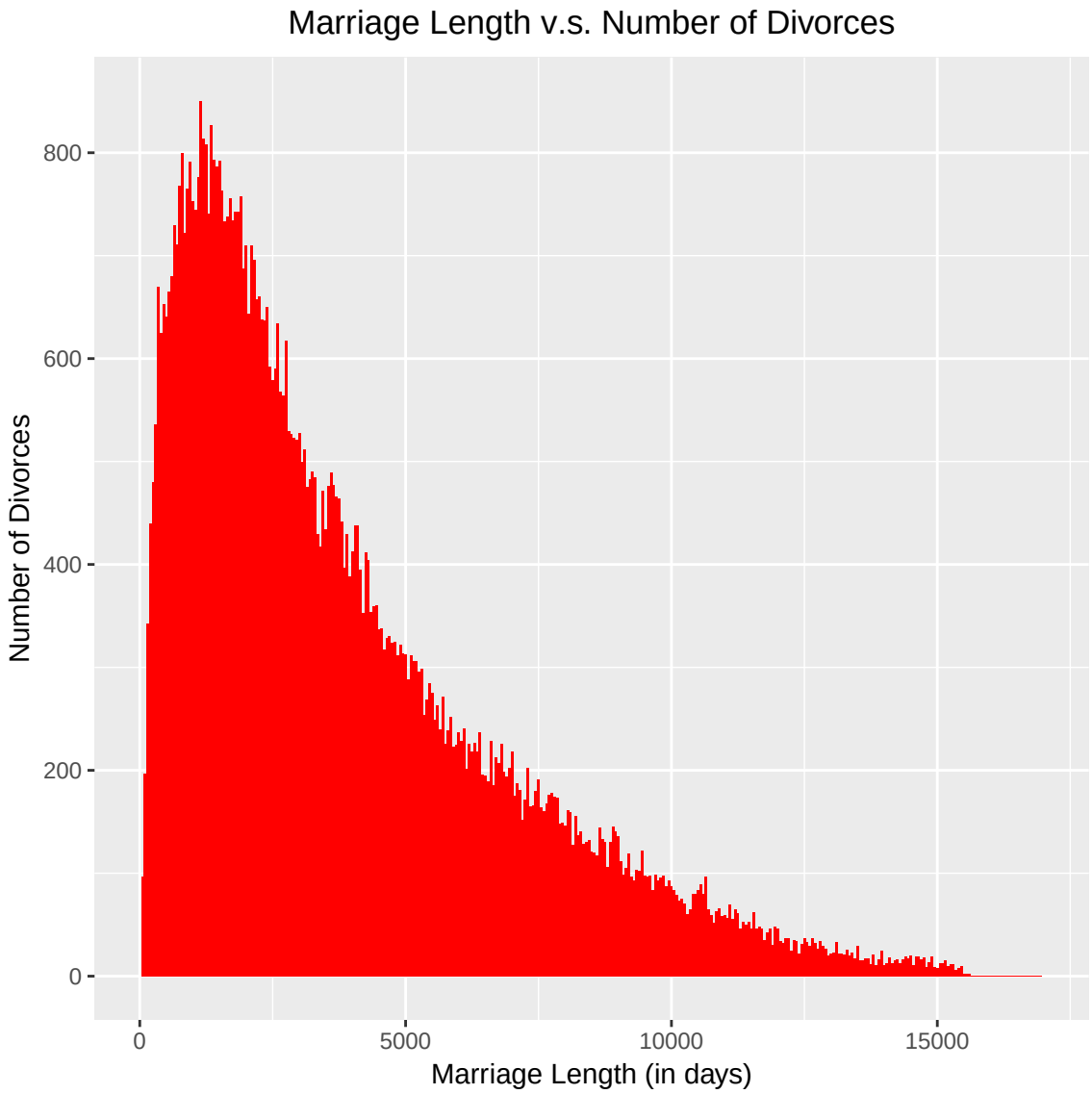
2011 - 2013 Divorce in Texas

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 6081 rows containing non-finite values (stat_bin).
```

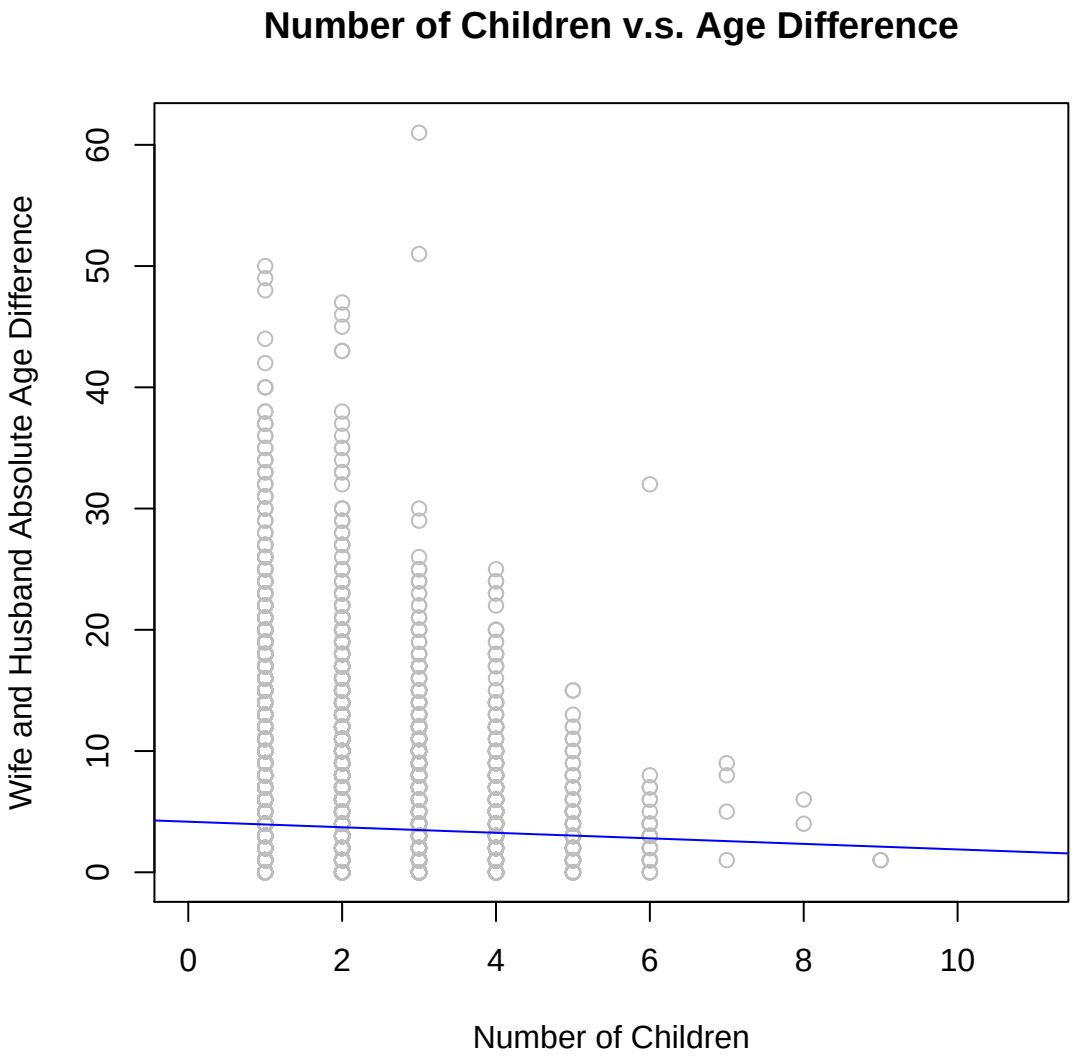


Analysis No.2: Number of Children v.s. Age Difference Between Husband and Wife

```
## Warning: NAs introduced by coercion

## Warning: NAs introduced by coercion

## Warning in `[.data.frame`(div11, div11$date_diff > 0 &
## as.numeric(as.character(div11$NUM_CHILD)) > : NAs introduced by coercion
```

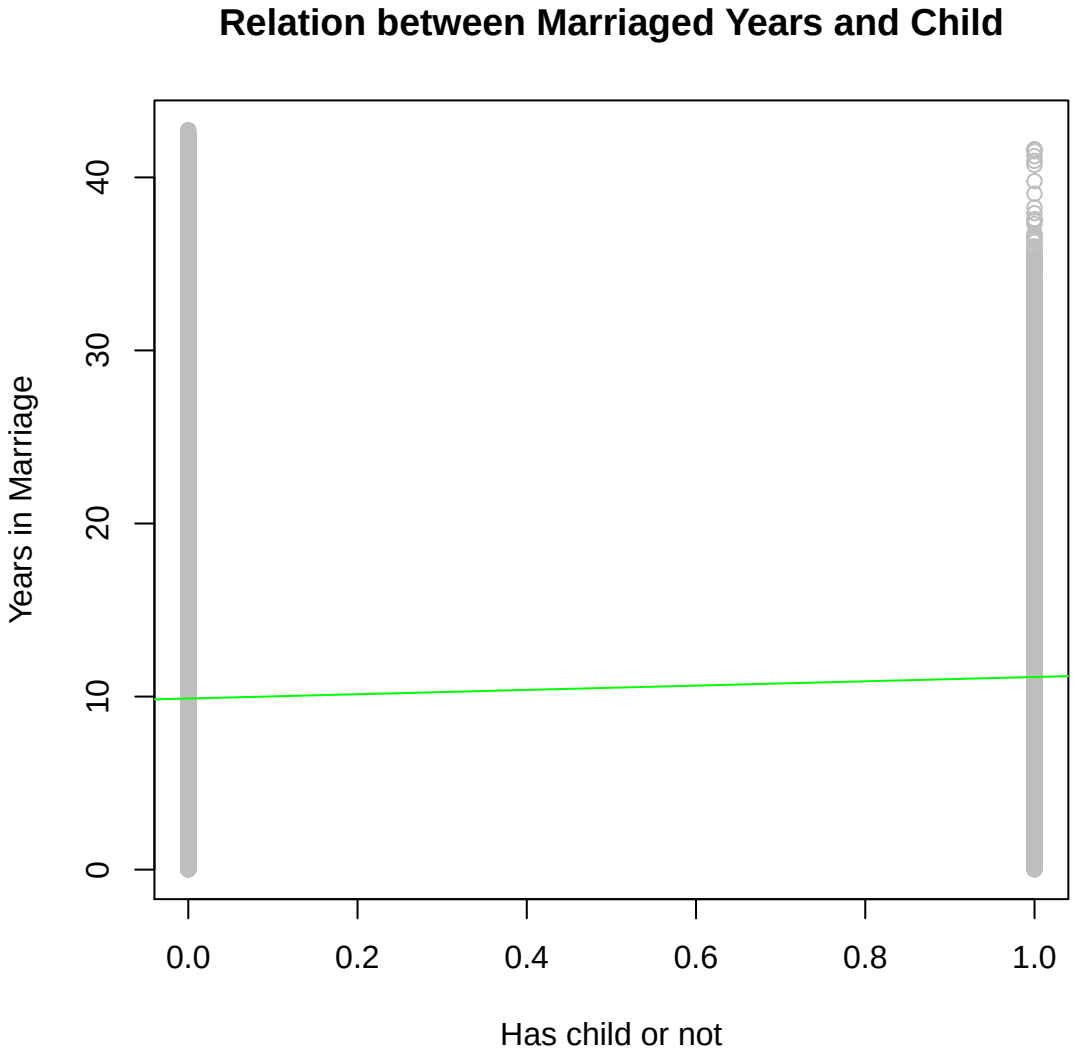


(Currently there exist a data entry for having 40 children, however we currently is investigating on whether it is a data entry error. We are also currently eliminating entries with 0 children for this graph since we want to show that.) We can see from the first graph that, as the regression line suggest, the more number of children that the family has, the less difference there is in their age.

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

```
## Warning: NAs introduced by coercion
```

```
##
## Call:
## lm(formula = y2 ~ x2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -11.127  -6.291  -2.448   4.092  32.851
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.88550    0.04059  243.57   <2e-16 ***
## x2TRUE       1.24467    0.05987   20.79   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.268 on 76789 degrees of freedom
##   (6147 observations deleted due to missingness)
## Multiple R-squared:  0.005597,   Adjusted R-squared:  0.005584
## F-statistic: 432.2 on 1 and 76789 DF,  p-value: < 2.2e-16
```



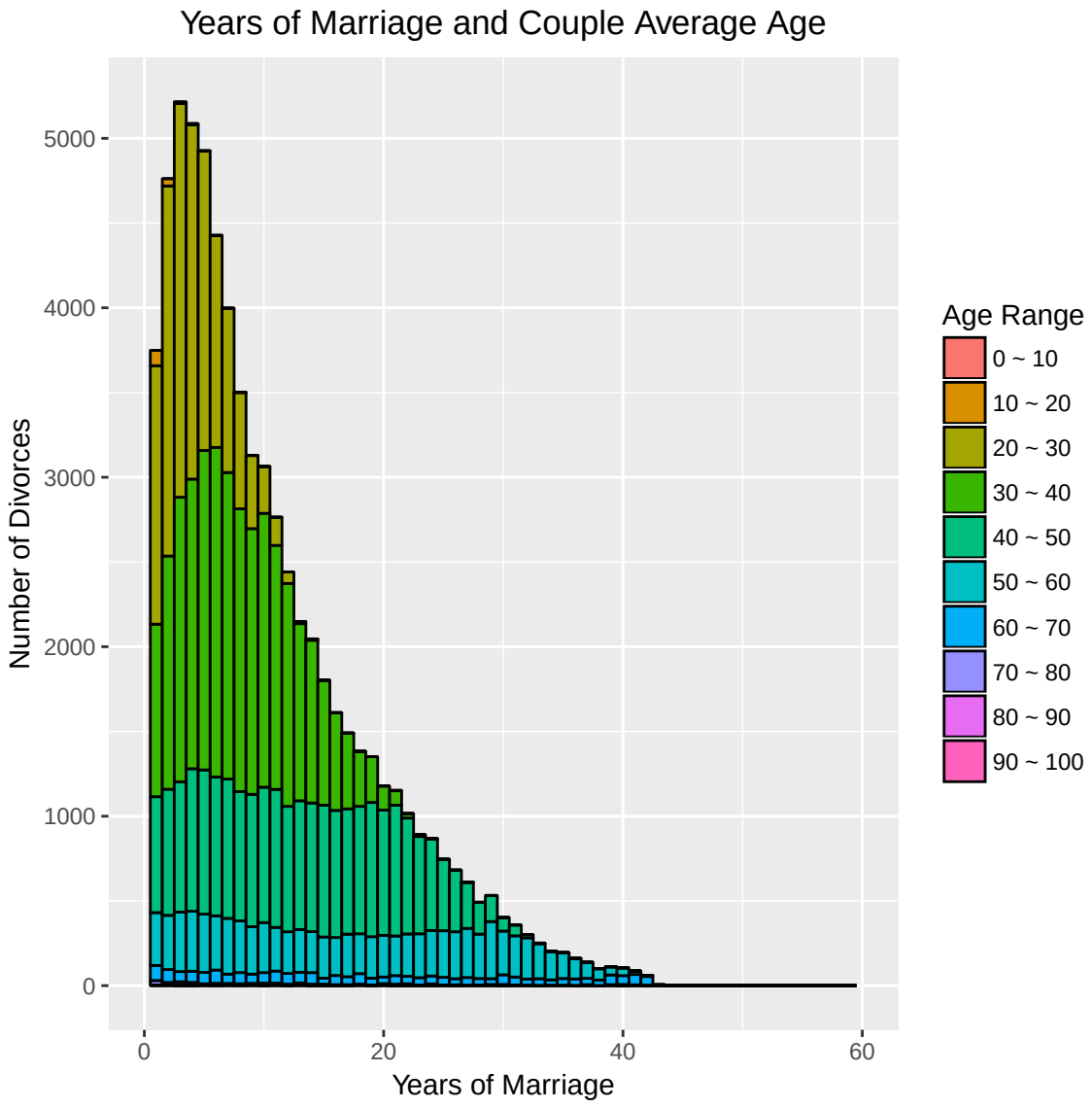
We can see from the regression that couple with longer marriaged years has a slightly bigger chance of having at least one child.

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

```
## Warning: NAs introduced by coercion
```

```
## Warning: NAs introduced by coercion
```

```
## Warning: Removed 4491 rows containing non-finite values (stat_bin).
```



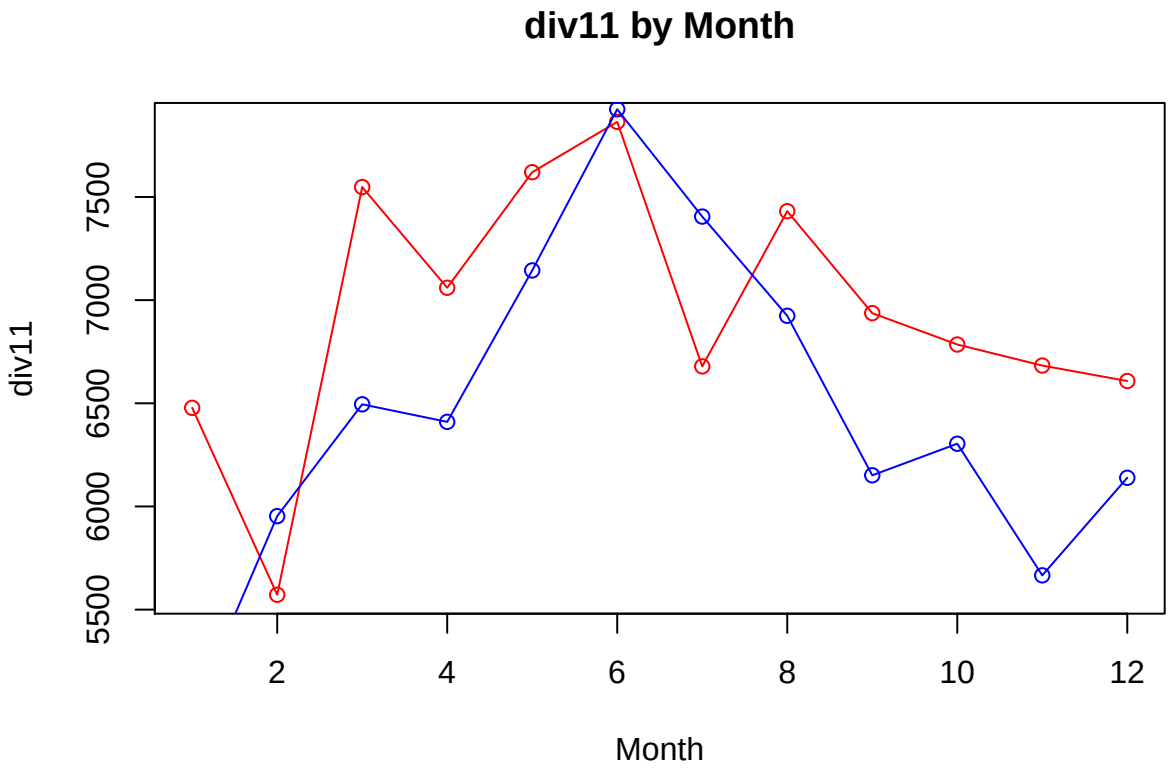
Marriage length of year 2013

Hypothesis: Younger couples who divorce in Texas in 2011 tend to have a shorter marriage, while elder couples who divorce will have a longer marriage.

Analysis: From the graph above, we can see that the average age of couples who divorce in Texas in 2011 mainly lies between 20 to 60. Obviously, couples with average age of 20-30 and 30-40 who divorce in Texas in 2011 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. However, the number of couples who divorce in Texas in 2011 and have an average age between 40-60 stays almost the same for different lengths of marriage, which contradicts the second half of our hypothesis.

Conclusion: Younger couples who divorce in Texas in 2011 tend to have a shorter marriage, while for elder couples who divorce in Texas in 2011, the length of marriage does not vary a lot among them.

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



Hypothesis??? Most people tend to choose to divorce and get married in the middle of the year.

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 June and least divorce happens in January. The distribution of divorce by month is similar to normal distribution with more divorce occuring in the middle of the year and less occuring at the beginning or end of the year.

2. About marriage: Most marriage happens in June and least marriage happens in Febrary. The graph indicates that marriage is not distributed evenly, and we cannot find an exact distribution for marriage based on month.

3. Similarities and differences: Both divorce and marriage happens most frequently in June. But unlike divorce, there is not a regular and exact distribution for marriage.

The analysis above inferred from the graph supports my hypothesis.

Conclusions to the Question