Young people's excess drinking behavior: the significance of alcohol consumption in the everyday lives of 'ordinary' young adult drinkers.

What are the potential long-term consequences of excess drinking behavior among 'ordinary' young adult drinkers, including impacts on physical health, mental well-being, and social relationships?

17.09.2023

Università Cattolica del Sacro Cuore

Abstract: This study delves into the glamorous world of young people and their relationship with alcohol, examining the importance of alcohol consumption in their everyday lives. I explore the experiences and behaviors of "normal" young alcoholics, with the aim of illuminating the complexity of this important period in life. My survey of diverse respondents covered many factors, from the frequency and quantity of alcohol consumed to the motivations behind drinking. I explore the fascinating interaction between alcohol, peer influence and emotional well-being, and dissect the effects – both positive and negative – of this social lubricant on my life. youth. Among my findings are learning about the prevalence of binge drinking and its effects, the role of peer pressure in shaping drinking habits, and the nuances of alcohol's influence to social interactions and emotional states. Additionally, I study awareness levels about safe drinking guidelines and assess concerns about the health consequences of excessive alcohol consumption. By observing the lives of these young people, I seek to provide valuable insights that can inform interventions, public health strategies, and support systems tailored to their needs. needs of this demographic group. Ultimately, my research humanizes statistical data, painting a vivid picture of young people's experiences with alcohol and its profound significance in their daily lives.

Introduction: Excessive alcohol consumption among young people is a serious problem that requires further attention. This is an issue that goes beyond simple statistics; it has a profound impact on the lives of each individual and society as a whole. Youth is the time to explore, discover yourself and pursue your aspirations. However, drinking too much alcohol can disrupt these developmental years, affecting your health, relationships and future prospects. The importance of this study lies in its aim to understand the experiences of young "mainstream" drinkers and the impact of their binge drinking behavior. By exploring this complex issue, I hope to provide valuable information that can inform prevention measures, interventions, and support systems to enhance youth well-being.

To collect comprehensive data on young people's binge drinking behavior and its importance in their lives, I used a user-friendly and accessible survey platform. My survey was designed using Google Forms, a flexible and widely used tool for collecting structured data. Google Forms allowed us to create a questionnaire specifically for our target demographic: young adults. The survey was distributed through a variety of channels, reaching different groups of respondents. I ensured the anonymity and confidentiality of the survey to encourage honest responses. By leveraging Google Forms capabilities, I can efficiently collect significant amounts of data, forming the basis for our in-depth analysis.

During data collection and analysis, I leveraged the power of data science tools and modern programming languages. Python, a popular general-purpose programming language, was used as the basis for my data analysis. I used essential libraries like NumPy and Pandas for data manipulation and cleaning, allowing me to manage survey data effectively. Matplotlib and Seaborn, powerful data visualization libraries, allow me to present our results in meaningful tables and graphs. These technologies have played a vital role in converting raw survey data into valuable information. In later sections of the study, I will delve deeper into the specifics of the data collection methods and analysis techniques used.

Literature Review: In my research, I wanted to gain a deeper understanding of the complex dynamics of young people's binge drinking behavior and its significance in their everyday lives. Through a carefully designed survey, I collected information directly from young people aged 20-50+, examining the prevalence of binge drinking, its impact on health their physical and mental health and its impact on relationships, personal and academic or professional activities. This research not only provides a comprehensive understanding of the challenges facing young adults, but also sheds light on the underlying risk factors and levels of alcohol consumption that hinder the achievement of one's life goals. Using Python, with libraries such as numpy, pandas, matplotlib and seaborn, I will analyze survey data, uncovering important trends and correlations to inform evidence-based strategies and interventions for young people to make healthier lifestyle choices.

The research conducted by Quigley and Marlatt (1) underscores the pressing concerns surrounding young adult drinking patterns. Their study highlights that young adults exhibit higher rates of alcohol consumption and binge drinking compared to other age groups. Furthermore, they grapple with more adverse consequences of drinking, including alcohol abuse and dependence, often leading to (3) alarming statistics in alcohol-related traffic fatalities. Over time, a distinctive trend emerges, wherein many young adults eventually transition away from heavy drinking patterns as they assume greater responsibilities in adulthood.

In parallel, the study discussed in the International Journal of Drug Policy (2) explores the multifaceted reasons behind the decline in alcohol consumption among young people. Qualitative interviews with young adults in Sweden unearthed social mechanisms, including a shift in the cultural significance of drinking and reduced peer pressure. These findings point to the evolving landscape of young adult drinking behavior.

My research seeks to complement these existing studies by providing a fresh perspective on the issue. Employing a pragmatic approach, I intend to bridge the gap between the broader trends identified in previous research and the specific experiences of young adults. Through a comprehensive survey and data analysis using Python, I aim to offer a nuanced understanding of the significance of alcohol consumption in the lives of young adults and the potential interventions needed to promote healthier choices.

Methodology: The methodology used in this research project includes three basic elements: data acquisition, data preprocessing, and exploratory data analysis (EDA). This comprehensive approach was taken to gain an in-depth understanding of young people's binge drinking behavior and its positive or negative importance in their daily lives.

A. **Data Collection:** The first phase of this study involved the development of a well-designed questionnaire consisting of 20 essential questions. Leveraging the reach and reach of modern technology, Google Forms serves as the primary data collection tool. To ensure a diverse and representative sample, the survey was shared widely across various social media platforms. These efforts yielded an impressive data set that included responses from 172 people. Additionally, to enhance the effectiveness of the survey and ensure it was consistent with best practices, I sought valuable information from educational YouTube videos that provide essential guidance throughout the design and implementation of the survey. To access the Google Forms survey, please click on the following link: https://docs.google.com/forms/d/e/1FAIpQLSfd1k7SWEbzD19_be1JMT6nRKx7_-wjCs06ioAVnyVfzyVdNw/viewform. Additionally, you can find YouTube video resources used to develop the survey here: https://www.youtube.com/watch?v=3lQztUFEPrE,https://www.youtube.com/watch?v=Xf4wtlWOPnI,https://www.youtube.com/watch?v=dsfHhIngCG0,https://www.youtube.com/watch?v=HoZiRDIQyIY

B. **Data Preprocessing**: Managing data quality was a critical aspect of this research. Some survey questions yielded a substantial number of missing values, with certain columns exhibiting missing data rates exceeding 80%. To ensure the robustness of subsequent analysis, a strategic decision was made to eliminate these heavily affected columns. For the remaining columns with partial missing data, two imputation techniques were employed: the most frequent response and mean imputation. This meticulous data preprocessing approach aimed to enhance the dataset's overall integrity and facilitate meaningful analysis.

C. **Exploratory Data Analysis (EDA)**: The third phase of this research revolved around conducting exploratory data analysis (EDA). In this stage, I focused on selected columns deemed central to the research objectives. The dataset was segmented based on three critical demographic variables: Age Range (20-30, 30-40, 40-50, 50+), Gender (Male, Female), and Educational Level (High School, College, University Students). This segmentation allowed for a nuanced examination of how different subpopulations perceive and cope with the adverse consequences of alcohol consumption. In the subsequent Appendix section, I will provide the detailed Python code and methodological procedures employed for the exploratory data analysis. This section will encompass data visualization techniques, statistical tests, and other analytical tools to extract valuable insights and patterns from the dataset.

Results and Discussion: This important part of the research project examines the results of in-depth data analysis performed on a dataset collected through a Google Forms survey. This section begins by presenting the main findings, presented visually through a series of carefully constructed graphs and diagrams. These visual representations provide a clear and concise overview of survey data and reveal meaningful patterns and trends. A discussion then follows, providing in-depth analysis of the meaning and interpretation of these results.
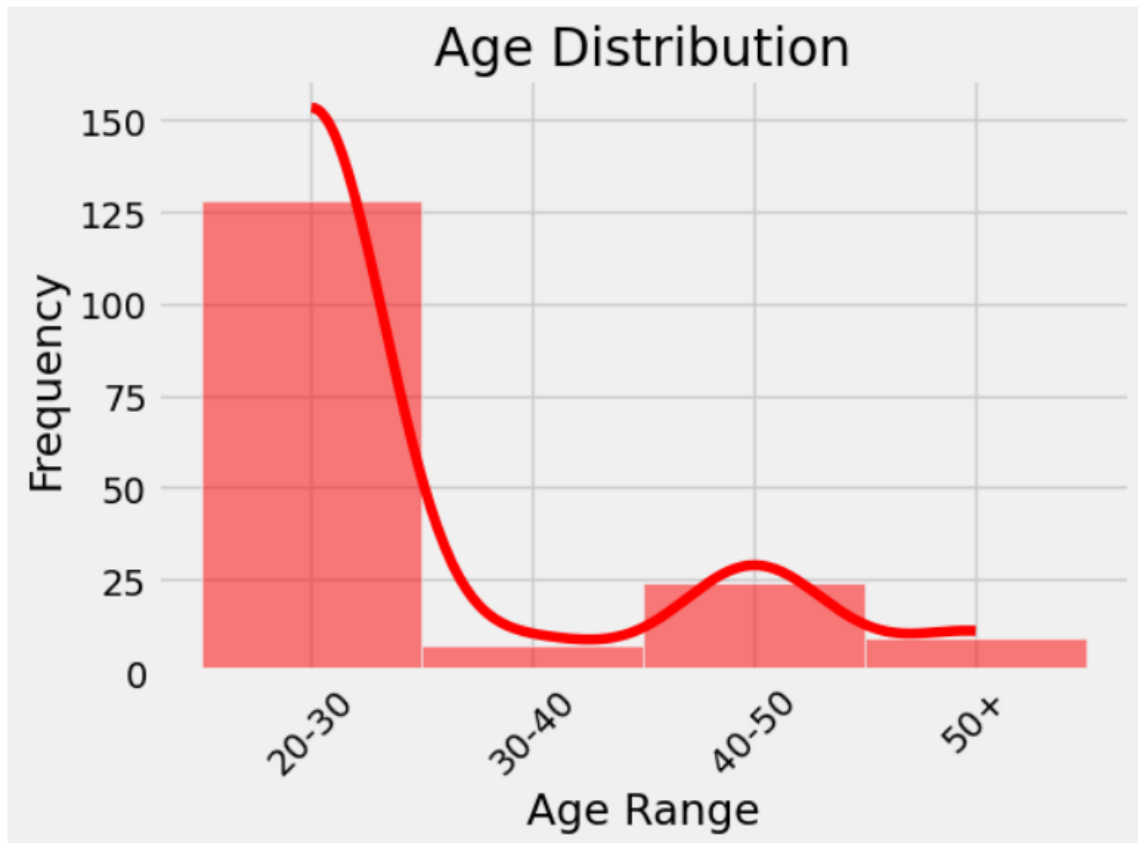

Figure 1: Age Distribution

Figure-1 provides an overview of the age distribution among the survey participants. The majority of respondents fall within the 20-30 age range, constituting the largest group with 128 individuals. Following this, there are 24 respondents aged between 40 and 50, 9 individuals aged 50 or above, and 7 participants between the ages of 30 and 40. A few respondents reported unique age ranges, such as '20-32,' '20-33,' '20-34,' and '20-35,' each represented by a single individual.
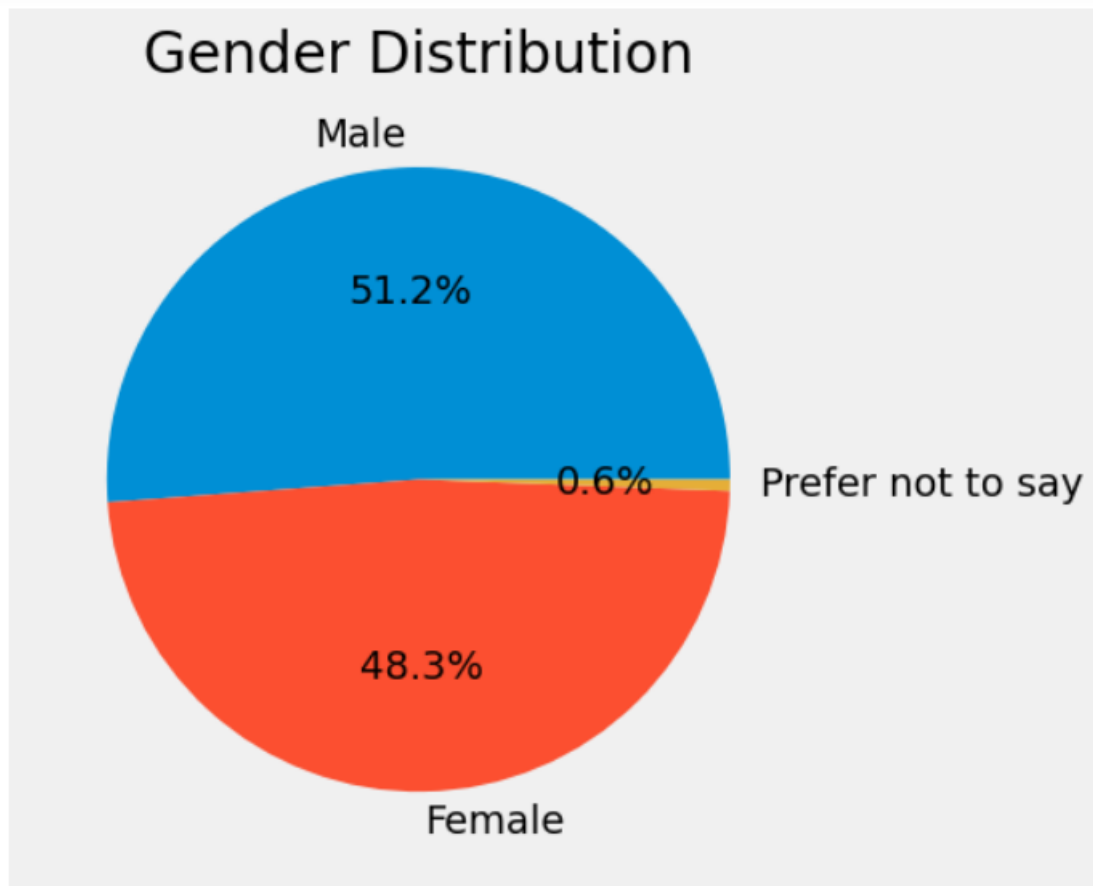
Figure-2: Gender Distribution

As depicted in Figure 2, the majority of survey participants identified as Male, making up approximately 51.8% of the total responses. Female participants accounted for a slightly lower percentage, comprising around 48.2% of the dataset. Additionally, a negligible percentage of participants (0.6%) chose not to specify their gender. Understanding the gender distribution is essential for analyzing how gender may influence excess drinking behavior in young adults. This demographic information provides valuable context for exploring potential gender-related patterns and disparities in alcohol consumption and its consequences.
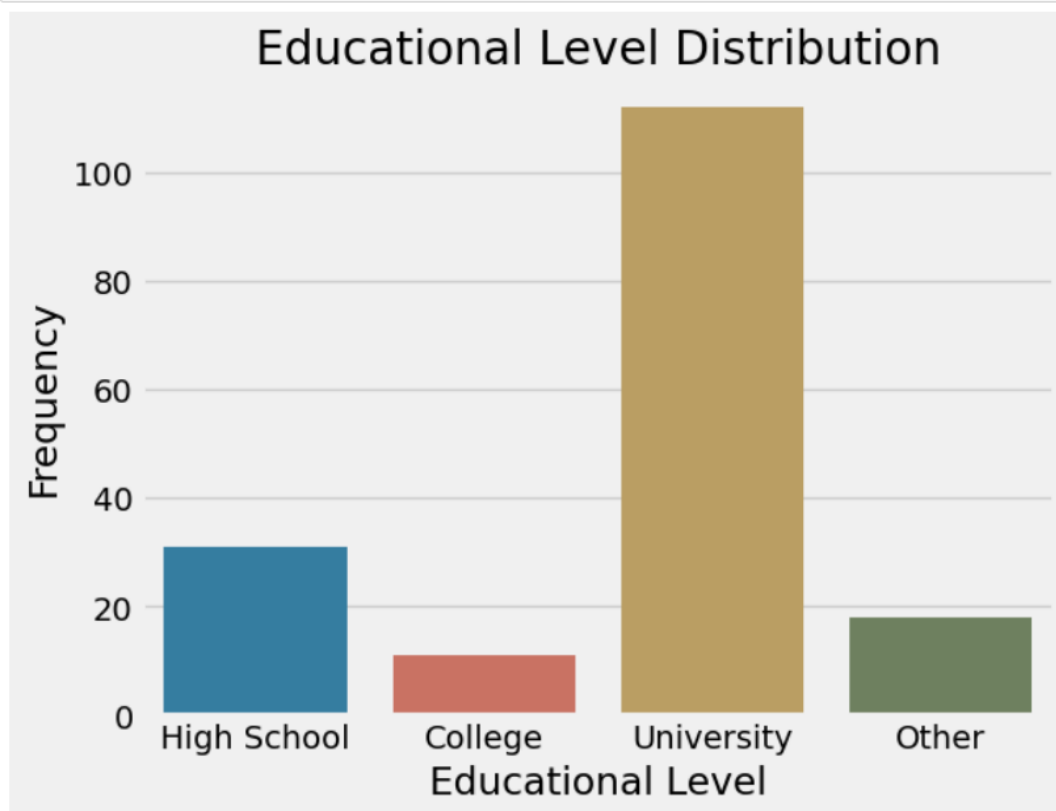
Figure-3: Educational Distribution

The survey participants' educational backgrounds were categorized into four groups (Figure-3): University, High School, Other, and College. The distribution of respondents across these educational levels provides insights into the participants' educational diversity and its potential impact on their excess drinking behavior. As indicated in Figure 3, the majority of survey participants (approximately 65.1%) reported being enrolled in a university. This suggests that a significant portion of the respondents were young adults pursuing higher education. High School accounted for the second-largest group, with approximately 18.0% of participants having completed or currently attending high school. A smaller percentage of participants, 10.4%, reported their educational level as Other, which might encompass various educational paths not explicitly categorized. Lastly, College was the educational level for approximately 6.3% of the respondents.
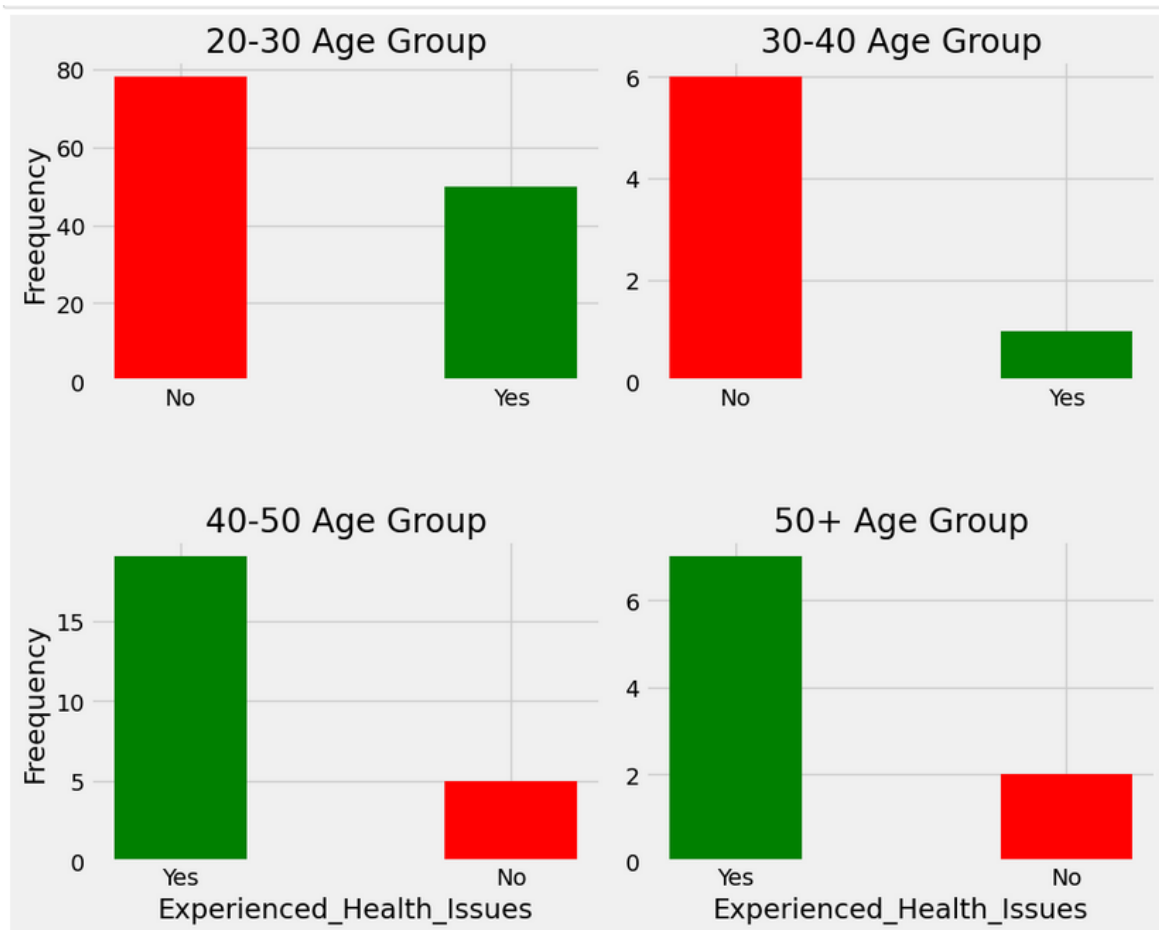
Figure-4: Experienced Health Issues Across Age Groups

I investigated the perception of health issues associated with alcohol consumption across various age groups shown in Figure-4. Among those aged 20-30, approximately 56.2% reported experiencing health problems related to alcohol, while 43.8% did not. In the 30-40 age group, the majority (85.7%) did not report any health issues, with only 14.3% indicating otherwise. In the 40-50 age bracket, about 82.6% acknowledged experiencing health problems, whereas 17.4% did not. Lastly, in the 50+ age group, roughly 77.8% reported health issues, with 22.2% not experiencing any. These findings highlight age-related variations in the perception of health issues linked to alcohol consumption
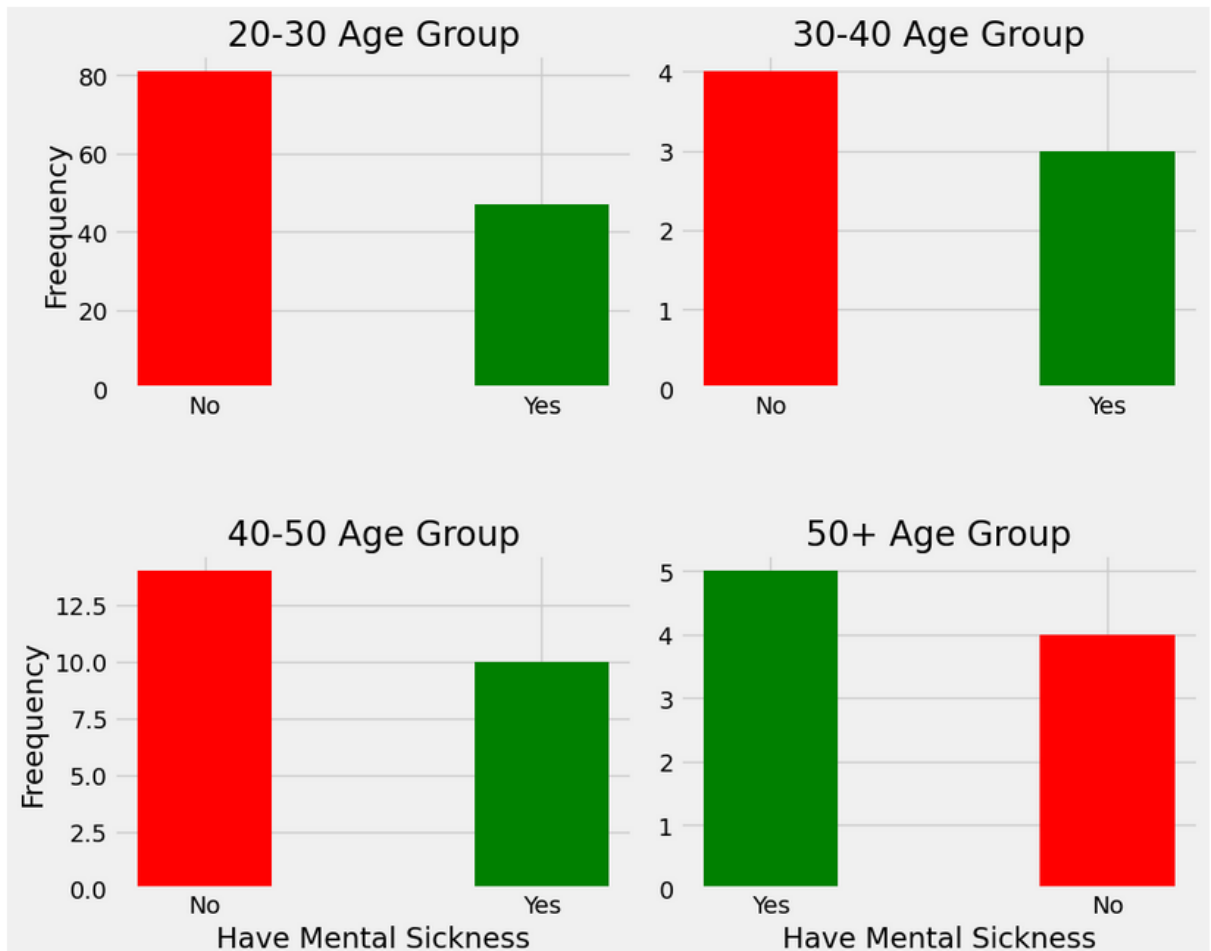
Figure-5: Have Mental Sickness Across Age Groups

I examined the prevalence of mental sickness across different age groups (Figure 5). In the 20-30 age group, 37.1% reported experiencing mental sickness, while 62.9% did not. Among those aged 30-40, 42.9% indicated having mental sickness, whereas 57.1% did not. In the 40-50 age category, 43.5% reported mental sickness, with 56.5% not experiencing it. For the 50+ age group, 55.6% acknowledged having mental sickness, while 44.4% did not. These results provide insights into the varying perceptions of mental health concerns among different age demographics.

Figure-6: Have Mental Sickness Across Gender Groups

Figure 6 displays the distribution of mental sickness among survey respondents categorized by gender. Among male participants, 63.6% reported not having experienced mental sickness, while 36.4% indicated that they had. On the othe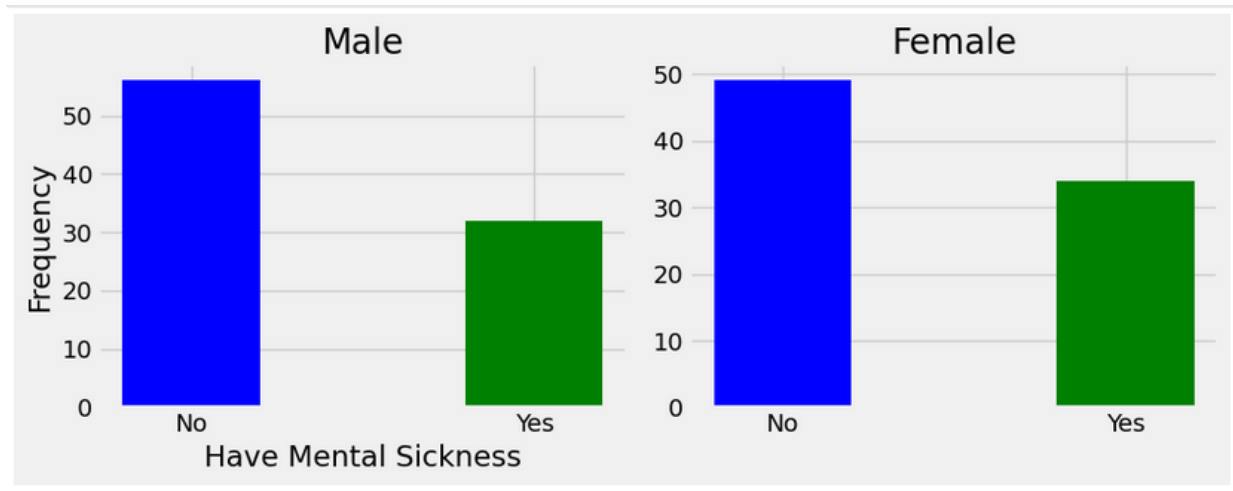r hand, among female participants, 59% reported no experience of mental sickness, while 41% stated they had experienced it. This breakdown by gender provides insights into the prevalence of mental sickness among different genders in the surveyed population.
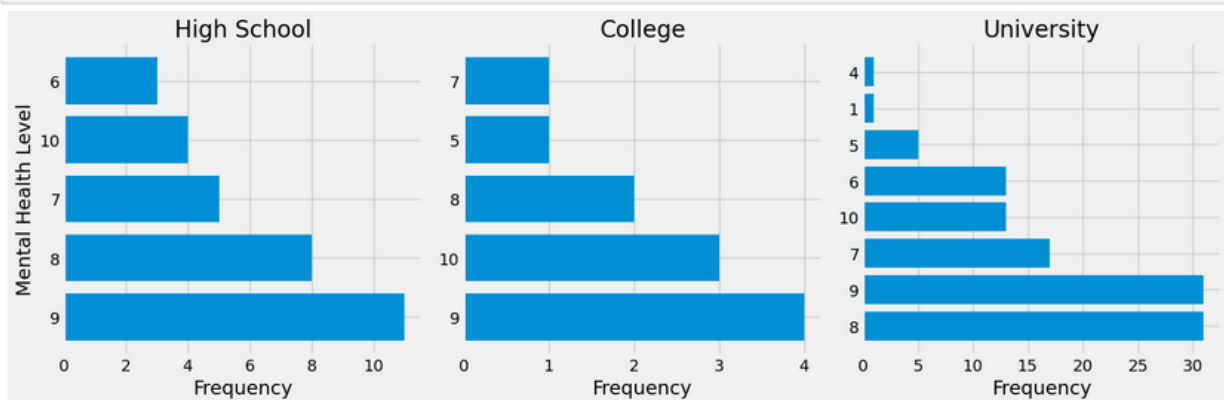
Figure-7: Mental Sickness Level Across Educational Level Groups

Figure 7 illustrates the distribution of mental health levels among respondents categorized by their educational levels. In the "High School" group, the majority of participants reported a mental health level of 9, followed by 8 and 7. Among "College" students, the most common mental health level was 9, followed by 10 and 8. In the "University" group, mental health levels 8 and 9 were the most prevalent, with several other levels also represented. This visualization provides insights into how different educational backgrounds may be associated with varying mental health levels among the survey participants.
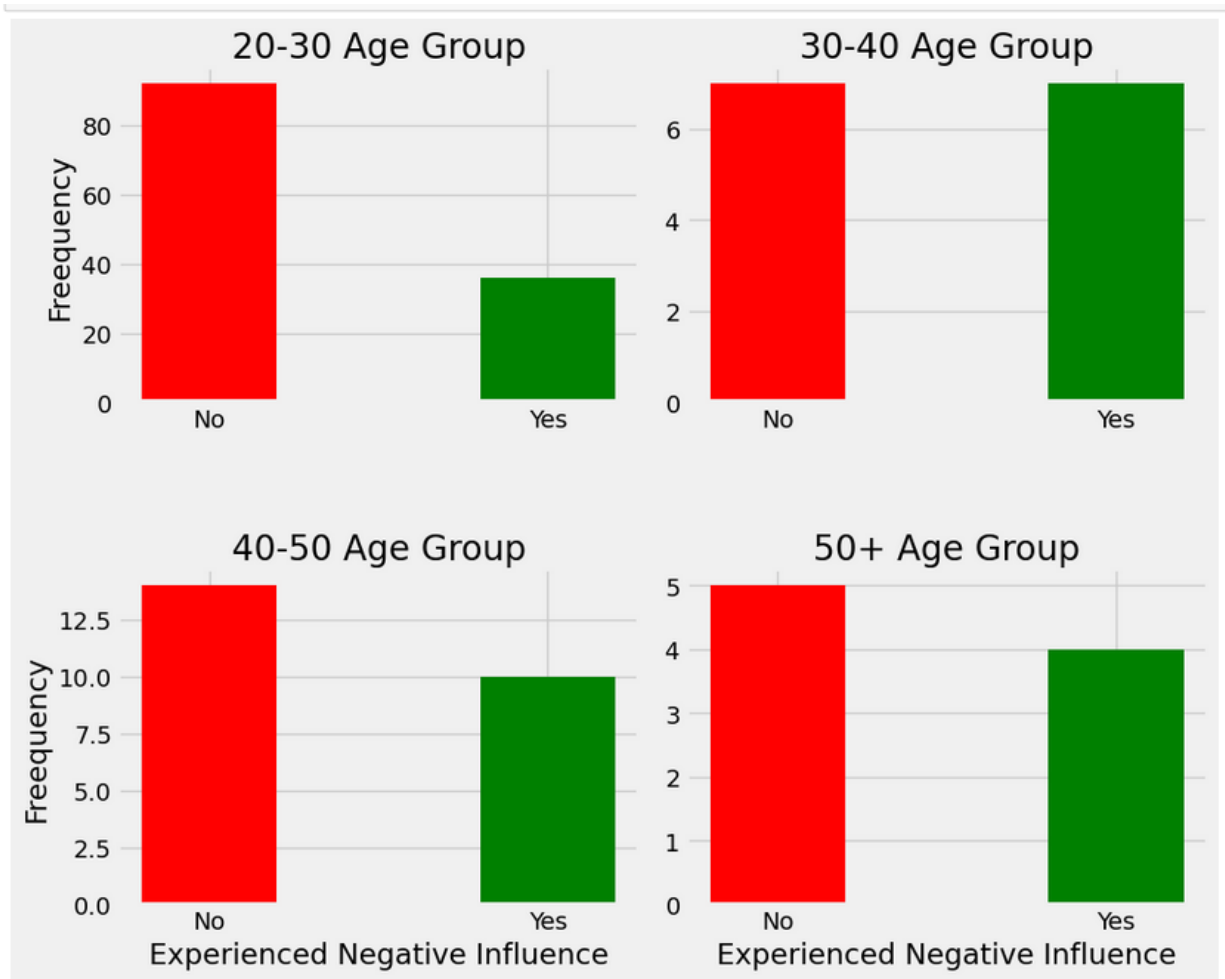
Figure-8: Experienced Negative Influence Across Age Groups

Figure 6 illustrates the impact of negative influences on academic and professional life across different age groups. Among those aged 20-30, 28.1% reported experiencing negative influences, while 71.9% did not. In the 30-40 age category, 100% indicated that they were not affected by negative influences. For the 40-50 age group, 41.7% reported experiencing negative influences, with 58.3% not affected. In the 50+ age group, 44.4% acknowledged negative influences, while 55.6% did not experience them. These findings provide insights into how different age demographics perceive the impact of negative influences on their academic and professional lives.
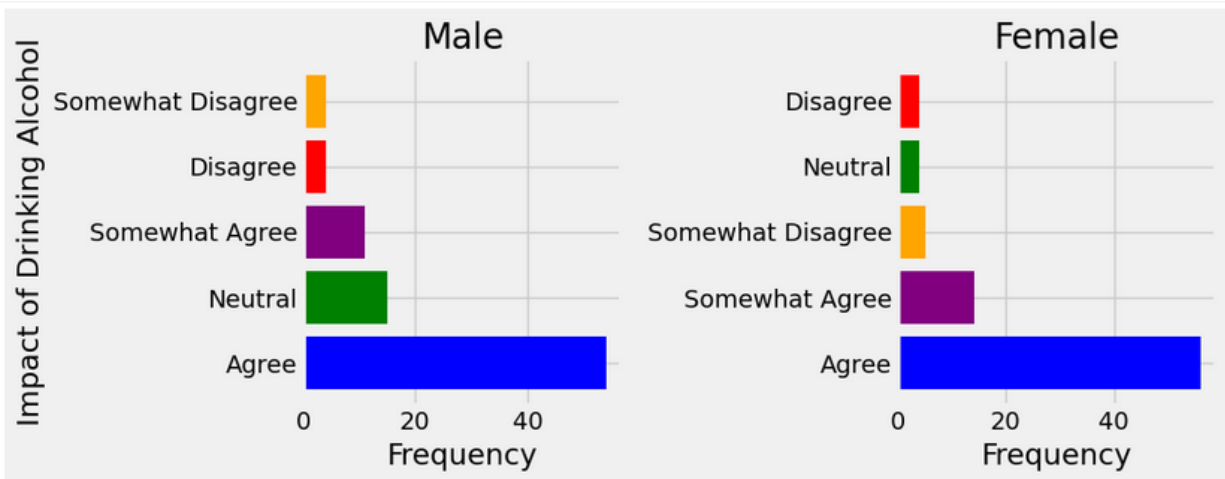
Figure -9: Impact of Drinking Alcohol Across Gender Group

Figure-9 displays the distribution of responses regarding the negative impact of alcohol consumption among male and female participants. In the "Male" group, the majority of respondents agreed that alcohol had a negative impact, while a smaller portion had neutral or differing opinions. Conversely, in the "Female" group, most participants agreed that alcohol had a negative impact, with fewer neutral and differing opinions. This visualization highlights the gender-based variations in perceptions of the negative effects of alcohol consumption among the survey participants.

Data analysis provides valuable insights into the relationship between age, gender, education level, mental health and alcohol consumption. The majority of survey participants were in their 20s and 30s, reflecting a majority of young people. Both men and women expressed concerns about the negative effects of alcohol, challenging gender stereotypes. Education level also plays an important role, with college students reporting more mental health problems and high school students reporting more alcohol-related negative effects. These differences highlight the need for appropriate interventions. The relationship between mental health and alcohol use is complex and tends to vary across age groups. Young people, especially those in their 20s and 30s, seem to be the most affected in terms of their academic and professional lives, possibly due to career difficulties.

In summary, this research shows the complex links between demographics, mental health and alcohol consumption, challenging stereotypes. Tailored interventions are needed, and additional research is needed to refine strategies and promote healthier drinking behaviors across different demographic groups.

Conclusion: In summary, this study sheds light on the complex interactions between demographic factors, mental health, and drinking habits in young adults. These findings highlight the need for nuanced approaches to addressing alcohol-related problems that challenge conventional assumptions. The data highlight that age, gender, and education level significantly influence alcohol-related perceptions and experiences. Young people, especially those in their 20s and 30s, seem to be especially vulnerable to the negative effects of alcohol, which can affect their academic and professional lives. The study also dismantled traditional gender stereotypes by revealing that both men and women expressed concerns about the negative consequences of alcohol. To effectively address these complex problems, appropriate interventions are required. This research serves as a basis for designing targeted programs that address the unique challenges faced by different demographic groups. It highlights the importance of promoting responsible consumption practices, supporting mental health and encouraging healthier choices among young people.

Looking forward, future research could delve deeper into the specific factors that influence alcohol use among young adults, exploring the role of peer pressure, social networks, and identity change. gender form. Additionally, longitudinal studies can provide insight into how drinking behavior changes over time, contributing to the development of more comprehensive intervention strategies. Ultimately, this research paves the way for a more comprehensive and informed approach to addressing alcohol-related concerns in youth.

References:

1. **Exploring the links between unhealthy eating behaviour and heavy alcohol use in the social, emotional and cultural lives of young adults (aged 18–25): A qualitative research study (https://www.sciencedirect.com/science/article/pii/S0195666318313904)**
2. **2.Why are young people drinking less than earlier? Identifying and specifying social mechanisms with a pragmatist approach (https://www.sciencedirect.com/science/article/pii/S0955395918303025)**
3. **Drinking Among Young Adults (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6876515/)**

Appendix:

```
In [1]: # IMPORTING NECESSARY LIBRRARIES
        import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        import seaborn as sns
        import math
```

```
In [2]: # FATCHING THE SURVEY RESPONSE DATA AND QUESTION SCHEMA
        df = pd.read_csv('Data/Survey - Form responsess.csv')
        schema_df = pd.read_csv('Data/Question_Schema.csv')
```

```
In [3]: df.head(10)
```

Out[3]:

	Age_Range	Gender	Educational_Level	Experienced_Health_Issues	Health_Issues_Type	Hangover_Experience	Mental_Sickness	Mental_Sickness_Level
0	20-30	Male	High School	No	NaN	5.0	No	NaN
1	20-30	Male	College	No	NaN	2.0	No	NaN
2	20-30	Female	University	No	NaN	8.0	Yes	7.0
3	20-30	Female	University	No	NaN	8.0	Yes	7.0
4	20-30	Female	University	No	NaN	8.0	Yes	7.0
5	20-30	Male	College	No	NaN	3.0	No	2.0
6	20-30	Female	College	Yes	hangover	9.0	Yes	5.0
7	20-30	Female	University	Yes	Hangover	7.0	No	NaN
8	20-30	Female	University	No	NaN	8.0	No	1.0
9	20-30	Female	College	No	NaN	1.0	No	1.0

```
In [4]: pd.set_option('display.max_colwidth', None)
        schema_df.head(20)
```

Out[4]:

	Question ID	Full_Question	Short_Form
0	1	Select the range of your age	Age_Range
1	2	Your gender	Gender
2	3	Select your educational level	Educational_Level
3	4	Have you experienced any health issues as a result of your excess drinking behavior?	Experienced_Health_Issues
4	5	If yes, please specify the health issues you have encountered due to excess drinking.	Health_Issues_Type
5	6	How frequently do you experience hangovers after a night of excessive drinking?	Hangover_Experience
6	7	Have you noticed any changes in your mental well-being, such as increased anxiety or depression, related to your excess drinking behavior?	Mental_Sickness
7	8	If yes, please describe how much excess drinking has affected your mental well-being.	Mental_Sickness_Level
8	9	Have your personal relationships, including friendships and romantic relationships, been impacted by your excess drinking behavior?	Relationship_Impact
9	10	If yes, please indicate how much your social relationships have been affected.	Relationship_Impact_Level
10	11	Do you believe that your academic or professional life has been negatively influenced by your excess drinking?	Negative_Influenced_Academic_Professional
11	12	Have you ever searched for medical treatment or counseling for alcohol-related issues?	Medical_Treatment
12	13	Have you tried to cut down or control your alcohol consumption because of concerns about its long-term consequences?	Control_CutDown
13	14	How would you rate your overall physical health on a scale of 1 to 10?	Physical _Health_Level
14	15	How would you rate your mental well-being on a scale of 1 to 10?	Mental_Health_Level
15	16	Have you ever participated in alcohol rehabilitation or support programs?	Participated_Alcohol_Rehabilitation
16	17	Are you aware of the potential risk factors associated with long-term excess drinking, such as liver disease or addiction?	Potential_Risk_Factor
17	18	Do you feel that excess drinking has hindered your ability to achieve your life goals and aspirations?	Achieve_Life_Goals
18	19	Have you lost friends or strained family relationships due to your drinking habits?	Friendship_Lost
19	20	Indicate the extent to which you believe alcohol has negative impact on the following aspects [Physical health]	Negative_Impact_of_Alcohol

DATA PRE-PROCESSING

```
In [5]: df.shape
```

```
Out[5]: (172, 20)
```

```
In [6]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 172 entries, 0 to 171
Data columns (total 20 columns):
 #   Column                                  Non-Null Count  Dtype
---  ------                                  --------------  -----
 0   Age_Range                               172 non-null    object
 1   Gender                                  172 non-null    object
 2   Educational_Level                       172 non-null    object
 3   Experienced_Health_Issues              170 non-null    object
 4   Health_Issues_Type                      81 non-null     object
 5   Hangover_Experience                     169 non-null    float64
 6   Mental_Sickness                         170 non-null    object
 7   Mental_Sickness_Level                   84 non-null     float64
 8   Relationship_Impact                     170 non-null    object
 9   Relationship_Impact_Level               86 non-null     float64
 10  Negative_Influenced_Academic_Professional  171 non-null  object
 11  Medical_Treatment                       171 non-null    object
 12  Control_CutDown                         169 non-null    object
 13  Physical _Health_Level                  171 non-null    float64
 14  Mental_Health_Level                     172 non-null    int64
 15  Participated_Alcohol_Rehabilitation    171 non-null    object
 16  Potential_Risk_Factor                   172 non-null    object
 17  Achieve_Life_Goals                      170 non-null    object
 18  Friendship_Lost                         169 non-null    object
 19  Negative_Impact_of_Alcohol              170 non-null    object
dtypes: float64(4), int64(1), object(15)
memory usage: 27.0+ KB
```

```
In [7]: df.describe()
```

Out[7]:

	Hangover_Experience	Mental_Sickness_Level	Relationship_Impact_Level	Physical _Health_Level	Mental_Health_Level
count	169.000000	84.000000	86.000000	171.000000	172.000000
mean	5.366864	5.964286	6.244186	8.128655	7.976744
std	2.718165	2.590280	2.877646	1.348571	1.442681
min	1.000000	1.000000	1.000000	1.000000	1.000000
25%	3.000000	4.000000	4.000000	8.000000	7.000000
50%	6.000000	7.000000	7.000000	8.000000	8.000000
75%	8.000000	8.000000	8.000000	9.000000	9.000000
max	10.000000	10.000000	10.000000	10.000000	10.000000

```
In [8]: # CHECKING MISSING VALUES
        df.isna().sum()
```

```
Out[8]: Age_Range                                   0
        Gender                                      0
        Educational_Level                           0
        Experienced_Health_Issues                   2
        Health_Issues_Type                         91
        Hangover_Experience                         3
        Mental_Sickness                             2
        Mental_Sickness_Level                      88
        Relationship_Impact                         2
        Relationship_Impact_Level                  86
        Negative_Influenced_Academic_Professional   1
        Medical_Treatment                           1
        Control_CutDown                             3
        Physical _Health_Level                      1
        Mental_Health_Level                         0
        Participated_Alcohol_Rehabilitation         1
        Potential_Risk_Factor                       0
        Achieve_Life_Goals                          2
        Friendship_Lost                             3
        Negative_Impact_of_Alcohol                  2
        dtype: int64
```

```
In [9]: # DROPPING COLUMNS WHICH HAS MORE THAN 80 MISSING VALUES
        df.drop(columns = {'Health_Issues_Type','Mental_Sickness_Level','Relationship_Impact_Level'},inplace = True)
```

FILLING THE MISSING VALUES WITH MEAN VALUE/FREQUENCEY

```
In [10]: df['Experienced_Health_Issues'].value_counts()
```

```
Out[10]: No      92
         Yes     78
         Name: Experienced_Health_Issues, dtype: int64
```

```
In [11]: # FILLING THE MISSING VALUE WITH MOST FREQUENT ANSWER
         df['Experienced_Health_Issues'].fillna('No',inplace = True)
```

```
In [12]: df['Hangover_Experience'].value_counts()
```

```
Out[12]: 8.0     25
         7.0     21
         2.0     20
         6.0     19
         5.0     18
         1.0     17
         9.0     15
         3.0     14
         4.0     13
         10.0     7
         Name: Hangover_Experience, dtype: int64
```

```
In [13]: # FILLING THE MISSING VALUE WITH MEAN VALUE
         df['Hangover_Experience'].fillna(int(df['Hangover_Experience'].mean()),inplace = True)
```

```
In [14]: df['Mental_Sickness'].value_counts()
```

```
Out[14]: No     104
         Yes     66
         Name: Mental_Sickness, dtype: int64
```

```
In [15]: # FILLING THE MISSING VALUE WITH MOST FREQUENT ANSWER
         df['Mental_Sickness'].fillna('No',inplace = True)
```

```
In [16]: df['Relationship_Impact'].value_counts()
```

```
Out[16]: No     108
         Yes     62
         Name: Relationship_Impact, dtype: int64
```

```
In [17]: # FILLING THE MISSING VALUE WITH MOST FREQUENT ANSWER
         df['Relationship_Impact'].fillna('No',inplace = True)
```

```
In [18]: df['Negative_Influenced_Academic_Professional'].value_counts()
```

```
Out[18]: No     121
         Yes     50
         Name: Negative_Influenced_Academic_Professional, dtype: int64
```

```
In [19]: # FILLING THE MISSING VALUE WITH MOST FREQUENT ANSWER
         df['Negative_Influenced_Academic_Professional'].fillna('No',inplace = True)
```

```
In [20]: df['Medical_Treatment'].value_counts()
```

```
Out[20]: No     150
         Yes     21
         Name: Medical_Treatment, dtype: int64
```

```
In [21]: # FILLING THE MISSING VALUE WITH MOST FREQUENT ANSWER
         df['Medical_Treatment'].fillna('No',inplace = True)
```

```
In [22]: df['Control_CutDown'].value_counts()
```

```
Out[22]: No     123
         Yes     46
         Name: Control_CutDown, dtype: int64
```

```
In [23]: # FILLING THE MISSING VALUE WITH MOST FREQUENT ANSWER
         df['Control_CutDown'].fillna('No',inplace = True)
```

```
In [24]: df['Physical _Health_Level']
```

```
Out[24]: 0       8.0
         1       5.0
         2       8.0
         3       8.0
         4       8.0
                 ...
         167    10.0
         168     NaN
         169     9.0
         170     6.0
         171     7.0
         Name: Physical _Health_Level, Length: 172, dtype: float64
```

```
In [25]: # FILLING THE MISSING VALUE WITH MEAN VALUE
         df['Physical _Health_Level'].fillna(int(df['Physical _Health_Level'].mean()),inplace = True)
```

```
In [26]: df['Participated_Alcohol_Rehabilitation'].value_counts()
```

```
Out[26]: No     166
         Yes      5
         Name: Participated_Alcohol_Rehabilitation, dtype: int64
```

```
In [27]: # FILLING THE MISSING VALUE WITH MOST FREQUENT ANSWER
         df['Participated_Alcohol_Rehabilitation'].fillna('No',inplace = True)
```

```
In [28]: df['Achieve_Life_Goals'].value_counts()
```

```
Out[28]: No     128
         Yes     42
         Name: Achieve_Life_Goals, dtype: int64
```

```
In [29]: # FILLING THE MISSING VALUE WITH MOST FREQUENT ANSWER
         df['Achieve_Life_Goals'].fillna('No',inplace = True)
```

```
In [30]: df['Friendship_Lost'].value_counts()
```

```
Out[30]: No     151
         Yes     18
         Name: Friendship_Lost, dtype: int64
```

```
In [31]: # FILLING THE MISSING VALUE WITH MOST FREQUENT ANSWER
         df['Friendship_Lost'].fillna('No',inplace = True)
```

```
In [32]: df['Negative_Impact_of_Alcohol'].value_counts()
```

```
Out[32]: Agree              109
         Somewhat Agree      25
         Neutral             19
         Somewhat Disagree    9
         Disagree             8
         Name: Negative_Impact_of_Alcohol, dtype: int64
```

```
In [33]: # FILLING THE MISSING VALUE WITH MOST FREQUENT ANSWER
         df['Negative_Impact_of_Alcohol'].fillna('Agree',inplace = True)
```

```
In [34]: # CHECKING AGAIN IF ANY MISSING VALUES
         df.isna().sum()
```

```
Out[34]: Age_Range                                0
         Gender                                   0
         Educational_Level                        0
         Experienced_Health_Issues               0
         Hangover_Experience                      0
         Mental_Sickness                          0
         Relationship_Impact                      0
         Negative_Influenced_Academic_Professional 0
         Medical_Treatment                        0
         Control_CutDown                          0
         Physical _Health_Level                   0
         Mental_Health_Level                      0
         Participated_Alcohol_Rehabilitation       0
         Potential_Risk_Factor                    0
         Achieve_Life_Goals                       0
         Friendship_Lost                          0
         Negative_Impact_of_Alcohol               0
         dtype: int64
```

EXPLORATORY DATA ANALYSIS

In [35]:
```python
# SETTING THE GRAPH STYLE
plt.style.use('fivethirtyeight')
```

In [36]:
```python
# CREATING a CATAGORICAL VARIABLE FOR BINS ORDER
bin_order = ['10-20', '20-30', '30-40', '40-50', '50+']
df['Age_Range'] = pd.Categorical(df['Age_Range'], categories=bin_order, ordered=True)

# SETTING THE WIDTH
bar_width = 0.2

# CREATING HISTOGRAM PLOT
sns.histplot(df['Age_Range'], bins=bin_order, kde=True, color='red', discrete=True)
plt.title("Age Distribution")
plt.xlabel("Age Range")
plt.ylabel("Frequency")
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```
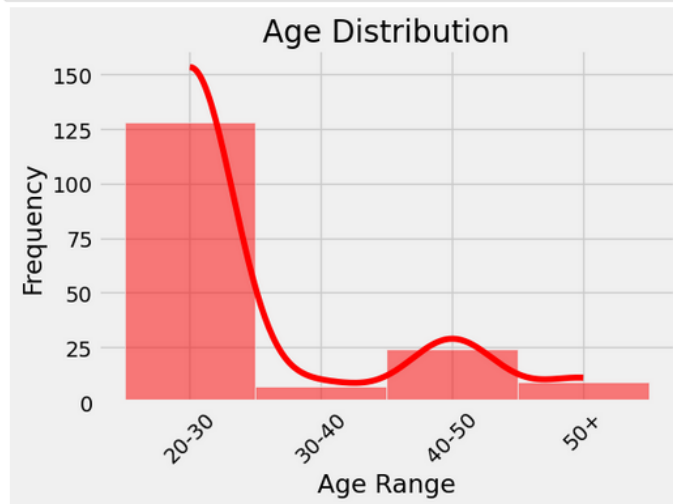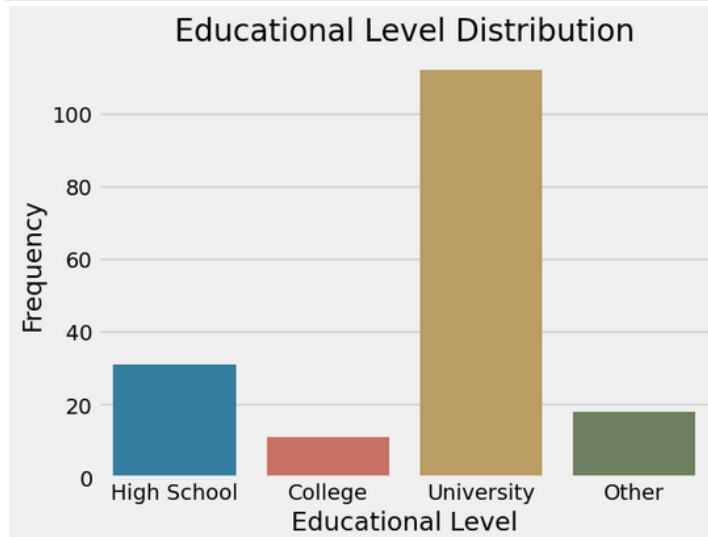

In [37]:
```python
plt.title('Gender Distribution')
gender_count = df['Gender'].value_counts()
plt.pie(gender_count, labels = gender_count.index, autopct = '%.1f%%')
plt.show()
```
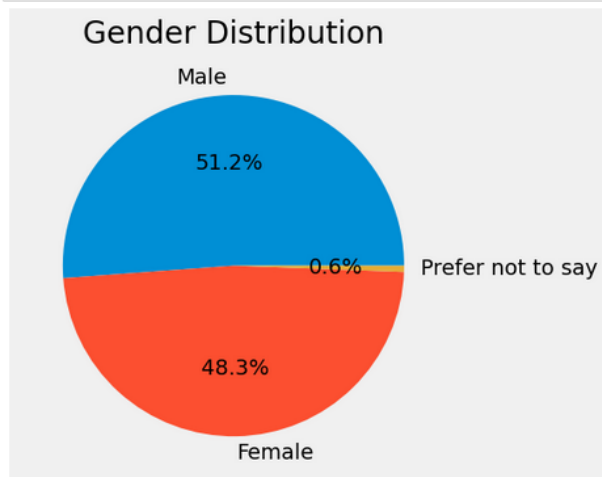
```
In [38]:  plt.title('Educational Level Distribution')
          sns.countplot(x= 'Educational_Level', data= df,saturation = .5)
          plt.xlabel("Educational Level")
          plt.ylabel("Frequency")
          plt.show()
```

Educational Level Distribution


```
In [39]:  # CREATING GROUP BASED ON AGE_RANGE
          age_grp = df.groupby(['Age_Range'])
```

HOW DIFFERENT AGE RANGE PERSONS EXPERIENCED HEALTH ISSUES

```
In [40]:  # CALCULATING VALUE COUNTS FOR EACH AGE GROUP
          count_20_30 = age_grp.get_group('20-30')['Experienced_Health_Issues'].value_counts()
          count_30_40 = age_grp.get_group('30-40')['Experienced_Health_Issues'].value_counts()
          count_40_50 = age_grp.get_group('40-50')['Experienced_Health_Issues'].value_counts()
          count_50_plus = age_grp.get_group('50+')['Experienced_Health_Issues'].value_counts()

          print(count_20_30)
          print(count_30_40)
          print(count_40_50)
          print(count_50_plus)

          No     78
          Yes    50
          Name: Experienced_Health_Issues, dtype: int64
          No     6
          Yes    1
          Name: Experienced_Health_Issues, dtype: int64
          Yes    19
          No      5
          Name: Experienced_Health_Issues, dtype: int64
          Yes    7
          No     2
          Name: Experienced_Health_Issues, dtype: int64
```

```
In [41]:  # ADJUSTING THE FIG SIZE AND BAR CHART WIDTH
          fig, axes = plt.subplots(2, 2, figsize=(10, 8))
          bar_width = 0.4

          # DEFINING COLOR FOR THE BAR
          colors = ['red', 'green']
          colors_rev = ['green', 'red']

          # SUBPLOT-1
          axes[0, 0].bar(['No', 'Yes'], count_20_30, width=bar_width, color = colors)
          axes[0, 0].set_title('20-30 Age Group')
          axes[0, 0].set_ylabel('Freequency')

          # SUBPLOT 2
          axes[0, 1].bar(['No', 'Yes'], count_30_40, width=bar_width, color = colors)
          axes[0, 1].set_title('30-40 Age Group')

          # SUBPLOT 3
          axes[1, 0].bar(['Yes', 'No'], count_40_50, width=bar_width, color = colors_rev)
          axes[1, 0].set_title('40-50 Age Group')
          axes[1, 0].set_xlabel('Experienced_Health_Issues')
          axes[1, 0].set_ylabel('Freequency')

          # SUBPLOT 4
          axes[1, 1].bar(['Yes', 'No'], count_50_plus, width=bar_width, color = colors_rev)
          axes[1, 1].set_title('50+ Age Group')
          axes[1, 1].set_xlabel('Experienced_Health_Issues')

          # ADJUSTING SPACE
          plt.tight_layout()
          plt.subplots_adjust(hspace=0.5)
```
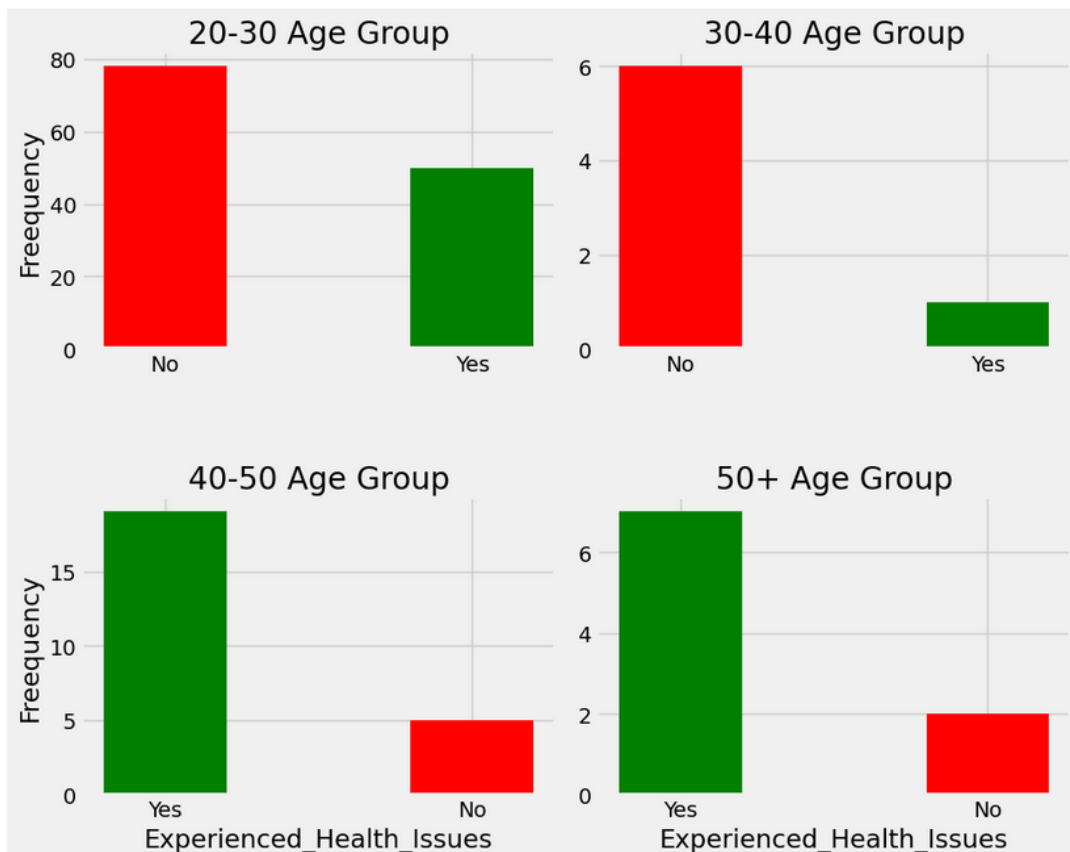
MENTAL SICKNESS OF DIFFERENT AGE RANGE PERSONS

In [42]:
```python
# CALCULATING VALUE COUNTS FOR EACH AGE GROUP
count_20_30 = age_grp.get_group('20-30')['Mental_Sickness'].value_counts()
count_30_40 = age_grp.get_group('30-40')['Mental_Sickness'].value_counts()
count_40_50 = age_grp.get_group('40-50')['Mental_Sickness'].value_counts()
count_50_plus = age_grp.get_group('50+')['Mental_Sickness'].value_counts()

print(count_20_30)
print(count_30_40)
print(count_40_50)
print(count_50_plus)
```

```
No     81
Yes    47
Name: Mental_Sickness, dtype: int64
No     4
Yes    3
Name: Mental_Sickness, dtype: int64
No     14
Yes    10
Name: Mental_Sickness, dtype: int64
Yes    5
No     4
Name: Mental_Sickness, dtype: int64
```

In [43]:
```python
# ADJUSTING THE FIG SIZE AND BAR CHART WIDTH
fig, axes = plt.subplots(2, 2, figsize=(10, 8))
bar_width = 0.4

# DEFINING COLOR FOR THE BAR
colors = ['red', 'green']
colors_rev = ['green', 'red']

# SUBPLOT-1
axes[0, 0].bar(['No', 'Yes'], count_20_30, width=bar_width, color = colors)
axes[0, 0].set_title('20-30 Age Group')
axes[0, 0].set_ylabel('Freequency')

# SUBPLOT 2
axes[0, 1].bar(['No', 'Yes'], count_30_40, width=bar_width, color = colors)
axes[0, 1].set_title('30-40 Age Group')

# SUBPLOT 3
axes[1, 0].bar(['No', 'Yes'], count_40_50, width=bar_width, color = colors)
axes[1, 0].set_title('40-50 Age Group')
axes[1, 0].set_xlabel('Have Mental Sickness')
axes[1, 0].set_ylabel('Freequency')

# SUBPLOT 4
axes[1, 1].bar(['Yes', 'No'], count_50_plus, width=bar_width, color = colors_rev)
axes[1, 1].set_title('50+ Age Group')
axes[1, 1].set_xlabel('Have Mental Sickness')

# ADJUSTING SPACE
plt.tight_layout()
plt.subplots_adjust(hspace=0.5)

plt.show()
```
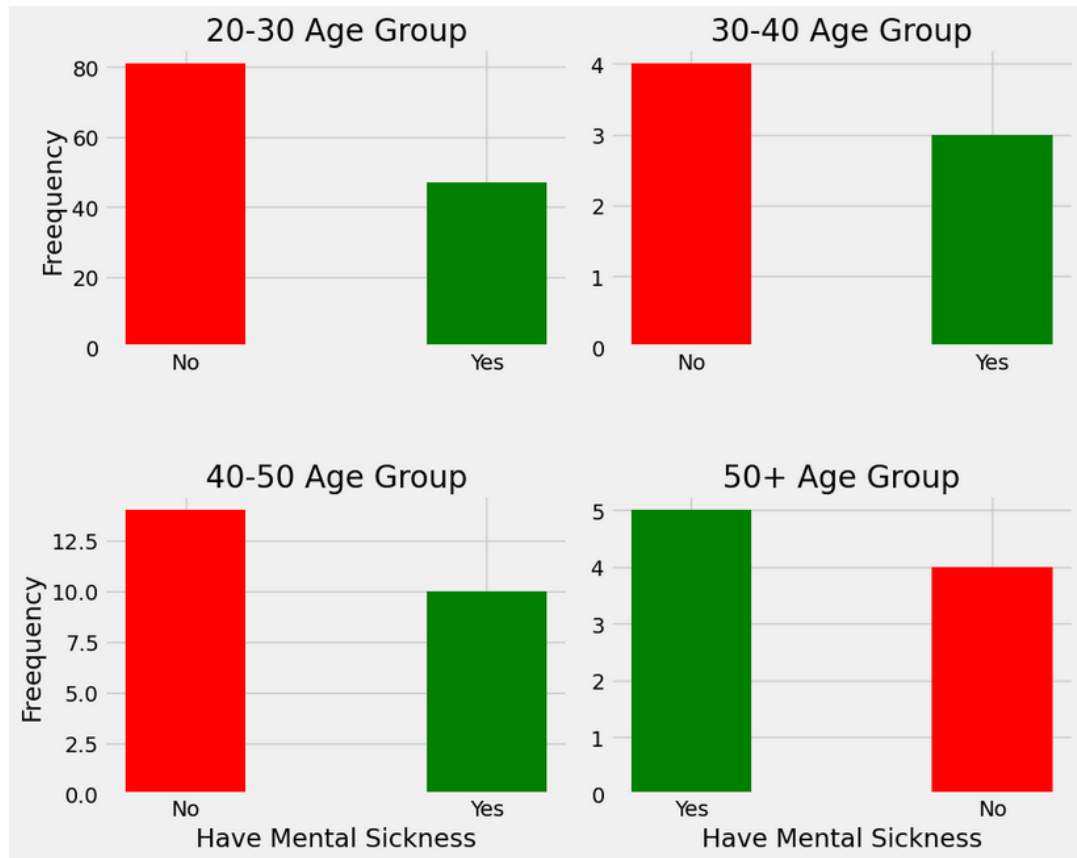
HOW DIFFERENT AGE RANGE PERSONS EXPERIENCED NEGATIVE INFLUENCE OF THEIR ACADEMIC OR PROFESSIONAL LIFE

In [44]:
```python
# CALCULATING VALUE COUNTS FOR EACH AGE GROUP
count_20_30 = age_grp.get_group('20-30')['Negative_Influenced_Academic_Professional'].value_counts()
count_30_40 = age_grp.get_group('30-40')['Negative_Influenced_Academic_Professional'].value_counts()
count_40_50 = age_grp.get_group('40-50')['Negative_Influenced_Academic_Professional'].value_counts()
count_50_plus = age_grp.get_group('50+')['Negative_Influenced_Academic_Professional'].value_counts()

print(count_20_30)
print(count_30_40)
print(count_40_50)
print(count_50_plus)
```

```
No      92
Yes     36
Name: Negative_Influenced_Academic_Professional, dtype: int64
No      7
Name: Negative_Influenced_Academic_Professional, dtype: int64
No      14
Yes     10
Name: Negative_Influenced_Academic_Professional, dtype: int64
No      5
Yes     4
Name: Negative_Influenced_Academic_Professional, dtype: int64
```

```python
# ADJUSTING THE FIG SIZE AND BAR CHART WIDTH
fig, axes = plt.subplots(2, 2, figsize=(10, 8))
bar_width = 0.4

# DEFINING COLOR FOR THE BAR
colors = ['red', 'green']

# SUBPLOT-1
axes[0, 0].bar(['No', 'Yes'], count_20_30, width=bar_width, color = colors)
axes[0, 0].set_title('20-30 Age Group')
axes[0, 0].set_ylabel('Freequency')

# SUBPLOT 2
axes[0, 1].bar(['No', 'Yes'], count_30_40, width=bar_width, color = colors)
axes[0, 1].set_title('30-40 Age Group')

# SUBPLOT 3
axes[1, 0].bar(['No', 'Yes'], count_40_50, width=bar_width, color = colors)
axes[1, 0].set_title('40-50 Age Group')
axes[1, 0].set_xlabel('Experienced Negative Influence')
axes[1, 0].set_ylabel('Freequency')

# SUBPLOT 4
axes[1, 1].bar(['No', 'Yes'], count_50_plus, width=bar_width, color = colors)
axes[1, 1].set_title('50+ Age Group')
axes[1, 1].set_xlabel('Experienced Negative Influence')

# ADJUSTING SPACE
plt.tight_layout()
plt.subplots_adjust(hspace=0.5)

plt.show()
```
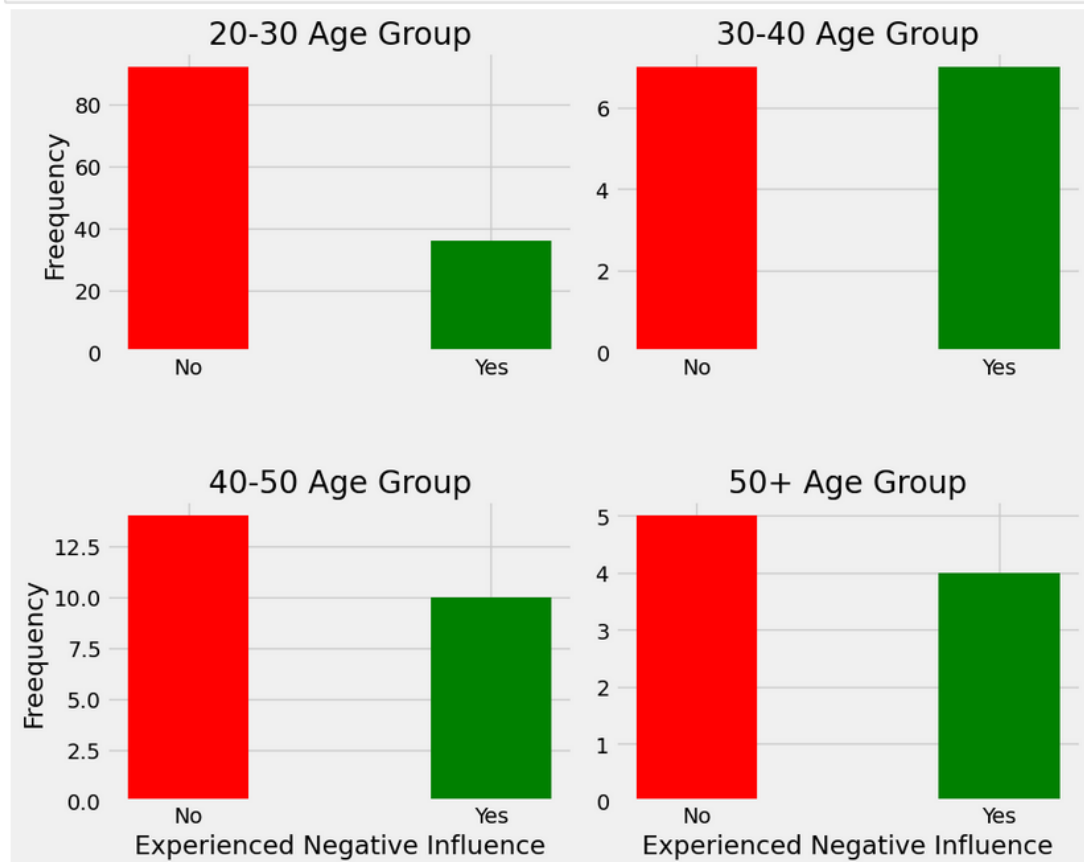
```
In [46]:   # CREATING GROUP BASED ON GENDER & EDUCATIONAL_LEVEL
           gender_grp = df.groupby(['Gender'])
```

GENDER BASED PERSONS MENTAL HEALTH PROBELEM

```
In [47]:   # CALCULATING VALUE COUNTS FOR EACH GENDER GROUP
           count_male = gender_grp.get_group('Male')['Mental_Sickness'].value_counts()
           count_female = gender_grp.get_group('Female')['Mental_Sickness'].value_counts()

           print(count_male)
           print(count_female)

           No     56
           Yes    32
           Name: Mental_Sickness, dtype: int64
           No     49
           Yes    34
           Name: Mental_Sickness, dtype: int64
```

```
In [48]:   # ADJUSTING THE FIGURE SIZE AND BAR CHART WIDTH
           fig, axes = plt.subplots(1, 2, figsize=(10, 4))   # 1-row, 2-column layout
           bar_width = 0.4

           # DEFINING COLOR FOR THE BAR
           colors = ['blue', 'green']

           # SUBPLOT-1
           axes[0].bar(['No', 'Yes'], count_male, width=bar_width, color=colors)
           axes[0].set_title('Male')
           axes[0].set_ylabel('Frequency')
           axes[0].set_xlabel('Have Mental Sickness')

           # SUBPLOT 2
           axes[1].bar(['No', 'Yes'], count_female, width=bar_width,color=colors)
           axes[1].set_title('Female')

           # ADJUSTING SPACE
           plt.tight_layout()

           plt.show()
```


GENDER BASED PERSONS MENTAL RELATIONSHIP IMPACT

```
In [49]:   # CALCULATING VALUE COUNTS FOR EACH GENDER GROUP
           count_male = gender_grp.get_group('Male')['Relationship_Impact'].value_counts()
           count_female = gender_grp.get_group('Female')['Relationship_Impact'].value_counts()

           print(count_male)
           print(count_female)

           No     56
           Yes    32
           Name: Relationship_Impact, dtype: int64
           No     53
           Yes    30
           Name: Relationship_Impact, dtype: int64
```

```
# ADJUSTING THE FIGURE SIZE AND BAR CHART WIDTH
fig, axes = plt.subplots(1, 2, figsize=(10, 4))  # 1-row, 2-column layout
bar_width = 0.4

# DEFINING COLOR FOR THE BAR
colors = ['blue', 'green']

# SUBPLOT-1
axes[0].bar(['No', 'Yes'], count_male, width=bar_width, color=colors)
axes[0].set_title('Male')
axes[0].set_ylabel('Frequency')
axes[0].set_xlabel('Relationship Impact')

# SUBPLOT 2
axes[1].bar(['No', 'Yes'], count_female, width=bar_width,color=colors)
axes[1].set_title('Female')

# ADJUSTING SPACE
plt.tight_layout()

plt.show()
```
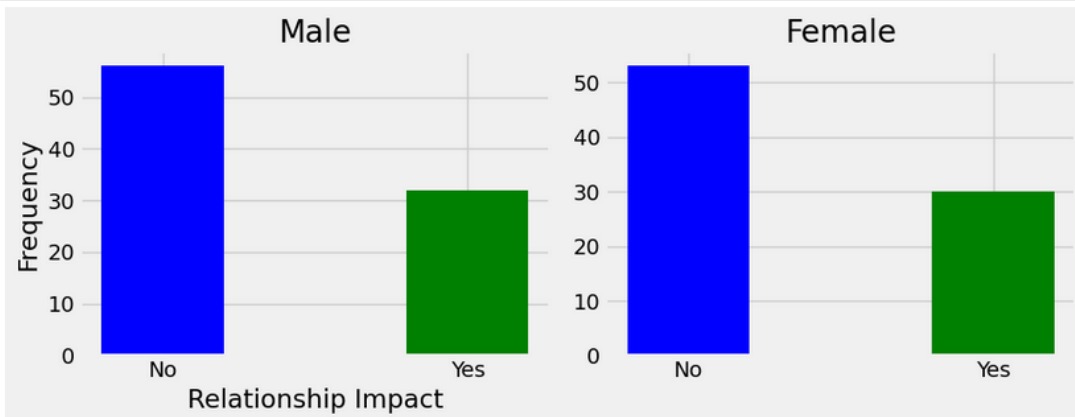

GENDER BASED THOUGHTS ABOUT NEGATIVE IMPACT OF DRINKING ALCOHOL

```
# CALCULATING VALUE COUNTS FOR EACH GENDER GROUP
count_male = gender_grp.get_group('Male')['Negative_Impact_of_Alcohol'].value_counts()
count_female = gender_grp.get_group('Female')['Negative_Impact_of_Alcohol'].value_counts()

print(count_male)
print(count_female)
```

```
Agree                54
Neutral              15
Somewhat Agree       11
Disagree              4
Somewhat Disagree     4
Name: Negative_Impact_of_Alcohol, dtype: int64
Agree                56
Somewhat Agree       14
Somewhat Disagree     5
Neutral               4
Disagree              4
Name: Negative_Impact_of_Alcohol, dtype: int64
```

```
In [52]: # ADJUSTING THE FIGURE SIZE
         fig, axes = plt.subplots(1, 2, figsize=(10, 4))

         # DEFINING COLOR FOR THE BAR
         colors = ['blue', 'green', 'purple', 'red', 'orange']
         colors_rev = ['blue', 'purple','orange', 'green', 'red']

         # SUBPLOT-1
         axes[0].barh(['Agree', 'Neutral', 'Somewhat Agree', 'Disagree', 'Somewhat Disagree'], count_male, color=colors)
         axes[0].set_title('Male')
         axes[0].set_ylabel('Impact of Drinking Alcohol')
         axes[0].set_xlabel('Frequency')

         # SUBPLOT 2
         axes[1].barh(['Agree', 'Somewhat Agree', 'Somewhat Disagree', 'Neutral', 'Disagree'], count_female, color=colors_rev)
         axes[1].set_title('Female')
         axes[1].set_xlabel('Frequency')

         # ADJUSTING SPACE
         plt.tight_layout()

         plt.show()
```
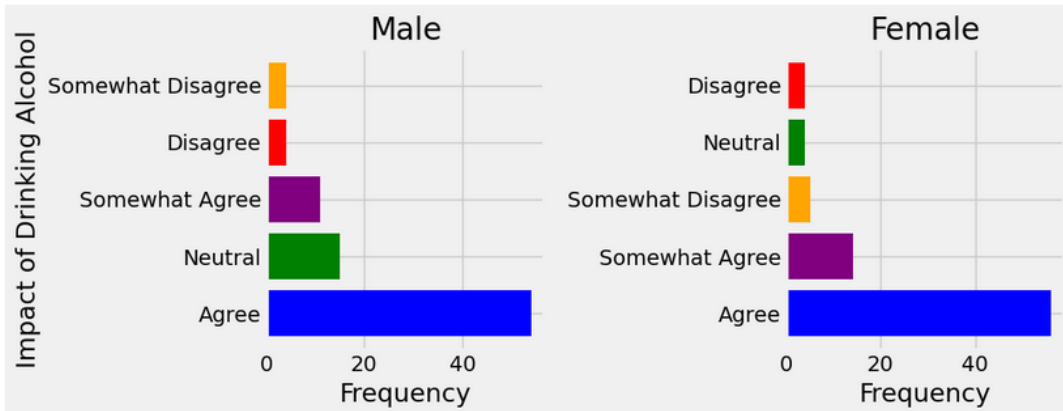
```
In [53]:   # CREATING GROUP BASED ON EDUCATIONAL_LEVEL
           education_grp = df.groupby(['Educational_Level'])
```

```
In [61]:   df['Educational_Level']
```

```
Out[61]:   0        High School
           1            College
           2         University
           3         University
           4         University
                       ...
           167       University
           168      High School
           169          College
           170       University
           171      High School
           Name: Educational_Level, Length: 172, dtype: object
```

MENTAL PRESSURE LEVEL OF DIFFERENT LEVEL OF STUDENT

```
In [67]:   # CALCULATING VALUE COUNTS FOR EACH GENDER GROUP
           count_school = education_grp.get_group('High School')['Mental_Health_Level'].value_counts()
           count_clg = education_grp.get_group('College')['Mental_Health_Level'].value_counts()
           count_uni = education_grp.get_group('University')['Mental_Health_Level'].value_counts()

           print(count_school)
           print(count_clg)
           print(count_uni)
```

```
           9     11
           8      8
           7      5
           10     4
           6      3
           Name: Mental_Health_Level, dtype: int64
           9      4
           10     3
           8      2
           5      1
           7      1
           Name: Mental_Health_Level, dtype: int64
           8     31
           9     31
           7     17
           10    13
           6     13
           5      5
           1      1
           4      1
           Name: Mental_Health_Level, dtype: int64
```

```
In [69]:  # Creating horizontal bar subplots for each educational level
          fig, axes = plt.subplots(1, 3, figsize=(15, 5))

          # Horizontal bar for 'High School'
          axes[0].barh(count_school.index.astype(str), count_school.values)
          axes[0].set_title('High School')
          axes[0].set_xlabel('Frequency')
          axes[0].set_ylabel('Mental Health Level')

          # Horizontal bar for 'College'
          axes[1].barh(count_clg.index.astype(str), count_clg.values)
          axes[1].set_title('College')
          axes[1].set_xlabel('Frequency')

          # Horizontal bar for 'University'
          axes[2].barh(count_uni.index.astype(str), count_uni.values)
          axes[2].set_title('University')
          axes[2].set_xlabel('Frequency')

          plt.tight_layout()
          plt.show()
```