

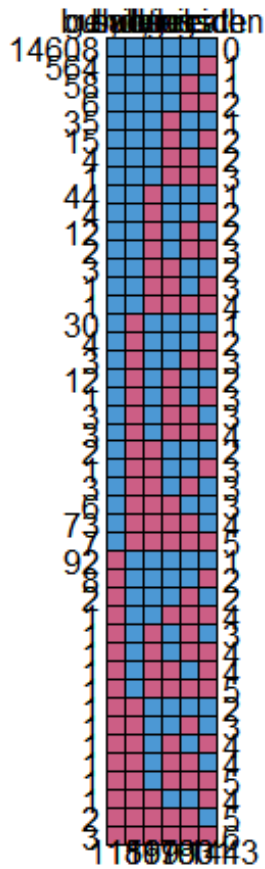
R codes

```
#Missing data patterns and MCAR test
```

```
>mice::md.pattern(ideation)
```

```
gender  bullying  elc sadness  bullyings  chc  suicide threaten
14608   1         1     1         1     1     1         0
564     1         1     1         1     1     0         1
58      1         1     1         1     0     1         1
6       1         1     1         1     0     0         2
35      1         1     1         0     1     1         1
15      1         1     1         0     1     0         2
4       1         1     1         0     0     1         2
1       1         1     1         0     0     0         3
44      1         1     0         1     1     1         1
4       1         1     0         1     1     0         2
12      1         1     0         1     0     1         2
2       1         1     0         1     0     0         3
3       1         1     0         0     1     1         2
1       1         1     0         0     1     0         3
1       1         1     0         0     0     0         4
30      1         0     1         1     1     1         1
4       1         0     1         1     1     0         2
3       1         0     1         1     0     0         3
12      1         0     1         0     1     1         2
1       1         0     1         0     1     0         3
3       1         0     1         0     0     1         3
3       1         0     1         0     0     0         4
2       1         0     0         1     1     1         2
1       1         0     0         1     1     0         3
3       1         0     0         1     0     1         3
6       1         0     0         0     1     1         3
73      1         0     0         0     0     1         4
7       1         0     0         0     0     0         5
92      0         1     1         1     1     1         1
8       0         1     1         1     1     0         2
2       0         1     1         1     0     1         2
1       0         1     1         0     0     0         4
1       0         1     0         1     0     1         3
1       0         1     0         1     0     0         4
1       0         1     0         0     1     0         4
1       0         1     0         0     0     0         5
1       0         0     1         1     1     1         2
1       0         0     1         1     0     1         3
1       0         0     1         0     1     0         4
1       0         0     1         0     0     1         4
1       0         0     1         0     0     0         5
1       0         0     0         1     1     0         4
2       0         0     0         0     0     1         5
3       0         0     0         0     0     0         6
      118      159  169      176  190 631 1443
```

```
>
```



```
library(MissMech)
> #Create a subset of the data
> ideation <- subset(yrbs,
+   select=c("csuicide",
+     "bullyingsch",
+     "bullyingelc",
+     "sadness",
+     "threaten",
+     "gender"))
> MissMech::TestMCARNormality(ideation)
```

Warning: 3 Cases with all variables missing have been removed

from the data.

```
Call:
MissMech::TestMCARNormality(data = ideation)
```

Number of Patterns: 13

Total number of cases used in the analysis: 15558

Pattern(s) used:

Csuicide	bullyingsch	bullyingelc	sadness	threaten	gender	Number of cases
group.1	1	1	1	1	1	14608

group.2	1	1	1	1	NA	1	564
group.3	NA	NANANA		1	1		73
group.4	1	1	NA	1	1	1	30
group.5	1	NA	1	1	1	1	35
group.6	1	1	1	NA	1	1	44
group.7	1	1	1	1	1	NA	92
group.8	NA	1	1	1	1	1	58
group.9	1	NA	NA	1	1	1	12
group.10	NA	NANANANA		1		7	
group.11	NA	1	1	NA	1	1	12
group.12	1	1	1	1	NA	NA	8
group.13	1	NA	1	1	NA	1	15

Test of normality and Homoscedasticity:

Hawkins Test:

P-value for the Hawkins test of normality and homoscedasticity: 2.507311e-258

Either the test of multivariate normality or homoscedasticity (or both) is rejected.

Provided that normality can be assumed, the hypothesis of MCAR is rejected at 0.05 significance level.

Non-Parametric Test:

P-value for the non-parametric test of homoscedasticity: 0

Hypothesis of MCAR is rejected at 0.05 significance level.

The multivariate normality test is inconclusive.

>

#Creating a frequency distribution table

```
>genderrtable<-table(yrbsregdata1c$genderr)
```

```
>genderrtable
```

```
Male Female
```

```
7026 7126
```

```
>#Creating a frequency barplot
```

```
>barplot(genderrtable,
```

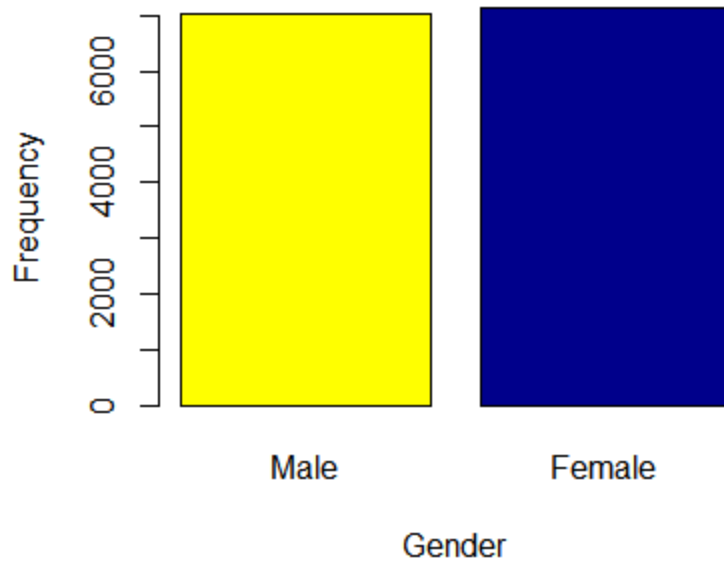
```
+ main="Gender Distribution",
```

```
+ ylab="Frequency",
```

```
+ xlab="Gender",
```

```
+ col=c("yellow","darkblue"))
```

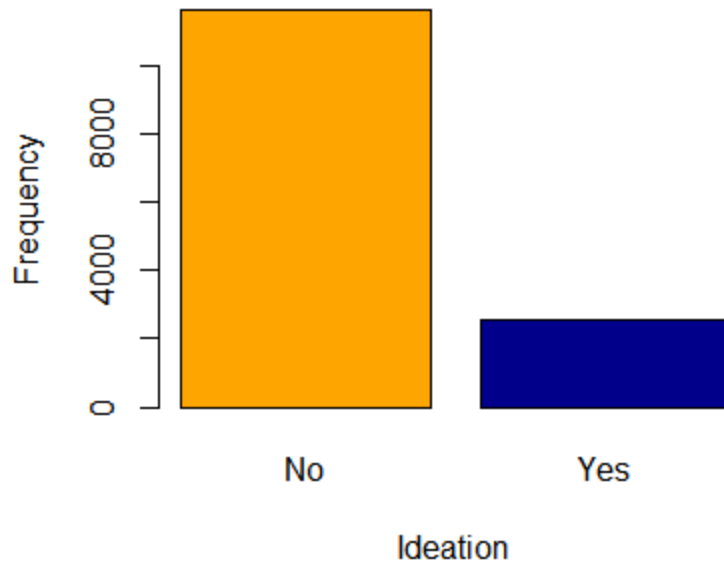
Gender Distribution



```
#Creating a frequency distribution table  
>csuicidertable<-table(yrbsregdata1c$csuicider)  
>csuicidertable
```

```
  No  Yes  
11608 2544  
> #Creating a frequency barplot  
>barplot(genderrtable,  
+   main="Suicadal ideation Distribution",  
+   ylab="Frequency",  
+   xlab="Ideation",  
+   col=c("orange","darkblue"))  
>
```

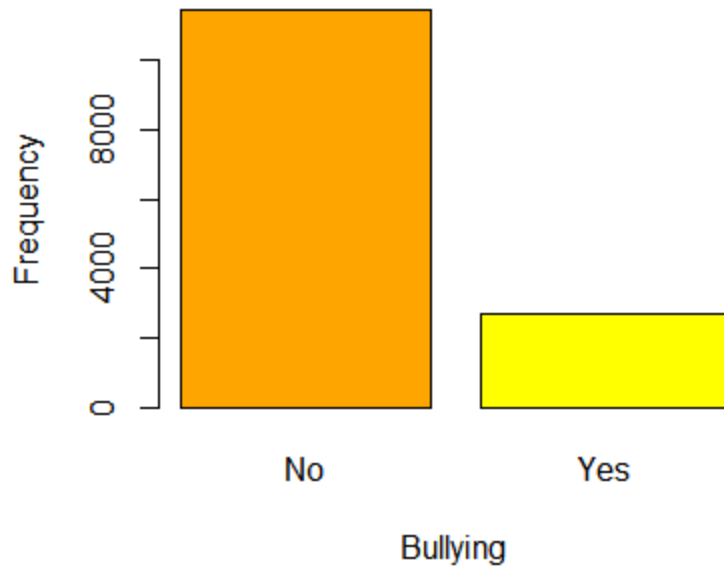
Suicidal ideation Distribution



```
#Creating a frequency distribution table  
>bullyingrtable<-table(yrbsregdata1c$bullyingschr)  
>bullyingrtable
```

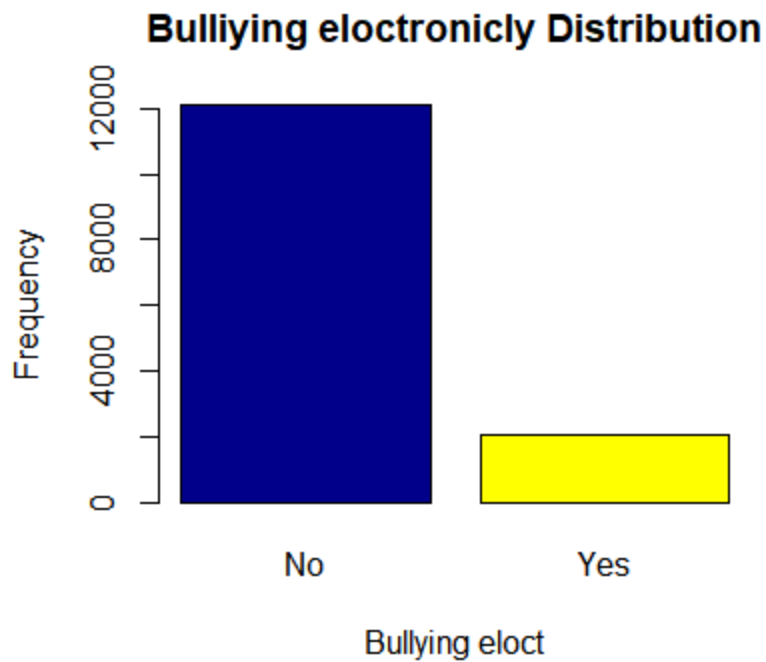
```
  No  Yes  
11438 2714  
> #Creating a frequency barplot  
>barplot(bullyingrtable,  
+   main="Bullying at school Distribution",  
+   ylab="Frequency",  
+   xlab="Bullying",  
+   col=c("orange", "yellow"))  
  
>
```

Bullying at school Distribution



```
#Creating a frequency distribution table  
>bullyingelcrtable<-table(yrbsregdata1c$bullyingelcr)  
>bullyingelcrtable
```

```
  No  Yes  
12095 2057  
> #Creating a frequency barplot  
>barplot(bullyingelcrtable,  
+   main="Bullyingelctronicly Distribution",  
+   ylab="Frequency",  
+   xlab="Bullying eloct",  
+   col=c("darkblue","yellow"))  
>
```



```
#Creating a frequency distribution table
>hopelessnessrtable<-table(yrbsregdata1c$sadnessr)
>hopelessnessrtable
```

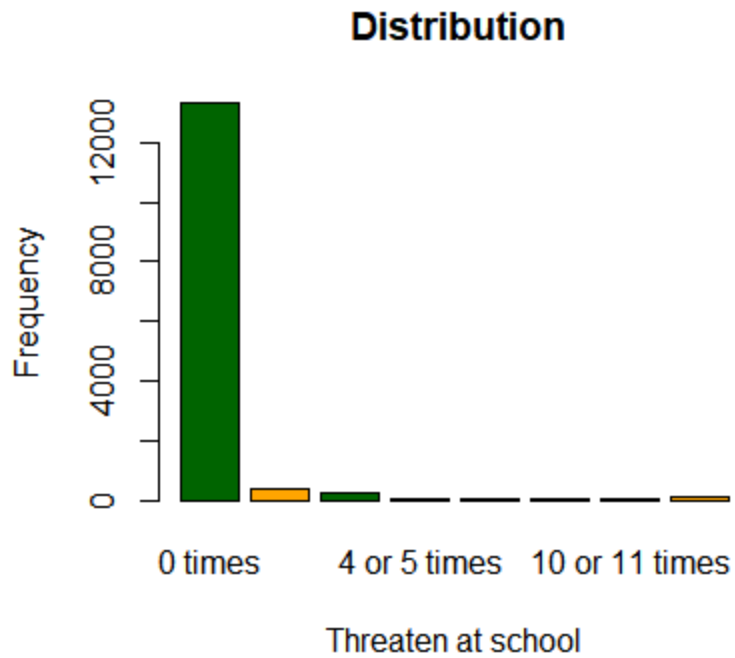
```
No Yes
9815 4337
> #Creating a frequency barplot
>barplot(hopelessnessrtable,
+   main=" Distribution",
+   ylab="Frequency",
+   xlab="Hopelessness",
+   col=c("darkblue","orange"))
>
```



```
#Creating a frequency barplot
>barplot(hopelessnessrtable,
+   main=" Distribution",
+   ylab="Frequency",
+   xlab="Hopelessness",
+   col=c("darkblue","orange"))
>#Creating a frequency distribution table
>threatentable<-table(yrbsregdata1c$threaten)
>threatentable
```

0 times	1 time	2 or 3 times	4 or 5 times	6 or 7 times
13311	391	214	61	36
8 or 9 times	10 or 11 times	12 or more times		
28	15	96		

```
>#Creating a frequency barplot
>barplot(threatentable,
+   main=" Distribution",
+   ylab="Frequency",
+   xlab="Threaten at school",
+   col=c("darkgreen","orange"))
>
```

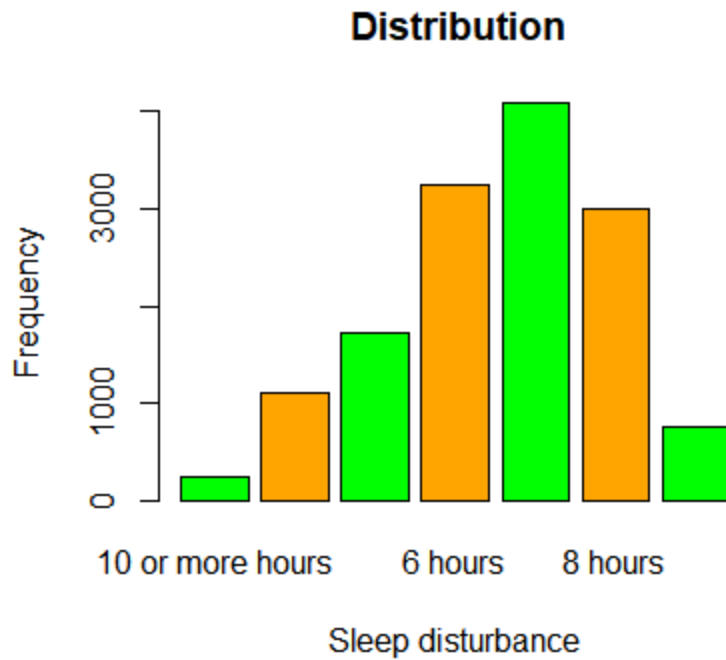



```
#Creating a frequency distribution table
>sleeprtable<-table(yrbsregdata1c$sleepr)
>sleeprtable
```

10 or more hours	4 or less hours	5 hours	6 hours	7 hours
247	1110	1726	3244	4087
8 hours	9 hours			
2994	744			

```
> #Creating a frequency barplot
>barplot(sleeprtable,
+ main=" Distribution",
+ ylab="Frequency",
+ xlab="Sleep disturbance",
+ col=c("green","orange"))
```

```
>
```



#Assessing multi co linearity among predictors

```
>car::vif(logreg2)
```

```
GVIF Df GVIF^(1/(2*Df))
```

```
genderr1.066264 1 1.032600
bullyingschr1.292564 1 1.136910
bullyingelcr1.297221 1 1.138956
sadnessr1.052699 1 1.026011
threaten 1.102205 7 1.006975
sleepr1.045451 6 1.003711
```

```
>
```

```
yrbs$csuicide<-factor(yrbs$Q27,
```

```
+
```

```
level=c(1,2),
```

```
+
```

```
labels=c("Yes", "No"))
```

```
>yrbs$gender<- factor(yrbs$Q2,
```

```
+
```

```
levels=c(1,2),
```

```
+
```

```
labels=c("Female",
```

```
+
```

```
"Male"))
```

```
>yrbs$bullyingsch<-factor(yrbs$Q24,
```

```
+
```

```
level=c(1,2),
```

```
+
```

```
labels=c("Yes", "No"))
```

```
>yrbs$bullyingelc<-factor(yrbs$Q25,
```

```
+
```

```

+           level=c(1,2),
+
+           labels=c("Yes", "No"))
>yrbs$sadness<-factor(yrbs$Q26,
+
+           level=c(1,2),
+
+           labels=c("Yes", "No"))
>yrbs$threaten<- factor(yrbs$Q17,
+           levels=c(1:8),
+           labels=c("0 times",
+                   "1 time",
+                   "2 or 3 times",
+                   "4 or 5 times",
+                   "6 or 7 times",
+                   "8 or 9 times",
+                   "10 or 11 times",
+                   "12 or more times"))
>yrbs$severalcahol<- factor(yrbs$Q41,
+           levels=c(1:7),
+           labels=c("0 days",
+                   "1 or 2 days",
+                   "3 to 9 days",
+                   "10 to 19 days",
+                   "20 to 39 days",
+                   "40 to 99 days",
+                   "100 or more days"))
>yrbs$sleepdisturbances<- factor(yrbs$Q88,
+           levels=c(1:7),
+           labels=c("4 or less hours",
+                   "5 hours",
+                   "6 hours",
+                   "7 hours",
+                   "8 hours",
+                   "9 hours",
+                   "10 or more hours"))
>yrbsregdata1<-dplyr::select(yrbs,
+           gender,
+           csuicide,
+           bullyingsch,
+           bullyingelc,
+           sadness,
+
+           threaten,
+           sleepdisturbances)
>yrbsregdata1c<-na.omit(yrbsregdata1)
>#Descriptive statistics for regression variables
>psych::describe(yrbsregdata1c)
  vars  n mean  sd median trimmed  mad min max range  skew
gender*   1 14152 1.50 0.50   1  1.50 0.00  1  2  1 0.01
csuicide*  2 14152 1.82 0.38   2  1.90 0.00  1  2  1 -1.67
bullying*  3 14152 1.81 0.39   2  1.89 0.00  1  2  1 -1.57
bullyingelc*  4 14152 1.85 0.35   2  1.94 0.00  1  2  1 -2.01
sadness*   5 14152 1.69 0.46   2  1.74 0.00  1  2  1 -0.84
threaten*  6 14152 1.14 0.75   1  1.00 0.00  1  8  7  6.99
sleepdisturbances*  7 14152 3.66 1.37   4  3.69 1.48  1  7  6 -0.10

```

	kurtosis	se
gender*	-2.00	0.00
csuicide*	0.78	0.00
bullyingsch*	0.45	0.00
bullyingelc*	2.05	0.00
sadness*	-1.30	0.00
threaten*	54.20	0.01
sleepdisturbances*	-0.38	0.01

Hmisc::describe(yrbsregdata1c)
yrbsregdata1c

7 Variables 14152 Observations

gender
n missing distinct
14152 0 2

Value	Female	Male
Frequency	7126	7026
Proportion	0.504	0.496

C suicide
n missing distinct
14152 0 2

Value	Yes	No
Frequency	2544	11608
Proportion	0.18	0.82

Bullyings ch
n missing distinct
14152 0 2

Value	Yes	No
Frequency	2714	11438
Proportion	0.192	0.808

Bullying elc
n missing distinct
14152 0 2

Value	Yes	No
Frequency	2057	12095
Proportion	0.145	0.855

sadness
n missing distinct
14152 0 2

Value	Yes	No
Frequency	4337	9815
Proportion	0.306	0.694

Sleep disturbances

```
n missing distinct
14152 0 7
```

```
lowest : 4 or less hours 5 hours 6 hours 7 hours 8 hours
highest: 6 hours 7 hours 8 hours 9 hours 10 or more hours
```

Value	4 or less hours	5 hours	6 hours	7 hours	8 hours
Frequency	1110	1726	3244	4087	2994
Proportion	0.078	0.122	0.229	0.289	0.212

Value	9 hours	10 or more hours
Frequency	744	247
Proportion	0.053	0.017

```
threaten
```

```
n missing distinct
14152 0 8
```

```
lowest : 0 times 1 time 2 or 3 times 4 or 5 times 6 or 7 times
highest: 4 or 5 times 6 or 7 times 8 or 9 times 10 or 11 times 12 or more times
```

Value	0 times	1 time	2 or 3 times	4 or 5 times	6 or 7 times
Frequency	13311	391	214	61	36
Proportion	0.941	0.028	0.015	0.004	0.003

Value	8 or 9 times	10 or 11 times	12 or more times
Frequency	28	15	96
Proportion	0.002	0.001	0.007

```
>
```

```
#Relevel the categorical predi
```

```
>yrebsregdata1c$csuicider<-relevel(yrebsregdata1c$csuicide, "No")
```

```
>yrebsregdata1c$genderr<-relevel(yrebsregdata1c$gender, "Male")
```

```
>yrebsregdata1c$bullyingschr<-relevel(yrebsregdata1c$bullyingsch, "No")
```

```
>yrebsregdata1c$bullyingelcr<-relevel(yrebsregdata1c$bullyingelc, "No")
```

```
>yrebsregdata1c$sadnessr<-relevel(yrebsregdata1c$sadness, "No")
```

```
yrebsregdata1c$sleepr<-relevel(yrebsregdata1c$sleepdisturbances, "10 or more hours")
```

```
#Fitting the null logistic regression model
```

```
>logreg0<-glm(data=yrebsregdata1c,
```

```
+ csuicider ~ 1,
```

```
+ family=binomial)
```

```
>summary(logreg0)
```

```
Call:
```

```
glm(formula = csuicider ~ 1, family = binomial, data = yrebsregdata1c)
```

```
Deviance Residuals:
```

```
Min 1Q Median 3Q Max
-0.6295 -0.6295 -0.6295 -0.6295 1.8526
```

Coefficients:

```
Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.51796 0.02189 -69.34 <2e-16 ***
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

```
Null deviance: 13332 on 14151 degrees of freedom
Residual deviance: 13332 on 14151 degrees of freedom
AIC: 13334
```

Number of Fisher Scoring iterations: 4

#Fitting the model with predictors

```
>logreg2<-glm(data=yrebsregdata1c,
+ csuicider ~ genderr+bullyingschr+bullyingelcr+ sadnessr+
+ threaten + sleepr,
+ family=binomial)
>summary(logreg2)
```

Call:

```
glm(formula = csuicider ~ genderr + bullyingschr + bullyingelcr +
sadnessr + threaten + sleepr, family = binomial, data = yrebsregdata1c)
```

Deviance Residuals:

```
Min 1Q Median 3Q Max
-2.2251 -0.3998 -0.2960 -0.2481 2.6444
```

Coefficients:

```
Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.1934550 0.2091875 -15.266 <2e-16 ***
genderrFemale0.3597897 0.0555610 6.476 9.44e-11 ***
bullyingschrYes0.6561910 0.0646722 10.146 <2e-16 ***
bullyingelcrYes0.5906671 0.0693367 8.519 <2e-16 ***
sadnessrYes2.3188035 0.0565138 41.031 <2e-16 ***
threaten1 time 0.4054919 0.1343609 3.018 0.00255 **
threaten2 or 3 times 0.3610385 0.1777684 2.031 0.04226 *
threaten4 or 5 times 0.8879535 0.3326075 2.670 0.00759 **
threaten6 or 7 times 0.4516117 0.4164738 1.084 0.27820
threaten8 or 9 times 1.9647434 0.4825244 4.072 4.67e-05 ***
threaten10 or 11 times 1.9845686 0.6555075 3.028 0.00247 **
threaten12 or more times 1.1250742 0.2710347 4.151 3.31e-05 ***
sleepr4 or less hours 0.7068623 0.2170534 3.257 0.00113 **
sleepr5 hours 0.1398138 0.2146861 0.651 0.51489
sleepr6 hours 0.0003281 0.2107866 0.002 0.99876
sleepr7 hours -0.1609465 0.2105939 -0.764 0.44472
sleepr8 hours -0.2722001 0.2142353 -1.271 0.20388
sleepr9 hours -0.2232387 0.2393832 -0.933 0.35105
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

```
Null deviance: 13332.1 on 14151 degrees of freedom
Residual deviance: 9546.1 on 14134 degrees of freedom
```

AIC: 9582.1

Number of Fisher Scoring iterations: 5

```
###Overall model
>options(scipen = 999)
>chidiff = logreg2$null.deviance - logreg2$deviance
>dfdifff = logreg2$df.null - logreg2$df.residual
>chidiff
[1] 3728.35
>dfdifff
[1] 10
>pchisq(chidiff,dfdifff,lower.tail=F)
[1] 0
> #Compute McFadden's pseudo R-square
> McFaddenR2 <- 1 -(logreg1$deviance/logreg0$null.deviance)
> McFaddenR2
[1] 0.2796516
>pchisq(chidiff, ddfdiff, lower.tail=F)
[1] 0

>

#Compute McFadden's pseudo R-square
>McFaddenR2 <- 1 -(logreg2$deviance/logreg0$null.deviance)
>McFaddenR2
[1] 0.2839744
>#Computing odds-ratios (effect sizes) for predictors
library(BaylorEdPsych)
> PseudoR2(logreg2)
McFadden Adj.McFaddenCox.SnellNagelkerkeMcKelvey.ZavoinaEfron
0.2839744 0.2811241 0.2347271 0.3846859 0.3769981 0.2909277
      Count      Adj.Count      AIC      Corrected.AIC
      0.8477954      0.1533019      9582.1438472      9582.1922446

>ORs<-exp(logreg2$coefficients)
>#rounded
>round(ORs, digits=3)
      (Intercept)  genderrFemalebullyingschrYes
      0.041        1.433        1.927
bullyingelcrrYessadnessrYes  threaten1 time
      1.805        10.164        1.500
  threaten2 or 3 times  threaten4 or 5 times  threaten6 or 7 times
      1.435        2.430        1.571
  threaten8 or 9 times  threaten10 or 11 times  threaten12 or more times
      7.133        7.276        3.080
  sleep4 or less hours  sleep5 hours  sleep6 hours
      2.028        1.150        1.000
  sleep7 hours  sleep8 hours  sleep9 hours
      0.851        0.762        0.800

## odds ratios and 95% CI
>exp(cbind(OR = coef(logreg2), confint(logreg2)))
```

Waiting for profiling to be done...

```
OR 2.5 % 97.5 %
(Intercept) 0.04102987 0.0269276 0.06120199
genderrFemale 1.43302806 1.2853330 1.59814995
bullyingschrYes 1.92743672 1.6975761 2.18746927
bullyingelcrYes 1.80519221 1.5755005 2.06763083
sadnessrYes 10.16350665 9.1035757 11.36159536
sleepdisturbancesr4 or less hours 2.02761922 1.3359138 3.13176435
sleepdisturbancesr5 hours 1.15005961 0.7613749 1.76843643
sleepdisturbancesr6 hours 1.00032820 0.6676117 1.52715371
sleepdisturbancesr7 hours 0.85133758 0.5684410 1.29930361
sleepdisturbancesr8 hours 0.76170187 0.5047908 1.17038938
sleepdisturbancesr9 hours 0.79992390 0.5030747 1.28729939
threaten1 time 1.50004017 1.1514332 1.95017933
threaten2 or 3 times 1.43481865 1.0110832 2.03063467
threaten4 or 5 times 2.43015121 1.2586986 4.64191298
threaten6 or 7 times 1.57084190 0.6892317 3.55067103
threaten8 or 9 times 7.13308207 2.7787966 18.68882712
threaten10 or 11 times 7.27590756 1.9259944 25.84880176
threaten12 or more times 3.08044536 1.8141856 5.25174179
```

>

```
#Comparing the null and full models
>anova(logreg0,logreg2,test="Chisq")
Analysis of Deviance Table
```

```
Model 1: csuicider ~ 1
```

```
Model 2: csuicider ~ genderr + bullyingschr + bullyingelcr + sadnessr +
threaten + sleepr
```

```
Resid. Df Resid. Dev Df Deviance Pr(>Chi)
```

```
1 14151 13332.1
```

```
2 14134 9546.1 17 3786 < 2.2e-16 ***
```

```
---
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
binarycorrect = ifelse(correct > 0.5,1,2)
```

```
>binarycorrect = factor(binarycorrect,
```

```
+ levels = c(1,2),
```

```
+ labels = c("Yes", "No"))
```

```
>table(yrbsregdata1c$csuicider, binarycorrect)
```

```
binarycorrect
```

```
Yes No
```

```
No 595 11013
```

```
Yes 985 1559
```

```
##Accuracy for each group
```

```
> 11013 / (595+11013)*100
```

```
[1] 94.87422
```

```
> 985 / (985+1559)*100
```

```
[1] 38.71855
```

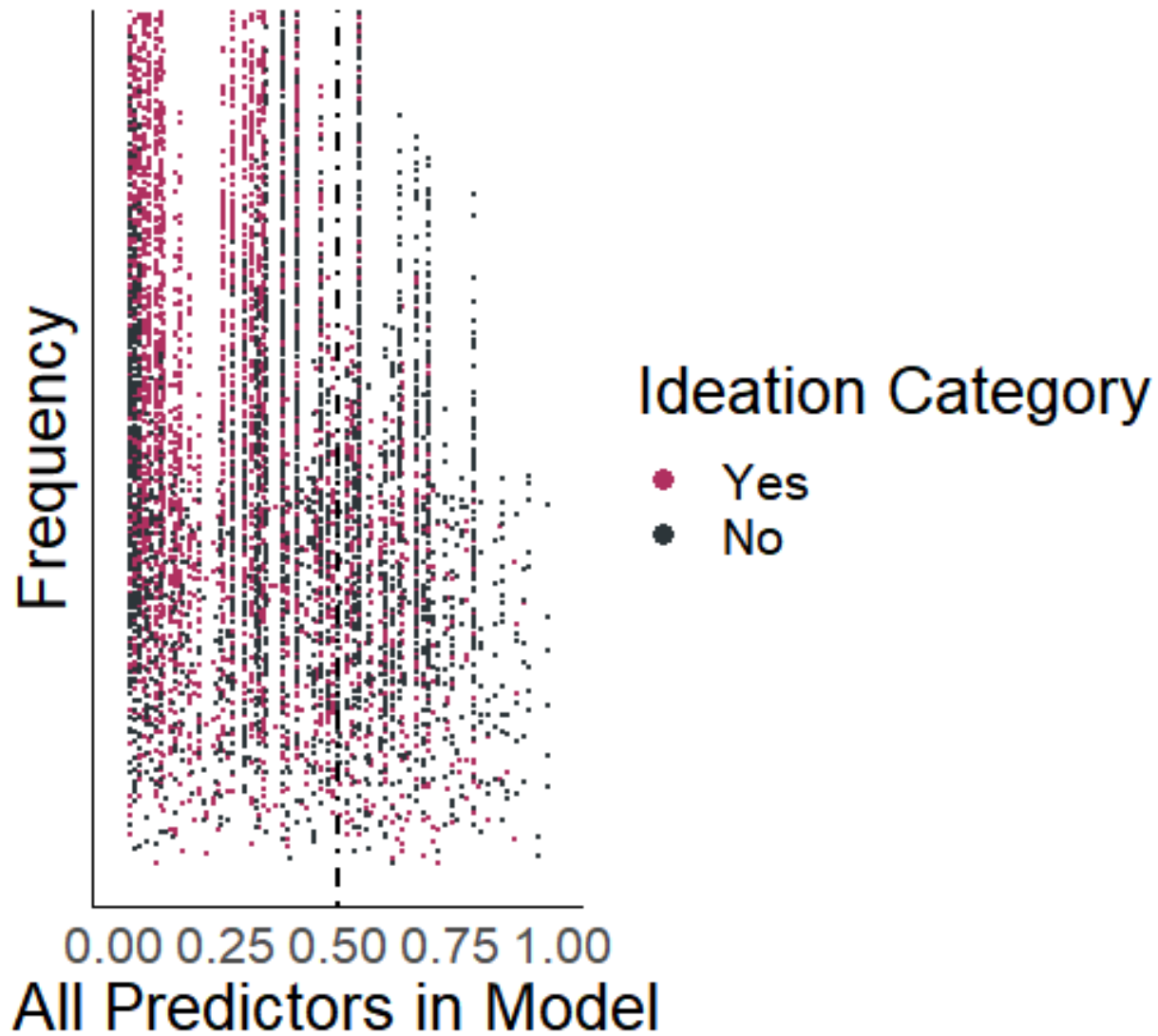


```
> (11013+985) / nrow(logreg1)*100
numeric(0)
> (11013+985) / nrow(yrbsregdata1c)*100
[1] 84.77954
```

```
>
```

```
theme = theme(panel.grid.major = element_blank(),
+             panel.grid.minor = element_blank(),
+             panel.background = element_blank(),
+             axis.text.y = element_blank(),
+             axis.ticks = element_blank(),
+             axis.line.x = element_line(color = "black"),
+             axis.line.y = element_line(color = "black"),
+             text = element_text(size=20),
+             legend.key = element_blank())
> hist = ggplot(yrbsregdata1c, aes(correct, color = csuicider, fill=csuicider))
> hist +
+ theme+
+ geom_dotplot(binwidth = .01, position = "jitter")+
+ coord_cartesian(xlim = c(1,2))+
+ xlab("All Predictors in Model")+
+ ylab("Frequency")+
+ scale_colour_manual(values = c("Maroon", "#2C3539"),
+                     labels = c("Yes", "No"),
+                     name = "Ideation Category")+
+ scale_fill_manual(values = c("Maroon", "#2C3539"),
+                   labels = c("Yes", "No"),
+                   name = "Ideation Category")+
+ geom_vline(xintercept=c(.50), linetype="dotted", size = 1)
> hist +
+ theme+
+ geom_dotplot(binwidth = .01, position = "jitter")+
+ coord_cartesian(xlim = c(0,1))+
+ xlab("All Predictors in Model")+
+ ylab("Frequency")+
+ scale_colour_manual(values = c("Maroon", "#2C3539"),
+                     labels = c("Yes", "No"),
+                     name = "Ideation Category")+
+ scale_fill_manual(values = c("Maroon", "#2C3539"),
+                   labels = c("Yes", "No"),
+                   name = "Ideation Category")+
+ geom_vline(xintercept=c(.50), linetype="dotted", size = 1)
```

```
>
```



```
#First we convert the Dep var into 0-1:
> library(varhandle)
> actual=as.data.frame(to.dummy(yrbsregdata1c$scuicider,"yes"))
> actual=actual$yes.Yes
> #Now install lib for H-L test and run:
> library(ResourceSelection)
```

```
hoslem.test(actual, fitted(logreg2), g=10)
```

Hosmer and Leme show goodness of fit (GOF) test

```
data: actual, fitted(logreg1)
X-squared = 15.3, df = 8, p-value = 0.05356
```

```
>
```

2808 total said yes of suicide.
 Among the participants who reported they considered committing suicide
 1058 (41%) of them declared they were bullied at school property
 905 (36%) of them declared they were electronically bullied
 2000 (79%) of them expressed they experienced sadness and hopelessness
 338 (13%) of them expressed they were victimized of threatening injury by a weapon at school
 1488 (58%) of the participants presented that they suffered from sleep deprivation sleeping 6 hours or less daily.

table(yrbssuisadility\$gender)

1852	Female	936	Male
------	--------	-----	------

table(yrbssuisadility\$bullyingsch)

Yes	No
1148	1634

table(yrbssuisadility\$bullyingelc)

Yes	No
994	1791

>

table(yrbssuisadility\$sadness)

Yes	No
2188	603

table(yrbssuisadility\$sadness)

Yes	No
2188	603

>

table(yrbssuisadility\$sleepdisturbances)

4 or less hours	5 hours	6 hours	7 hours	8 hours
460	433	634	579	354
9 hours	10 or more hours			
101	44			

>