Mental Health Analysis

Code

Hide

```
file.exists("C:\\Users\\Yash\\Desktop\\Mental Health Data\\Mental Health Analysis\\mental_health_finaldata_1.csv")
```

```
[1] TRUE
```

Hide

```
df <- read.csv("mental_health_finaldata_1.csv")
```

#Exploring the Dataset:

Hide

```
head(df)
```

	Age <chr>	Gen… <chr>	Occupation <chr>	Days_Indoors <chr>	Growing_Stress <chr>	Quarantine_Frustrations <chr>
1	20-25	Female	Corporate	1-14 days	Yes	Yes
2	30-Above	Male	Others	31-60 days	Yes	Yes
3	30-Above	Female	Student	Go out Every day	No	No
4	25-30	Male	Others	1-14 days	Yes	No
5	16-20	Female	Student	More than 2 months	Yes	Yes
6	25-30	Male	Housewife	More than 2 months	No	Yes

6 rows | 1-7 of 13 columns

#Summarising the Dataset:

Hide

```
summary(df)
```

```
     Age               Gender           Occupation        Days_Indoors       Growing_Stress    
Quarantine_Frustrations Changes_Habits     Mental_Health_History
 Length:824         Length:824         Length:824         Length:824         Length:824        
Length:824              Length:824         Length:824        
 Class :character   Class :character   Class :character   Class :character   Class :character  
Class :character        Class :character   Class :character  
 Mode  :character   Mode  :character   Mode  :character   Mode  :character   Mode  :character  
Mode  :character        Mode  :character   Mode  :character  
 Weight_Change      Mood_Swings        Coping_Struggles   Work_Interest      Social_Weakness   
 Length:824         Length:824         Length:824         Length:824         Length:824        
 Class :character   Class :character   Class :character   Class :character   Class :character  
 Mode  :character   Mode  :character   Mode  :character   Mode  :character   Mode  :character  
```

#Checking for NA Values:

Hide

```
str(df)
```

```
'data.frame':   824 obs. of  13 variables:
 $ Age                    : chr  "20-25" "30-Above" "30-Above" "25-30" ...
 $ Gender                 : chr  "Female" "Male" "Female" "Male" ...
 $ Occupation             : chr  "Corporate" "Others" "Student" "Others" ...
 $ Days_Indoors           : chr  "1-14 days" "31-60 days" "Go out Every day" "1-14 days" ...
 $ Growing_Stress         : chr  "Yes" "Yes" "No" "Yes" ...
 $ Quarantine_Frustrations: chr  "Yes" "Yes" "No" "No" ...
 $ Changes_Habits         : chr  "No" "Maybe" "Yes" "Maybe" ...
 $ Mental_Health_History  : chr  "Yes" "No" "No" "No" ...
 $ Weight_Change          : chr  "Yes" "No" "No" "Maybe" ...
 $ Mood_Swings            : chr  "Medium" "High" "Medium" "Medium" ...
 $ Coping_Struggles       : chr  "No" "No" "Yes" "No" ...
 $ Work_Interest          : chr  "No" "No" "Maybe" "Maybe" ...
 $ Social_Weakness        : chr  "Yes" "Yes" "No" "Yes" ...
```

Hide

```
sum(is.na(df))
```

```
[1] 0
```

#Importing the ggplot and tidyverse libraries:

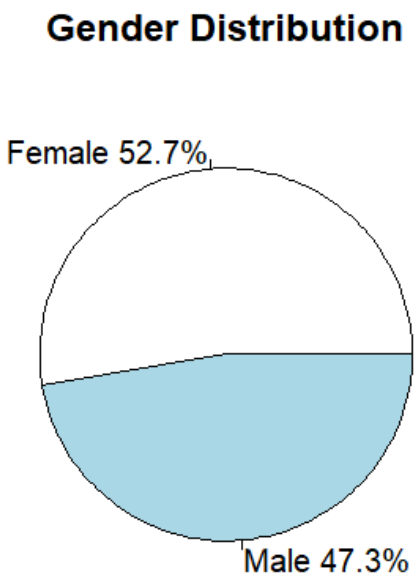
Hide

```
library(ggplot2,tidyverse)
```

#Gender distribution:

Hide

```
gender_counts <- table(df$Gender)
gender_percentages <- prop.table(gender_counts) * 100
pie(gender_counts, labels = paste(names(gender_counts), sprintf("%.1f%%", gender_percentage
s)), main = "Gender Distribution")
```

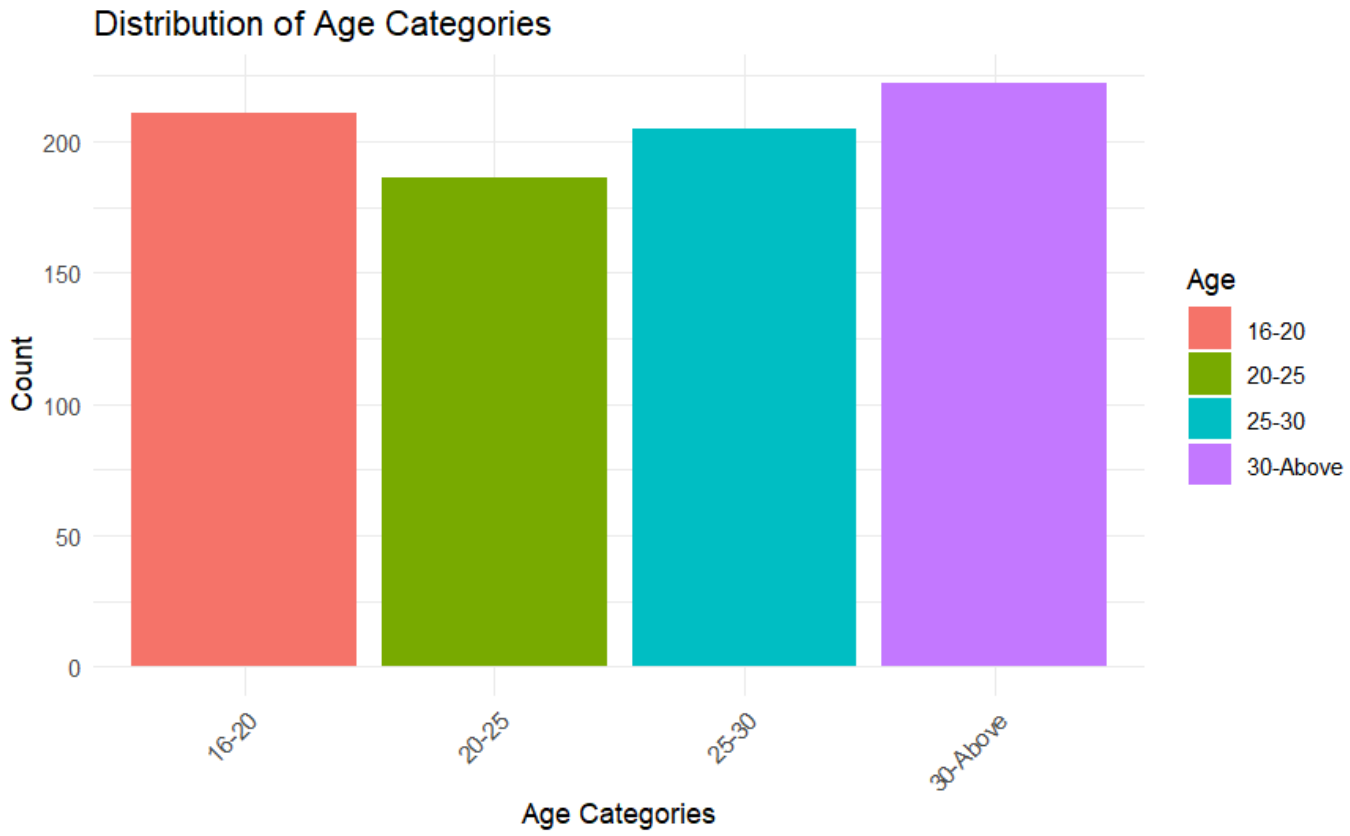


#*Observation*: The data set is more female oriented, with females taking up to 52.7% of the data set.

#Age Category distribution:

Hide

```
ggplot(df, aes(x = factor(Age, levels = c("16-20", "20-25", "25-30", "30-Above")), fill = Ag
e)) +
  geom_bar() +
  labs(title = "Distribution of Age Categories", x = "Age Categories", y = "Count") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



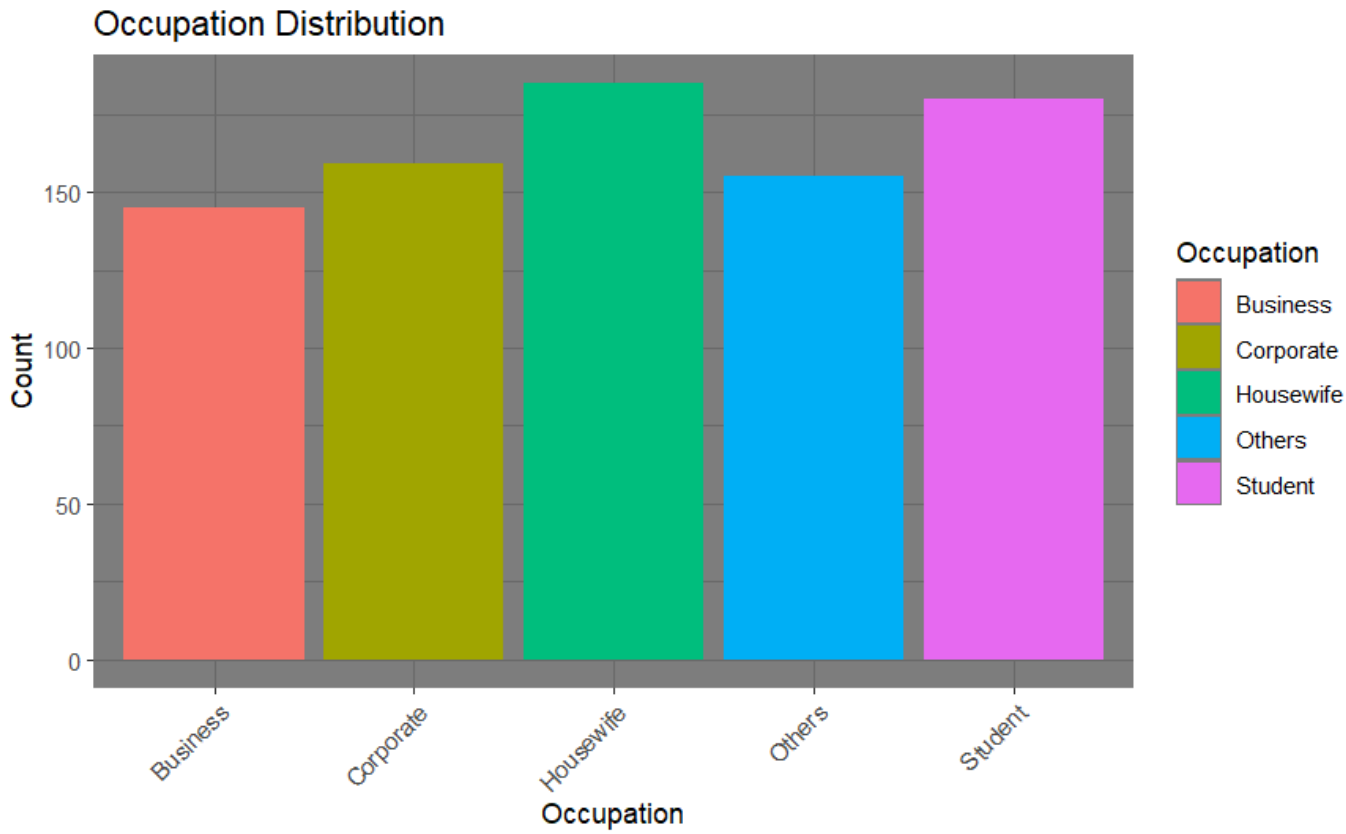
#*Observation*: The data set comprises of people aged 30 and above.

We can also observe that the focus is also on the youth, with age bracket 16-20 being involved heavily in the data mining/collection.

#Distribution by Occupation:

Hide

```
ggplot(df, aes(x = factor(Occupation, levels = c("Business","Corporate","Housewife","Other
s","Student")), fill = Occupation))+
  geom_bar()+
  labs(title = "Occupation Distribution", x = "Occupation", y = "Count")+
  theme_dark()+
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



#Observation:

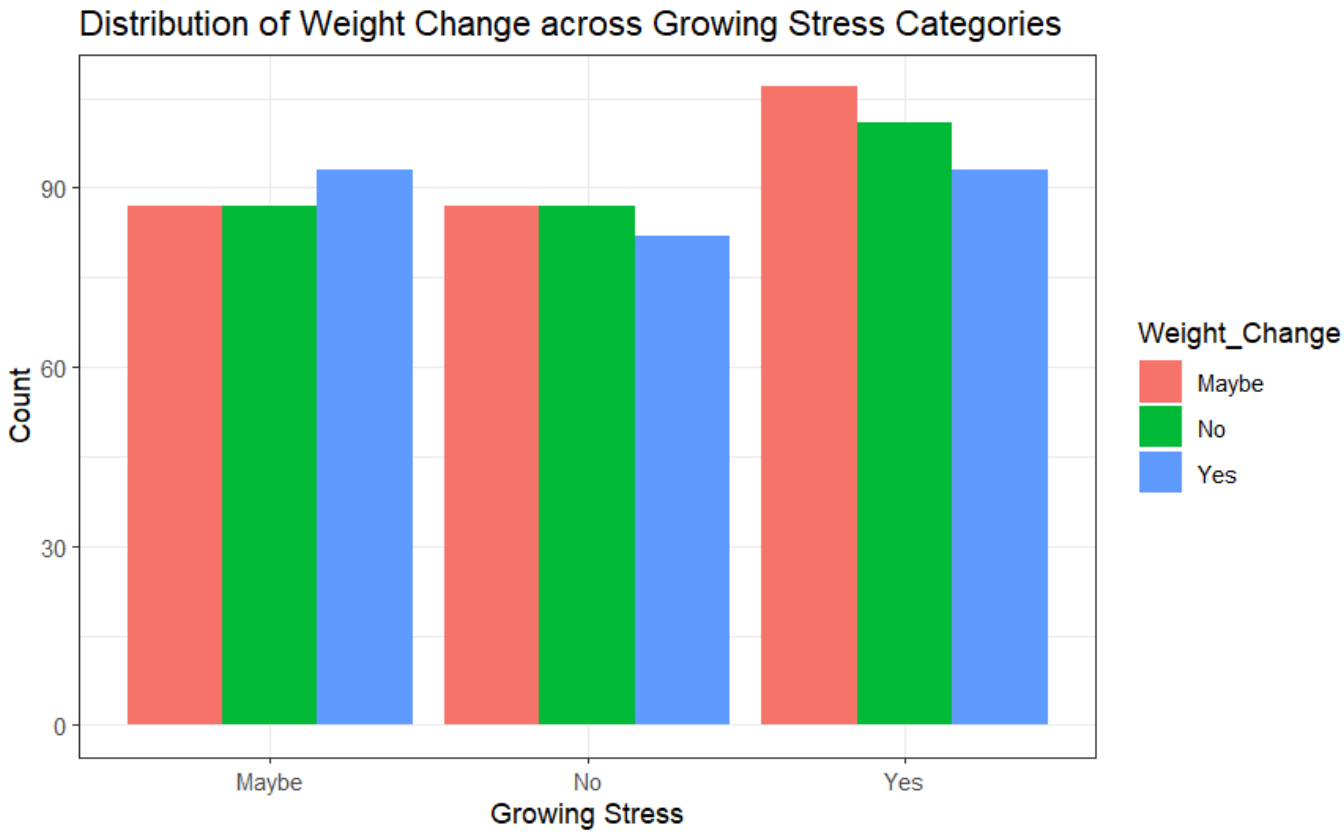
It can be observed that majority of the people of the data set are housewives.

It also includes students, as observed earlier that the data set also includes youth more.

#Visual relationship between stress and weight gain:

Hide

```
ggplot(df, aes(x = Growing_Stress, fill = Weight_Change))+
  geom_bar(position = "dodge")+
  labs(title = "Distribution of Weight Change across Growing Stress Categories", x = "Growing
Stress", y = "Count")+
  theme_bw()
```



#Observation: The weight change appears to be normally distributed around zero, with most data points falling within the range of -30 to 30.

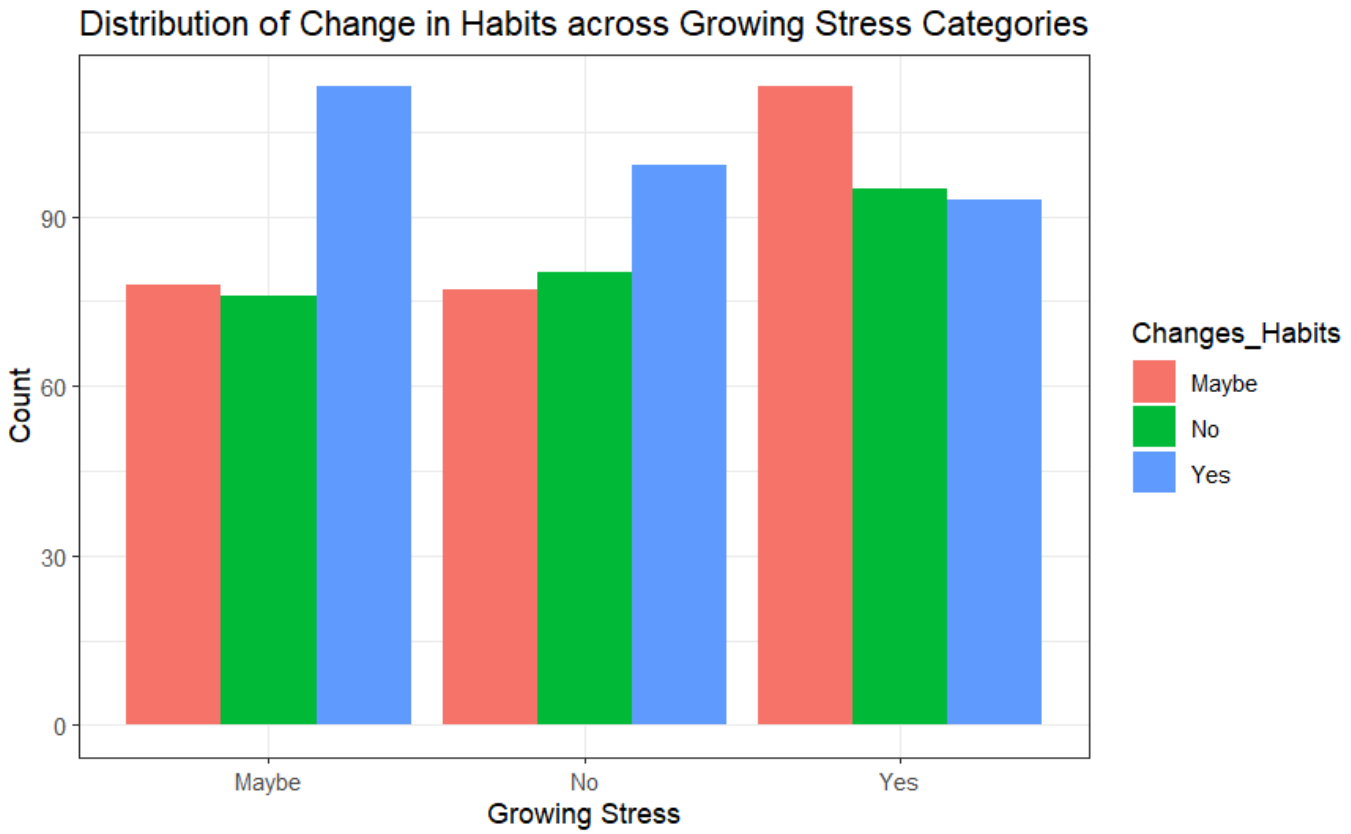
There is a small but noticeable skew towards negative weight change, meaning that there are slightly more data points showing weight loss than weight gain.

The distribution of weight change appears to be similar across all three growing stress categories (“Maybe”, “No”, and “Yes”).

#Visual representation of Habit changes in accordance with Growing Stress:

Hide

```
ggplot(df, aes(x = Growing_Stress, fill = Changes_Habits))+
  geom_bar(position = "dodge")+
  labs(title = "Distribution of Change in Habits across Growing Stress Categories", x = "Grow
ing Stress", y = "Count")+
  theme_bw()
```



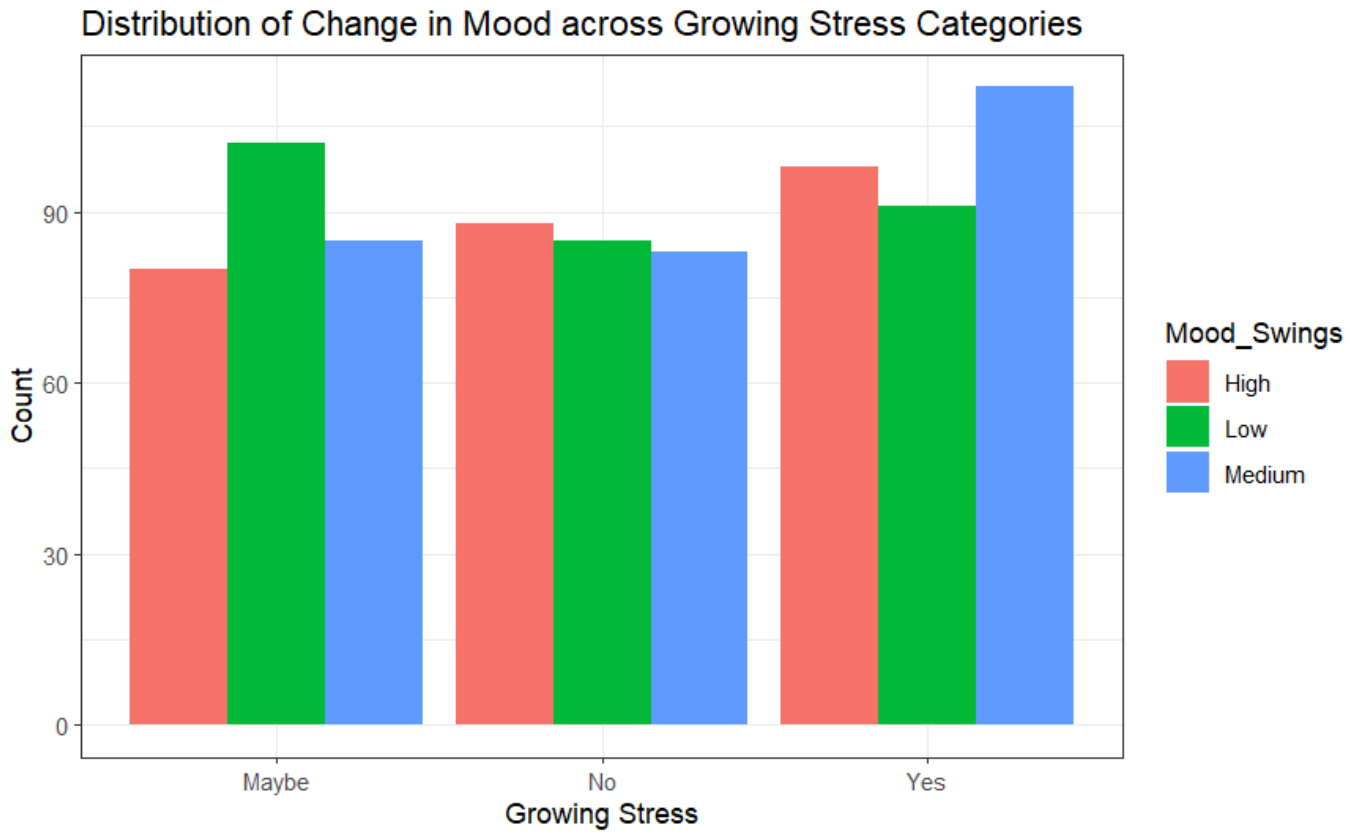
#Observations: The bar for “Maybe” in the “Yes” stress category is the tallest, which means that there are more data points where people reported significant changes in habits when they perceived high stress.

there is a relationship between stress and changes in habits. People who perceive more stress are more likely to report changes in their habits, and these changes are more likely to be significant.

#Visual representation of mood swings in accordance with growing stress:

Hide

```
ggplot(df, aes(x = Growing_Stress, fill = Mood_Swings))+
  geom_bar(position = "dodge")+
  labs(title = "Distribution of Change in Mood across Growing Stress Categories", x = "Growin
g Stress", y = "Count")+
  theme_bw()
```



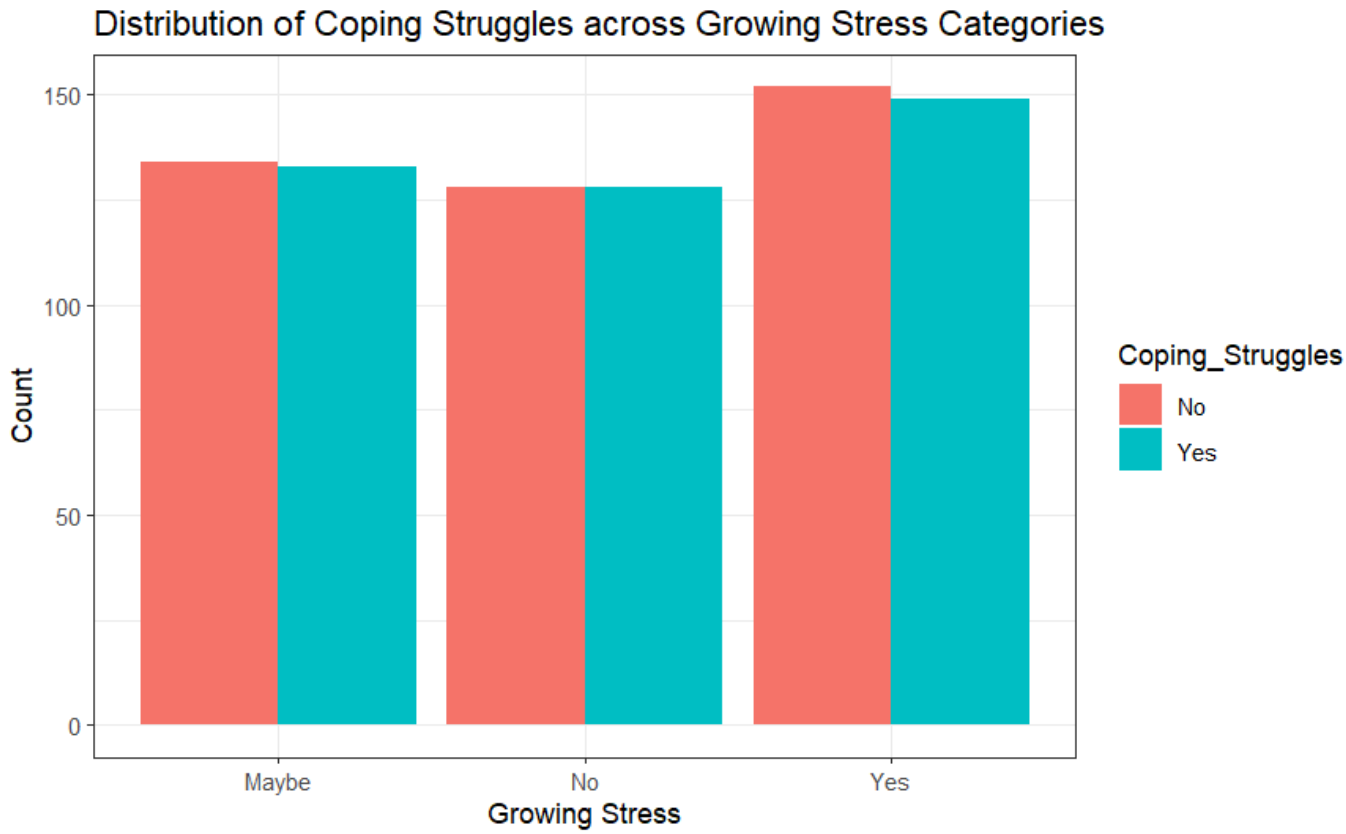
#Observations: The bar for “High” in “Yes” suggests that there are more people who reported feeling much more negative when they perceived high stress.

there is a relationship between growing stress and mood swings. People who perceive more stress are more likely to report feeling more negative moods.

#Distribution of Struggles across growing stress:

Hide

```
ggplot(df, aes(x = Growing_Stress, fill = Coping_Struggles))+
  geom_bar(position = "dodge")+
  labs(title = "Distribution of Coping Struggles across Growing Stress Categories", x = "Grow
ing Stress", y = "Count")+
  theme_bw()
```

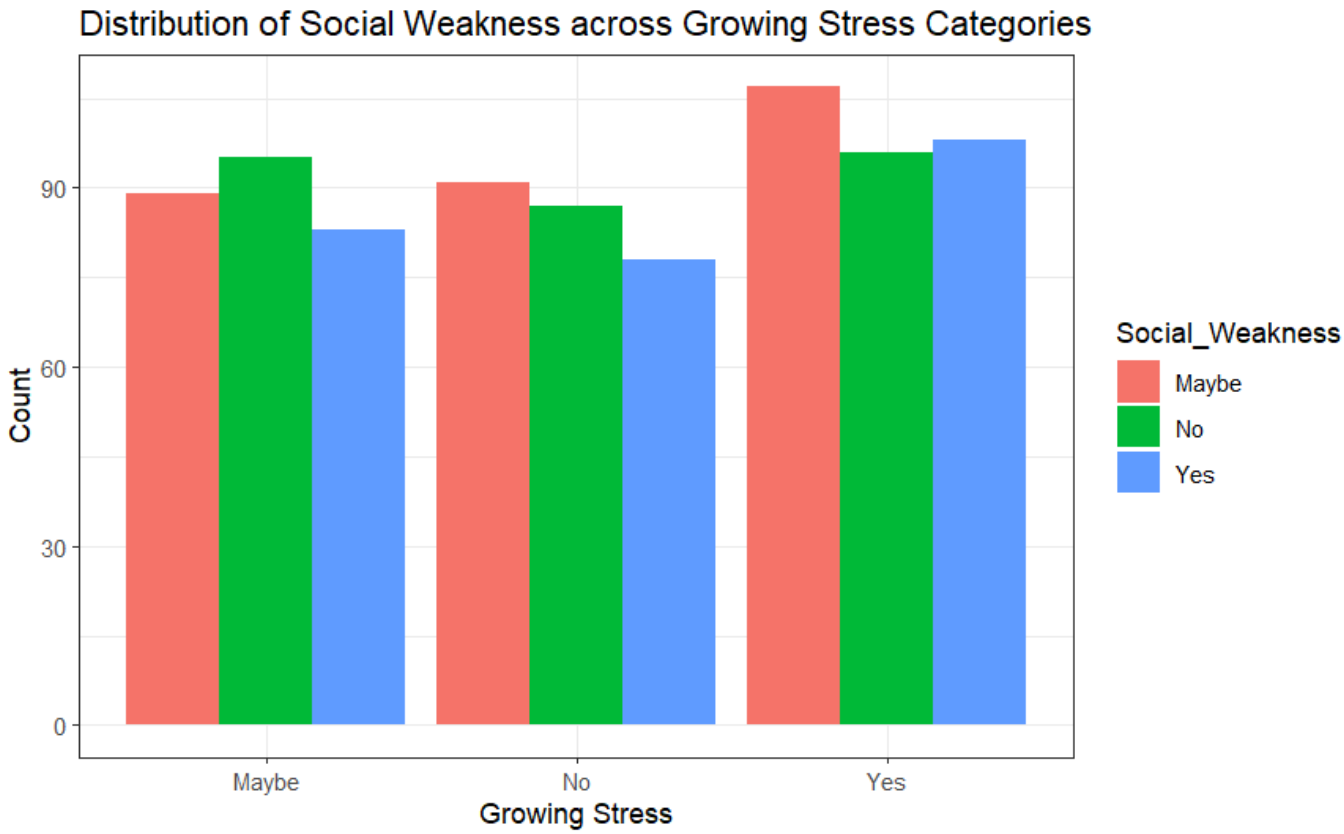


#Observations: There is a positive relationship between coping struggles and stress. People who perceive more stress are more likely to report experiencing coping struggles.

#Distribution of social weakness across growing stress:

Hide

```
ggplot(df, aes(x = Growing_Stress, fill = Social_Weakness))+
  geom_bar(position = "dodge")+
  labs(title = "Distribution of Social Weakness across Growing Stress Categories", x = "Growi
ng Stress", y = "Count")+
  theme_bw()
```

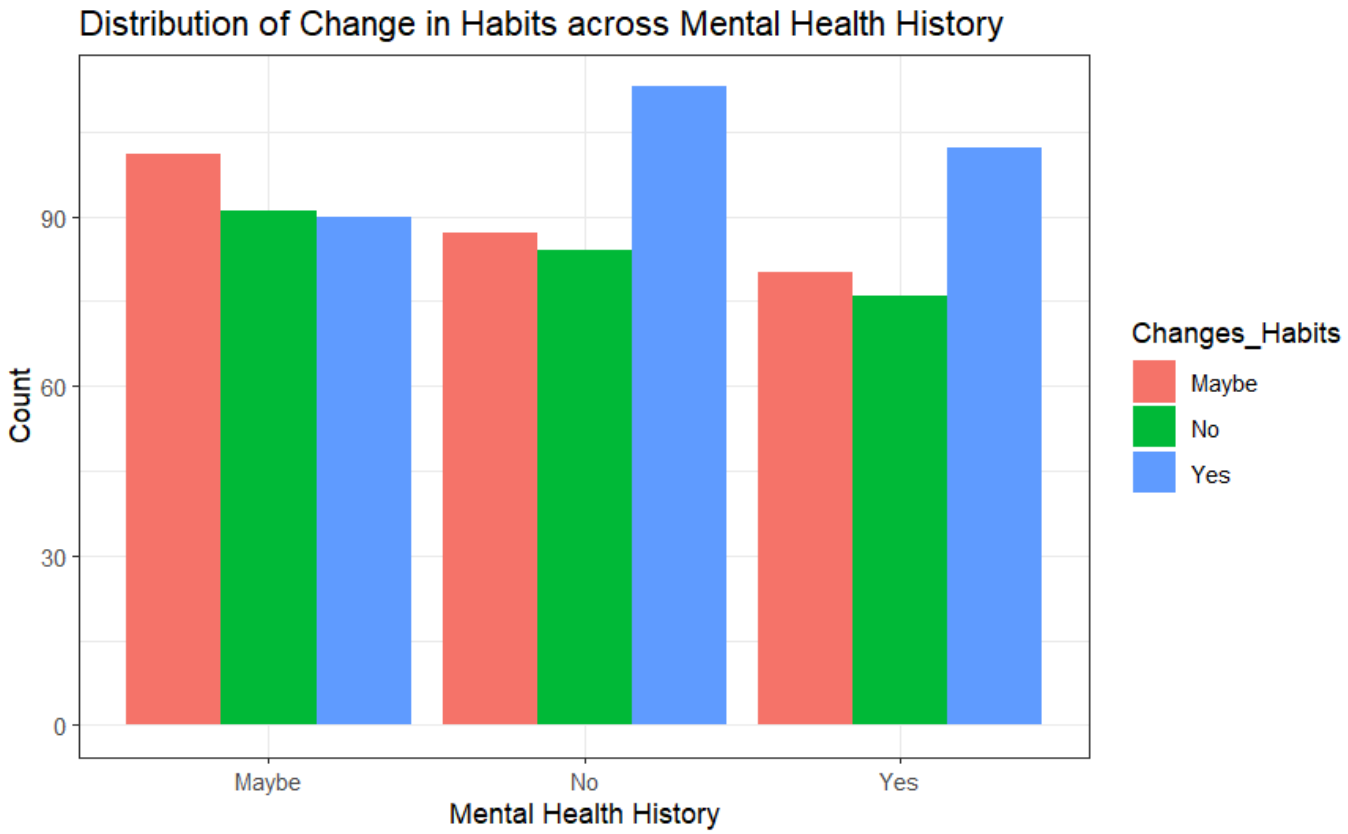


#Observations: There is a positive relationship between stress and social weakness. People who experience more stress tend to report feeling more socially weak.

#Distribution of change in habits across mental health history:

Hide

```
ggplot(df, aes(x = Mental_Health_History, fill = Changes_Habits))+
  geom_bar(position = "dodge")+
  labs(title = "Distribution of Change in Habits across Mental Health History", x = "Mental H
ealth History", y = "Count")+
  theme_bw()
```

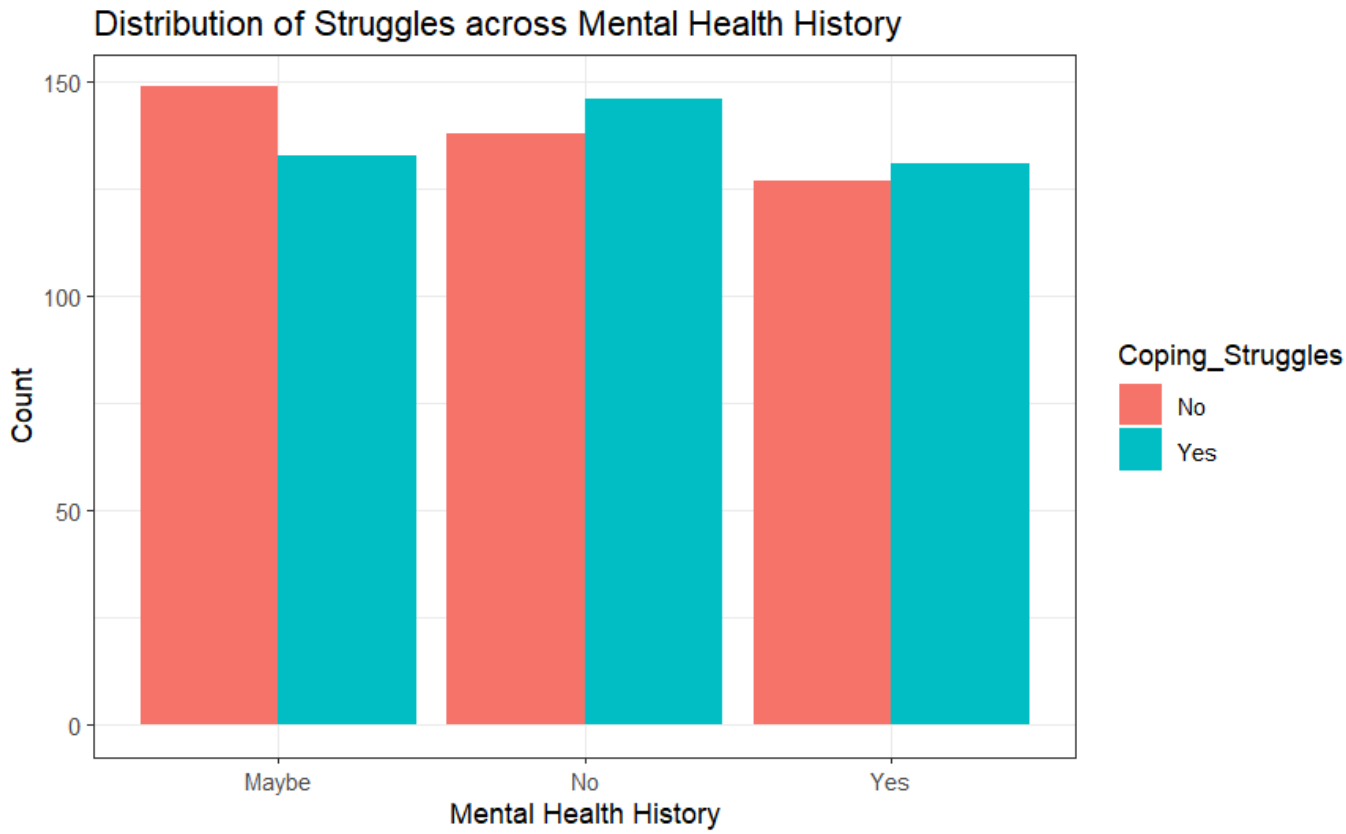


#Observations: The distribution of changes in habits appears to be similar across all three mental health history categories. There is no clear pattern to suggest that people with a history of mental health problems are more or less likely to report changes in their habits.The graph does not provide strong evidence for a relationship between mental health history and changes in habits.

#Distribution of struggles across mental health history:

Hide

```
ggplot(df, aes(x = Mental_Health_History, fill = Coping_Struggles))+
  geom_bar(position = "dodge")+
  labs(title = "Distribution of Struggles across Mental Health History", x = "Mental Health H
istory", y = "Count")+
  theme_bw()
```

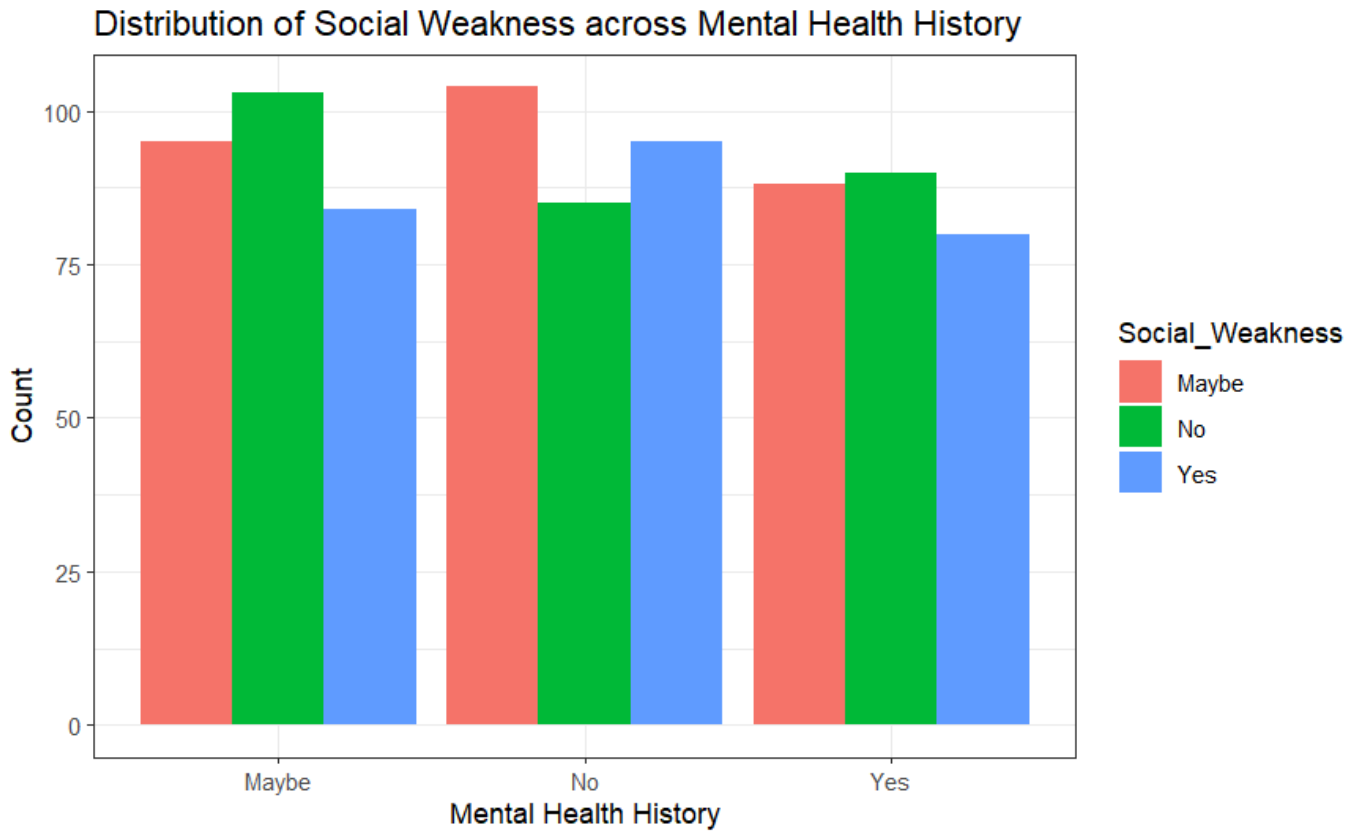


#Observations: There is a positive relationship between struggles and mental health issues. Someone with a history of mental health issues is struggling in daily life.

#Distribution of social weakness across mental health history:

Hide

```
ggplot(df, aes(x = Mental_Health_History, fill = Social_Weakness))+
  geom_bar(position = "dodge")+
  labs(title = "Distribution of Social Weakness across Mental Health History", x = "Mental He
alth History", y = "Count")+
  theme_bw()
```

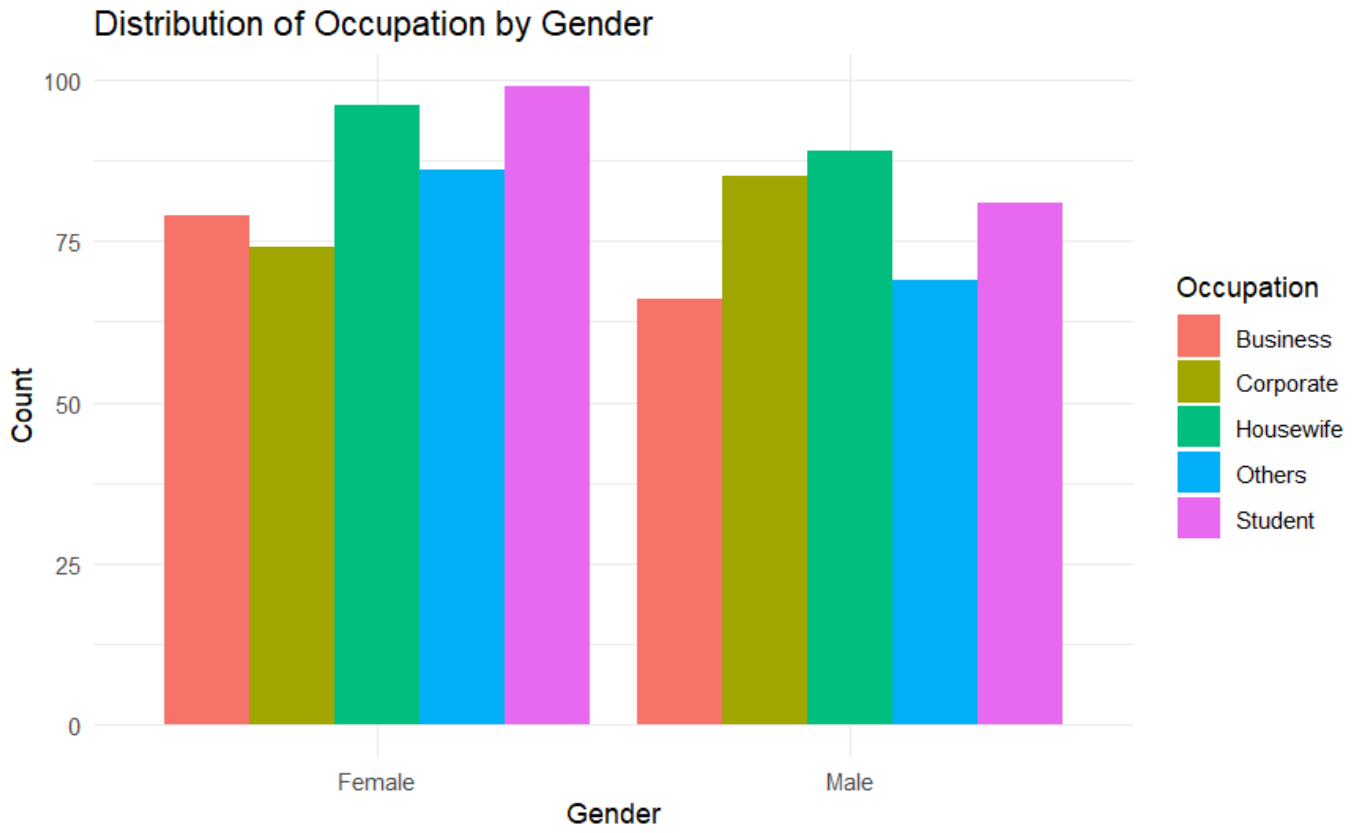


#Observations: There is a potential correlation between a history of mental health issues and increased social weakness. People with a histry of menatal health issues struggle socially and are weak in their social life.

#Distribution of occupation by gender:

Hide

```
ggplot(df, aes(x = Gender, fill = Occupation))+
  geom_bar(position = "dodge")+
  labs(title = "Distribution of Occupation by Gender",x = "Gender", y = "Count")+
  theme_minimal()
```



#Observations: Women - most of the women surveyed for this data set are students and housewives. Men - as compared to women, surprisingly, women are more into business than men. The data set shows that most of the men are into corporate jobs and also help at home with household.