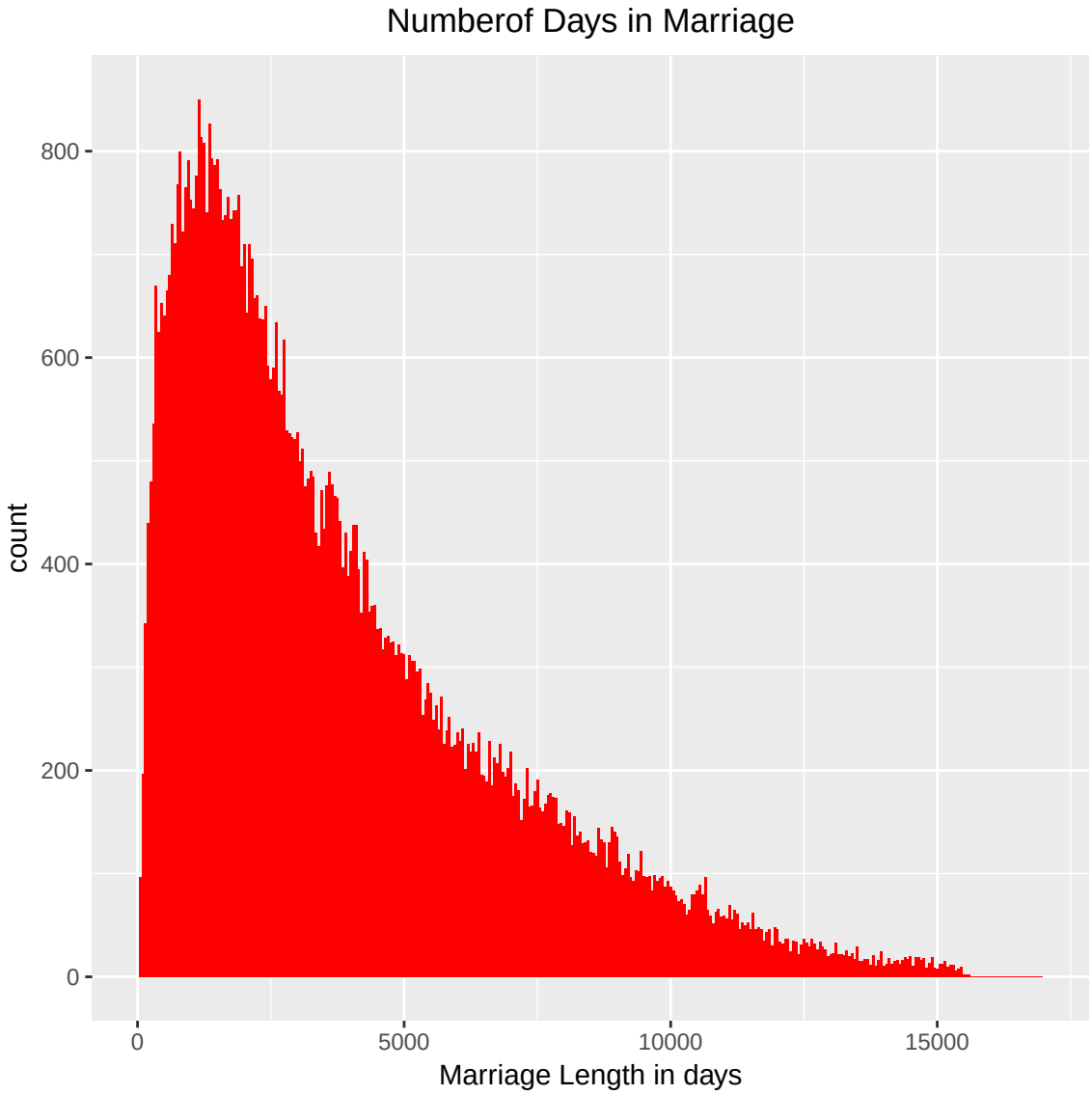
2011 - 2013 Divorce in Texas

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
```

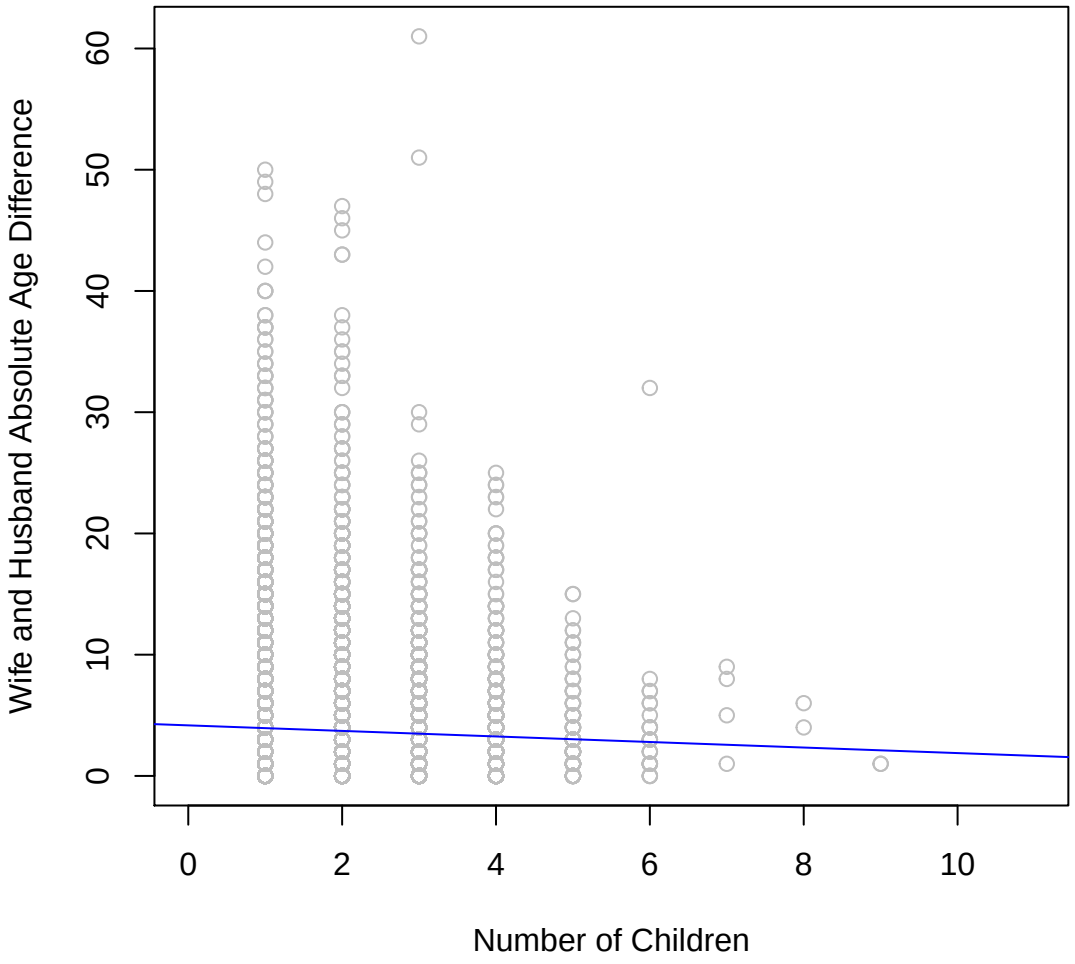
```
## Warning: Removed 6081 rows containing non-finite values (stat_bin).
```




```
## 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

##
## Call:
## lm(formula = y ~ x)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -3.941 -2.712 -0.941  1.288 57.517
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.16913    0.05053  82.501   <2e-16 ***
## x           -0.22863    0.02632  -8.687   <2e-16 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.058 on 32810 degrees of freedom
##   (2500 observations deleted due to missingness)
## Multiple R-squared:  0.002295,   Adjusted R-squared:  0.002264
## F-statistic: 75.47 on 1 and 32810 DF,  p-value: < 2.2e-16
```

tween absolute age difference between wife and husband and nu

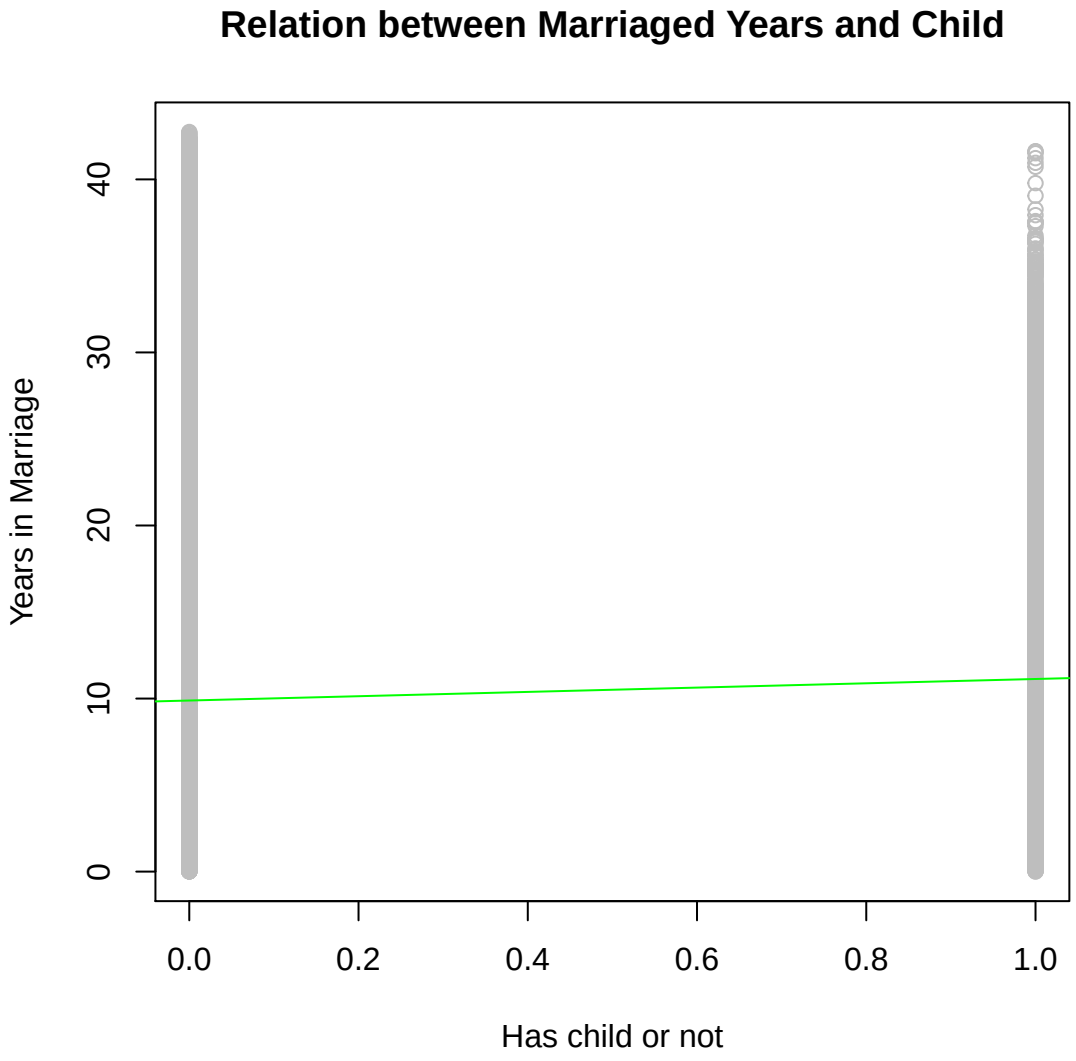


We can see from the first graph that, as the regression line suggest, the more number of children that the family has, the less difference there is in their age.

```
## 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
```

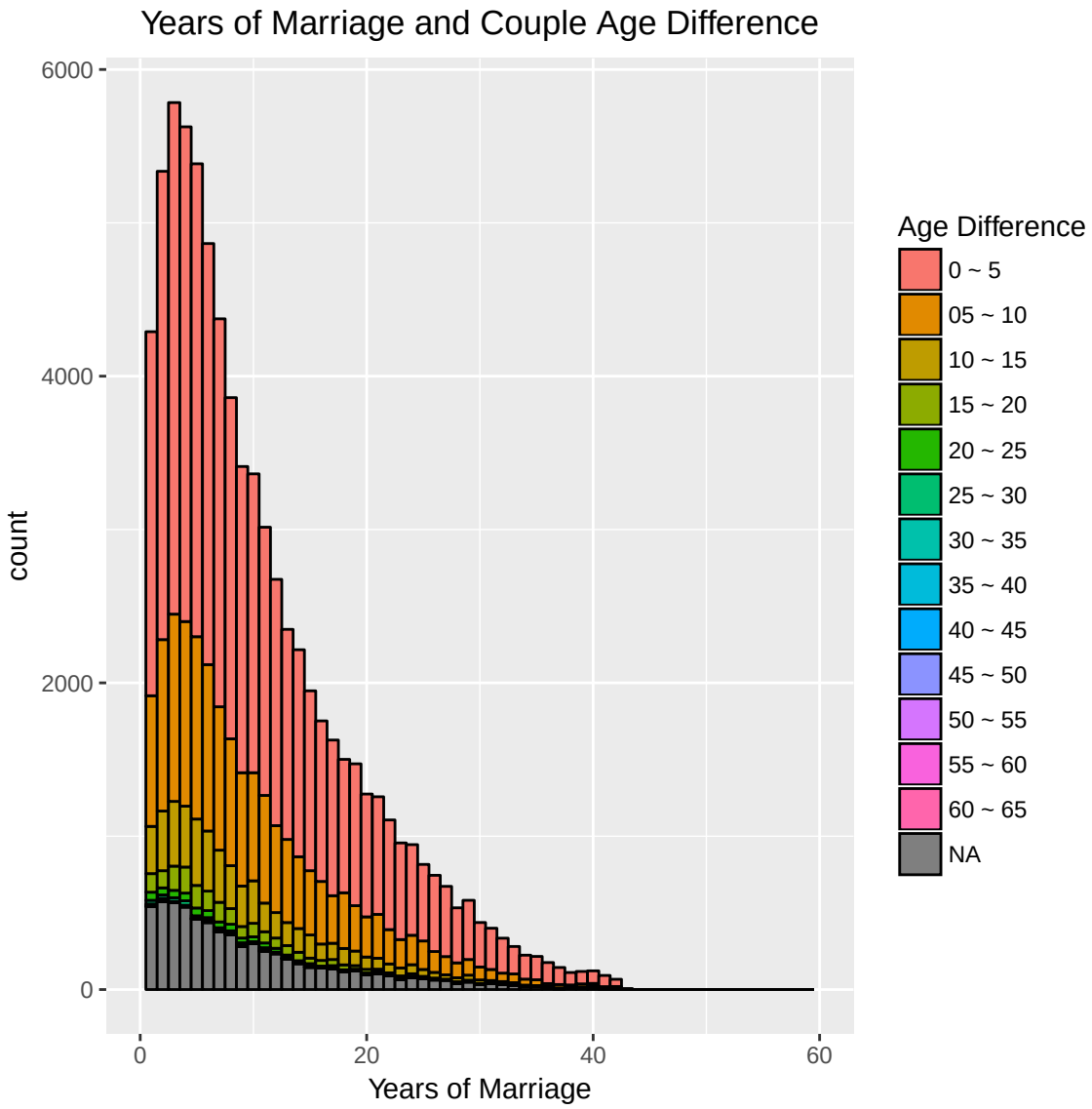


We can see from the regression that couple with longer marriaged years has a slightly bigger chance of having at least one child.

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

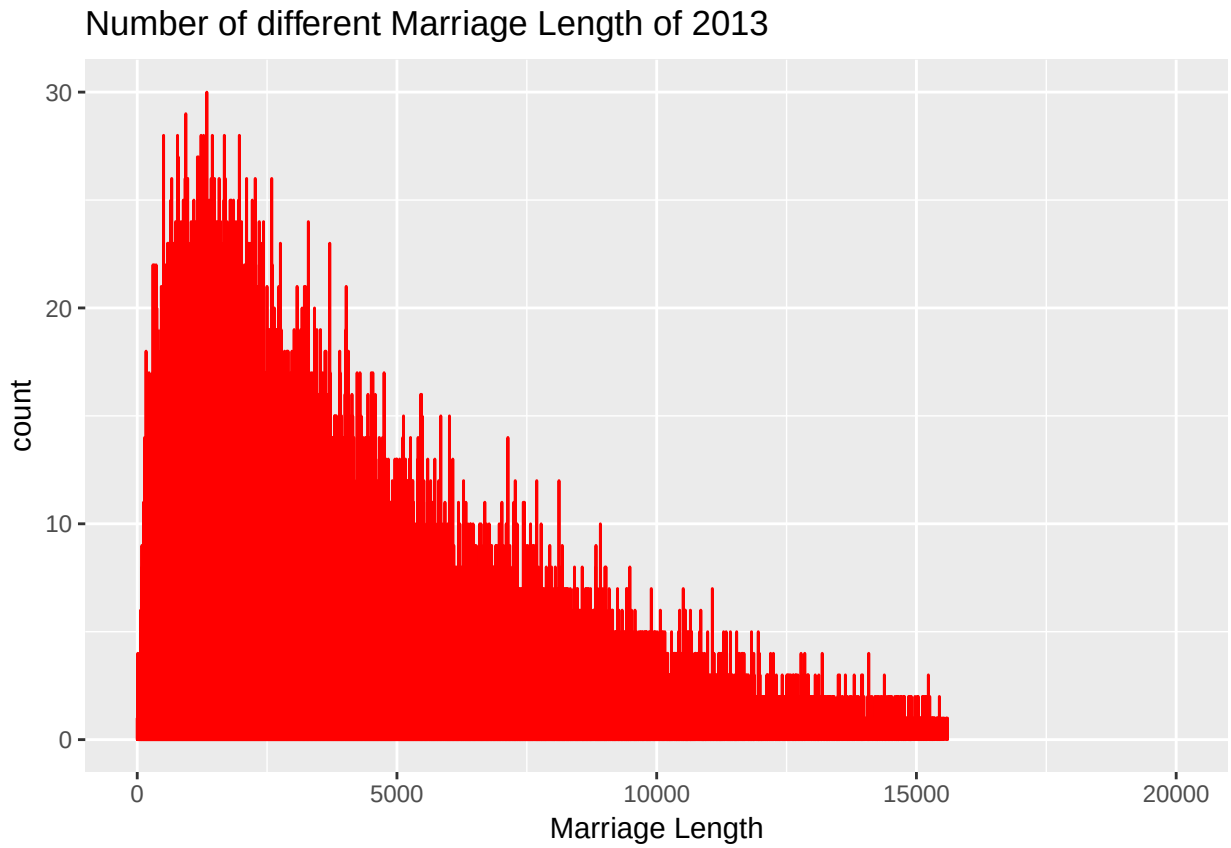
```
## Warning: NAs introduced by coercion
```

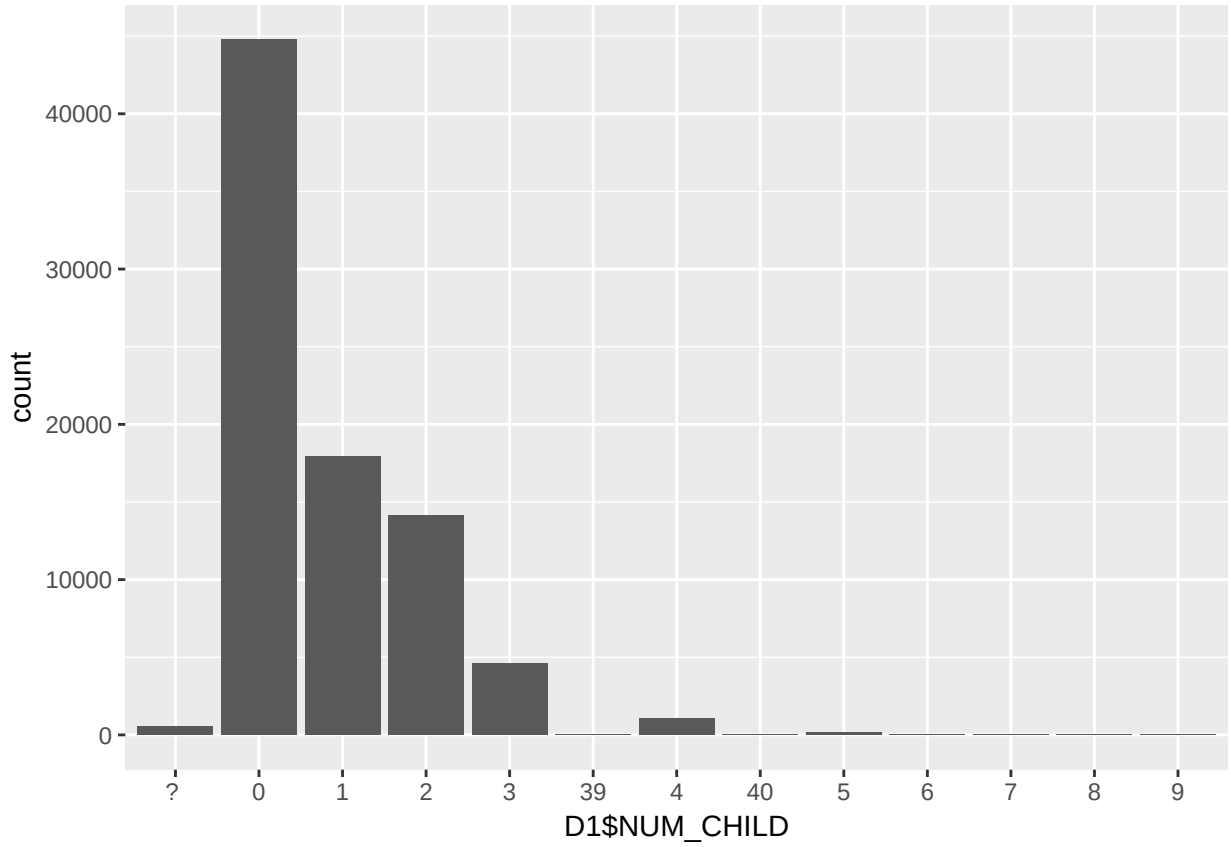
```
## Warning: Removed 6081 rows containing non-finite values (stat_bin).
```

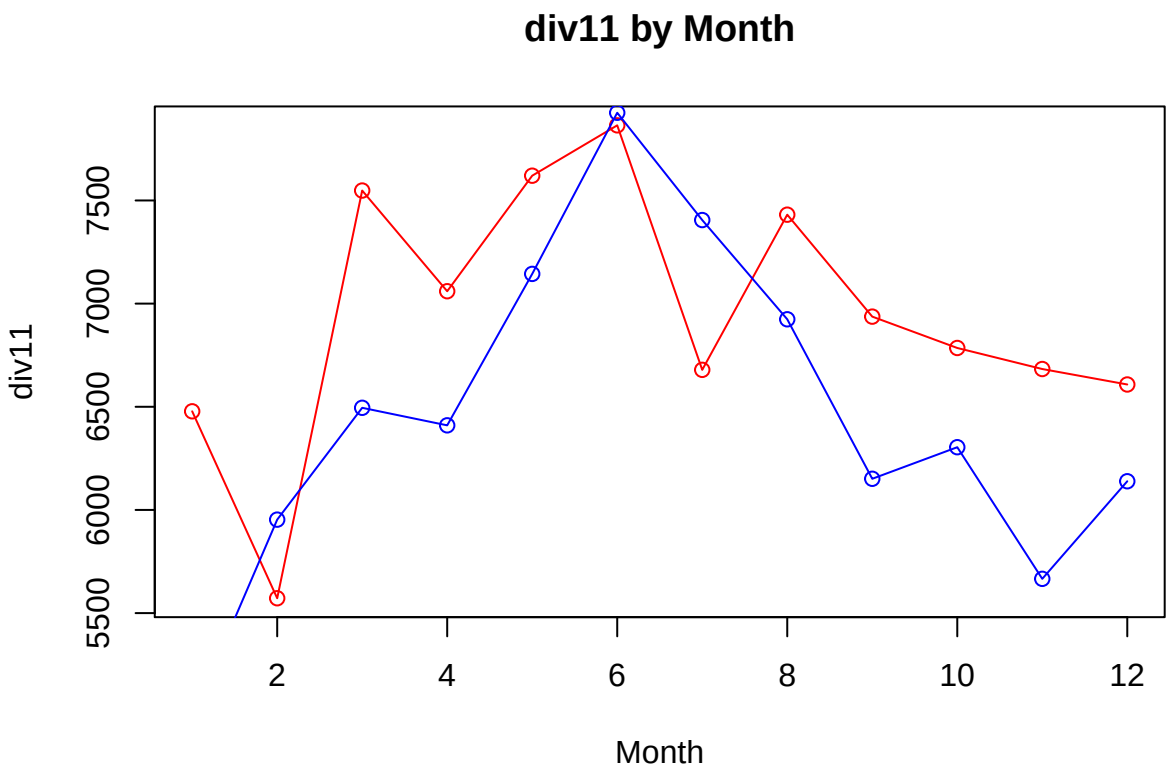


Marriage length of year 2013

```
## Warning: Removed 6081 rows containing non-finite values (stat_count).
```


Number of different Marriage Length of 2013
count
30
20
10
0
0
5000
10000
15000
20000
Marriage Length


count
40000
30000
20000
10000
0
?
0
1
2
3
39
4
40
5
6
7
8
9
D1$NUM_CHILD


div11 by Month
div11
5500
6000
6500
7000
7500
2
4
6
8
10
12
Month