Assignment 3

```r
rm(list=ls())
library(foreign)
library(ggplot2)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --
```

```
## v tibble  3.0.6     v dplyr   1.0.4
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1
## v purrr   0.3.4
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library(haven)
library(MASS)
```

```
##
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':
##
##     select
```

```r
library(dplyr)
library(ISLR)
library(glmnet)
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'Matrix'
```

```
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
```

```
## Loaded glmnet 4.1-1
```

```r
library(arm)
```

```
## Loading required package: lme4
```

```
##
## arm (Version 1.11-2, built: 2020-7-27)
```

```
## Working directory is /home/enminz/Graduate/Data-2020/Assignment 3
```

```r
library(coefplot)
```

```
##
## Attaching package: 'coefplot'
## The following objects are masked from 'package:arm':
##
##     coefplot, coefplot.default, invlogit
```

NAME: Your Name
DUE DATE: March 9th, 11:59pm

Problem 1 (100 pts)

In the folder Assignment 3, you will find the data set called FF_wave6_2020v2.dta. This data set is from the Fragile Family Data Set, and it includes many different variables (socio-demographic, economics, and health status) of teenagers (15 years old) and their parents. The codebook (ff_wave6_codebook.txt) associated with the data set is on Canvas (folder Assignment 3).

(a) (20 points) Consider the variable *doctor diagnosed youth with depression/anxiety*. In the data set, the name of this variable is *p6b5*. Then consider in the data set these variables: *p6b10*, *p6b35*, *p6b55*, *p6b60*, *p6c21*, *p6f32*, *p6f35*, *p6h74*, *p6h102*, *p6i7*, *p6i8*, *p6i11*, *p6j37*, *k6b21a*, *k6b22a*, *k6c1*, *k6c4e*, *k6c28*, *k6d37*, *k6f63*, *ck6cbmi*, *k6d10*. Now, you have a data set with 4898 subjects and 23 variables. Clean the data in these three steps. 1- Each variable has a value with a number and a text (for example, a value for the variable *p6b5* is *2 No*). Remove the text from all the variables in the data set (hint: use the function sub for each column). 2- Transform each variable in numeric (hint: use the function as.numeric for each column). 3- Transform all the values less than 0 in NA and then remove all your NA values from the data set. Show the dimensions of the cleaned data and print the first 6 rows.

```
data = read_dta('FF_wave6_2020v2.dta', col_select = c(p6b5, p6b10,p6b35,p6b55,p6b60,p6c21,p6f32,p6f35,p6
attach(data)
cols = c(1:23)
data[,cols] = apply(data[,cols], 2, function(x) as.numeric(x));
data[,cols][data[,cols]<0] <- NA
data <- na.omit(data)
print(dim(data))
```

```
## [1] 488  23
```

```
head(data)
```

```
## # A tibble: 6 x 23
##    p6b5 p6b10 p6b35 p6b55 p6b60 p6c21 p6f32 p6f35 p6h74 p6h102  p6i7  p6i8 p6i11
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  <dbl> <dbl> <dbl> <dbl>
## 1     1     2     2     1     1     2     2     1     1      1     2     2     3
## 2     2     2     1     1     1     2     2     2     2      2     2     2     4
## 3     1     2     1     2     1     2     2     2     2      2     2     2     4
## 4     2     2     1     2     1     2     2     2     1      2     1     1     4
## 5     2     2     1     1     1     2     2     2     1      2     2     2     4
## 6     2     2     1     1     1     2     2     2     2      2     2     3     4
## # ... with 10 more variables: p6j37 <dbl>, k6b21a <dbl>, k6b22a <dbl>,
## #   k6c1 <dbl>, k6c4e <dbl>, k6c28 <dbl>, k6d10 <dbl>, k6d37 <dbl>,
## #   k6f63 <dbl>, ck6cbmi <dbl>
```

(b) (20 points) Now call the variables with an appropriate name (for example *p6b5* can become *Depression*). Perform a logistic regression using the variable *Depression* as the outcome and the remaining variables as the covariates. Be careful: the variable *Depression* has value 1 and 2, you should transform in 0,1 before running the logistic regression in R (1 for Yes, 0 for No). What are the important and significant covariates for the depression? For these, what can you say about the standard error? Perform the binned residual plot by using the library ggplot2 in R.

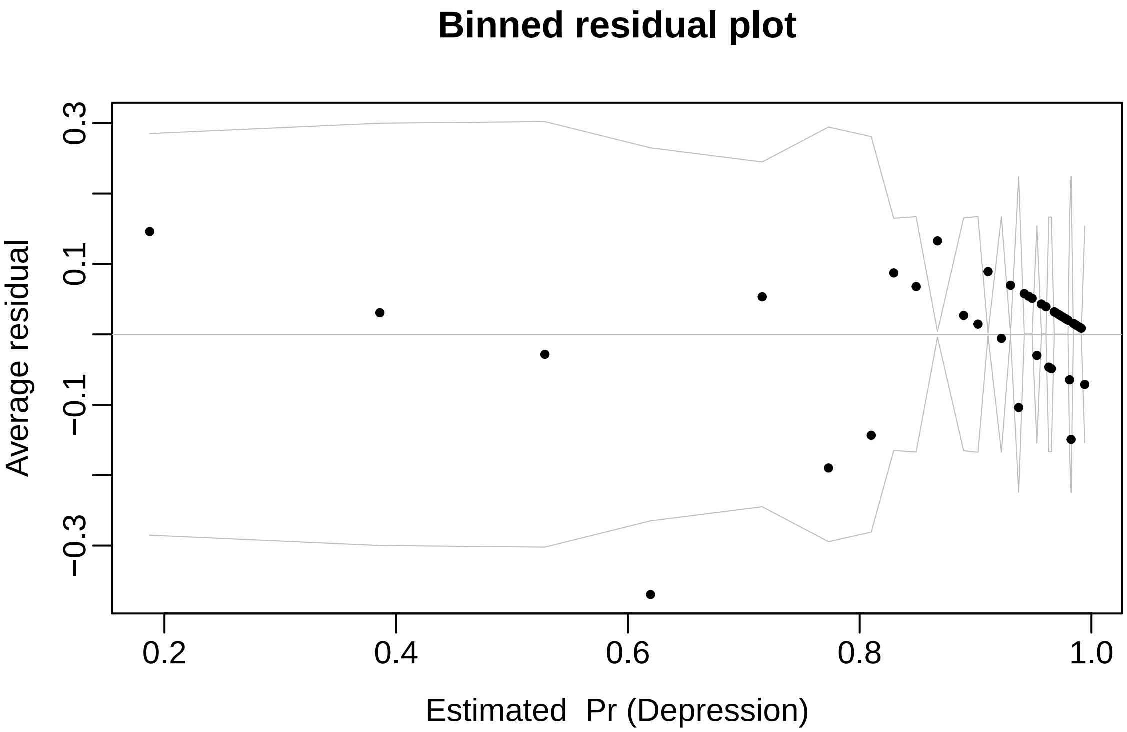
Answer b: ADD, trouble_sleeping, suspend, trouble_attention are significant. For these significant covariates, the magnitudes of standard errors are smaller than magnitudes of estimates, while others do not.

```
data = data %>% rename(Depression=p6b5, ADD=p6b10,
                cruel=p6b35,  trouble_sleeping=p6b55,
                run_away=p6b60, suspend=p6c21,
                drug=p6f32, parent_jail=p6f35, smoke=p6h74, jail=p6h102, helpful_neighborhood=p6i7, clos
                gangs_neighborhood=p6i11, receive_free_food=p6j37, trouch_attention=k6b21a, athletic=k6b
                atmosphere_calm=k6c4e, close_with_father=k6c28, age_menstruated=k6d10, physically_active
                BMI=ck6cbmi
                )
data[,1] <- data[,1] - 1
attach(data)
fit.2 <- glm(Depression ~ ., data=data[,-1],  family=binomial(link="logit"))
summary(fit.2)
```

```
##
## Call:
## glm(formula = Depression ~ ., family = binomial(link = "logit"),
##     data = data[, -1])
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.1486   0.1772   0.2570   0.3921   2.0580
##
## Coefficients:
##                                Estimate Std. Error z value Pr(>|z|)
## (Intercept)                    3.603900   3.291053   1.095 0.273490
## ADD                            1.677897   0.471019   3.562 0.000368 ***
## cruel                         -0.295209   0.337778  -0.874 0.382132
## trouble_sleeping              -1.453627   0.270194  -5.380 7.45e-08 ***
## run_away                      -0.678764   0.740884  -0.916 0.359586
## suspend                       -1.041740   0.458242  -2.273 0.023006 *
## drug                           0.891316   0.481588   1.851 0.064200 .
## parent_jail                   -0.587508   0.474254  -1.239 0.215418
## smoke                          0.424382   0.357364   1.188 0.235016
## jail                           0.362720   0.651642   0.557 0.577785
## helpful_neighborhood           0.164381   0.275061   0.598 0.550097
## close_knit_neighborhood       -0.066191   0.234927  -0.282 0.778133
## gangs_neighborhood            -0.008095   0.185011  -0.044 0.965101
## receive_free_food              0.863758   0.447409   1.931 0.053535 .
## trouch_attention              -0.731812   0.239586  -3.054 0.002254 **
## athletic                      -0.014254   0.111795  -0.128 0.898542
## biological_parent_relationship -0.015920   0.161038  -0.099 0.921249
## atmosphere_calm               -0.599335   0.309494  -1.936 0.052807 .
## close_with_father              0.070795   0.173388   0.408 0.683049
## age_menstruated               -0.010842   0.130825  -0.083 0.933952
## physically_active              0.100177   0.091585   1.094 0.274038
## marijuana                      0.360426   0.385711   0.934 0.350074
## BMI                           -0.039232   0.026002  -1.509 0.131354
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
```

```
##     Null deviance: 347.79  on 487  degrees of freedom
## Residual deviance: 245.05  on 465  degrees of freedom
## AIC: 291.05
##
## Number of Fisher Scoring iterations: 6
```

```r
binned.resids <- function (x, y, nclass=sqrt(length(x))){
  breaks.index <- floor(length(x)*(1:(nclass-1))/nclass)
  breaks <- c (-Inf, sort(x)[breaks.index], Inf)
  output <- NULL
  xbreaks <- NULL
  x.binned <- as.numeric (cut (x, breaks))
  for (i in 1:nclass){
    items <- (1:length(x))[x.binned==i]
    x.range <- range(x[items])
    xbar <- mean(x[items])
    ybar <- mean(y[items])
    n <- length(items)
    sdev <- sd(y[items])
    output <- rbind (output, c(xbar, ybar, n, x.range, 2*sdev/sqrt(n)))
  }
  colnames (output) <- c ("xbar", "ybar", "n", "x.lo", "x.hi", "2se")
  return (list (binned=output, xbreaks=xbreaks))
}
pred.2 <- fit.2$fitted.values
br.2 <- binned.resids (pred.2, Depression-pred.2, nclass=40)$binned
plot(range(br.2[,1]), range(br.2[,2],br.2[,6],-br.2[,6]), xlab="Estimated  Pr (Depression)", ylab="Avera
abline (0,0, col="gray", lwd=.5)
lines (br.2[,1], br.2[,6], col="gray", lwd=.5)
lines (br.2[,1], -br.2[,6], col="gray", lwd=.5)
points (br.2[,1], br.2[,2], pch=19, cex=.5)
```



(c) (20 points) Use the forward step procedure to detect the important covariates. Then, only for estimates that are greater than 0, draw with ggplot a plot similar to Figure 1. So in the x-axis, you should have each beta (beta1, beta2, etc.). In the y-axis, the estimate greater than 0 with the correspondent standard error. Be careful this plot is taken from another data set, so do not expect similar results. Take special care of the legend and the label. What can you say about this plot?

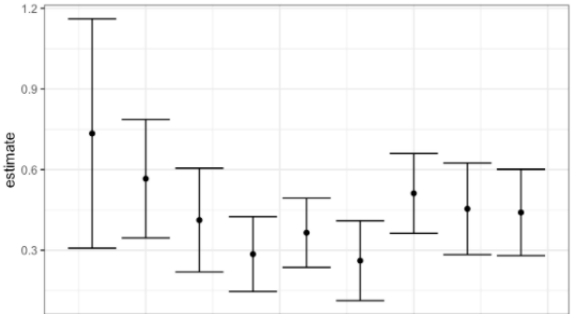


Figure 1: Estimate

Answer c: Except for the Intercept, covariates in this plot are close to zero with a small standard error.

```
fit_for1 <- glm(Depression~., data=data)
fit_for2 <- glm(Depression ~ 1, data=data)
print("FORWARD SELECTION")
```

```
## [1] "FORWARD SELECTION"
```

```r
model_forward <- stepAIC(fit_for2,direction="forward",scope=list(upper=fit_for1,lower=fit_for2))
```

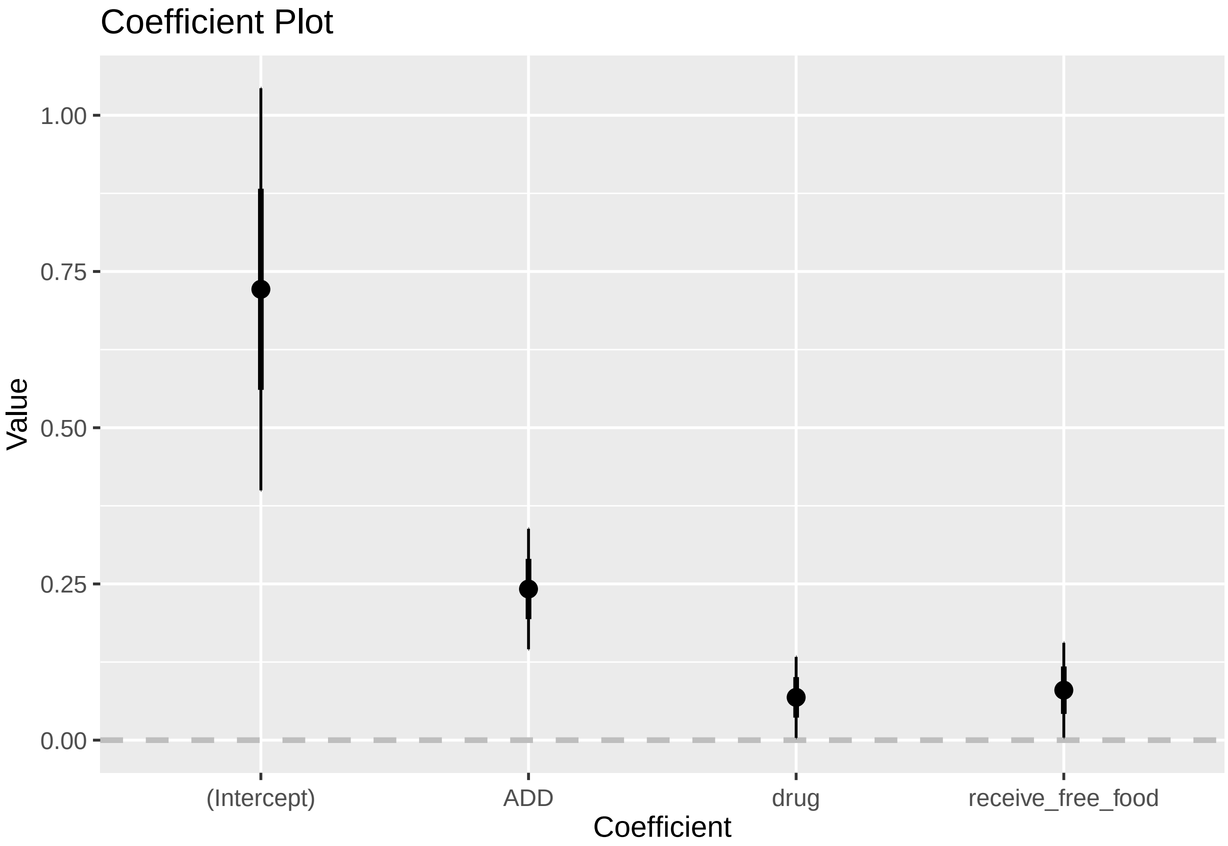
```
## Start:  AIC=272.9
## Depression ~ 1
##
##                                  Df Deviance    AIC
## + trouble_sleeping                1   41.969 193.64
## + ADD                             1   45.048 228.18
## + trouch_attention                1   47.550 254.56
## + receive_free_food               1   48.649 265.71
## + smoke                           1   48.672 265.94
## + drug                            1   48.747 266.69
## + athletic                        1   48.906 268.28
## + cruel                           1   49.007 269.29
## + run_away                        1   49.007 269.29
## + BMI                             1   49.248 271.68
## + marijuana                       1   49.282 272.02
## + jail                            1   49.282 272.02
## <none>                                49.574 272.90
## + physically_active               1   49.444 273.62
## + close_with_father               1   49.459 273.76
## + biological_parent_relationship  1   49.489 274.07
## + atmosphere_calm                 1   49.499 274.16
## + close_knit_neighborhood         1   49.535 274.52
## + helpful_neighborhood            1   49.539 274.56
## + age_menstruated                 1   49.542 274.58
## + parent_jail                     1   49.552 274.69
## + gangs_neighborhood              1   49.572 274.89
## + suspend                         1   49.573 274.89
##
## Step:  AIC=193.64
## Depression ~ trouble_sleeping
##
##                                  Df Deviance    AIC
## + ADD                             1   39.721 168.77
## + trouch_attention                1   40.510 178.37
## + drug                            1   41.524 190.43
## + receive_free_food               1   41.649 191.90
## + smoke                           1   41.721 192.74
## + run_away                        1   41.759 193.19
## + cruel                           1   41.766 193.26
## + athletic                        1   41.779 193.41
## + marijuana                       1   41.789 193.53
## <none>                                41.969 193.64
## + BMI                             1   41.802 193.69
## + physically_active               1   41.835 194.07
## + jail                            1   41.881 194.60
## + biological_parent_relationship  1   41.939 195.29
## + suspend                         1   41.944 195.35
## + atmosphere_calm                 1   41.949 195.40
## + age_menstruated                 1   41.953 195.44
## + parent_jail                     1   41.955 195.47
## + close_with_father               1   41.955 195.47
## + helpful_neighborhood            1   41.965 195.59
## + close_knit_neighborhood         1   41.968 195.62
## + gangs_neighborhood              1   41.969 195.63
##
## Step:  AIC=168.77
## Depression ~ trouble_sleeping + ADD
##
##                                 Df Deviance    AIC
## + trouch_attention               1   38.846 159.90
## + drug                           1   39.296 165.52
## + receive_free_food              1   39.344 166.12
## + smoke                          1   39.506 168.12
## <none>                               39.721 168.77
## + suspend                        1   39.592 169.18
## + run_away                       1   39.603 169.32
## + atmosphere_calm                1   39.617 169.49
## + BMI                            1   39.628 169.62
## + athletic                       1   39.638 169.75
## + physically_active              1   39.639 169.76
## + jail                           1   39.648 169.87
## + marijuana                      1   39.683 170.30
## + cruel                          1   39.695 170.44
## + helpful_neighborhood           1   39.704 170.56
## + biological_parent_relationship  1   39.711 170.65
## + gangs_neighborhood             1   39.714 170.68
## + close_with_father              1   39.715 170.69
## + age_menstruated                1   39.717 170.72
## + parent_jail                    1   39.718 170.72
## + close_knit_neighborhood        1   39.721 170.77
##
## Step:  AIC=159.9
## Depression ~ trouble_sleeping + ADD + trouch_attention
##
##                                 Df Deviance    AIC
## + drug                           1   38.471 157.17
## + receive_free_food              1   38.519 157.78
## + smoke                          1   38.664 159.61
## + atmosphere_calm                1   38.675 159.75
## + suspend                        1   38.676 159.76
## <none>                               38.846 159.90
## + BMI                            1   38.745 160.63
## + physically_active              1   38.782 161.09
## + run_away                       1   38.783 161.10
## + athletic                       1   38.786 161.14
## + jail                           1   38.800 161.32
## + marijuana                      1   38.823 161.61
## + helpful_neighborhood           1   38.834 161.75
## + cruel                          1   38.844 161.87
## + parent_jail                    1   38.845 161.88
## + biological_parent_relationship  1   38.845 161.88
## + gangs_neighborhood             1   38.845 161.89
## + close_with_father              1   38.846 161.89
## + close_knit_neighborhood        1   38.846 161.90
## + age_menstruated                1   38.846 161.90
##
```

```
## Step:  AIC=157.17
## Depression ~ trouble_sleeping + ADD + trouch_attention + drug
##
##                                  Df Deviance    AIC
## + receive_free_food               1   38.219 155.96
## + suspend                         1   38.255 156.41
## + atmosphere_calm                 1   38.287 156.82
## <none>                                38.471 157.17
## + parent_jail                     1   38.325 157.31
## + smoke                           1   38.330 157.37
## + BMI                             1   38.370 157.88
## + athletic                        1   38.394 158.18
## + physically_active               1   38.408 158.36
## + run_away                        1   38.417 158.48
## + jail                            1   38.433 158.68
## + biological_parent_relationship  1   38.445 158.84
## + close_with_father               1   38.446 158.84
## + marijuana                       1   38.455 158.96
## + helpful_neighborhood            1   38.460 159.03
## + cruel                           1   38.469 159.14
## + close_knit_neighborhood         1   38.471 159.17
## + age_menstruated                 1   38.471 159.17
## + gangs_neighborhood              1   38.471 159.17
##
## Step:  AIC=155.96
## Depression ~ trouble_sleeping + ADD + trouch_attention + drug +
##     receive_free_food
##
##                                  Df Deviance    AIC
## + suspend                         1   37.944 154.44
## + atmosphere_calm                 1   38.001 155.17
## + parent_jail                     1   38.058 155.89
## <none>                                38.219 155.96
## + smoke                           1   38.119 156.68
## + athletic                        1   38.128 156.80
## + BMI                             1   38.132 156.84
## + physically_active               1   38.150 157.08
## + run_away                        1   38.165 157.27
## + jail                            1   38.183 157.49
## + biological_parent_relationship  1   38.186 157.54
## + helpful_neighborhood            1   38.194 157.64
## + close_with_father               1   38.203 157.75
## + marijuana                       1   38.206 157.79
## + gangs_neighborhood              1   38.213 157.88
## + cruel                           1   38.216 157.92
## + close_knit_neighborhood         1   38.216 157.92
## + age_menstruated                 1   38.219 157.96
##
## Step:  AIC=154.43
## Depression ~ trouble_sleeping + ADD + trouch_attention + drug +
##     receive_free_food + suspend
##
##                                  Df Deviance    AIC
## + atmosphere_calm                 1   37.713 153.45
## <none>                               37.944 154.44
## + athletic                        1   37.815 154.78
## + parent_jail                     1   37.824 154.89
## + BMI                             1   37.842 155.12
## + physically_active               1   37.845 155.16
## + smoke                           1   37.853 155.26
## + run_away                        1   37.853 155.27
## + marijuana                       1   37.902 155.89
## + jail                            1   37.904 155.91
## + cruel                           1   37.914 156.04
## + biological_parent_relationship  1   37.914 156.05
## + helpful_neighborhood            1   37.927 156.22
## + close_with_father               1   37.928 156.22
## + gangs_neighborhood              1   37.943 156.42
## + close_knit_neighborhood         1   37.943 156.42
## + age_menstruated                 1   37.944 156.43
##
## Step:  AIC=153.45
## Depression ~ trouble_sleeping + ADD + trouch_attention + drug +
##     receive_free_food + suspend + atmosphere_calm
##
##                                  Df Deviance    AIC
## <none>                               37.713 153.45
## + parent_jail                     1   37.586 153.80
## + BMI                             1   37.594 153.91
## + physically_active               1   37.599 153.97
## + athletic                        1   37.602 154.01
## + smoke                           1   37.610 154.11
## + run_away                        1   37.612 154.15
## + marijuana                       1   37.654 154.68
## + cruel                           1   37.668 154.87
## + jail                            1   37.672 154.91
## + biological_parent_relationship  1   37.693 155.19
## + helpful_neighborhood            1   37.695 155.22
## + close_with_father               1   37.704 155.34
## + close_knit_neighborhood         1   37.710 155.42
## + gangs_neighborhood              1   37.711 155.43
## + age_menstruated                 1   37.713 155.45
```

```
out <- summary(model_forward)
out
```

```
##
## Call:
## glm(formula = Depression ~ trouble_sleeping + ADD + trouch_attention +
##     drug + receive_free_food + suspend + atmosphere_calm, data = data)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.05820  -0.00910   0.04796   0.11645   0.73893
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.72152    0.16090   4.484 9.16e-06 ***
## trouble_sleeping  -0.18706    0.02424  -7.716 6.99e-14 ***
## ADD                 0.24188    0.04823   5.015 7.46e-07 ***
## trouch_attention  -0.06940    0.02043  -3.397 0.000739 ***
## drug                0.06850    0.03239   2.114 0.034996 *
## receive_free_food  0.08000    0.03793   2.109 0.035459 *
## suspend           -0.05705    0.02978  -1.916 0.056002 .
## atmosphere_calm   -0.03676    0.02143  -1.715 0.086992 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.07856892)
##
##     Null deviance: 49.574  on 487  degrees of freedom
## Residual deviance: 37.713  on 480  degrees of freedom
## AIC: 153.45
##
## Number of Fisher Scoring iterations: 2
```

```
ss <- coef(out)
coefplot(model_forward, horizontal=TRUE, coefficients=c("(Intercept)","ADD", "drug", "receive_free_food
         color='black', fillcolor='grey')
```



(d) (20 points) Perform a bootstrap of 1000 samples for beta 1 (ADD or *p6b10*), beta 2 (sleep or *p6b55*), and beta 3 (attention at school or *k6b21a*) with a model that contains all the coefficients obtained in the forward procedure in point c. Plot these three bootstrapped beta coefficients that you have obtained with a boxplot in the ggplot (similar to Figure 2). (make sure not to use the default colors but rather choose your own). What can you say about these three distributions obtained?

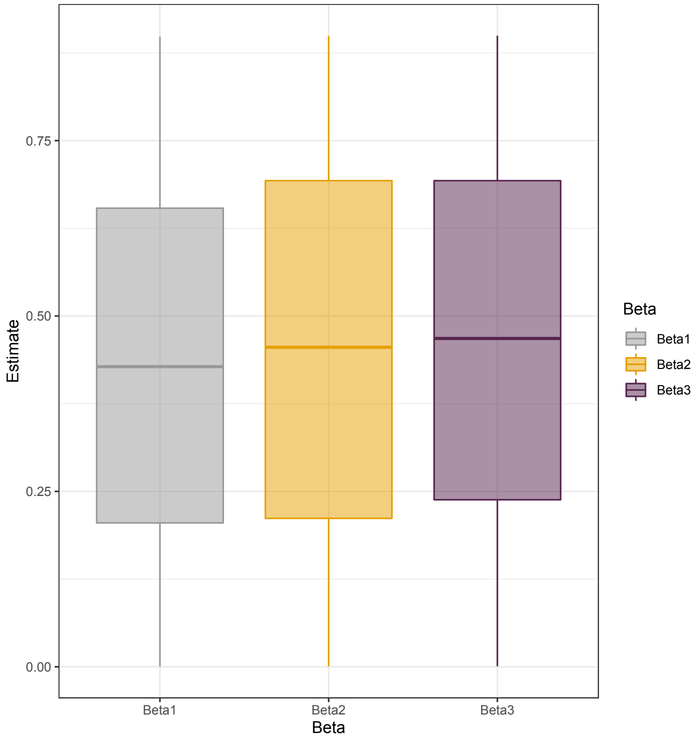


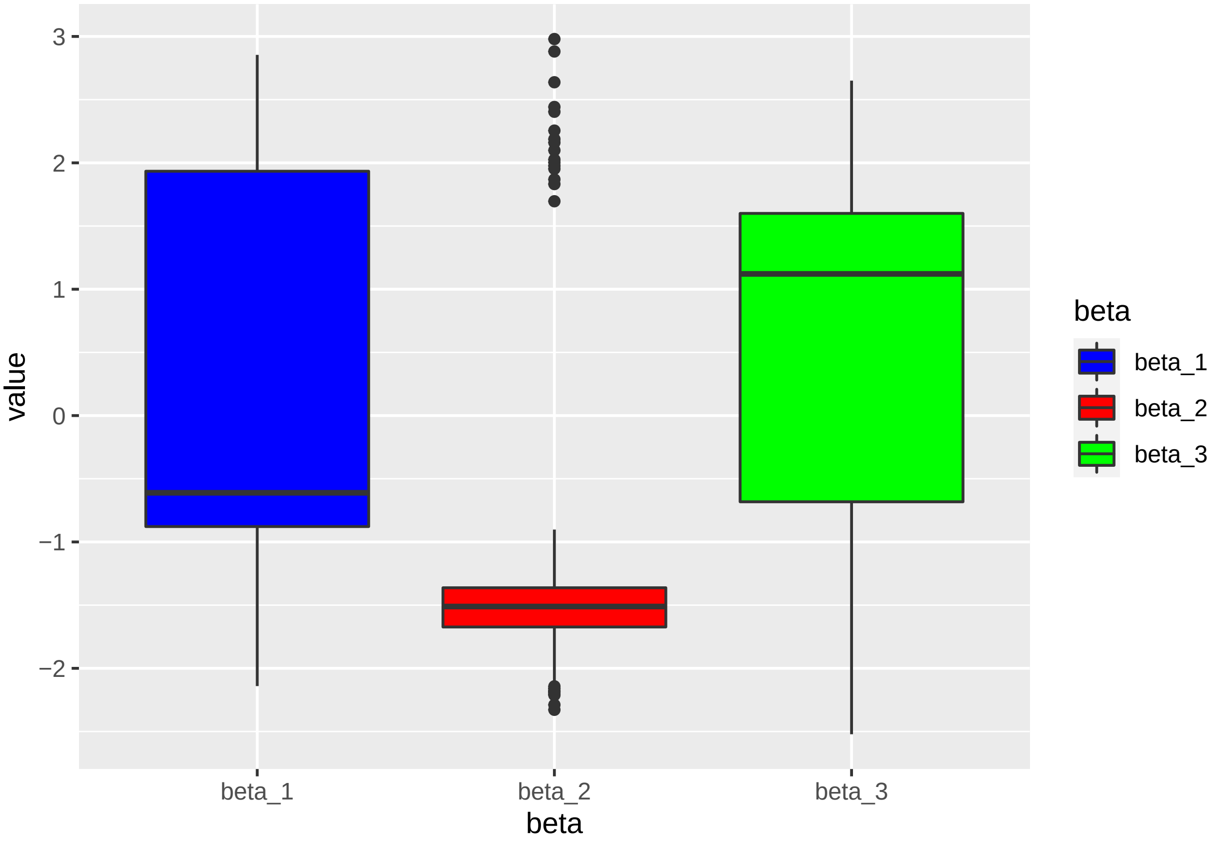
Figure 2: Boxplot

Answer d: The beta 1 and beta 3 have fewer outliers and the standard deviation between Q1 and Q3 are large. The beta2 has more outliers but the standard deviation between Q1 and Q2 is small.

```
n <- 1000
coef_boot <- matrix(NA, n, 3)
for (i in 1:n){
  s_boot <- sample(c(1:dim(data)[1]), n, replace=TRUE)
  data_boot <- data[s_boot,]
  fit4.1 <- glm(Depression ~ ., data=data_boot, family=binomial(link="logit"))
  fit4.2 <- glm(Depression ~ 1, data=data_boot, family=binomial(link="logit"))
  mod_d = stepAIC(fit4.2, direction="forward",scope=list(upper=fit4.1,lower=fit4.2), trace=FALSE)
  coef_boot[i,] <- mod_d$coefficients[2:4]
}
mod_d$coefficients
```

```
##               (Intercept)        trouble_sleeping        trouch_attention
##                2.83453758             -1.37712328             -0.98649911
##                       ADD                   smoke        atmosphere_calm
##                1.79655504              0.91893886             -0.88280246
##                      drug       close_with_father              marijuana
##                0.98778847              0.33047940              0.71104636
## close_knit_neighborhood       physically_active                suspend
##                0.27359606              0.09746124             -0.79258054
##                     cruel                     BMI      receive_free_food
##               -0.45566853             -0.03603124              0.59220306
##          age_menstruated
##               -0.13849686
```

```
df4 <- data.frame(value = c(coef_boot[,2], coef_boot[,1], coef_boot[,3]),
                        beta = rep(c("beta_1","beta_2","beta_3"), each = n))
p<-ggplot(df4, aes(x=beta, y=value, fill=beta)) + geom_boxplot() + scale_fill_manual(values=c("blue", "
p
```



(e) (20 points) Perform the Lasso method for the full model. Choose λ with the cross-validation. Then perform the lasso with the best λ obtained. Plot the results in ggplot. Describe the results you obtained. Are the coefficients obtained with the lasso procedure similar to the coefficients obtained with the forward procedure? Explain!

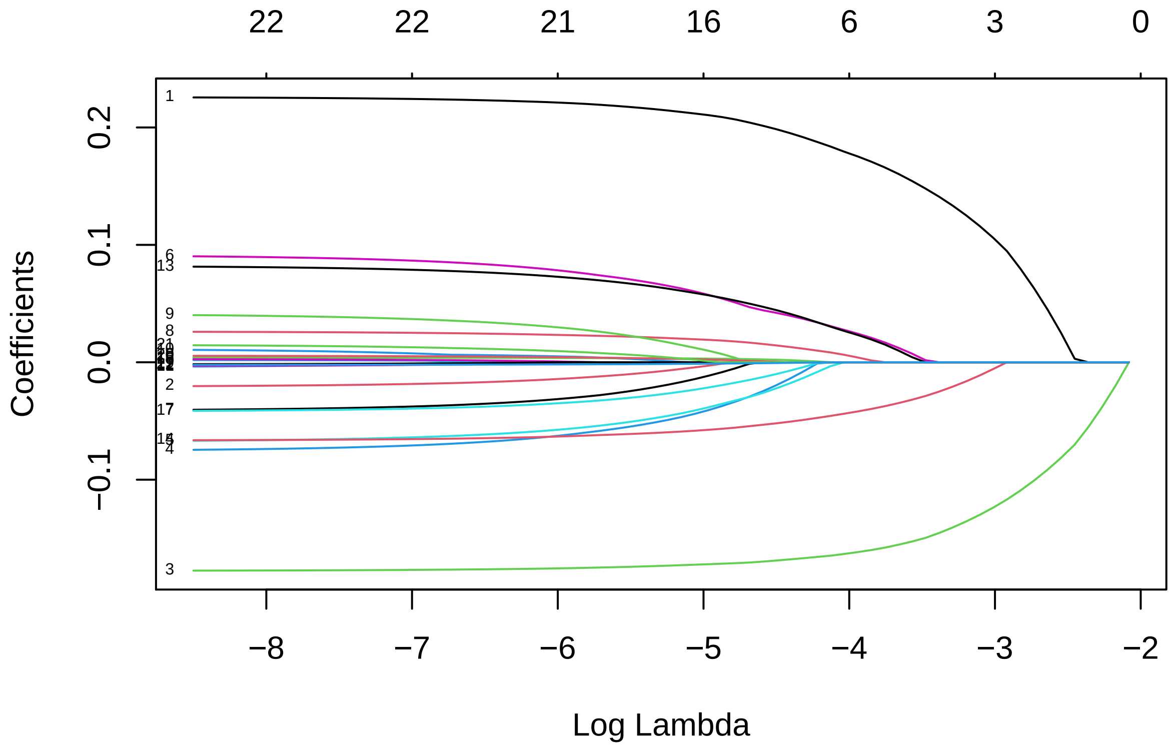
Asnwer e: No, lasso reduces the coefficients of many covariates to zero while forward procedure in d does not. For non zero covariates, the magnitude of the lasso coefficients are much smaller than those of the forward procedure in d.

```
x = model.matrix(Depression~., data)[,-1]
y = data %>%
dplyr::select(Depression) %>%
unlist() %>%
as.numeric()
train <- data %>% sample_frac(0.5)
test = data %>% setdiff(train)
x_train = model.matrix(Depression~., train)[,-1]
x_test = model.matrix(Depression~., test)[,-1]
y_train = train %>% dplyr::select(Depression) %>% unlist() %>% as.numeric()
y_test = test %>% dplyr::select(Depression) %>% unlist() %>% as.numeric()
set.seed(1)
cv.out = cv.glmnet(x_train, y_train, alpha = 1)
```

```
bestlam = cv.out$lambda.min
grid = 10^seq(10, -2, length = 100)
out = glmnet(x, y, alpha = 1, lambda = grid)
lasso_coef=predict(out, type="coefficients",s=bestlam)[1:20,]
print(lasso_coef)
```

```
##                    (Intercept)                            ADD
##                    0.730068226                    0.179710301
##                          cruel               trouble_sleeping
##                    0.000000000                   -0.163403978
##                       run_away                        suspend
##                    0.000000000                    0.000000000
##                           drug                    parent_jail
##                    0.027846548                    0.000000000
##                          smoke                           jail
##                    0.006567634                    0.000000000
##           helpful_neighborhood        close_knit_neighborhood
##                    0.000000000                    0.000000000
##             gangs_neighborhood              receive_free_food
##                    0.000000000                    0.026974792
##               trouch_attention                       athletic
##                   -0.043964518                    0.000000000
## biological_parent_relationship                atmosphere_calm
##                    0.000000000                    0.000000000
##              close_with_father                age_menstruated
##                    0.000000000                    0.000000000
```

```
lasso_mod <-  glmnet(x, y, alpha = 1)
lasso.mod =glmnet(x,y, alpha =1)#this will give 80 values of lambda
beta=coef(lasso.mod)
plot(lasso.mod, "lambda", label = TRUE)
```


22
22
21
16
6
3
0
Coefficients
0.2
0.1
0.0
−0.1
1
6
13
9
8
21
10
2
17
15
5
4
3
−8
−7
−6
−5
−4
−3
−2
Log Lambda