of children on divorce risk. *The Journal of Chinese Sociology*, 1. <https://doi.org/10.1186/s40711-015-0003-0>