

Key Indicators for Depression Based on Age Groups

Colin Clapham

2017/11/30

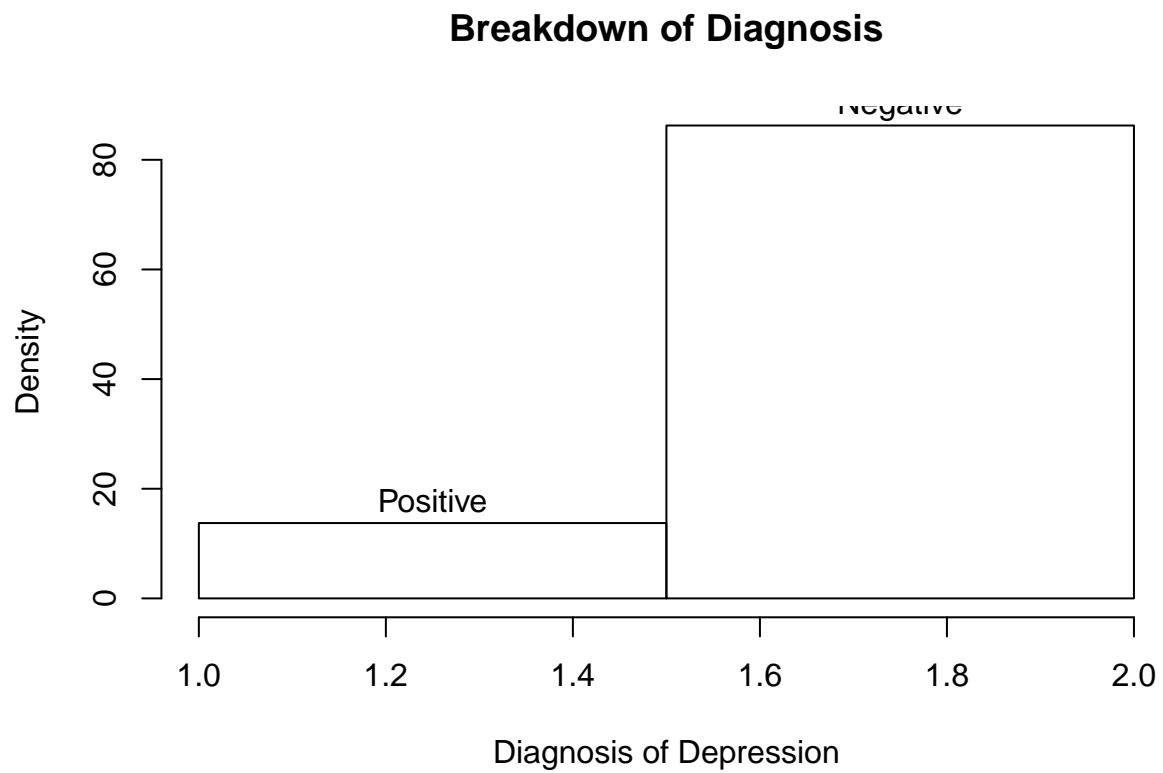
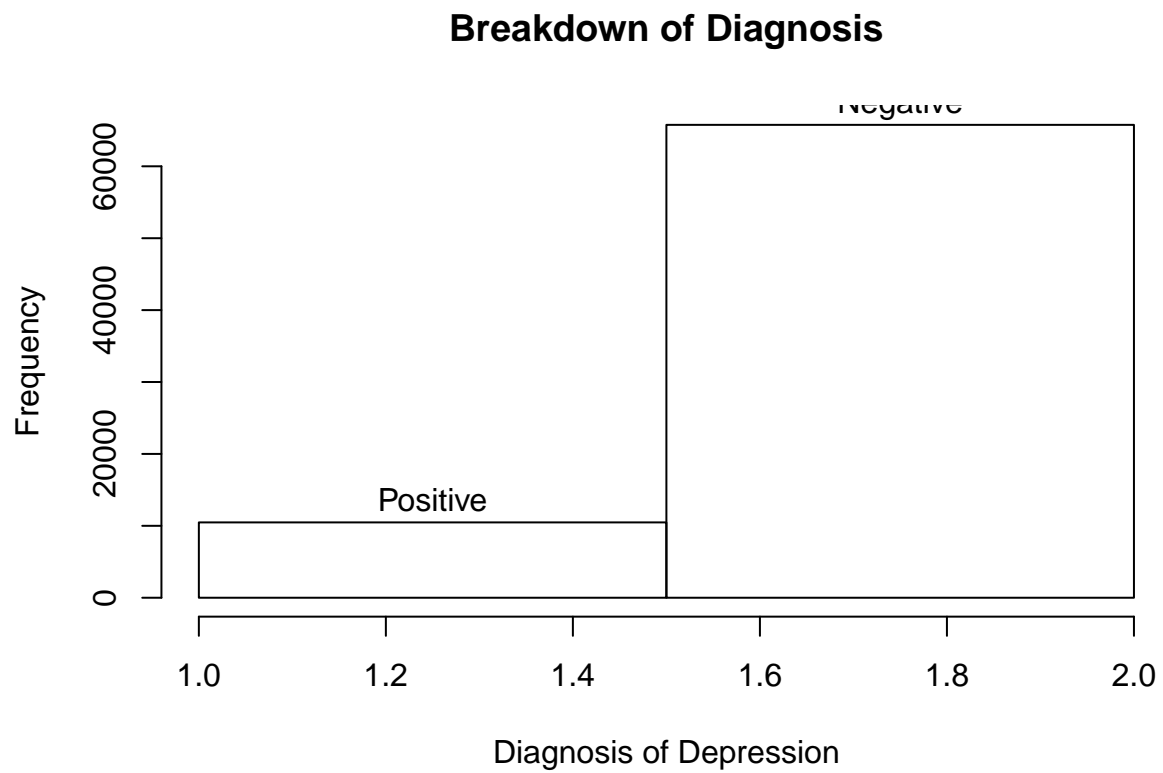
Project Background

Depression is a mental health disorder that affects the daily lives of those suffering from the illness. In the US alone, it is estimated that over three million people are diagnosed with depression every year. While the disorder can cause significant impairment to daily life, it has become very common and treatments are available. Causes of depression have been linked to biological, psychological, and social sources of distress. Trauma from one of these categories can lead to many symptoms, most commonly a depressed mood combined with a lack of interest in daily activities. Most often, treatment involves medication and/or therapy.

Research Objectives

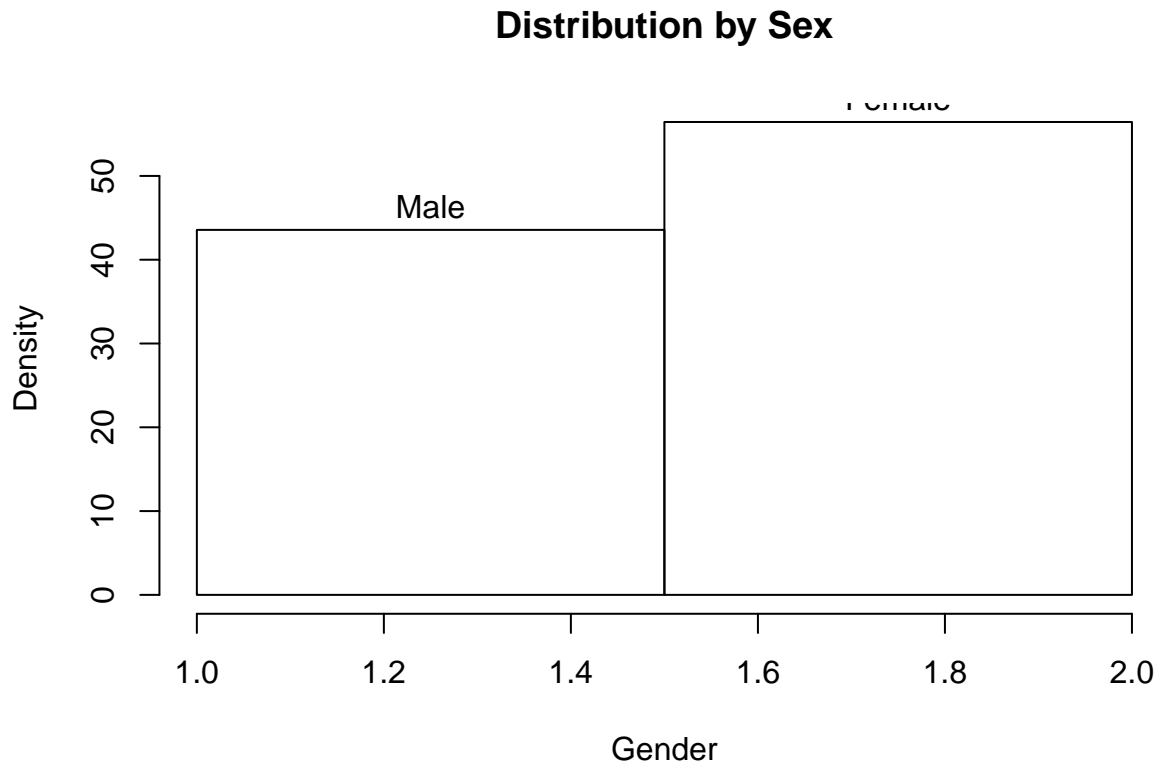
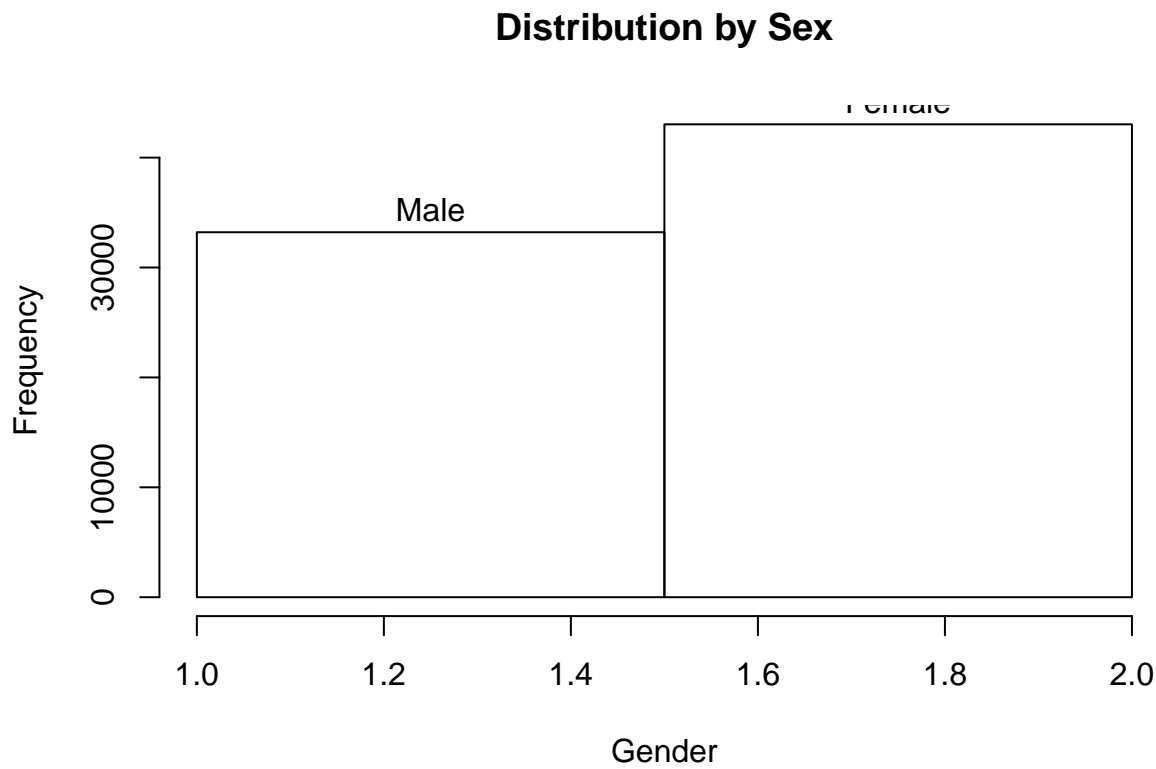
Depression is a disease that affects all ages from early teens all the way to the end of life. The disorder also goes undiagnosed for many with a number of people refusing care or even receiving improper diagnosis. It is extremely important to understand the factors that lead to the breakdown of the psyche. Through this research, I intend to dig deeper into the social sources of distress and their effect on patients across multiple age brackets. There may be social constructs which affect people with depression in different ways based on their age category. I intend to look at variables such as education levels, family status, activity levels, etc. and analyze their contribution to whether someone was diagnosed with depression.

Distribution of Positive versus Negative Diagnosis



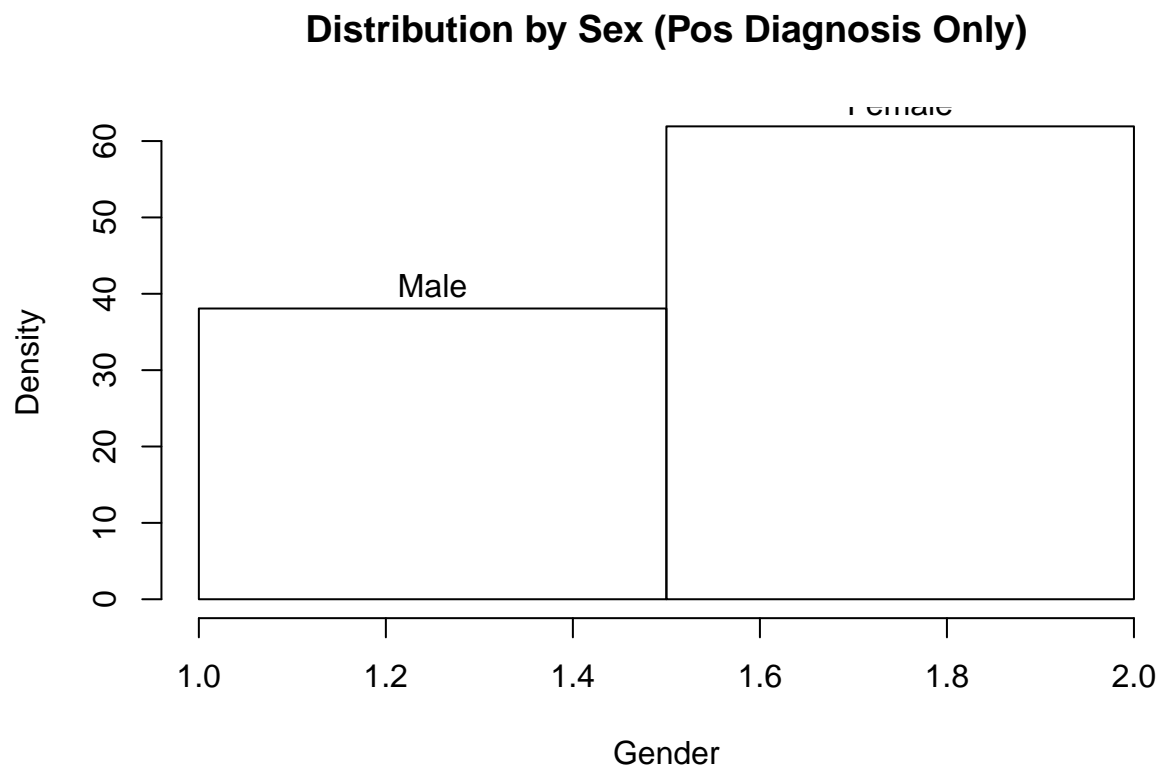
The histogram indicates that out of our ~686,000 patients, around 1/7 of them have a diagnosis of depression. Compared to the diagnosis rate of the United States, our sample closely matches the population parameter of around 1% diagnosed per year.

Distribution of Gender (Positive and Negative Diagnosis)



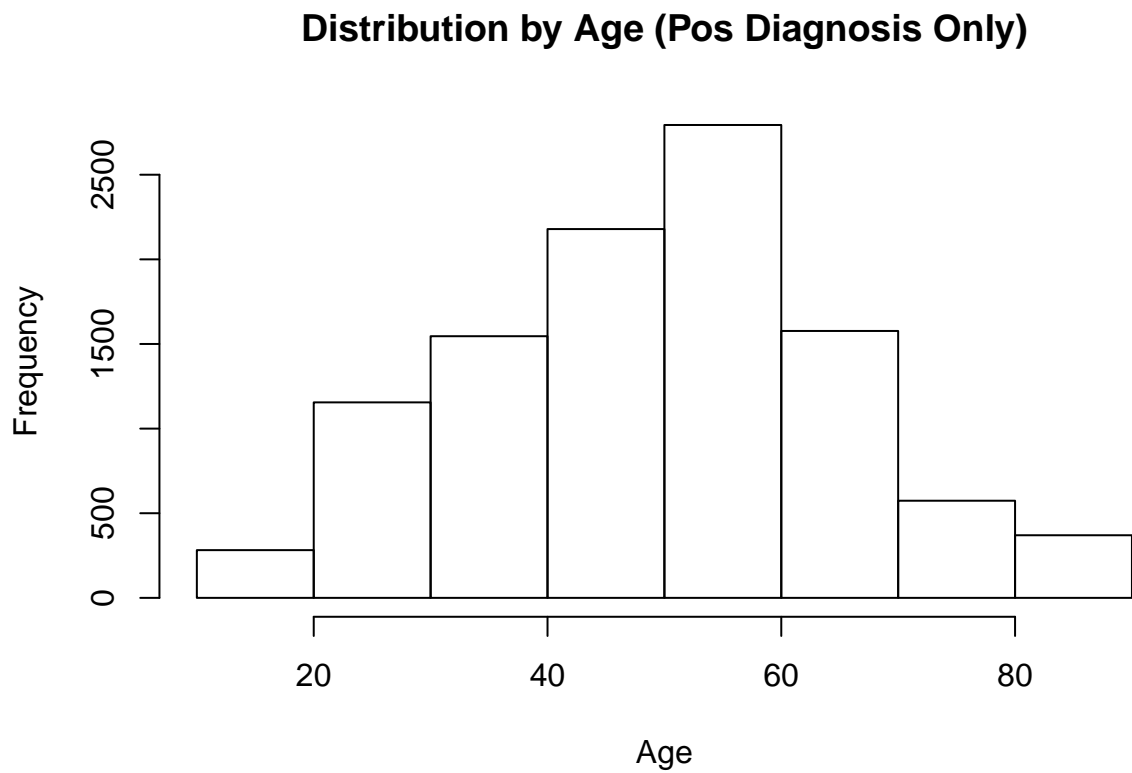
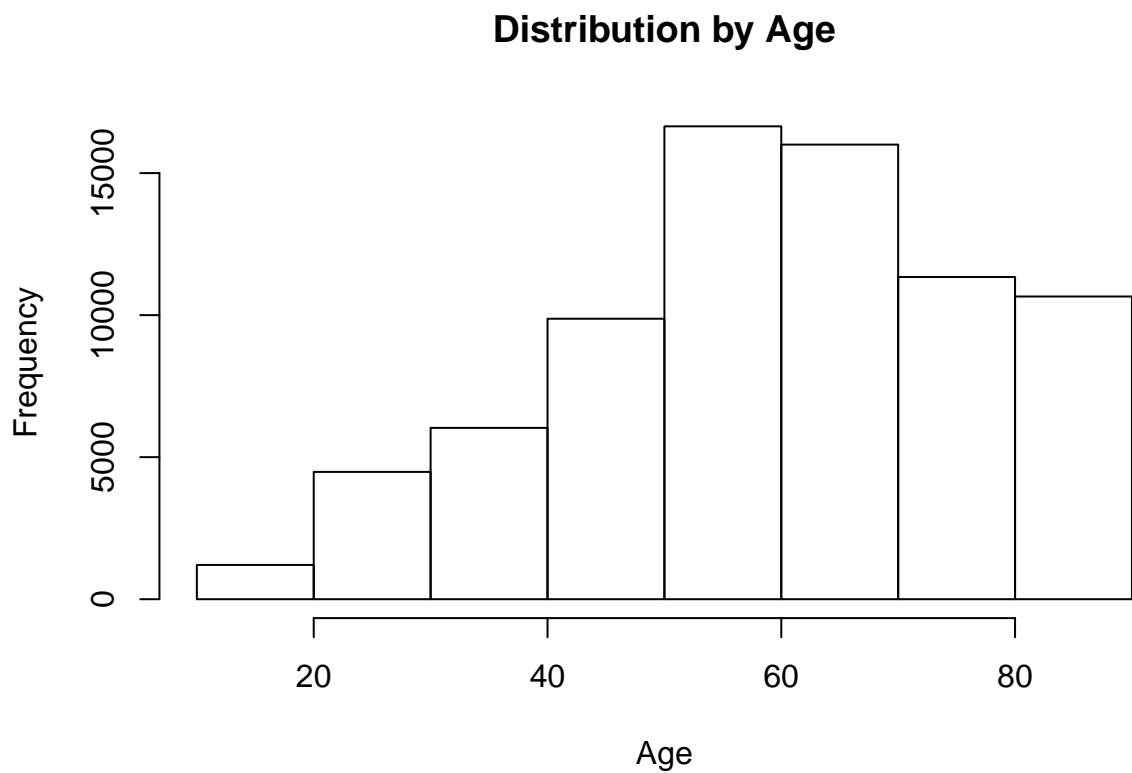
Distribution by Sex (Pos Diagnosis Only)





According to the Anxiety and Depression Association of America, women are twice as likely to be depressed as men. The subset of data looking only at those with a diagnosis reflects a truer representation of the population of those with depression than the master survey.

Distribution of Age (Positive and Negative Diagnosis)

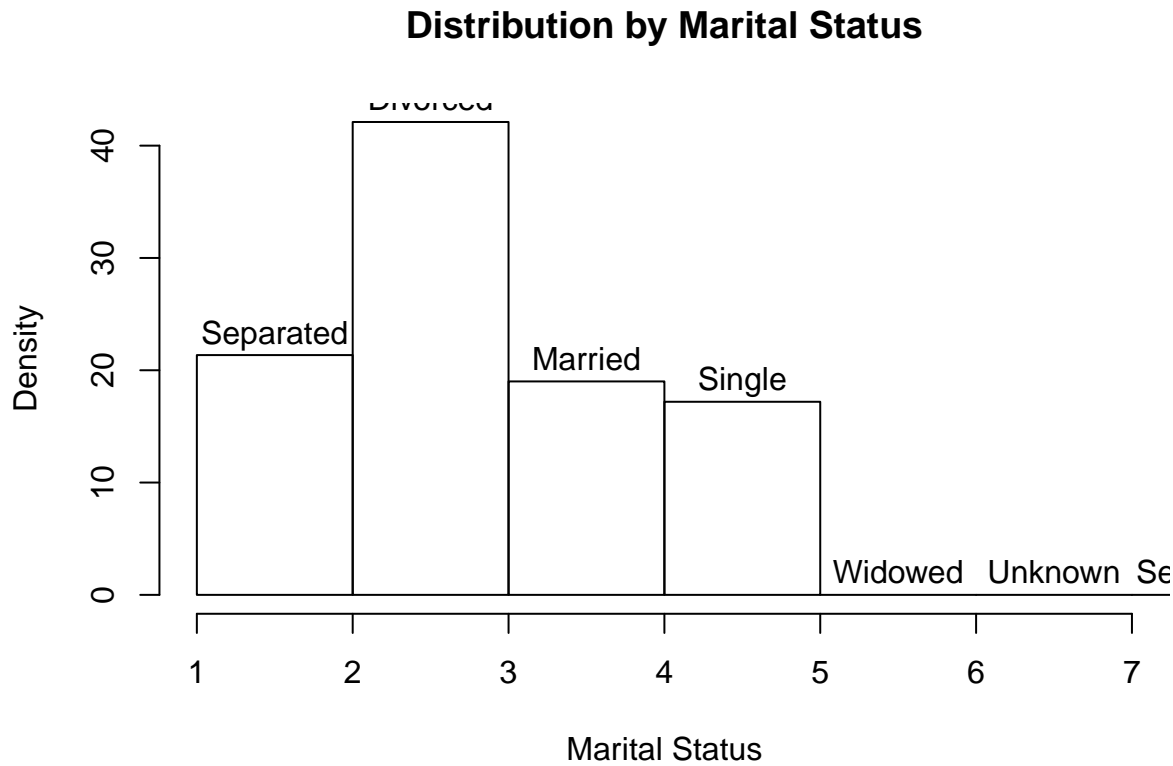
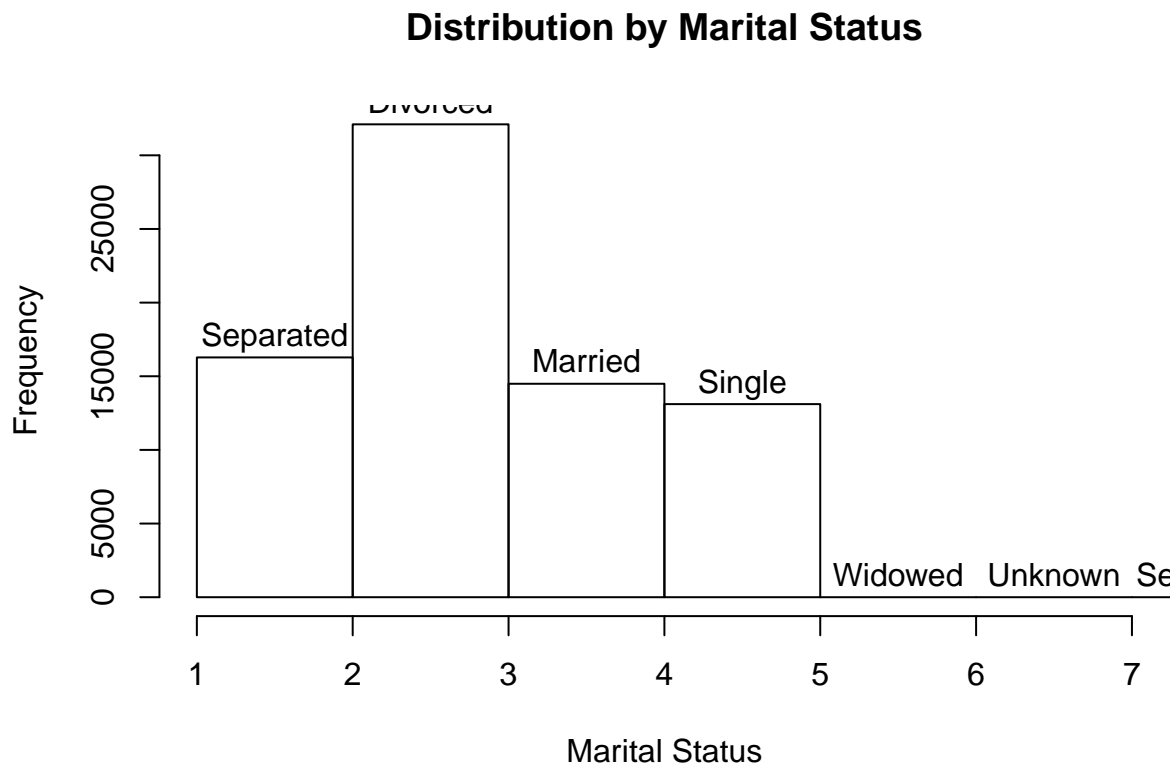


Median onset age of Major Depressive Disorder is 32.5. While the master survey reflects an even spread of ages across the data, the subset of those diagnosed with depression includes older individual that may be a close approximation of the overall population.

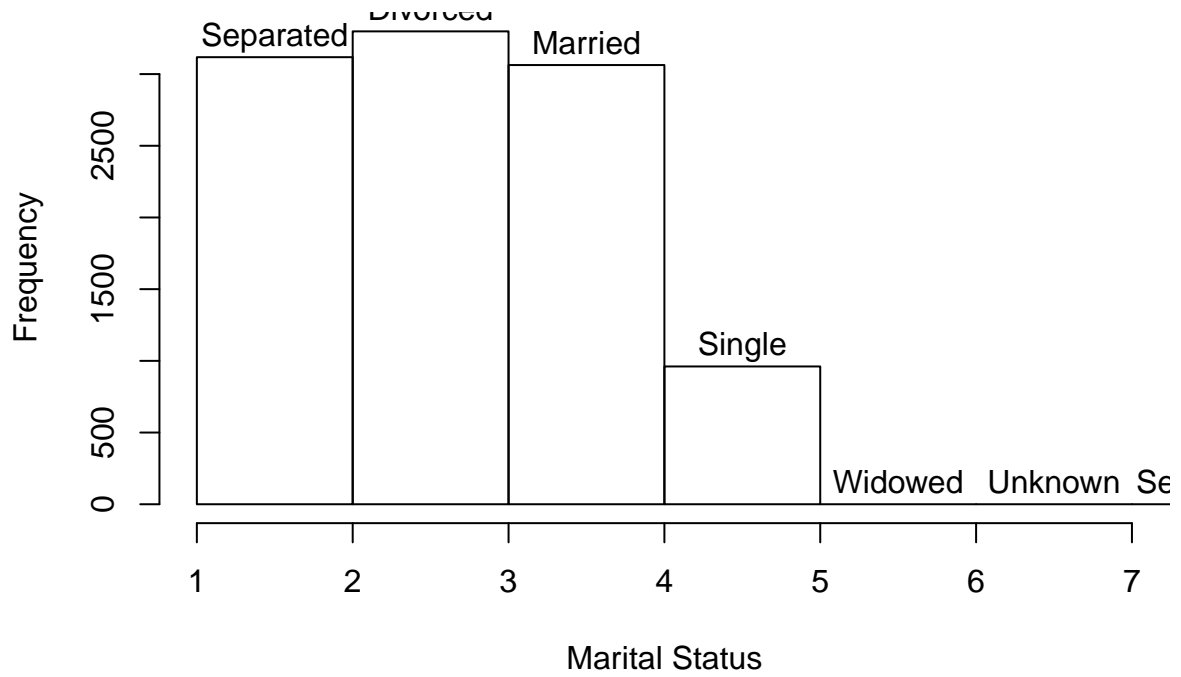
Social Factors

Given the nature of this study, I will be digging into social factors that may contribute to a diagnosis of depression. These include marriage status, family status, social limitations, education level, and job status.

Distribution of Marital Status (Positive and Negative Diagnosis)



Distribution by Marital Status(Pos Diagnosis Only)



Distribution by Marital Status (Pos Diagnosis Only)



Fitting a Model

```
## Loading required package: Matrix
## Warning: package 'arm' was built under R version 3.4.2
## Loading required package: MASS
##
## Attaching package: 'MASS'
## The following object is masked from 'package:dplyr':
##
##   select
##
## arm (Version 1.9-3, built: 2016-11-21)
## Working directory is C:/Users/Colin/Documents/MSSP/Multiple Regression/Final
## Warning in optwrap(optimizer, devfun, getStart(start, rho$lower, rho$pp), :
## convergence code 1 from bobyqa: bobyqa -- maximum number of function
## evaluations exceeded
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control
## $checkConv, : unable to evaluate scaled gradient
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control
## $checkConv, : Model failed to converge: degenerate Hessian with 1 negative
```

```
## eigenvalues

## lmer(formula = Diagnosis ~ SEX + CDCMSTAT + PARENTS + EDUC1 +
##       WRKHRS2 + (1 + CDCMSTAT | Age_Range) + (1 + PARENTS | Age_Range) +
##       (1 + EDUC1 | Age_Range) + (1 + WRKHRS2 | Age_Range), data = Master)
##           coef.est coef.se
## (Intercept)  0.85      0.14
## SEX          -0.03      0.01
## CDCMSTAT      0.01      0.00
## PARENTS       0.00      0.00
## EDUC1         0.00      0.00
## WRKHRS2       0.00      0.00
##
## Error terms:
##   Groups      Name      Std.Dev. Corr
## Age_Range    (Intercept) 0.15
##              CDCMSTAT    0.00   -1.00
## Age_Range.1  (Intercept) 0.19
##              PARENTS     0.01   -1.00
## Age_Range.2  (Intercept) 0.17
##              EDUC1       0.00    1.00
## Age_Range.3  (Intercept) 0.24
##              WRKHRS2     0.00   -1.00
## Residual                    0.32

## Warning in optwrap(optimizer, devfun, x@theta, lower = x@lower, calc.derivs
## = TRUE, : convergence code 1 from bobyqa: bobyqa -- maximum number of
## function evaluations exceeded

## ---
## number of obs: 15260, groups: Age_Range, 8
## AIC = 8854.6, DIC = 8656
## deviance = 8736.3

##      (Intercept)      SEX      CDCMSTAT      PARENTS      EDUC1
## 10  0.8892345 -0.03421712 0.006333714 0.0047141863 0.0006205978
## 20  0.8083067 -0.03421712 0.006658716 0.0063292136 0.0006198744
## 30  0.7236773 -0.03421712 0.006998582 0.0009242721 0.0006188967
## 40  0.7155473 -0.03421712 0.007031232 0.0068344653 0.0006204288
## 50  1.2929666 -0.03421712 0.004712350 0.0035758868 0.0006216567
## 60  0.6940682 -0.03421712 0.007117491 0.0050133264 0.0006184099
## 70  0.7506197 -0.03421712 0.006890383 0.0057158438 0.0006189299
## 80  0.9419760 -0.03421712 0.006121907 0.0040444471 0.0006203400
##      WRKHRS2
## 10 0.0013596532
## 20 0.0009858918
## 30 0.0011951121
## 40 0.0006963340
## 50 0.0014985321
## 60 0.0003350513
## 70 0.0002857923
## 80 0.0006982405
```

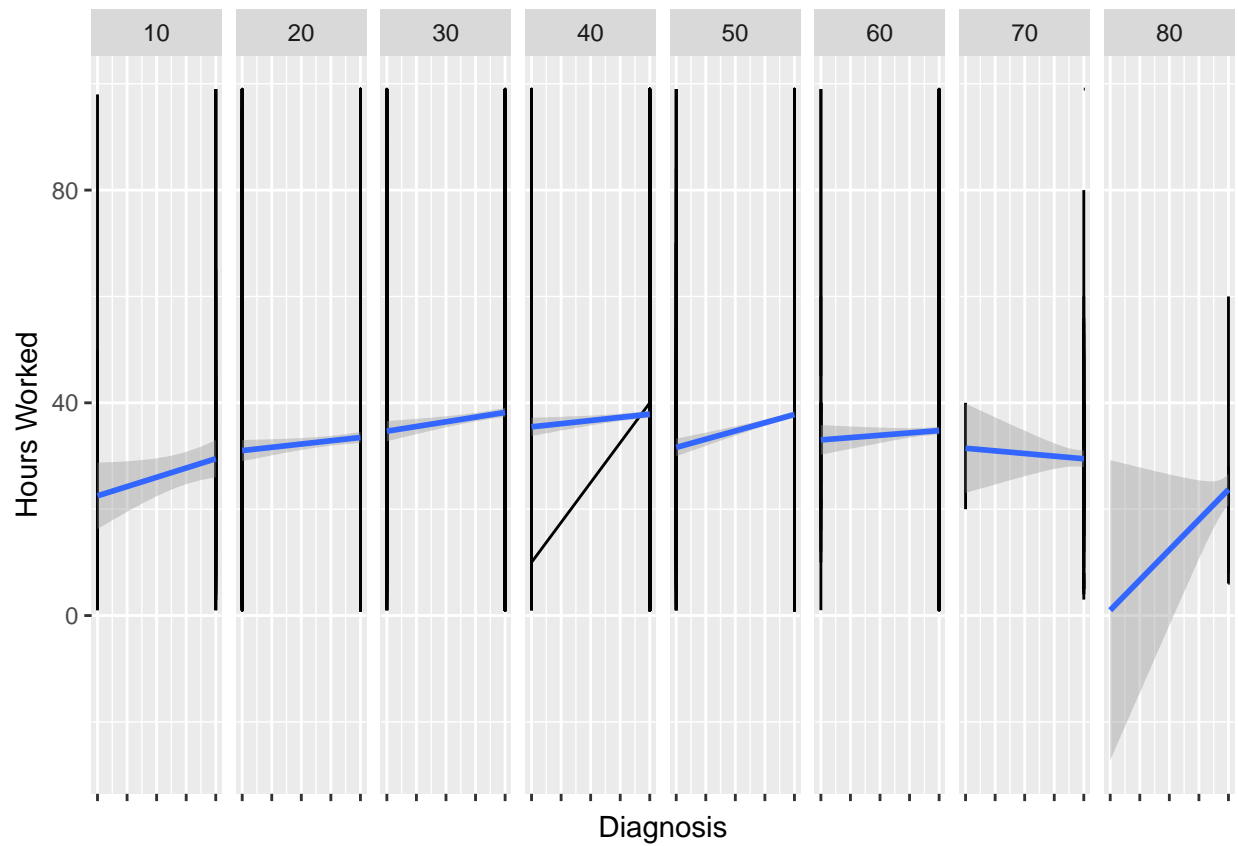
Fitting a multi-level, mixed-effect model, we can see that the possibility of diagnosis of depression varies by age. In addition, the social factors observed effect different ages to varying degrees. People in their 50's seem to have the highest chance of being diagnosed with depression (independent of social factors). This may have

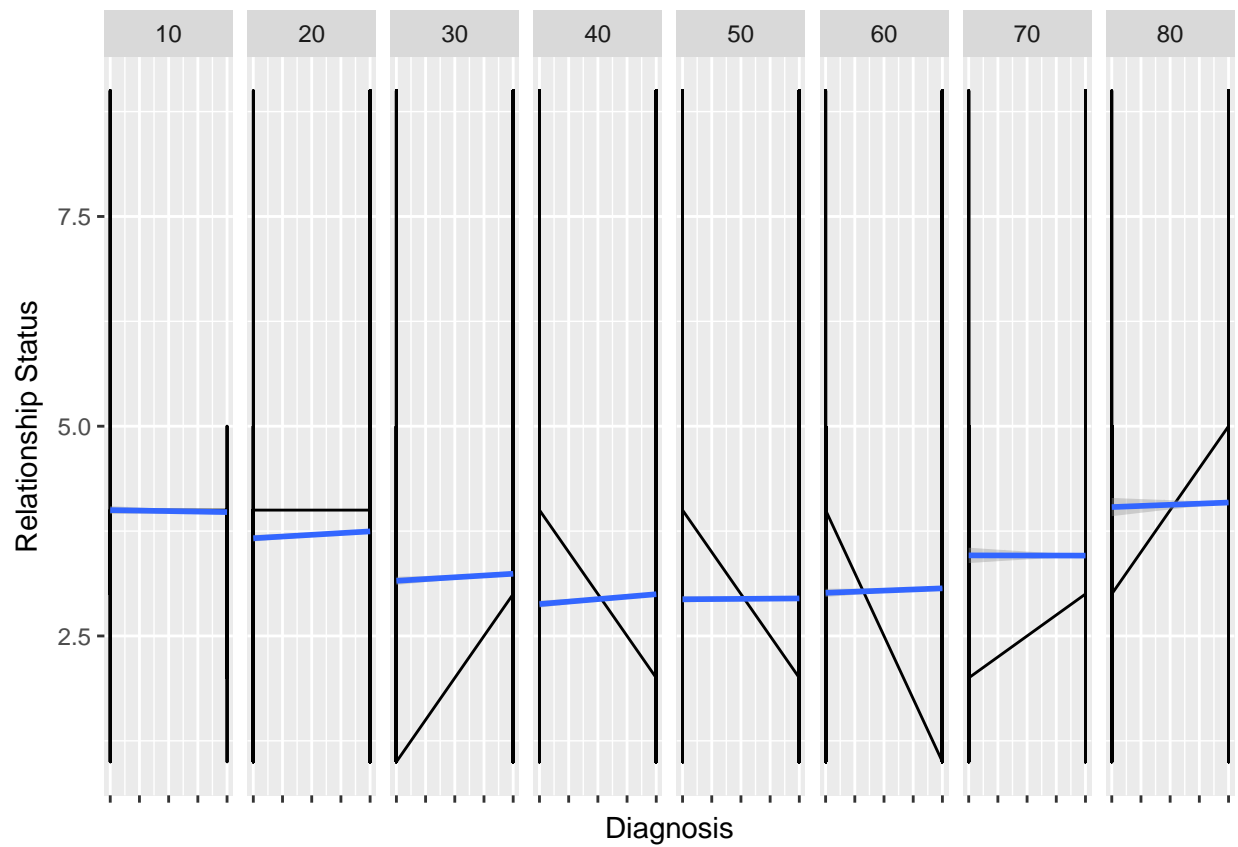
something to do with the effect of a mid-life crisis. In fact, of the four social factors observed, people in their fifties are effected to the greatest degree by two of the four (education level and hours worked per week). For the other social factors, people in their 60's seem most effected by relationship status while people in their 40's are most effected by family status.

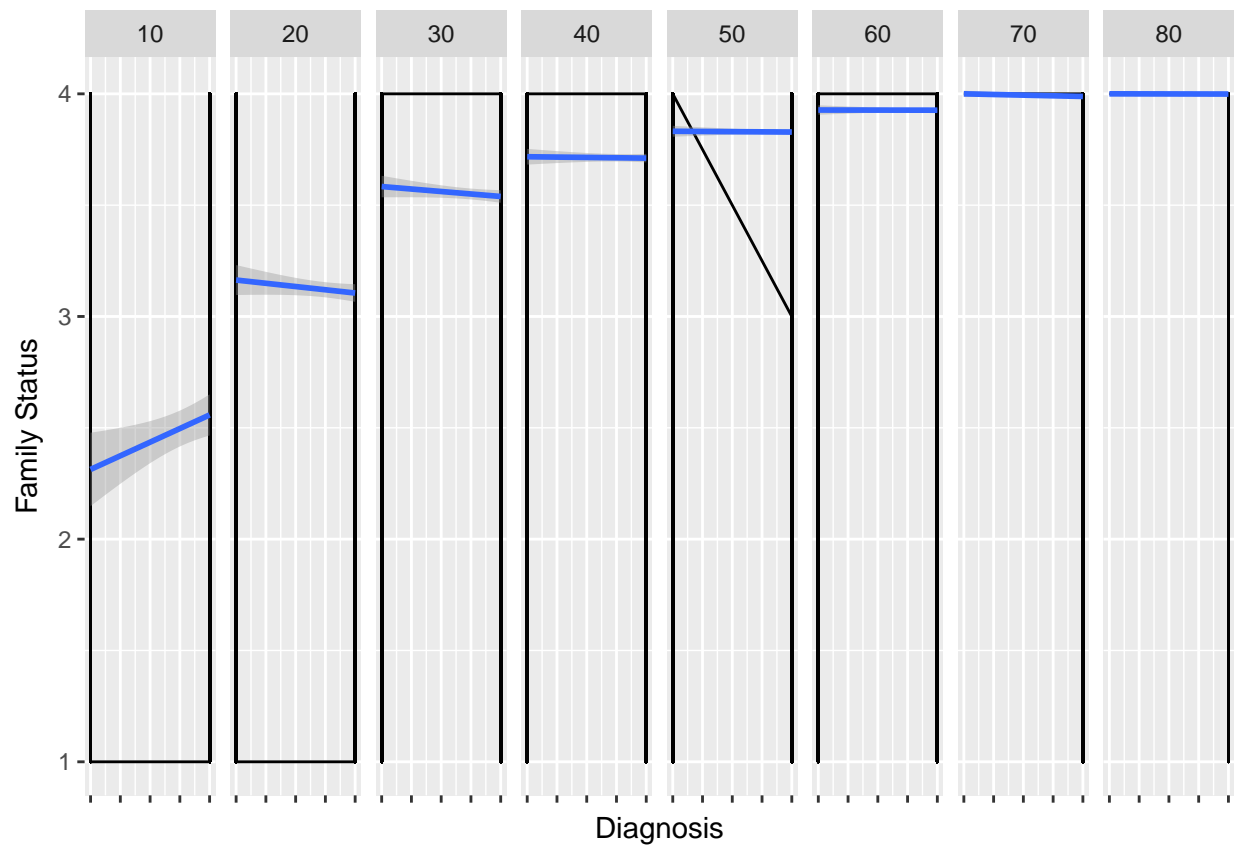
Graphs of Social Effects and Depression

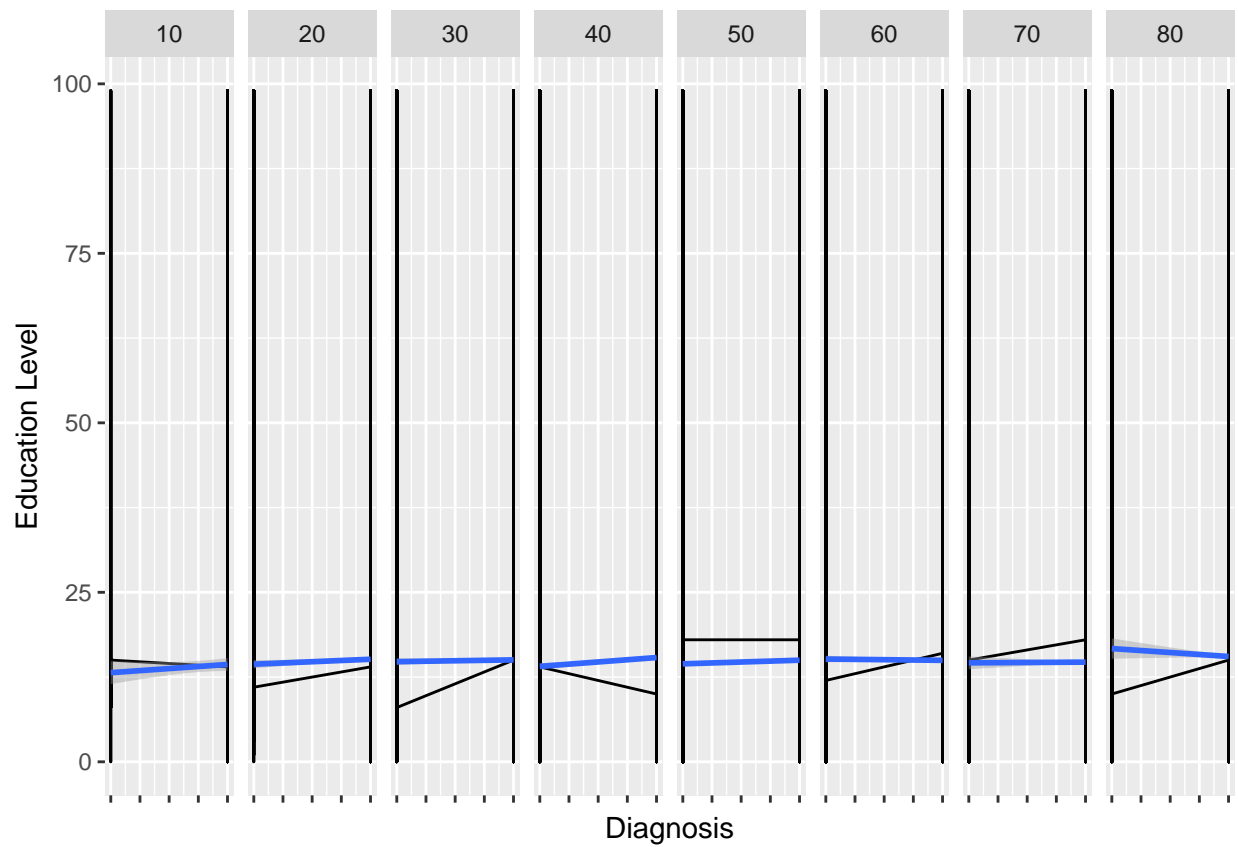
```
## Warning: Removed 60983 rows containing non-finite values (stat_smooth).
```

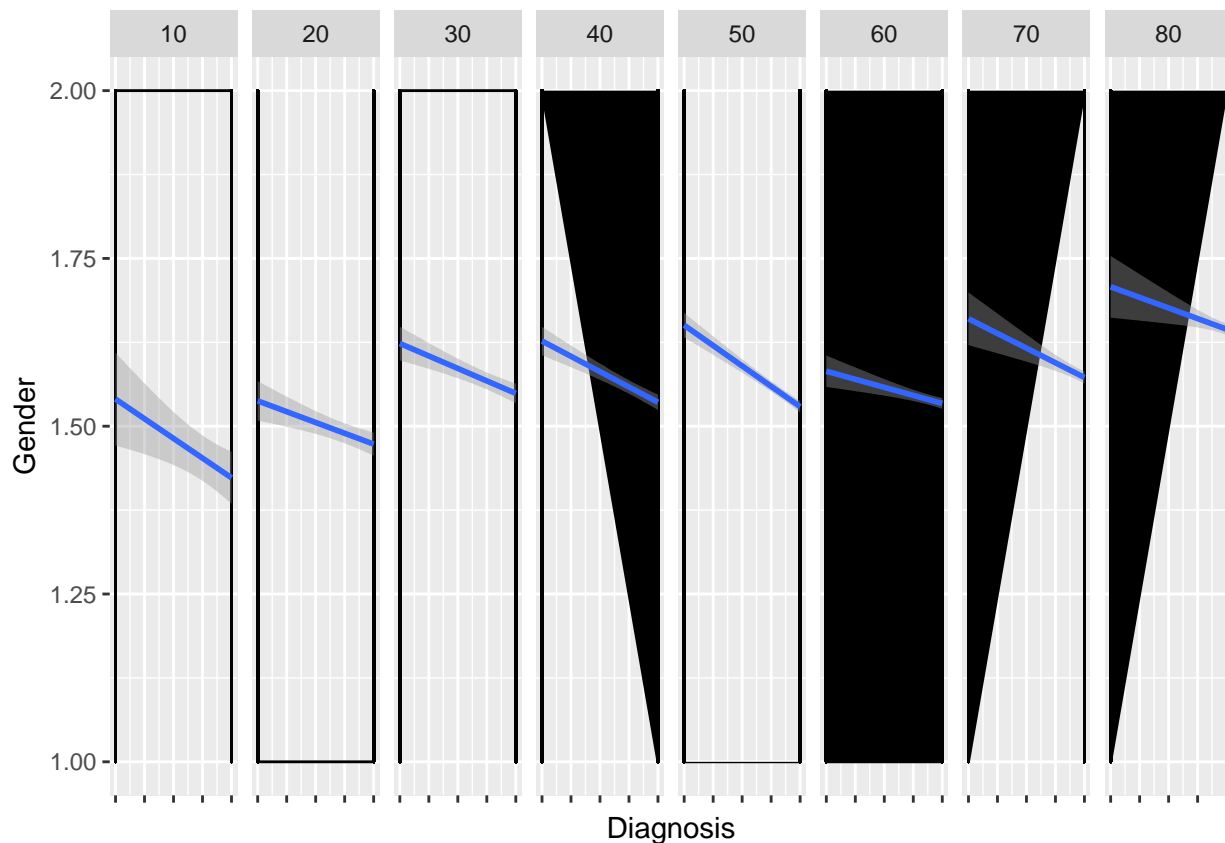
```
## Warning: Removed 70 rows containing missing values (geom_path).
```











A. The amount of hours worked increases into the 40's and begins to drop off at age 50 and beyond. This is expected as people begin to enter retirement age. The slopes begin more steep as age increases with work hours having a greater effect on depression diagnosis. At ages in the 40's and 50's, we see the steepest slope (people in their teens and 80's have smaller sample sizes).

B. As people approach mid-life, they are more likely to be in (or have been) in a relationship. Younger people tend to be single while older people tend to be widowed. Relationship status seems to have the biggest positive effect in people in their 30's and 80's.

C. Initially, I thought that parental status would have the greatest effect on diagnosis of depression, especially among younger patients. However, given the data, it appears that parental status does little to change the probability of receiving a diagnosis of depression. The data trends in an expected fashion (older respondents have lost their father and mother by the time they have taken the survey).

D. Education level seems to have a greater effect on people in their 30's and 40's. This may relate to someone in their 30's who is still in school not knowing what they want to do for a career. Interestingly, education level in the 40's age bracket decreases depression diagnosis. It is possible people in their 40's may be pursuing further education to revitalize or change careers.

E. As age increases, it becomes more probable that women receive a positive diagnosis of depression. The biggest jump (besides 70's to 80's) appears to be from the 20's to 30's. This may be due to issues with fertility and the closing window to have children.