



The University of Burdwan

Project name = Students' Mental Health Analysis and
Visualization using R programming

Name = Xeviers koner

Roll No. = BUR/ST/2021/026

Reg. No. = 201801015980 of 2018-19

Subject = Statistics

Semester = 3rd Semester

Course Name = Community engagement Programme

Course Code = MSST307

Session = 2021-23

Students' Mental Health Analysis and Visualization using R programming

Abstract:

The aim of this project "Students' Mental Health Analysis and Visualization using R programming" is to analyze and visualize the mental health of students using various statistical techniques and R programming tools. Firstly these are primary data . I have collected this data through a straight questionnaire using a Google form randomly from a 25 of students with different backgrounds. In our data there are many parameters like – depression , anxiety, number of incoming members , yearly family income , schooling , gender , age last semester mark and alone time spent . So first I collected this data then I the parameters are visualized using ggplot() or plotly() function . Then I used Generalized Linear Model(GLM) . So we know that ,GLM models have a defined relationship between the expected variance and the mean. This relationship can be used to evaluate the model's goodness of fit to the data. The deviance can be used for this goodness of fit check.

This project aims to provide insights into the mental health of students and identify areas where interventions can be made to improve their mental health and well-being.

Introduction:

- **What is MENTAL HEALTH ?**

Mental health is a state of well-being in which a person understands his or her own abilities, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to his or her community .Both physical and mental health are the result of a complex interplay between many individual and environmental factors, including: family history of illness and disease/genetics , lifestyle and health behaviors (e.g., smoking, exercise, substance use), levels of personal and workplace stress ,exposure to toxins ,exposure to trauma ,personal life circumstances and history , access to supports (e.g., timely healthcare, social supports), coping skills etc.. Since here we are talking about students' mental health , so will focus on it .

While there is limited empirical evidence to understand the full extent of the pandemic's impacts on mental health, several studies highlight the generally poor state of mental well-being for Indian youth and the broader stigma around it. According to UNICEF's survey findings for 21 countries in *The State of the World's Children 2021* report, every one in seven youngsters between 15 and 24 years in India reported some form of poor mental health such as feelings of depression or disinterest. and only 41% felt there was a need to reach out for support when people experienced mental health issues. This was the lowest across the 21 countries surveyed. The mental health crisis in India is further compounded by a severe shortage of trained professionals. A study from the *Indian Journal of Psychiatry* estimates that India has 0.75 psychiatrists per 1,00,000 patients, which is much lower than the recommended three psychiatrists per 1,00,000.

Objective:

Here our objective is first we input our data. then we will visualize our data using `plotly()` function in R programming . Here we will conduct our coding in R studio. We visualize age distribution of our data , gender distribution of data , schooling distribution of our data , depression ,depression vs gender stacked bar graph , depression vs age stacked bar graph , distribution of last semester marks , marks vs depression , income of family , depression vs income . And lastly I will conduct logistic regression to check how much our response variable depends upon our predictor variables. So here we will Generalized Linear Model (GLM).This function uses the following syntax:

`glm(formula, family=Gaussian, data, ...)`

Where,

- **formula:** The formula for the linear model (e.g. $y \sim x1 + x2$)
- **Family:** The statistical family to use to fit the model. Default is Gaussian but other options include binomial, Gamma, and Poisson among others.
- **data:** The name of the data frame that contains the data

In practice, this function is used most often to fit logistic regression models by specifying the 'binomial' family.

- **Data:**

So here is our primary data for our analysis:

name	schooling	income of source	family income	anxiety	depression	alone time(hrs)	gender	Age	last semester mark
Ritika Roy									
Chowdhury	regular	both	2L-5L	1	0	9	female	23	9.16
Mainak Mondal	regular	father	2L-5L	1	1	13	male	23	7.5
Priyanka Dutta	regular	both	5L+	1	1	12	female	22	8
sunanda	regular	father	5L+	1	0	2	female	23	7.5
Tiyasa Das	regular	father	5L+	1	1	17	female	23	8.34
Jit Mondal	regular	father	2L	0	0	9	male	18	8.5
Chiranjit Biswas	regular	both	2L	0	1	1	male	27	6.56
ASIF AHMED									
MOLLA	boarding	father	2L	1	0	8	male	25	8.3
SOUMALLYA KOLEY	boarding	both	2L	0	1	12	male	26	9
Najma Sahana	regular	father	2L	1	1	1	female	20	7.9
Arnab Biswas	regular	both	5L+	1	1	3	male	23	8.5
Debarati Jash	regular	both	5L+	1	1	20	female	19	7.5
Madhurima									
Bhaumik	regular	both	5L+	1	1	20	female	19	7.77
Ankita das	regular	father	5L+	1	1	18	female	20	8.32
Souparni									
Mazumder	regular	father	2L-5L	1	1	10	female	19	7.5
WRITU MUKHERJEE	regular	mother	2L	1	1	5	female	24	7.83
Disa Roy	boarding	mother	2L	1	1	2	female	22	6.4
Annesha Roy	regular	father	2L-5L	0	0	7	female	21	8.2
Pritam Kabiraj	regular	father	2L	0	1	10	male	25	8
Surya	boarding	father	2L	1	1	6	male	18	7.6
Jayeeta Pal	regular	father	2L-5L	0	0	8	female	20	9
Abir Biswas	boarding	both	5L+	0	0	10	male	19	8.6
Moini Ray	boarding	father	5L+	1	1	3	female	23	7.9
Arghya Mondal	regular	mother	2L-5L	0	1	12	male	28	8.2
Joy Hazra	boarding	father	2L	1	1	8	male	23	7.87

• Questions asked –

- 1) What is your name ? 2) What is your gender? 3) Your age?
- 4) What type of schooling? 5) Income of source?
- 6) Your family income (yearly)? 7) Do you have anxiety?
- 8) Do you have depression? 9) How much time you spend alone?
- 10) Your last semester mark?

➤ Analysis:

- We will call our data :-

R input :

```
library(plotly)
```

```
library(tidyverse)
```

```
library(dplyr)
```

```
## calling our data
```

```
project <- read.csv("C:\\Users\\User\\Documents\\mental  
health.csv")
```

```
project
```

➤ Visualizing our data:-

Now we will visualize our data using plotly() function.

- Age distribution of our data :-

We have age parameter in our data. First we will call our data that was defined by project() function , then we will summarize it then we will use plotly() function to construct a bar chart “Age” will lie on the X axis and “count” will lie on Y axis . Here is the R code for:

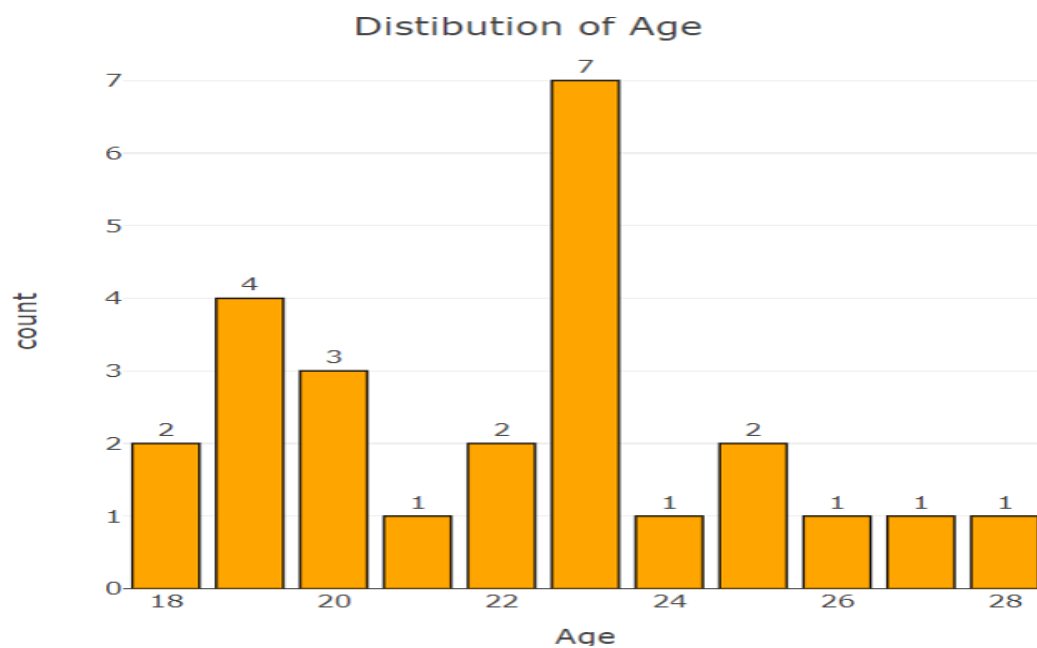
R input:

```
## Data visualization
```

```
## Age distribution of our data
```

```
project %>%  
  
  group_by(Age)%>%  
  
  summarize(count = n()) %>%  
  
  plot_ly(x=~Age, y=~count, type = 'bar',  
          text = ~count,  
          textposition = 'outside',  
          marker = list(color = 'orange',  
                        line = list(color = 'black',  
                                   width = 1.0))) %>%  
  
  layout(title = 'Distibution of Age')
```

Bar plot:



- **Gender distribution of our data :-**

We have gender parameter in our data. First we will call our data that was defined by project() function , then we will summarize it then we will use plotly() function to construct a pie chart. Here is the R code for:

R Input:

```
## Gender distribution of our data
```

```
project_SummaryStat <- project %>%
```

```
  group_by(gender) %>%
```

```
  summarise(count = n(),
```

```
            percentage = round((n()/ nrow(project)), digits = 4))
```

```
project_SummaryStat
```

```
colors <- c("yellow","red")
```

```
Gender_PieChart <- plot_ly(data = project_SummaryStat, labels = ~gender,  
  values = ~percentage,
```

```
    type = 'pie', sort = F,
```

```
    textposition = 'inside',
```

```
    textinfo = 'label+percent',
```

```
    insidetextfont = list(color = 'White'),
```

```
    hoverinfo = 'text',
```

```
    text = ~count,
```

```
    marker = list(colors = colors,
```



```
line = list(color = 'Black', width = 1)),
```

```
showlegend = TRUE)
```

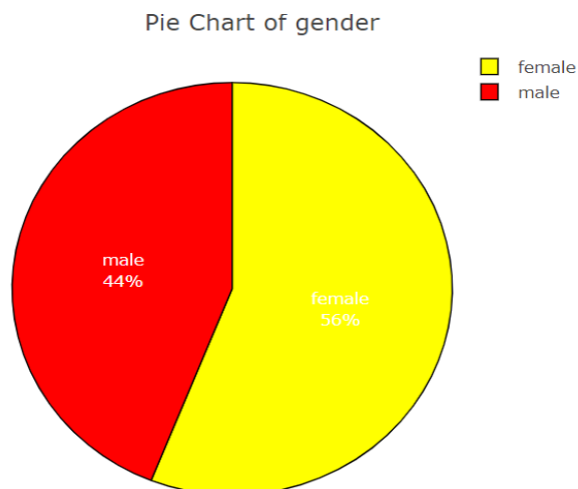
```
Gender_PieChart <- Gender_PieChart %>% layout(title = 'Pie Chart of gender')
```

```
Gender_PieChart
```

R Output:

```
# A tibble: 2 × 3
  gender count percentage
  <chr>   <int>         <dbl>
1 female     14         0.56
2 male      11         0.44
```

Pie chart:



So we can see from our output and pie chart that 14 females and 11 males i.e., 56% female and 44% male participated in this survey. The females are identified by yellow and the males are identified by red color in our pie chart. So, majority of our data belongs to females.

- **Schooling distribution of our data :-**

We have schooling parameter in our data. First we will call our data that was defined by project() function , then we will summarize it then we will use plotly() function to construct a pie chart. Here is the code for:

R Input:

```
##Schooling distribution of our data

project_SummaryStat1 <- project %>%

group_by(schooling) %>%

  summarise(count = n(),

             percentage = round((n()/ nrow(project)), digits = 4))

project_SummaryStat1

colors <- c("pink")

schooling_PieChart <- plot_ly(data = project_SummaryStat1, labels = ~schooling,
                              values = ~percentage,

                              type = 'pie', sort = F,

                              textposition = 'inside',

                              textinfo = 'label+percent',

                              insidetextfont = list(color = 'black'),

                              hoverinfo = 'text',

                              text = ~count,
```

```

marker = list(colors = colors,

               line = list(color = 'Black', width = 2)),

showlegend = TRUE)

schooling_PieChart <- schooling_PieChart %>% layout(title = 'Pie Chart of
schooling')

schooling_PieChart

```

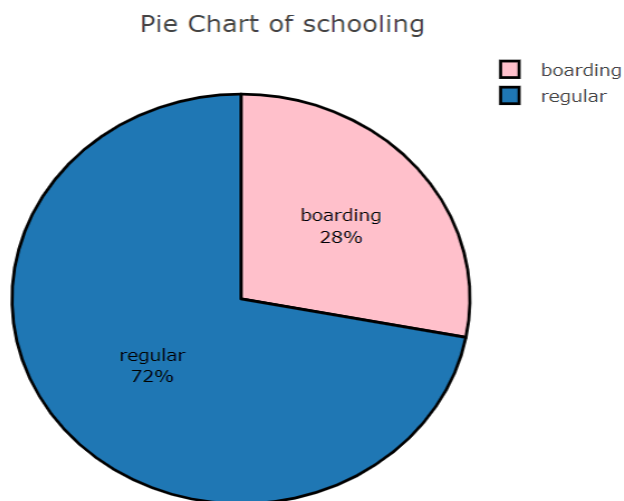
R Output:

```

# A tibble: 2 × 3
  schooling count percentage
  <chr>      <int>      <dbl>
1 boarding     7      0.28
2 regular    18      0.72

```

Pie chart:



So here from the output we can see that among the 25 students 18 students had completed their regular schooling and 7 students had completed their boarding schooling. And from the pie chart we see

that the percentage of regular schooling student is 72% which is identified by blue color in our pie chart and the percentage of bearding schooling student is 28% which is identified by pink color in our pie chart.

- **Distribution of Depression in our data:-**

We have depression parameter in our data. First we will our data that was defined by project() function ,then we will summarize our data then we will use plotly() function to construct a pie chart . Here is the R code for :

R input:

```
## Depression

project_SummaryStat2 <- project %>%

group_by(depression) %>%

summarise(count = n(),

           percentage = round((n()/ nrow(project)), digits = 4))

project_SummaryStat2

colors <- c("green" ,"red")

depression_PieChart <- plot_ly(data = project_SummaryStat2,           labels =
~depression, values = ~percentage,

                                type = 'pie', sort = F,

                                textposition = 'inside',

                                textinfo = 'label+percent',

                                insidetextfont = list(color = 'black'),

                                hoverinfo = 'text',
```

```
text = ~count,
```

```
marker = list(colors = colors,
```

```
line = list(color = 'Black', width = 2)),
```

```
showlegend = TRUE)
```

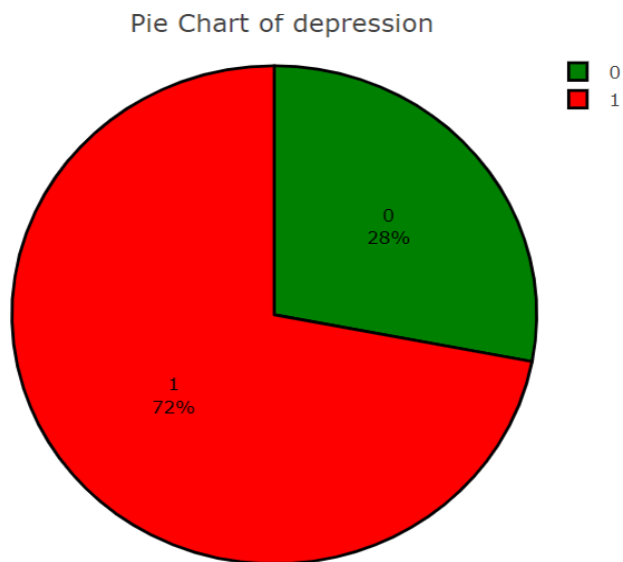
```
depression_PieChart <- depression_PieChart %>% layout(title = 'Pie Chart of  
depression')
```

```
depression_PieChart
```

R output:

```
# A tibble: 2 × 3  
  depression count percentage  
    <int>    <int>      <dbl>  
1         0         7      0.28  
2         1        18      0.72
```

Pie chart:



So We can see from our output and pie chart that 72% of our data suffer in depression that are identified by red color in pie chart and 28 % not that are identified by green color in our pie chart .

- **Depression vs. Gender:**

Now we will see that if there is any relationship between depression and gender. That means we will check that depression belongs to which of the amongst genders. So for this we will construct a stacked bar plot using plotly() function . Here is the R code for :

R Input:

```
## depression vs gender
```

```
project %>%
```

```
  count(gender, depression, sort = F) %>%
```

```
  group_by(gender) %>%
```

```
  mutate(prop = round((n / sum(n)),digits = 4)) %>%
```

```
  plot_ly(x = ~gender, y=~prop, color = ~depression, type = "bar",
```

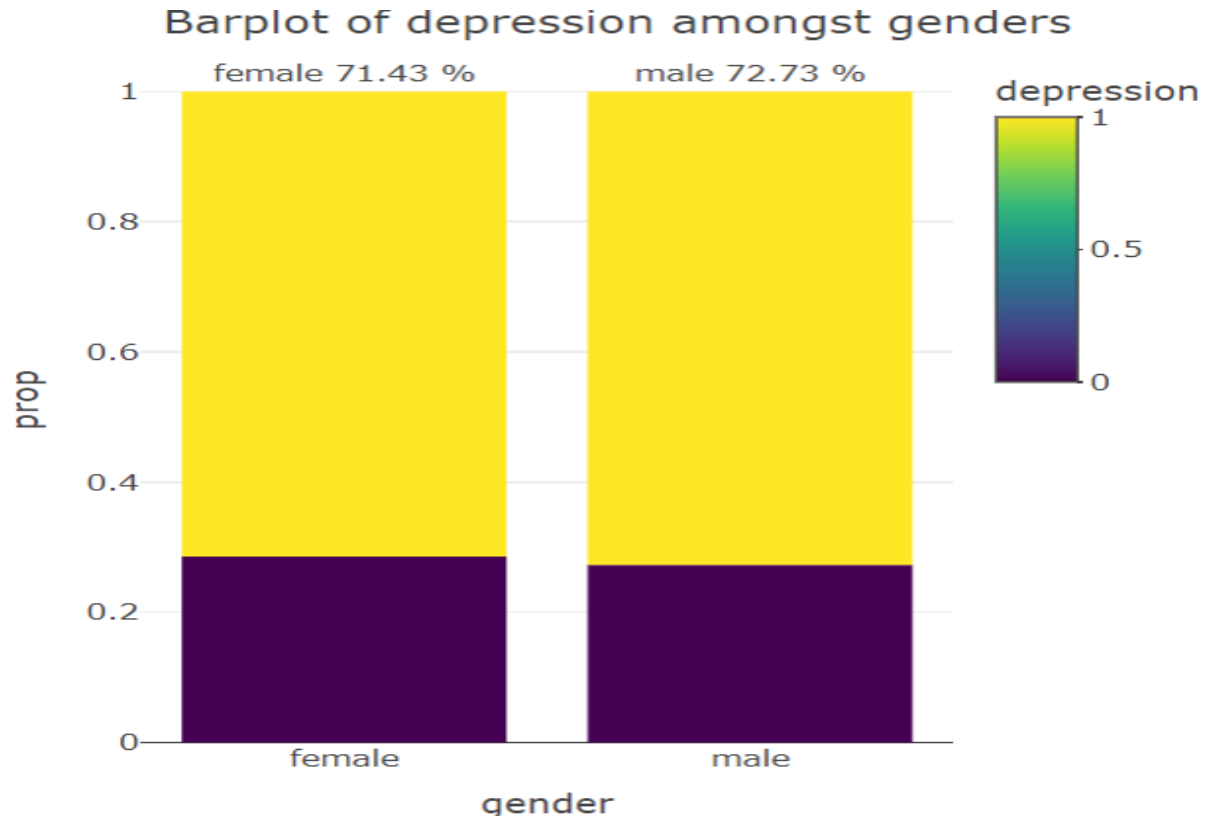
```
          text = ~paste(gender, prop*100 ,'%'),
```

```
          textposition = 'outside') %>%
```

```
  layout(barmode = 'Stacked',
```

```
          title = 'Barplot of depression amongst genders')
```

Bar Plot:



So we can see from our stacked bar plot that the percentage of depression belongs to female gender is 71.43% and on the other hand this percentage is 72.73% to male genders. Although in our data the percentage female participants are more than the male participants, the percentage of depression is quite more for male genders than females. Here already given our depression parameter in our bar plot aside. As much as close to yellow color means the depression is close to 1 (which is 'yes' in our questionnaire) and as much as close to purple color means that he/she has no depression which is close to 0.

- **Depression vs. Age:**

Now we will see that if there is any relationship between depression and age. That means we will check that depression belongs to which of

the year group or period. So for this we will construct a stacked bar plot using `plotly()` function . Here is the R code for :

R Input:

```
## depression vs age

project %>%

  count(Age, depression, sort = F) %>%

  group_by(Age) %>%

  mutate(prop = round((n / sum(n)),digits = 4)) %>%

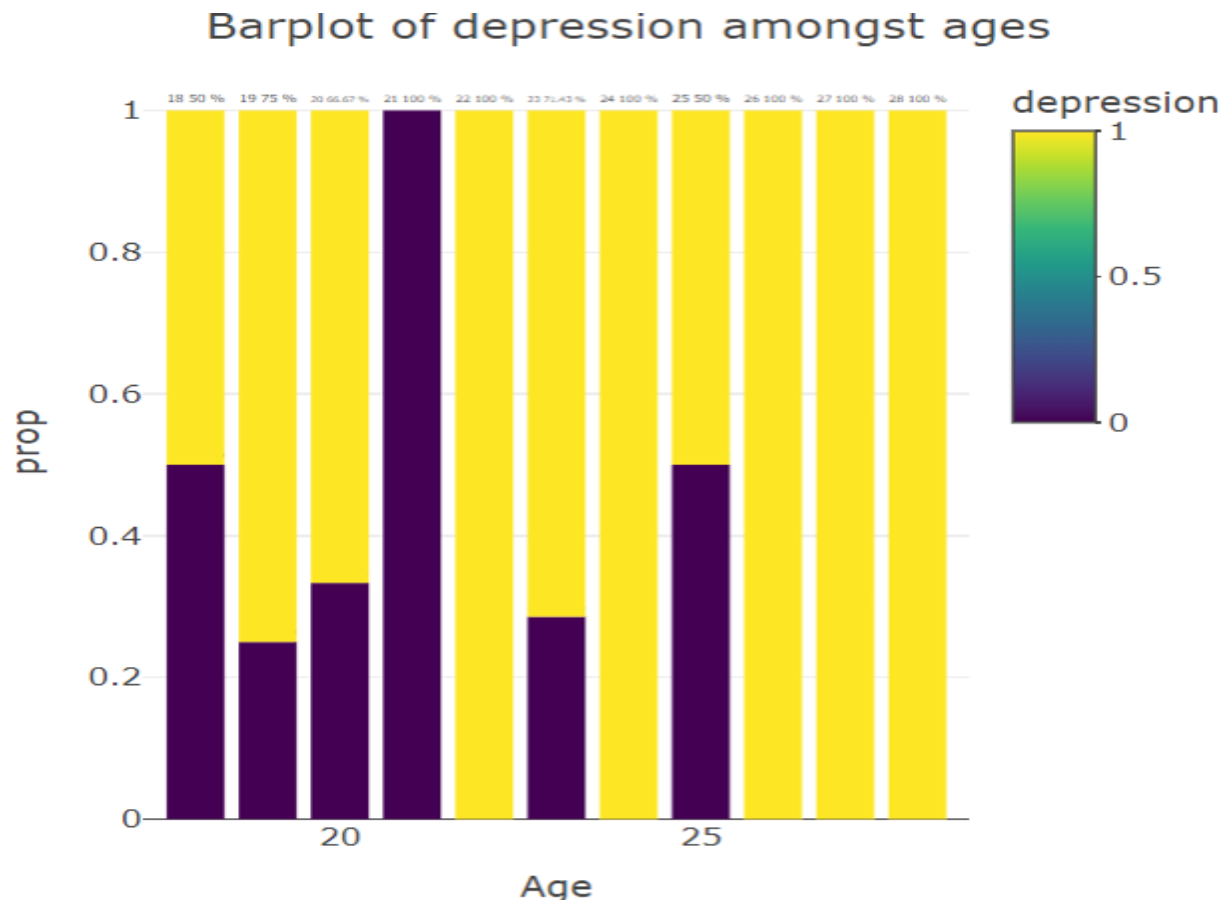
  plot_ly(x = ~Age, y=~prop, color = ~depression, type = "bar", text = ~paste(Age,
prop*100 ,'%'),

          textposition = 'outside') %>%

  layout(barmode = 'group', title = 'Barplot of depression amongst ages')
```

R Output:

So, from our stacked bar plot we can see that the distribution of depression among the ages of students. Here the depression parameter is given aside our bar plot. Close to yellow color means close to positive depression and close to purple means close to no depression. So we can see that since we have no 21 years old aged student in our data so the rate of depression is more from 22 years old and it is high after 25 years old .



- **Distribution of last semester marks:**

We have last semester marks parameter in our data. First we will call our data that was defined by `project()` function ,then we will summarize our data then we will use `plotly()` function to construct a bar chart respective for each students .Here is the R code for :

R input:

```
## distribution of last semester marks
```

```
project %>%
```

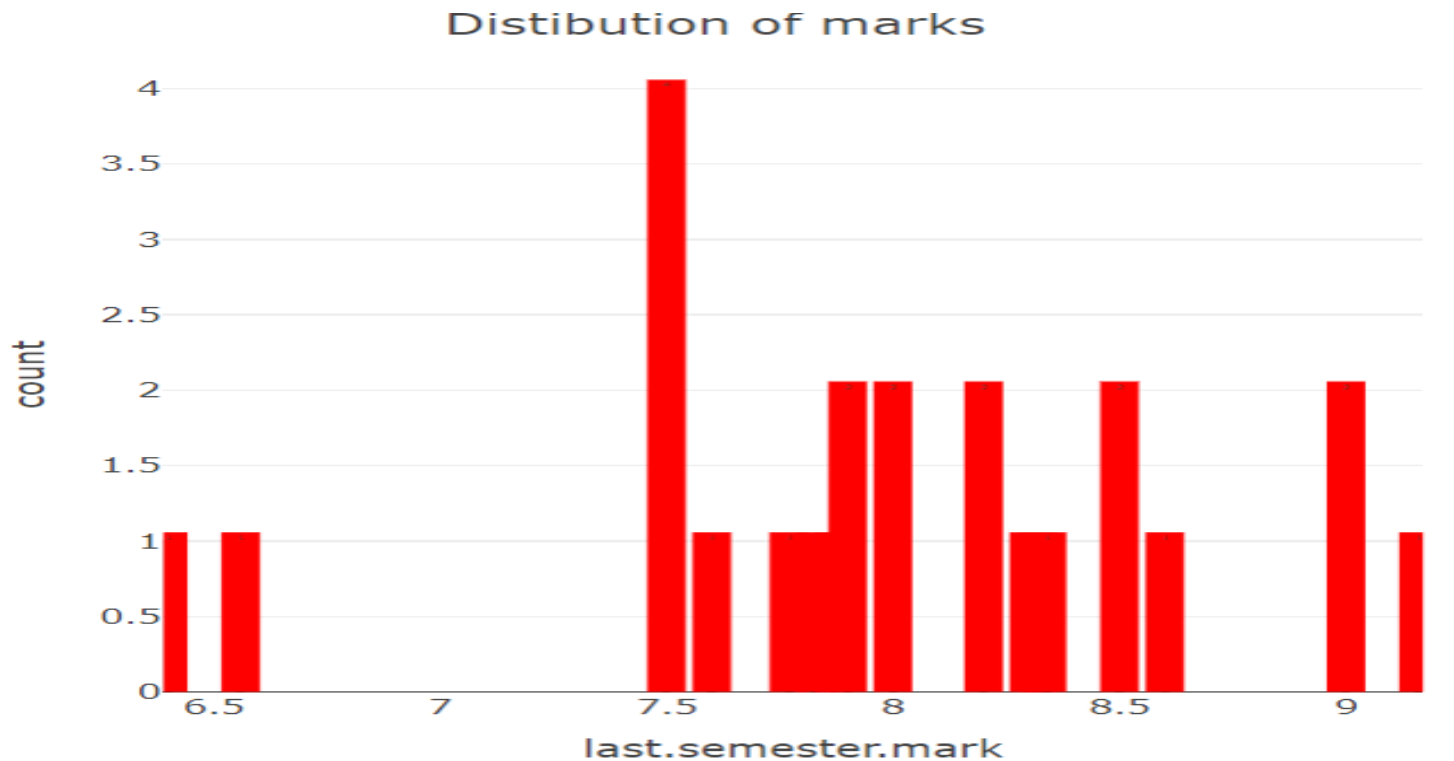
```
  group_by(last.semester.mark)%>%
```

```
  summarize(count = n()) %>%
```

```
plot_ly(x=~last.semester.mark, y=~count, type = 'bar', text = ~count,  
textposition = 'outside', marker = list(color = 'black', line = list(color = 'red', width  
= 10.0))) %>%
```

```
layout(title = 'Distibution of marks')
```

Bar Chart:



So we can see from our bar chart that most of the students obtained marks between 7.5 to 8.5 cgpa in their respective last semester examination. And above 9.00 and less than 7.5 cgpa the number of students is few in our data.

- **Depression vs. marks :**

Now we will see that if there is any relationship between depression and semester obtained marks. That means we will check that if

depression depends upon the last semester marks respect to each students. So for this we will construct a stacked bar plot using plotly() function . Here is the R code for:

R Input:

```
##Marks vs depression
```

```
project %>%
```

```
  count(last.semester.mark, depression, sort = F) %>%
```

```
  group_by(last.semester.mark) %>%
```

```
  mutate(prop = round((n / sum(n)),digits = 4)) %>%
```

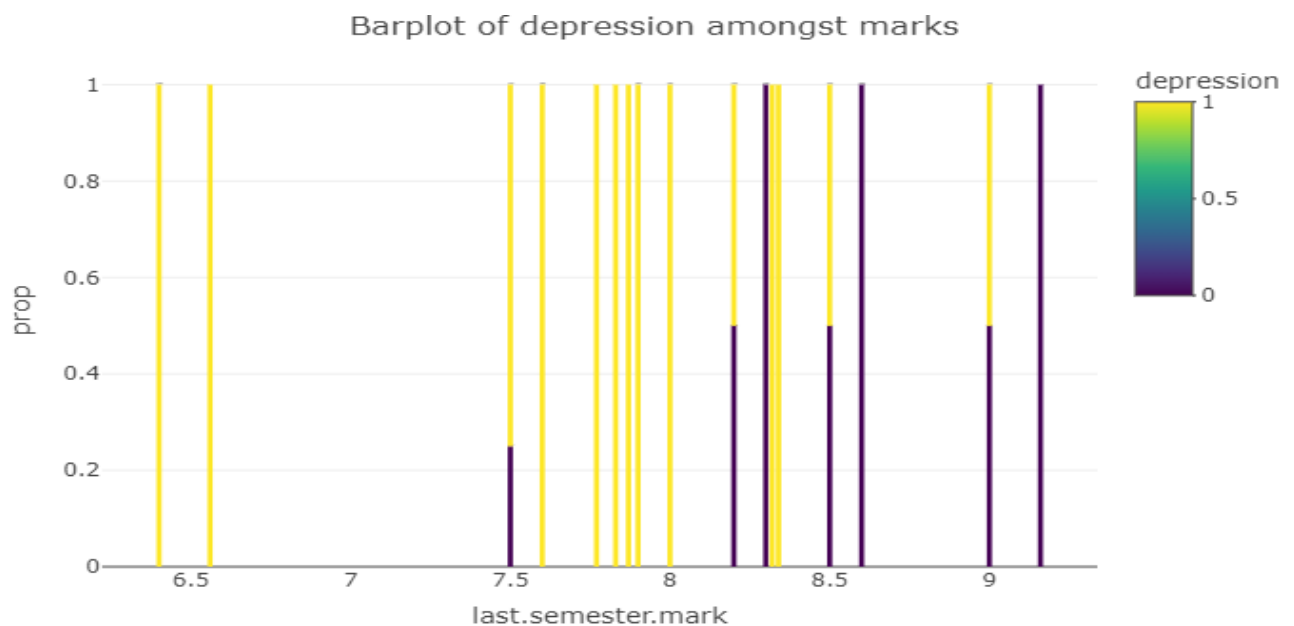
```
  plot_ly(x = ~last.semester.mark, y=~prop, color = ~depression, type = "bar",
```

```
    text = ~paste(last.semester.mark, prop*100 ,'%'),
```

```
    textposition = 'outside' , width=8.0) %>%
```

```
  layout(barmode = 'group', title = 'Barplot of depression amongst marks')
```

Bar plot:



So can see from this stacked bar graph that the students who obtained cgpa between 7.5 and 8.5 has lot depression than the students who obtained more cgpa more than 8.5 and less than 7.5. Here in our bar chart the depression parameter is given .

- **Distribution of Family Income:**

We have last family income parameter in our data. First we will our data that was defined by project() function ,then we will summarize our data then we will use plotly() function to construct a bar chart respective for each students .Here is the R code for :

R Input:

```
## Family income

project %>%

  group_by(family.income)%>%

  summarize(count = n()) %>%

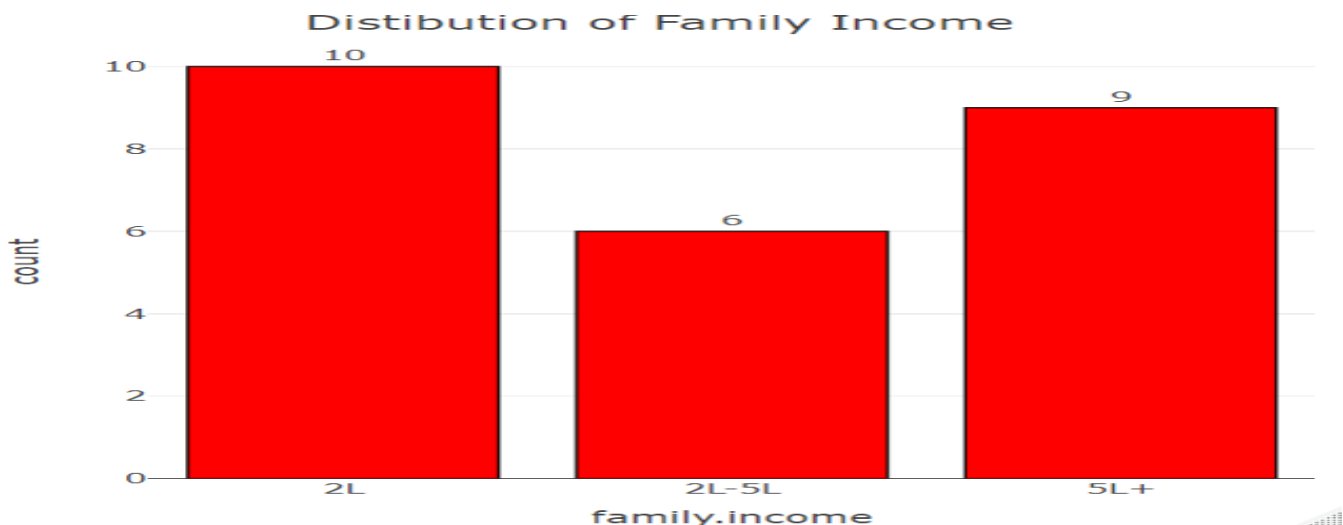
  plot_ly(x=~family.income, y=~count, type = 'bar',text = ~count,

          textposition = 'outside',marker = list(color = 'red',line = list(color =
'black', width = 1.0))) %>%

  layout(title = 'Distibution of Family Income')
```

R output:

So, we can see from our below bar graph that most of the students' family income is below 2lakh(2L) rupees and the number is 10 .The number of students whose family income is between 2lakh and 5lakh(2L-5L) rupees is 6 and the number of students whose family income is above 5lakh(5L+) is 9 .



- **Depression vs. Family Income:**

Now we will see that if there is any relationship between depression and family income of each respective students. That means we will check that how much depression depends upon their family income or not. So for this we will construct a stacked bar plot using `plotly()` function . Here is the R code for:

R Input:

```
## Family Income vs depression
```

```
project %>%
```

```
  count(family.income, depression, sort = F) %>%
```

```
  group_by(family.income) %>%
```

```
  mutate(prop = round((n / sum(n)), digits = 4)) %>%
```

```
  plot_ly(x = ~family.income, y=~prop, color = ~depression, type = "bar",
```

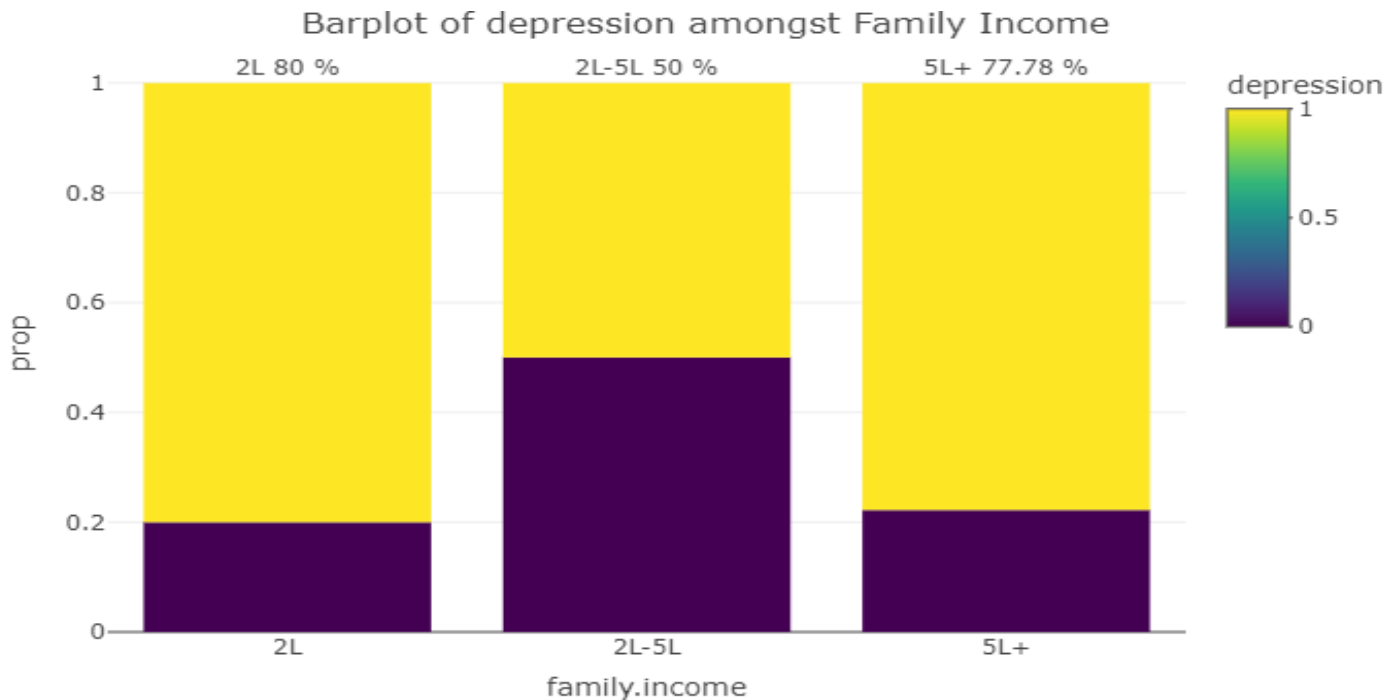
```
    text = ~paste(family.income, prop*100 ,'%'),
```

```
    textposition = 'outside' , width=8.0) %>%
```

```
  layout(barmode = 'group',
```

```
  title = 'Barplot of depression amongst Family Income')
```

Bar Plot:



Here we can see from our bar plot that those students have most depression whose family income is less than 2lakh(2L) and the percentage is 80% . Then those students have 2nd most depression whose family is more than 5 lakh(5L+) and the percentage is 77.78% , then the students have last most depression whose family income is between 2lakh and 5lakh(2L-5L) and the percentage is 50%.

- **Analysis using Logistic Regression:**

Logistic regression is a simple but powerful model to predict binary outcomes. That is, whether something will happen or not. It's a type of classification model for supervised machine learning. Logistic regression is used in almost every industry- marketing, healthcare, social sciences, and others.

Logistic regression finds the best possible fit between the predictor and target variables to predict the probability of the target variable belonging to a labeled class/category.

Linear regression tries to find the best straight line that predicts the outcome from the features. It forms an equation like:

```
y_predictions = intercept + slope * features
```

So here we will use Generalized Linear Model (GLM). In practice, this function is used most often to fit logistic regression models by specifying the 'binomial' family.

Here, we will assume that the response variable ,Y = depression and the predictor variables are , X =

R Input:

```
## Logistic regression in our data
```

```
library(caTools) # For Logistic regression
```

```
library(ROCR) # For ROC curve to evaluate model
```

```
project <- read.csv("C:\\Users\\User\\Documents\\mental health.csv")
```

```
project
```

```
split <- sample.split(project, SplitRatio = 0.8) # Splitting dataset
```

```
split
```

```
train_reg <- subset(project, split == "TRUE")
```

```
test_reg <- subset(project, split == "FALSE")
```

```
logistic_model <- glm(depression ~ alone.time.hrs. +  
last.semester.mark+family.income+Age,data = train_reg, family = "binomial") #  
Training model
```

```
logistic_model
```

```
summary(logistic_model) # Summary
```

R Output:

```
## Logistic regression in our data
>
> library(caTools)
> library(ROCR)
> project <- read.csv("C:\\Users\\User\\Documents\\mental health.csv")
> project
```

	name	schooling	income.of.source	family.income	anxiety
1	Ritika Roy Chowdhury	regular	both	2L-5L	1
2	Mainak Mondal	regular	father	2L-5L	1
3	Priyanka Dutta	regular	both	5L+	1
4	sunanda	regular	father	5L+	1
5	Tiyasa Das	regular	father	5L+	1
6	Jit Mondal	regular	father	2L	0
7	Chiranjit Biswas	regular	both	2L	0
8	ASIF AHMED MOLLA	boarding	father	2L	1
9	SOMALLYA KOLEY	boarding	both	2L	0
10	Najma Sahana	regular	father	2L	1
11	Arnab Biswas	regular	both	5L+	1
12	Debarati Jash	regular	both	5L+	1
13	Madhurima Bhaumik	regular	both	5L+	1
14	Ankita das	regular	father	5L+	1
15	Souparni Mazumder	regular	father	2L-5L	1
16	WRITU MUKHERJEE	regular	mother	2L	1
17	Disa Roy	boarding	mother	2L	1
18	Annesha Roy	regular	father	2L-5L	0
19	Pritam Kabiraj	regular	father	2L	0
20	Surya	boarding	father	2L	1
21	Jayeeta Pal	regular	father	2L-5L	0
22	Abir Biswas	boarding	both	5L+	0
23	Moini Ray	boarding	father	5L+	1
24	Arghya Mondal	regular	mother	2L-5L	0
25	Joy Hazra	boarding	father	2L	1

	depression	alone.time.hrs.	gender	Age	last.semester.mark
1	0	9	female	23	9.16
2	1	13	male	23	7.50
3	1	12	female	22	8.00
4	0	2	female	23	7.50
5	1	17	female	23	8.34
6	0	9	male	18	8.50
7	1	1	male	27	6.56
8	0	8	male	25	8.30
9	1	12	male	26	9.00
10	1	1	female	20	7.90
11	1	3	male	23	8.50
12	1	20	female	19	7.50
13	1	20	female	19	7.77
14	1	18	female	20	8.32
15	1	10	female	19	7.50
16	1	5	female	24	7.83
17	1	2	female	22	6.40
18	0	7	female	21	8.20
19	1	10	male	25	8.00
20	1	6	male	18	7.60
21	0	8	female	20	9.00
22	0	10	male	19	8.60
23	1	3	female	23	7.90
24	1	12	male	28	8.20
25	1	8	male	23	7.87


```
> split <- sample.split(project, SplitRatio = 0.8)
> split
[1] TRUE TRUE TRUE TRUE TRUE TRUE FALSE TRUE TRUE FALSE
> train_reg <- subset(project, split == "TRUE")
> test_reg <- subset(project, split == "FALSE")
> logistic_model <- glm(depression ~ alone.time.hrs. + last.semester.mark+family.income+Age,
+ data = train_reg,
+ family = "binomial")
> logistic_model
```

```
Call: glm(formula = depression ~ alone.time.hrs. + last.semester.mark +
family.income + Age, family = "binomial", data = train_reg)
```

```
Coefficients:
(Intercept) alone.time.hrs. last.semester.mark
5.90958 0.40639 -2.75187
family.income2L-5L family.income5L+ Age
0.01359 1.61989 0.59502
```

```
Degrees of Freedom: 20 Total (i.e. Null); 15 Residual
```

```
Null Deviance: 26.73
```

```
Residual Deviance: 14.05 AIC: 26.05
```

```
>
> # Summary
> summary(logistic_model)
```

```
Call:
glm(formula = depression ~ alone.time.hrs. + last.semester.mark +
family.income + Age, family = "binomial", data = train_reg)
```

```
Deviance Residuals:
Min 1Q Median 3Q Max
-1.7952 -0.2949 0.1221 0.6376 1.6012
```

```
Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept) 5.90958 13.50777 0.437 0.6618
alone.time.hrs. 0.40639 0.26485 1.534 0.1249
last.semester.mark -2.75187 1.57386 -1.748 0.0804 .
family.income2L-5L 0.01359 2.23269 0.006 0.9951
family.income5L+ 1.61989 2.10496 0.770 0.4416
Age 0.59502 0.43879 1.356 0.1751
---
```

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

```
(Dispersion parameter for binomial family taken to be 1)
```

```
Null deviance: 26.734 on 20 degrees of freedom
```

```
Residual deviance: 14.047 on 15 degrees of freedom
```

```
AIC: 26.047
```

```
Number of Fisher Scoring iterations: 6
```

```
>
```

- **Interpretation of our model:**

The coefficient estimate in the output indicates the average change in the log odds of the response variable associated with a one unit increase in each predictor variable.

The standard error gives us an idea of the variability associated with the coefficient estimate. We then divide the coefficient estimate by the standard error to obtain a z value.

The p-value $P(>|z|)$ tells us the probability associated with a particular z value. This essentially tells us how well each predictor variable is able to predict the value of the response variable in the model.

the null deviance in the output tells us how well the response variable can be predicted by a model with only an intercept term.

The residual deviance tells us how well the response variable can be predicted by the specific model that we fit with p predictor variables. The lower the value, the better the model is able to predict the value of the response variable.

To determine if a model is “useful” we can compute the Chi-Square statistic as:

$\chi^2 = \text{Null deviance} - \text{Residual deviance with } p \text{ degrees of freedom.}$

We can then find the p-value associated with this Chi-Square statistic. The lower the p-value, the better the model is able to fit the dataset compared to a model with just an intercept term.

The Akaike information criterion (**AIC**) is a metric that is used to compare the fit of different regression models. The lower the value, the better the regression model is able to fit the data.

- **Conclusion:**

In the summary as the p-value in the last column is more than 0.05 for all the explanatory variables. So all the explanatory variables are insignificant in contributing to the value of the variable 'depression'.

Here , Null deviance =26.734 with df=20

And Residual deviance = 14.047 with df=15

So, Chi-square = Null deviance – Residual deviance

$$= 26.734 - 14.047$$

$$= 12.687 \text{ with df } 5$$

Now we can use the p value table with chi-square value that X^2 value of 12.687 with 5 degrees of freedom has a p value of 0.026495. Since this p-value is much less than 0.05, we would conclude that the model is highly useful.

Acknowledgement

Presentation, Inspiration and Motivation have always played a key role in the success any venture. Primarily I would like to thank the supreme power the Almighty God who is obviously the one who has always guided me to work on the right path of life. I am greatly oblige in taking the opportunity to sincerely thanks our professors for helping and encouraging in this project even in this pandemic situation of covid19. Lastly, I thank almighty, my parents, brother and friends for their constant encouragement without which this project not be possible. I have no valuable words to express my thanks, but my heart is still full of favors receive from every person.
