

2011 - 2013 Divorce in Texas

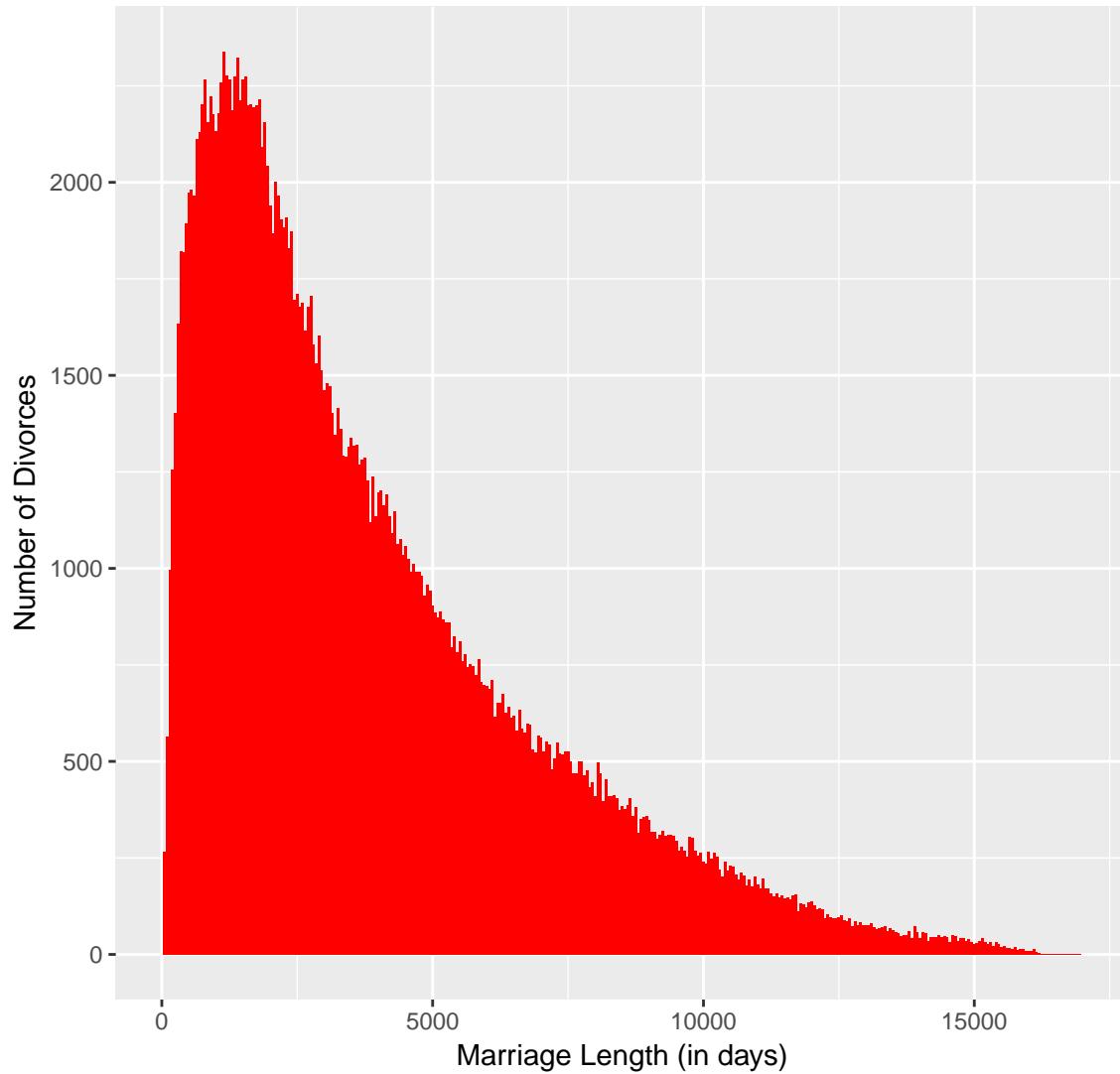
This is the format of our raw divorce record data:

```
##      SFN                      H_NAME H_AGE
## 1  56849 ,ARQIEZ EROCL KA,ES          32
## 2  51757 AANEB YAHRA                 40
## 3  48778 AARON BILLY RANDALL JR       25
## 4  109632 AARON CHARLES WAYNE        34
## 5  140705 AARON FIDEL                 34
## 6  139134 AARON HUGHES MICHEAL        37
##                      W_NAME W_AGE NUM_CHILD MARR_DATE DIV_DATE
## 1 LACEE                         24      2 12/15/07  6/8/11
## 2 CANDAC LYNETTE                  30      1 ?           3/27/11
## 3 TAYLOR ANN                     21      2 6/28/09  5/23/11
## 4 BRANDI DENISE                  27      2 3/15/03  11/1/11
## 5 HELEN ANDREA                   31      3 10/31/98 6/17/11
## 6 SALLY ANN                      30      1 10/13/03 11/18/11
##      COUNTY_ID COUNTY_NAME date_diff year_diff
## 1      227     TRAVIS 1271 days  3.482192
## 2      57      DALLAS   NA days      NA
## 3      252     YOUNG   694 days  1.901370
## 4      14      BELL   3153 days  8.638356
## 5      101     HARRIS  4612 days 12.635616
## 6       1    ANDERSON 2958 days  8.104110
```

Analysis No.1: Marriage Length v.s. Number of Divorces

Warning: Removed 17889 rows containing non-finite values (stat_bin).

Marriage Length v.s. Number of Divorces



Hypothesis: We assumed that the longer a couple is in marriage, the less likely that they will divorce. Thus, our hypothesis is that as the marriage length increases, the number of divorces will constantly decrease.

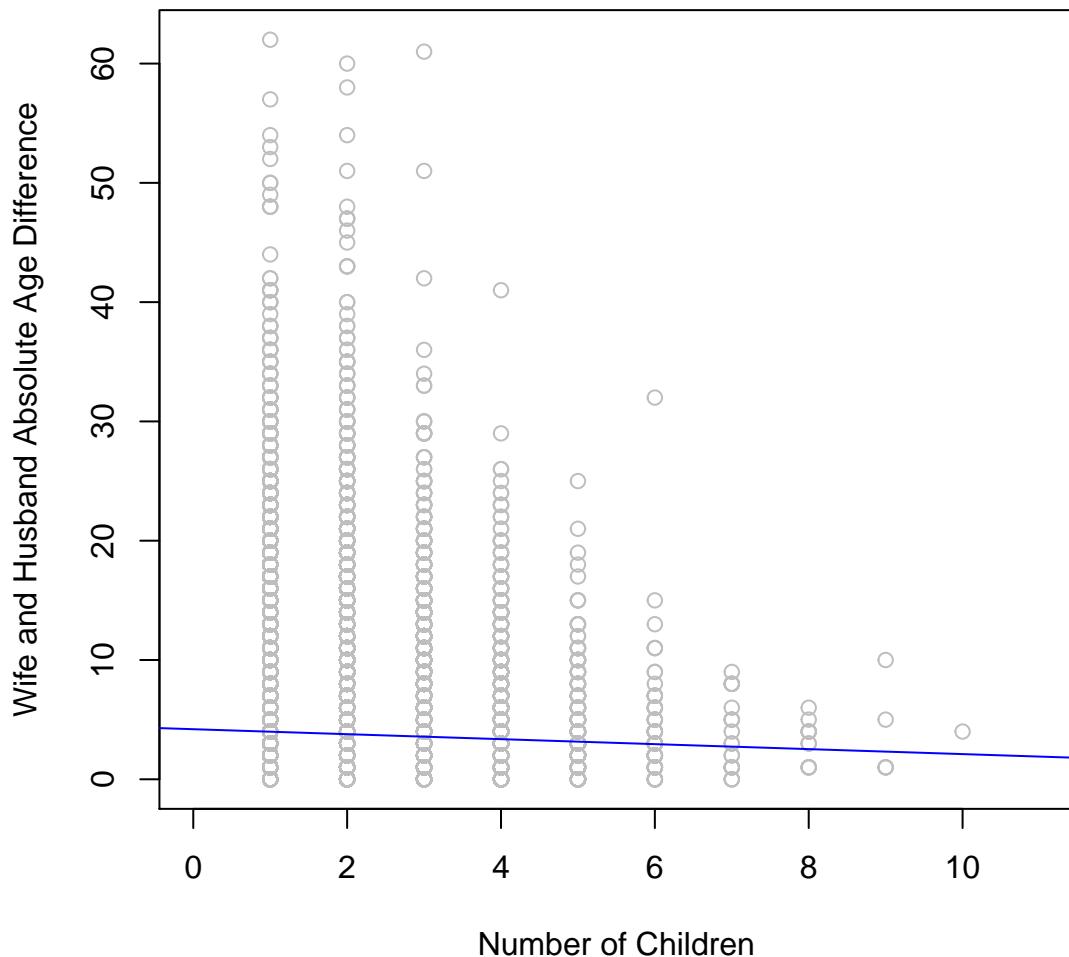
Analysis: The graph shows that when marriage length is shorter than 1200 days (approximately), the number of divorce increases as the marriage length increases. When the marriage length is longer than 1200 days (approximately), the number of divorces decreases as the marriage length increases.

Conclusion: The second part of the analysis corresponds to the hypothesis because the number of divorces decreases as marriage length increases when marriage length is larger than 1200 days. However, the first part of the analysis contradicts with our hypothesis because the number of divorces increases as marriage length increases when marriage length is shorter than 1200 days.

Analysis No.2: Number of Children v.s. Age Difference Between Husband and Wife

```
## Warning: NAs introduced by coercion  
## Warning: NAs introduced by coercion  
## Warning in ` [.data.frame` (div11_13, div11_13$date_diff > 0 &  
## as.numeric(as.character(div11_13$NUM_CHILD)) > : NAs introduced by coercion
```

Number of Children v.s. Age Difference



*Note on the data: Currently there exist a data entry for having 40 children, however we currently is investigating on whether it is a data entry error. We are also currently eliminating entries with 0 children for this graph since we want to show the relation between having more child and longer marriage.

Hypothesis: The number of children that the family has has a negative relation to the difference in the couples age.

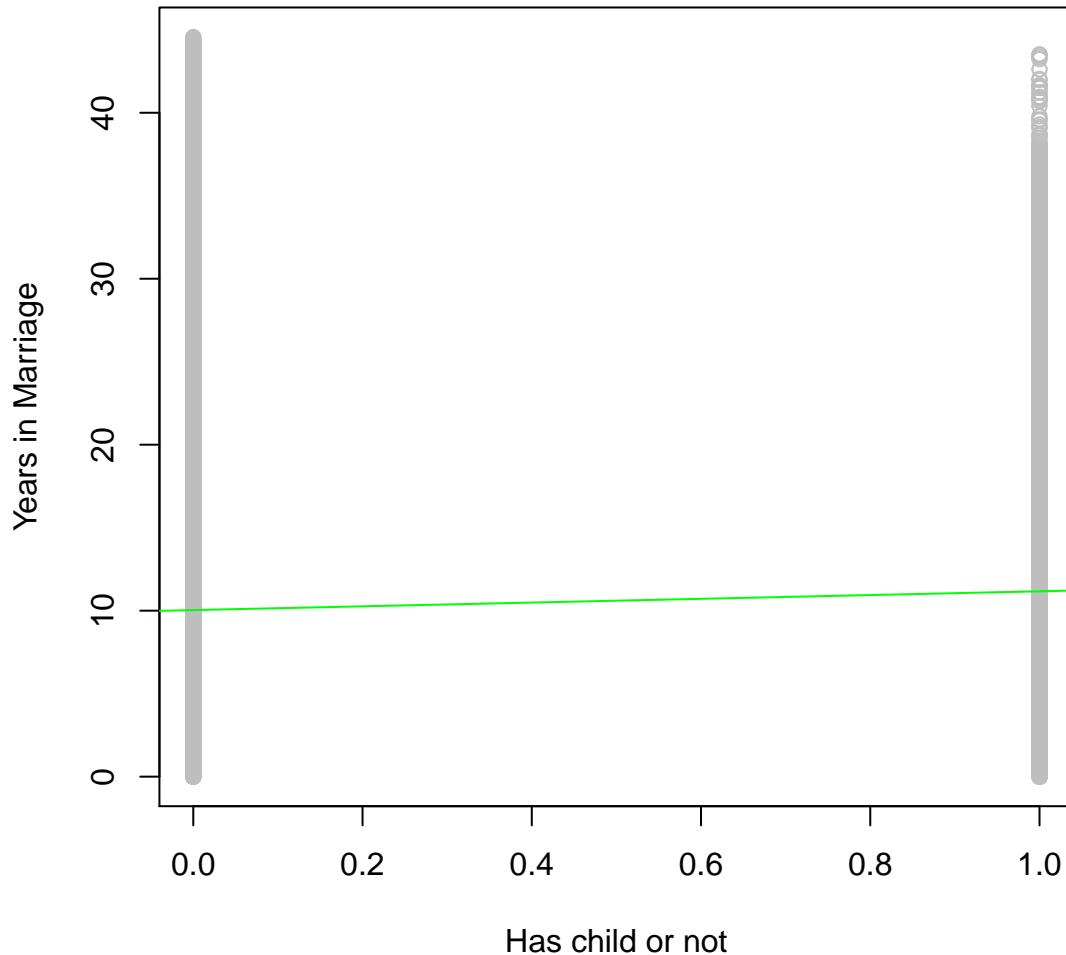
Analysis: From the regression line we can see that the more number of children that the family has, the less difference there is in their age. However, the relation is not very strong. We can see that there are primarily two problems: 1. There are a lot more data points for family that has less than three child than more than three child. 2. The variance of age difference is usually very large for each bar.

Conclusion: We cannot reject or agree with the hypothesis since the relation is not very strong as seen from the graph.

Analysis No.3: Has Child or Not v.s. Marriage Length (in years)

Warning: NAs introduced by coercion

Relation between Married Years and Child



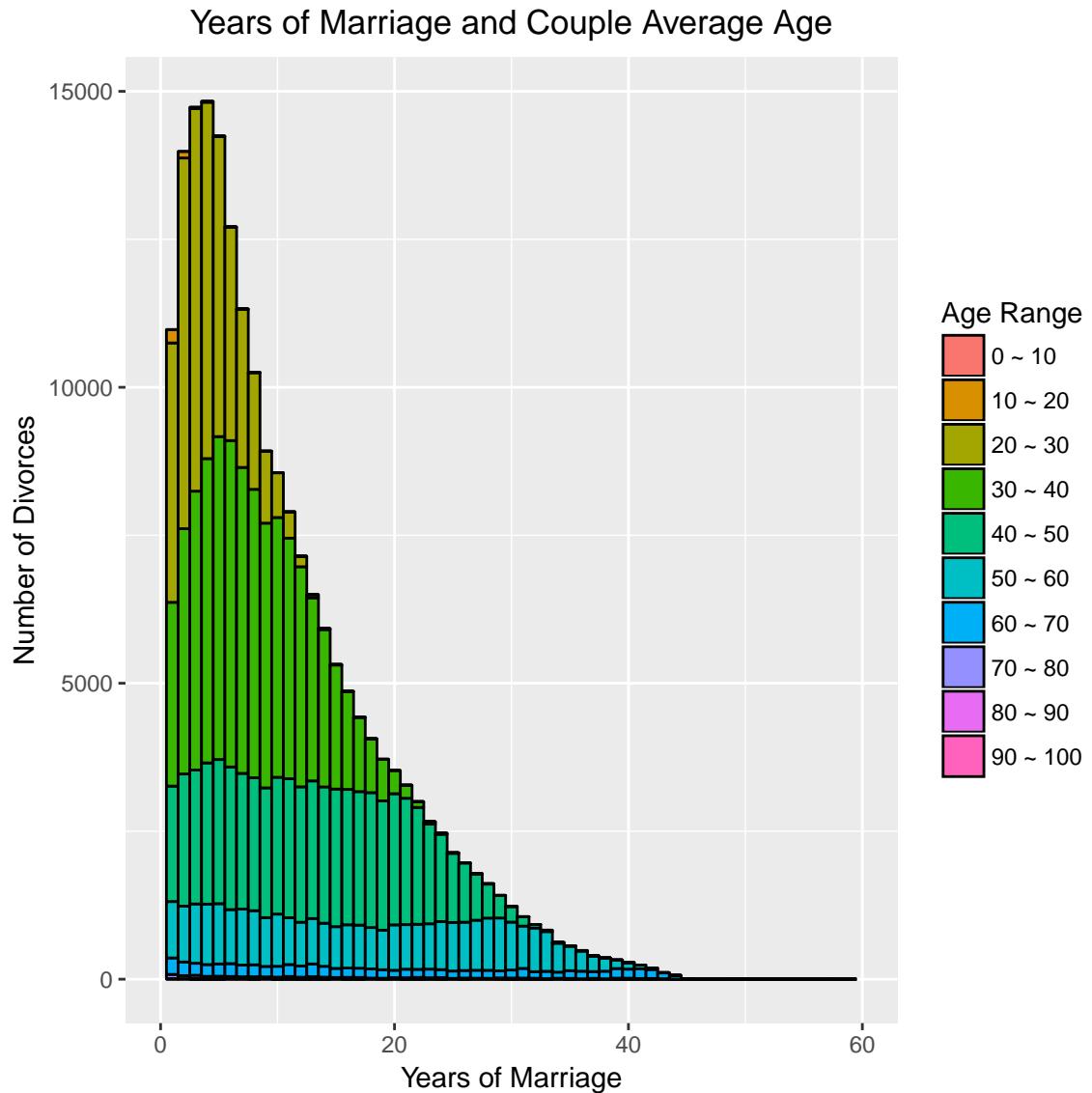
Hypothesis: Our hypothesis is that the couple with at least one child is more likely to have a longer marriage.

Analysis: We can see from the regression line that couple with at least one child has a slightly longer marriage.

Conclusion: The analysis supports the hypothesis, but a bit weakly. The relationship is not that strong.

Analysis No.4: Couple Average Age v.s. Number of Divorces

```
## Warning: NAs introduced by coercion  
## Warning: NAs introduced by coercion  
## Warning: Removed 13171 rows containing non-finite values (stat_bin).
```



Hypothesis: Younger couples tend to have a shorter marriage length, while elder couples who divorce will have a longer marriage length.

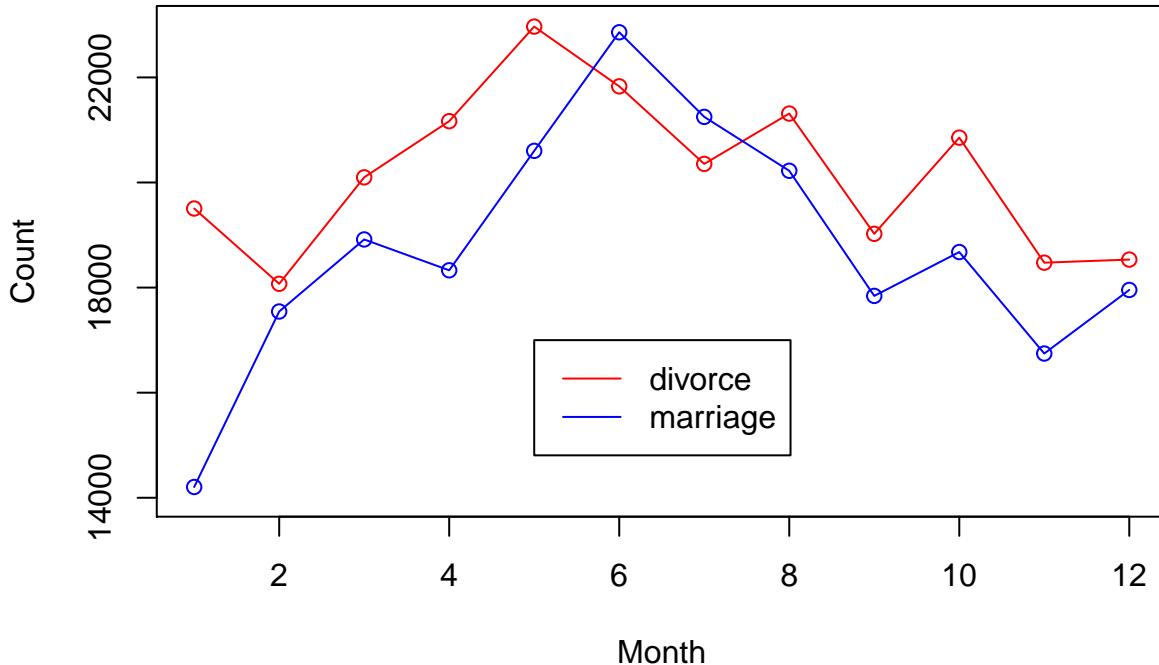
Analysis:

1. Obviously, couples with average age of 20-30 and 30-40 have a relatively shorter period of marriage than those in their 40s, 50s, or above. This also makes sense intuitively, as younger people have experienced a shorter period of time since they became legal to marry than those who are elder.
2. However, the number of couples who have an average age between 40-60 stays almost the same for different marriage lengths, which contradicts the second half of our hypothesis.
3. When marriage length is less than 20 years, for fixed marriage length, most divorce cases happen to couples who are 20 to 40 years old.

Conclusion: Younger couples tend to have a shorter marriage, while for elder couples, the length of marriage does not vary too much among them. When marriage length is less than 20 years, most divorce cases happen to couples who are 20 to 40 years old.

Analysis No.5: Month in a Year v.s. Number of Divorces

Divorce and Marriage by Month



Hypothesis: Most people tend to choose to get married in the middle of the year and to divorce in January or December.

(We get to this hypothesis based on the following: a. In the middle of the year, the weather is very nice, and people tend to be in good mood. Therefore, people are willing to hold wedding ceremony in the middle of the year. b. January and December are very cold. And people are very likely to be depressed by the coldness. Therefore, such depression might trigger divorce.)

Analysis: This graph is about the number of divorce and marriage in every month. In the above graph, the red line represents number of divorce and the blue line represents number of marriage. The x-axis is month and the y-axis is number of divorce or marriage. From the above graph, we can find that: 1. About divorce: Most divorce happens in May and least divorce happens in February. The number of divorces increases from February to May, and oscillates between 18000 and 21000 after May. The graph indicates that divorce is not distributed evenly, and we cannot find an exact distribution for divorce based on month.

2. About marriage: Most marriage happens in June and least marriage happens in January. The number of marriage increases from January to June and decreases after June (oscillates after September). The distribution of marriage by month is a bit similar to normal distribution with more marriage occurring in the middle of the year and less occurring at the beginning or end of the year.
3. Similarities and differences: Both divorce and marriage happen most frequently in the middle of the year and least at the beginning of the year.

Conclusion: The analysis above inferred from the graph supports my hypothesis on marriage but does not support my hypothesis on divorce (Most people divorce in May not at the beginning or end of the year).

Conclusions to the Question

Re-cap of the project question: How are different factors (including length of marriage, average age of the couple, age difference between husband and wife of the couple, number of children the couple has, and which month in a year the couple got married and got divorced) related to the number of divorces in Texas throughout 2011-2013? And are there any relationship between some of those factors?

According to our five analysis above, we've found that:

1. Relationship between marriage length and number of divorces. When the marriage length is less than 1200 days (approximately 3 years), the number of divorces increases as the marriage length increases, which means that when a couple has been married for less than 3 years, it becomes more and more likely for this couple to get divorced as their marriage continues. However, when the marriage length is more than 1200 days, the number of divorces decreases as the marriage length increases, which means that when a couple has been married for more than 3 years, it's less and less likely for them to get divorced as their marriage continues.
2. Relationship between the two factors, "number of children" and "age difference between husband and wife", is not strong enough.
3. Relationship between the two factors, "has a child or not" and "marriage length", is not strong enough.
4. Younger couples tend to have a shorter marriage, while for elder couples, the length of marriage does not vary too much among them.
5. Most people tend to get married and get divorced in the middle of the year.

To-Do List for future.

Find the optimal secondary data set for further analysis. Currently we have two data sets available for us to use: - Airbnb listing data by location - Weather data by zip code Using these data we can do further analysis such as the influence of airbnb listings and divorce indexes, the relation between good/bad weather and number of divorce/marriage records on that date, etc.