STAT 209 Project 1

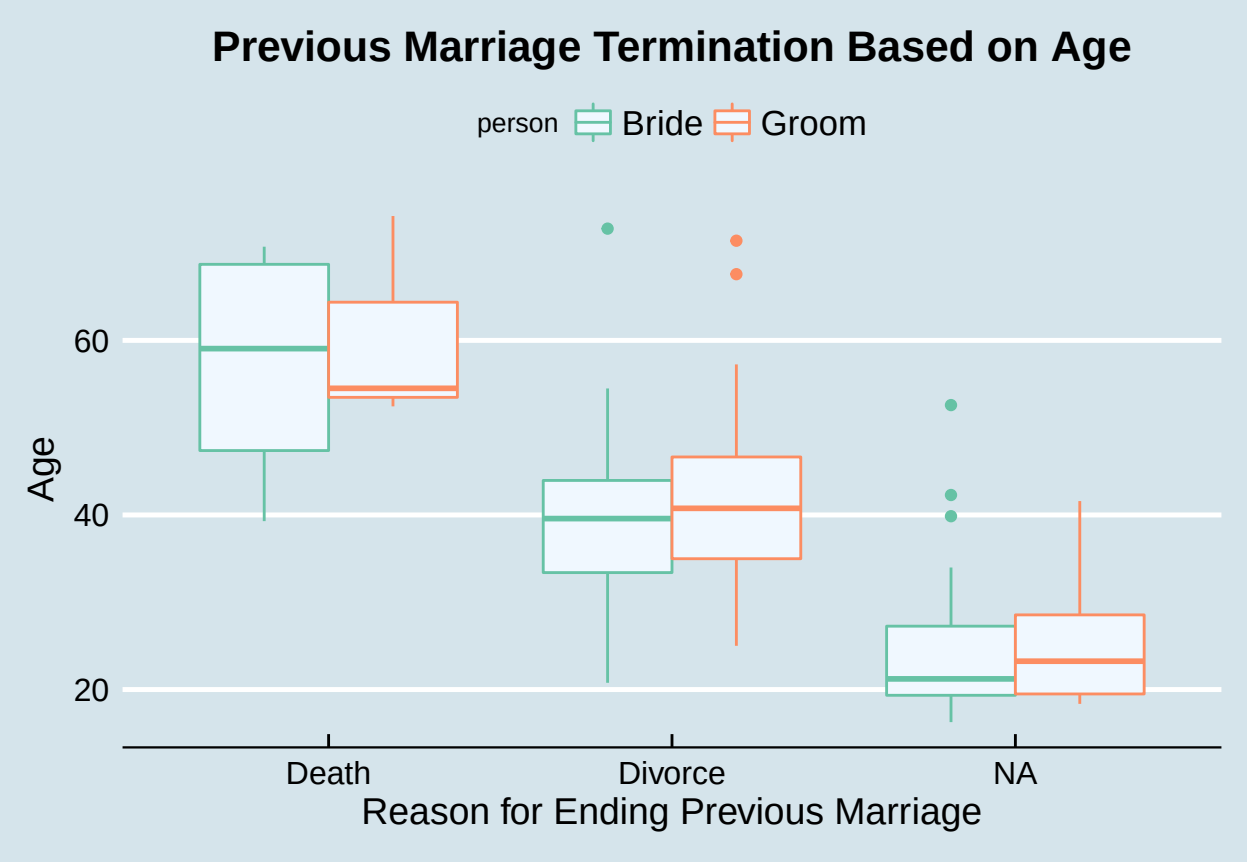
Jared Steinberg, Dizzy Farbanish, Julian Grabarek, Luis Solorzano

Introduction

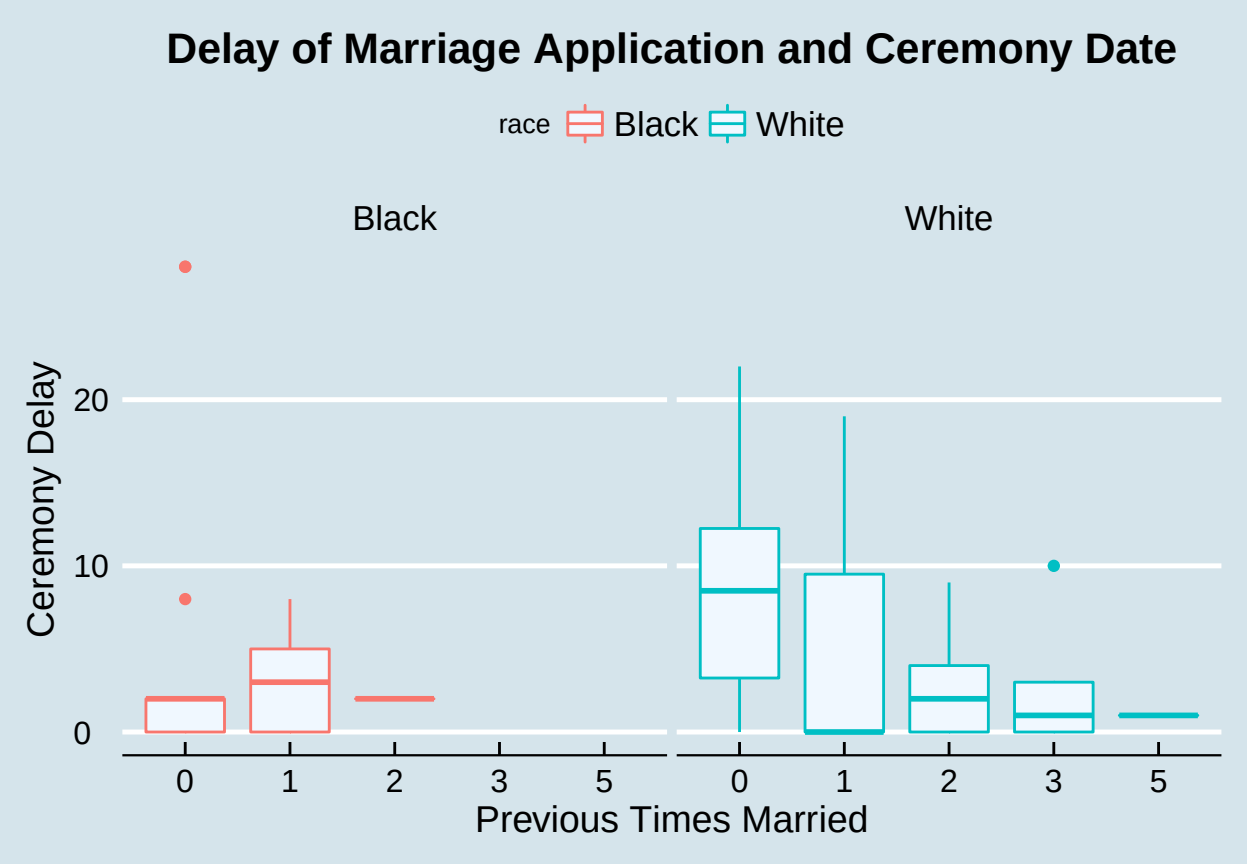
We are exploring the `Marraige` dataset provided in R package `mosaicData`. It provides marriage records from the Mobile County, AL probate court between 11/9/96 and 2/6/99. Variables recorded include:

- **age - age of person**
- **college - number of years in college**
- **person - a factor with levels Bride and Groom**
- **prevconc - way the last marriage ended**
- **prevcount - number of previous marriages**
- **race - race of person, being American Indian, Black, Hispanic, and White**

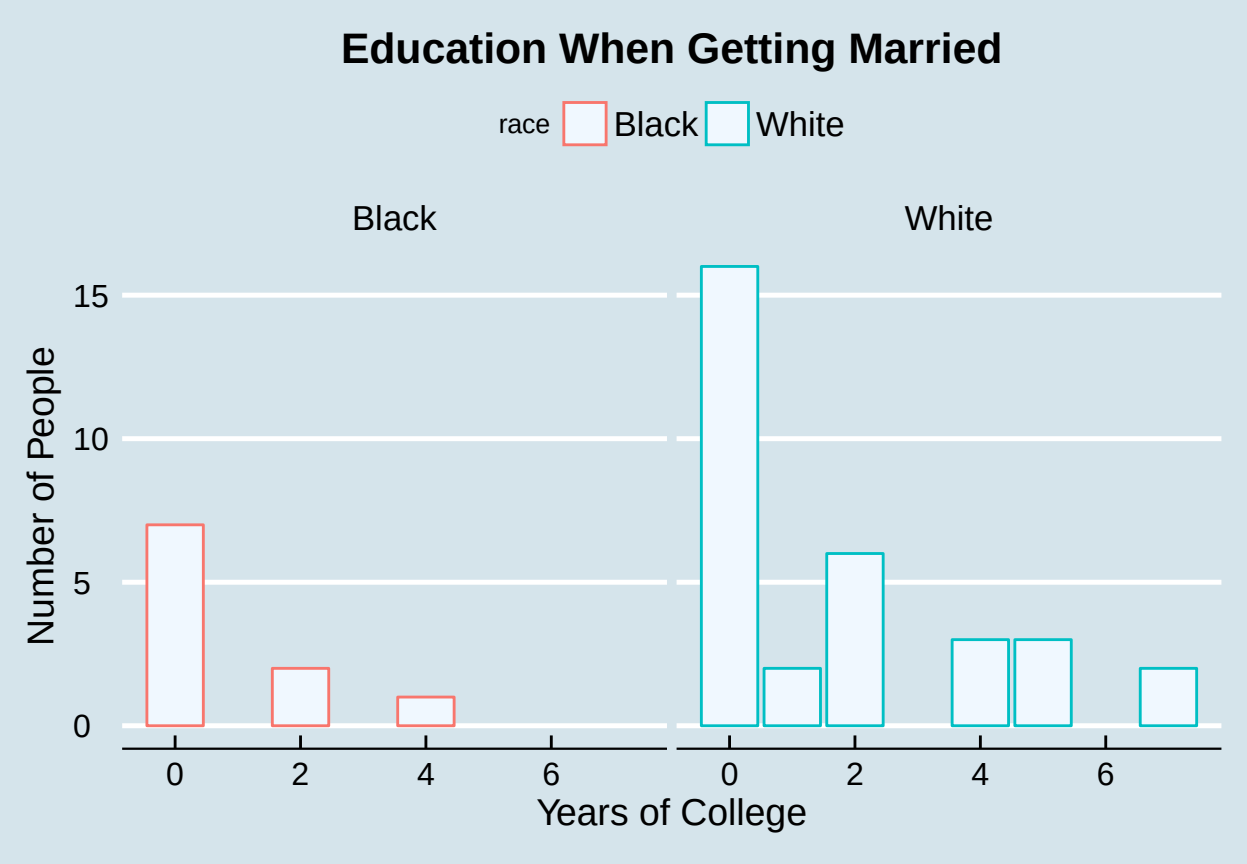
Results



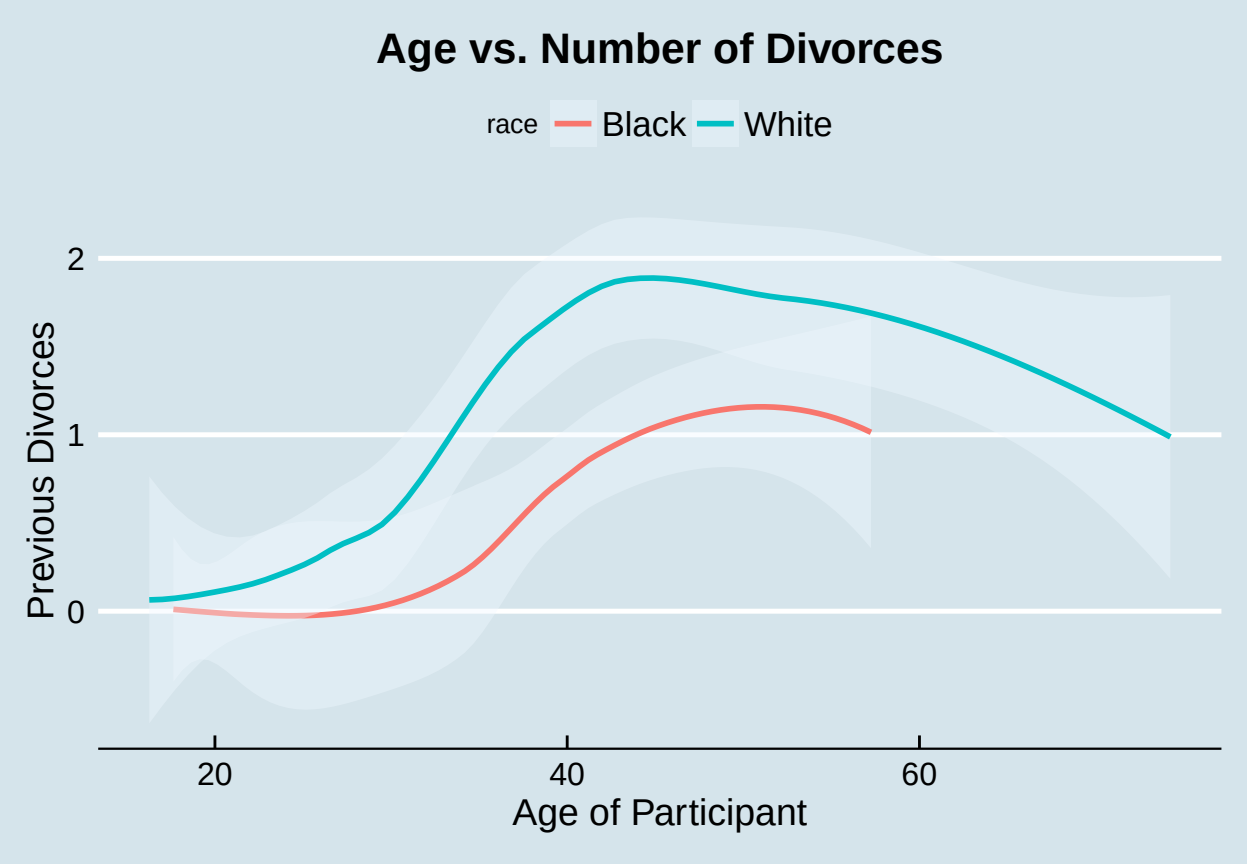
This graph illustrates how reasons for ending previous marriages change by age. The data shows that, for death being the reason the previous marriage ended, women's age vary more than men's, however have a lower median age. For divorcees, both intervals for men and women are approximately the same size and median age, probably because people in a relationship are usually the same age. The N/A category shows that people who have never had a previous marriage are typically young, regardless of sex.



In this graph, we see the time between marriage ceremonies and the application day is relative to how many times a person was previously married. Originally, we were looking to see if racial discrimination would cause more delays for Blacks, however the opposite was true. Note that for both races, the median is delay is fairly low and constant. However, For first time whites getting married, the delay was longer. The smaller inner quartile ranges as previous marriages increases is probably due to fewer people with 2 or more previous marriages.



This bar graph depicts the number of people that have completed a certain number of years upon eloping. Then the visual is separated by race to distinguish between black and white brides/grooms and their education level upon marriage. Those that did not attend college are placed in the 0 category for years. The data that this graph was derived did not include partial years such as semesters, or quarters. After analyzing the data, it is clear that most people, regardless of race, are not attending college upon getting married. A caveat to this data is that the sample size is extremely small.



Data shows that older generations have actually had less divorces than younger generations. This suggests that societal norms surrounding divorce have changed from the older to the younger generations. It is expected that the divorces averages should trend up with age but it is unexpected that divorce would be lower for older people from previous generations. Additionally it is interesting to note that divorces among african americans are lower on average than white people. The difference is statistically significant at 40, as the 95% confidence intervals of each line do not overlap.

Methodology

In general, we used boxplots for catagorical independent variables with a continuous dependent variable. We chose a bar graph when counting the quantity in a catagorical variable. Finally, we chose `geom_smooth()` in the 4th graph because we were looking for a trend line as people age rather than which ages have more divorces. When comparing race in multiple graphs, we wanted to visually show the same distinction throughout by having the colors correspond in each graph.