ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM

AFFILIATED TO MAHATMA GANDHI UNIVERSITY, KOTTAYAM

ANXIETY AND DEPRESSION ANALYSIS OF YOUTH

PROJECT REPORT

In partial fulfilment of the requirements for the award of the degree of

B.VOC SOFTWARE DEVELOPMENT

Submitted By

INDHU JOSHY

Register No: VB21SWD014

III B.VOC SOFTWARE DEVELOPMENT

Under the guidance of

Dr. DHANYA R

DEPARTMENT OF COMPUTER APPLICATIONS

2021-24

ST. TERESA'S COLLEGE
(AUTONOMOUS)
ESTD 1925

CERTIFICATE

This is to certify that the project entitled "**ANXIETY AND DEPRESSION ANALYSIS OF YOUTH**" is a bona-fide record of the work done by **INDHU JOSHY** (VB21SWD014) during the year 2023-2024 and submitted in partial fulfilment of the requirement for the degree of **B.Voc Software Development** under Mahatma Gandhi University.

ST. TERESA'S COLLEGE DEPT. OF COMPUTER SCIENCE ERNAKULAM

21/3/24

Head of the Department

Internal Examiner

External Examiner

21/3/2024

Date:- 21/03/2024

DECLARATION

I, **Indhu Joshy** (Register no: VB21SWD014), B. Voc Software Development final year student of St. Teresa's College (Autonomous), Ernakulam, hereby declare that the dissertation submitted for the Bachelor's of Vocational Degree in Software Development is my original work. I further declare that the said work has not previously been submitted to any other university or academic body.

Date:- 21-3-2024

INDHU JOSHY

Place:- ERNAKULAM

ACKNOWLEDGMENT

First and foremost, I thank God Almighty for his blessings. I wish to express my sincere gratitude to the **Provisional Superior and Manager Rev. Dr. Sr. Vinitha CSST, Director Rev. Sr. Emeline CSST, and Principal Dr. Alphonsa Vijaya Joseph** for providing all the facilities. I deeply express my sincere thanks to my guide **Dr. Dhanya R** for her proper guidance and support throughout the project work. I thank all other faculty of the department for their exemplary guidance and appreciation throughout the project. I convey my hearty thanks to my parents, relatives, and all my friends for encouraging me and for providing all the opportunities for the completion of this project.

INDHU JOSHY

ABSTRACT

Two of the most common mental health conditions in the world that impact people of all ages and genders are anxiety and depression. This paper provides a thorough analysis of anxiety and depression in people of different ages, with a particular emphasis on those in the 15–17, 18–22, and 23–25 age ranges. The research makes use of a dataset that includes more than 300 answers to a Google Form poll. The study's main goal is to examine the prevalence of anxiety and depression in each of these age groups and determine which ones, if any, have particularly high rates of these mental health issues.

In this project, my objective is to evaluate the frequency and severity of anxiety and depression in three age groups based on various features like Age, Gender, Beck depression inventory, and Generalized anxiety disorder 7-item (GAD-7) scale. This study specifically seeks to determine which age group experiences moderate, low, and greater levels of anxiety and depression. To accomplish this goal, the dataset was subjected to Python code, which determined each respondent's anxiety and depression score based on their answers. After that, the results were combined and examined to ascertain general patterns among various age groups. The analysis's findings showed that the 15–17 age range had the highest levels of anxiety and depression, pointing to a troublingly high frequency of mental health problems among young people.

This study examines the causes of the elevated levels of anxiety and depression in the 15–17 age range as well as the ramifications of these findings. In order to properly address these issues, it also highlights how crucial early intervention and mental health care are for teenagers. The study adds important new information to the field of mental health research and activism by highlighting the value of data-driven techniques in comprehending and addressing mental health issues across various age demographics.

CONTENTS

1. INTRODUCTION

1.1 About Project

This study aims to explore the prevalence and factors associated with depression and anxiety among youth. The research employs the Statistical Package for the Social Sciences (SPSS) to analyze data collected through surveys.

The findings provide valuable insights into the mental health challenges faced by today's youth, highlighting the need for targeted interventions and support systems.

1.2 Objective Of Project

The objective of this study is to compare the levels of anxiety and depression in three different age groups: 15–17, 18–22, and 23–25. Finding out which age group has higher, moderate, and lower levels of anxiety and depression is the main goal. Understanding these differences can help develop focused interventions and support plans for mental health.

Mental health practitioners can customize interventions to meet the unique needs of each population by determining which age group has higher, moderate, and lower levels of anxiety and depression.

2. SYSTEM ANALYSIS

2.1 Literature Review

1. A Comparative Study on Depression, Anxiety, and Stress among Medical and Engineering Students in North-East India (2021).

- Authors - Nabarun Karmakar, Jayati Saha, Anjan Datta, Kaushik Nag, Kaushik Tripura, Partha Bhattacharjee.
- Aimed to assess and compare the level of depression, anxiety, and stress among medical and engineering students.
- Merits – Large sample size, Quantitative Analysis
- Demerits – The study has limited external validity as it only focuses on medical students and may not apply to non-medical students.

2. Factors associated with anxiety and depression in young adults in metropolitan city (2023).
    - Authors – Dian Asih, Andi Nurhafidzah, Hidayah Askin, and Hoirun Nisa.
    - The study identifies that working status and duration of social media use are significantly associated with depression and anxiety in young adults.
    - Merits – The study uses a quantitative analytic observational approach with a cross-sectional design.
    - Demerits – The study has a relatively small sample size of 285 young adult respondents, which may limit the generalizability of the findings.

3. Prevalence and associated factors of depression, anxiety, and stress among high school students in, Northwest Ethiopia, 2021.
    - Authors – Girum Nakie, Tesfaye Segon, Mamaru Melkam, Getachew Tesfaw Desalegn, Tadele Amare Zeleke.
    - Bivariate and multivariate analysis is used.
    - Merits – Stratified multi-stage sampling technique.
    - Demerits – It lacks specific plans for expanding mental health services in high schools.

4. Frequency of depression, anxiety, and stress among university students (2017).
    - Authors – Saba Asif, Azka Mudassar, Talala Zainab Shahzad, Mobeen Raouf, and Tehmina Pervaiz.
    - Meta-analysis, Logistic Regression.
    - Merits – Providing a broader understanding of depression, anxiety, and stress in a specific population.
    - Demerits – Logistic regression analysis can only identify associations and not causal relationships between risk factors and stress, anxiety, and depression.

5. Stress, Anxiety, and Depression among a Cohort of health sciences undergraduate students: The prevalence and risk factors (2021).
    - Authors – Muhammad Faris Fauzi, Tengku Shahrul Anuar, Lay Kek The, Wai Feng Lim, Richard Johari James, Rohana Ahmad, Mawarni Mohamed, Sahol Hamid Abu Bakar, Farida Zuraina Mohd Yusof and Mohd Zaki Salleh.
    - Logistic Regression, Depression, Anxiety, and Stress Scale-21 (DASS-21).
    - Merits – Self-report questionnaire that provides a standardized measure of depression, anxiety, and stress levels.
    - Demerits – DASS-21 relies on self-reporting, which may be subject to response and social desirability biases. Participants may underreport or overreport their symptoms leading to potential inaccuracies.

6. Depression and anxiety among university students in Hong Kong (2018).
    - Authors – Kevin WC Lun, CK Chan, Patricia KY Ip, Samantha YK Ma, WW Tsai, CS Wong, Christie HT Wong, TW Wong, D Yan.
    - Binary logistic regression and correlation analysis were used to analyze.
    - Merits – The paper provides valuable insights into the prevalence and correlates of mental health problems among university students in Hong Kong, specifically focusing on depression, anxiety, and stress.
    - Demerits – The cross-sectional design of the study limits the ability to establish causal relationships between the variables.

2.2 Existing System

A Study of Anxiety and Depression among School and College going Adolescents – A Comparative Study (2015).

- Depression is among adolescents as 54.9% of the group has shown this tendency followed by anxiety i.e., only 31%.
- The study only focused on adolescents from Mysore city, which may limit the generalizability of the findings to a broader Population.
- The Study used the Multiphasic Personality Questionnaire (MPQ) by Murthy H.N to assess anxiety and depression levels, but it did not provide information on the reliability and validity of this questionnaire.

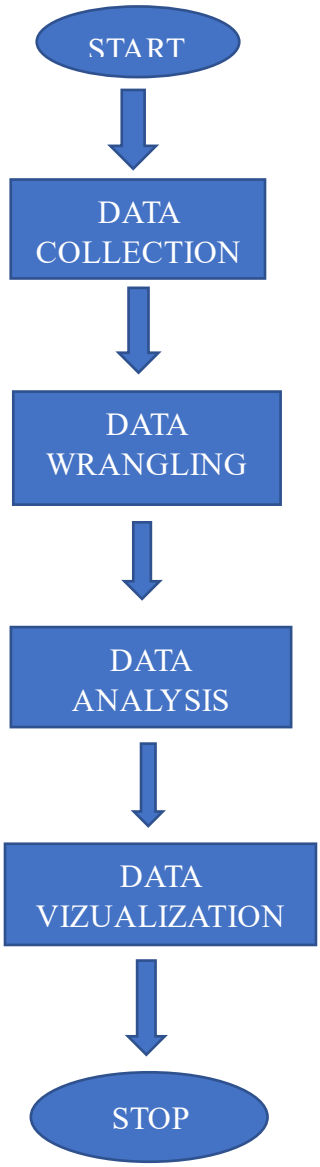
2.3 Proposed System

The suggested methodology for this research is a comparative analysis of three distinct age groups' degrees of anxiety and depression to determine which age group has higher levels, as well as who has moderate and lower levels. I employed chi-square to ascertain the gender and two-way ANOVA to discover the age groups with greater rates of anxiety and depression in this. The study also aims to investigate the effects of anxiety and depression on mental and physical health. Through the application of sophisticated statistical tools and a thorough study of pertinent data, this project seeks to offer important insights into the nature of anxiety and depression and how it affects families.

2.4 Drawback of the proposed system

It is essential to identify any possible drawbacks or limitations of a proposed system to assess its feasibility, effectiveness, and potential areas of improvement. However, one of the challenges that may arise during this process is response bias, which refers to the possibility that participants may not answer questions truthfully due to various reasons such as social desirability bias, fear of judgment, or privacy concerns. When participants hesitate to respond truthfully, it can lead to incomplete or inaccurate data, which could impact the validity and reliability of the findings. Therefore, it is crucial to account for response bias when conducting research or evaluating a system to ensure accurate and reliable results.

3. SYSTEM DESIGN



The method consisted of collecting data, selecting an appropriate method and algorithm for data analysis, data cleaning, data wrangling, analyzing our obtained data, and finally visualizing it.

4. SYSTEM DEVELOPMENT

4.1 Data Collection

Questionnaires are a popular method for collecting data in social and healthcare research. The questionnaire, which was given to more than 300 people in Ernakulam, was made to gather data on age, sex, and specifics regarding the experiences and symptoms of psychiatric conditions like anxiety and depression. Labels like the depression score and anxiety score are used for the analysis of anxiety and depression.

The process of data collection through a questionnaire involves the following steps:

- Design the Questionnaire: The questionnaire is designed based on the research objectives and the information required from the participants. The questionnaire is based on the Generalized Anxiety Disorder item-7 (GAD-7) and the Beck Depression Inventory.
  The assessment consists of 28 items that ask about events and symptoms related to anxiety and depression. The questionnaire was created using the Beck Depression Inventory and the GAD-7 to make sure the questions were correct and pertinent to the research.
- Select the Sample: The sample of participants is selected based on the research objectives and the population of interest. In this case, the sample consisted of 300 + individuals from Ernakulam.
- Collect the Responses: The responses from the participants were collected and recorded. In the case of a Google form survey, the responses were automatically recorded in a database as a CSV file.
- Data Cleaning: The collected data were cleaned to remove any inconsistencies or errors useful in the form to analyze our data. We used Python for data cleaning.

4.2 Methods Used

After reviewing relevant studies and using established research methods to investigate how anxiety and depression affect mental health in different age groups. By applying statistical techniques, such as Two-way ANOVA and chi-square tests, we have examined the connection between anxiety, depression, age groups, and Gender.

4.2.1 Pearson's Chi-square test for independence

This statistical test is commonly employed to determine whether there is a significant association between two categorical variables. In this case, the variables being analyzed are "Female" and "Male," and the test aims to ascertain whether there is a relationship between Anxiety and Depression.

The formula for Pearson’s chi-square (X2) test of independence is:

$$X^2 = \sum \frac{(O - E)^2}{E}$$

Where X2 is the chi-square test statistic

- Σ is the summation operator (it means "take the sum of")
- O is the observed frequency
- E is the expected frequency

The chi-square test can be used to analyze the impact of anxiety and depression. Steps include formulating hypotheses, preparing data, calculating expected frequencies, computing the chi-square statistic, determining degrees of freedom, setting a significance level, comparing the chi-square statistic with the critical value, and making a conclusion based on the results. To determine if there is an association between two categorical variables, you can calculate the chi-square statistic and corresponding p-value. The null hypothesis of the chi-square test is that the variables are independent. If the p-value is less than 0.05, it suggests that there is a significant association between the variables. For example, if we apply this to anxiety and depression, a low p-value would indicate that anxiety and depression has an impact on mental health.

4.2.2 Two-way ANOVA

A statistical technique called two-way ANOVA (Analysis of Variance) is used to examine how two categorical independent variables, or factors, affect a continuous dependent variable. In addition to determining if there is an interaction effect between the two independent variables, it evaluates whether there are significant changes in the means of the dependent variable across the levels of each independent variable.

The two-way ANOVA model's general formula is as follows:

$$Yijk=\mu+Ai+Bj+(AB)ij+\epsilon ijk$$

Here, k is the number of observations on the ith level of component A and jth level of factor B, respectively. - is the mean for the whole. - is the outcome of factor A's ith level.

Here's an illustration of how to apply two-way ANOVA to examine the effects of depression and anxiety on mental health:
Assume the following variables are present in the dataset we have:
The dependent variable Score for mental health
Separate variables:
Level of anxiety (low, medium, high)
Level of depression (low, medium, high)

4.3 Data Wrangling

Data wrangling is the process of cleaning, transforming, and preparing raw data for analysis. In the case of prediction and analysis, the data collected through the questionnaire had several categorical variables such as age, gender, and educational qualifications etc. To analyze this data, the categorical variables were converted to numeric values using dummy variables using Python code.

To convert categorical variables to numeric values using dummy variables, you need to follow these steps:

1. Load the Data: First, you need to load the raw data collected through the questionnaire into Python Jupyter Notebook.
2. Create Dummy Variables: Next, you need to create dummy variables for each categorical variable. In Python, you can use the pandas library to create dummy variables. For each category of the variable, a new column is created, and the values are set to either 0 or 1 depending on whether the category is present or not.
3. Merge Dummy Variables with Original Data: After creating the dummy variables, you need to merge them with the original data. You can use the pandas concat function to do this, which concatenates the original data with the newly created dummy variables.
4. Remove Original Categorical Variables: Once the dummy variables have been merged with the original data, you can remove the original categorical variables from the dataset. To do this, you can use the pandas drop function to remove the columns from the dataset.
5. Data Standardization: With the help of built-in functions the dataset is standardized and prepared.
6. Save the Processed Data: Finally, you need to save the processed data in a new file for further analysis.

4.4 Data Analysis

We have used Python to analyze our data in this project.

To determine which age groups—15–17, 18–22, 23–25—have moderate, severe, and low levels of anxiety and depression. The questionnaire was created using the Beck Depression Inventory and the Generalized Anxiety Disorder item-7 (GAD-7) scale.

In this task, a dataset related to anxiety and depression is loaded from a CSV file. A function called generate_outcome is created to generate an outcome variable based on these factors. Descriptive statistics and mean values are calculated to better understand the data and identify trends. The feature variables are standardized using StandardScaler to ensure consistency and eliminate outliers.

For the impact analysis and identification, Pearson's chi-square (X^2) test of independence is used. In the test first Null hypothesis (H0) is set as "There is association between gender and anxiety and depression. The chi-square test in Python involves importing the chi2_contingency function from the scipy. stats module and using it to calculate the chi-square statistic, p-value, degrees of freedom, and expected frequencies. A contingency table is

created using the Pandas crosstab function to calculate the frequency of occurrences for each combination of values between two columns. The results of the chi-square test are stored in variables such as chi2_stat, p_val, dof, and expected, which are then printed out to determine if there is a statistically significant association between the two variables.

Data visualization is a crucial aspect of data analysis as it helps in understanding patterns, trends, and relationships within the data. Python and QDA Miner Lite are used for a better understanding and visualization of the study.

5. SOURCE CODE

```python
import pandas as pd
import statsmodels.api as sm
from statsmodels.formula.api import ols

# Read data from Excel file
df = pd.read_excel('Book1_with_scores.xlsx')

# Fit the model
model = ols('Total_Depression_Score ~ C(Age) + C(Gender) + C(Age):C(Gender)', data=df).fit()

# Perform the ANOVA
anova_table = sm.stats.anova_lm(model, typ=2)

print(anova_table)


print(anova_table)
```

```
                     sum_sq     df          F        PR(>F)
C(Age)           44464.392573    2.0  106.589195  9.025023e-37
C(Gender)          923.218580    1.0    4.426244  3.611369e-02
C(Age):C(Gender)   471.599158    2.0    1.130509  3.240621e-01
Residual         71959.524142  345.0         NaN           NaN
                     sum_sq     df          F        PR(>F)
C(Age)           44464.392573    2.0  106.589195  9.025023e-37
```

```python
import pandas as pd
import statsmodels.api as sm
from statsmodels.formula.api import ols

# Read data from Excel file
df = pd.read_excel('Book1_with_scores.xlsx')

# Fit the model
model = ols('Total_Depression_Score ~ C(Age) + C(Gender) + C(Age):C(Gender)', data=df).fit()

# Perform the ANOVA
anova_table = sm.stats.anova_lm(model, typ=2)

print(anova_table)
```

```
                     sum_sq     df          F        PR(>F)
C(Age)           44464.392573    2.0  106.589195  9.025023e-37
C(Gender)          923.218580    1.0    4.426244  3.611369e-02
C(Age):C(Gender)   471.599158    2.0    1.130509  3.240621e-01
Residual         71959.524142  345.0         NaN           NaN
```

```
df = pd.DataFrame(data)

# Bin the continuous variables into categories
df['Age_Group'] = pd.cut(df['Age'], bins=[0, 30, 50, 100], labels=['Young', 'Middle-aged', 'Old'])
df['Depression_Group'] = pd.cut(df['Total_Depression_Score'], bins=[0, 10, 20, 30], labels=['Low', 'Medium', 'High'])
df['Anxiety_Group'] = pd.cut(df['Total_Anxiety_Score'], bins=[0, 10, 20, 30], labels=['Low', 'Medium', 'High'])

# Create a contingency table
contingency_table = pd.crosstab([df['Age_Group'], df['Depression_Group']], df['Anxiety_Group'])

# Perform the chi-square test
chi2, p, dof, expected = chi2_contingency(contingency_table)

# Print the results
print(f"Chi-Square: {chi2}")
print(f"P-value: {p}")
print(f"Degrees of Freedom: {dof}")
print("Expected Frequencies:")
print(expected)
```

```
Chi-Square: 17.770811355261017
P-value: 0.00013839402639050712
Degrees of Freedom: 2
Expected Frequencies:
[[31.63291139 10.36708861]
 [42.93037975 14.06962025]
```

Snipping Tool

Screenshot copied t
Select here to mark

Code for identifying anxiety score and depression score.

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Read the data from the Excel file
df = pd.read_excel('Book1.xlsx')

# Calculate the total scores for each row by summing up the values in each row
df['Total_Score'] = df.sum(axis=1)

# Define a function to classify the scores based on the provided scale
def classify_score(score):
    if score <= 4:
        return 'Minimal anxiety'
    elif score <= 9:
        return 'Mild anxiety'
    elif score <= 14:
        return 'Moderate anxiety'
    elif score <= 21:
        return 'Severe anxiety'

# Apply the classification function to the 'Total_Score' column
df['Classification'] = df['Total_Score'].apply(classify_score)

# Create a bar plot showing the distribution of classifications
plt.figure(figsize=(10, 6))
```

```
# Define a function to classify the scores based on the provided scale
def classify_score(score):
    if score <= 4:
        return 'Minimal anxiety'
    elif score <= 9:
        return 'Mild anxiety'
    elif score <= 14:
        return 'Moderate anxiety'
    elif score <= 21:
        return 'Severe anxiety'

# Apply the classification function to the 'Total_Score' column
df['Classification'] = df['Total_Score'].apply(classify_score)

# Create a bar plot showing the distribution of classifications
plt.figure(figsize=(10, 6))
sns.countplot(x='Classification', data=df, order=['Minimal anxiety', 'Mild anxiety', 'Moderate anxiety', 'Severe anxiety'])
plt.xlabel('Classification')
plt.ylabel('Count')
plt.title('Distribution of Classifications')
plt.xticks(rotation=45)
plt.show()
```

```
<ipython-input-8-91697d707691>:9: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric_only=None') is depr
```

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Create a bar plot showing the distribution of classifications
plt.figure(figsize=(10, 6))
sns.countplot(x='Classification', data=df, order=['Normal', 'Mild mood disturbance', 'Borderline clinical depression', 'Moderate depression', 'Severe depression
plt.xlabel('Classification')
plt.ylabel('Count')
plt.title('Distribution of Classifications')
plt.xticks(rotation=45)
plt.show()
```

```
[ ] import pandas as pd

    # Read the data from the Excel file
    df = pd.read_excel('Book1.xlsx')

    # Define the questions for depression and anxiety
    depression_questions = [
        'Sadness', 'Pessimism', 'Past failure', 'Loss of Pleasure', 'Guilty feelings',
        'Punishment feelings', 'Self- Dislike', 'Self- Criticalness', 'Suicidal thoughts or wishes',
        'Crying', 'Agitation', 'Loss of interest', 'Indecisiveness', 'Worthlessness',
        'Loss of energy', 'Changes in Sleeping Pattern', 'Irritability', 'Changes in Appetite',
        'Concentration difficulty', 'Tiredness of fatigue', 'Loss of interest in sex'
    ]

    anxiety_questions = [
        'Feeling nervous, anxious, or on edge.', 'Not being able to stop or control worrying.',
        'Worrying too much about different things.', 'Trouble relaxing',
        'Being so restless that it is hard to sit still.', 'Becoming easily annoyed or irritable.',
        'Feeling afraid, as if something awful might happen.'
    ]
```

```
# Calculate total depression score
df['Total_Depression_Score'] = df[depression_questions].sum(axis=1)

# Calculate total anxiety score
df['Total_Anxiety_Score'] = df[anxiety_questions].sum(axis=1)

# Concatenate the scores with the original data
df_concatenated = pd.concat([df, df[['Total_Depression_Score', 'Total_Anxiety_Score']]], axis=1)

# Save the concatenated data to a new Excel file
df_concatenated.to_excel('Book1_with_scores.xlsx', index=False)
```

```
import matplotlib.pyplot as plt

# Create a bar plot for the total depression and anxiety scores
plt.figure(figsize=(10, 6))
plt.bar(['Total Depression Score', 'Total Anxiety Score'], [df['Total_Depression_Score'].mean(), df['Total_Anxiety_Score'].mean()])
plt.xlabel('Score Type')
plt.ylabel('Mean Score')
plt.title('Mean Total Depression and Anxiety Scores')
plt.show()
```

```
import pandas as pd
import matplotlib.pyplot as plt

# Sample data
data = {
    'Age': [1, 1, 2, 2],
    'Gender': [1, 2, 1, 1],
    'Total_Depression_Score': [3, 7, 13, 8],
    'Total_Anxiety_Score': [3, 6, 24, 8]
}

# Create a DataFrame
df = pd.DataFrame(data)

# Group by Age and Gender
grouped = df.groupby(['Age', 'Gender']).agg({
    'Total_Depression_Score': 'mean',
    'Total_Anxiety_Score': 'mean'
}).reset_index()

# Visualize the grouped data
fig, ax = plt.subplots(figsize=(12, 6))

# Plot mean total depression scores
ax.bar(grouped.index, grouped['Total_Depression_Score'], width=0.4, label='Depression Score')

# Plot mean total anxiety scores
ax.bar(grouped.index + 0.4, grouped['Total_Anxiety_Score'], width=0.4, label='Anxiety Score')
```

```
# Plot mean total depression scores
ax.bar(grouped.index, grouped['Total_Depression_Score'], width=0.4, label='Depression Score')

# Plot mean total anxiety scores
ax.bar(grouped.index + 0.4, grouped['Total_Anxiety_Score'], width=0.4, label='Anxiety Score')

# Set x-ticks and labels
ax.set_xticks(grouped.index + 0.2)
ax.set_xticklabels(grouped['Age'].astype(str) + ' - ' + grouped['Gender'].astype(str))

# Set labels and title
ax.set_xlabel('Age - Gender')
ax.set_ylabel('Mean Score')
ax.set_title('Mean Total Depression and Anxiety Scores by Age Group and Gender')
ax.legend()

# Rotate x-tick labels for better readability
plt.xticks(rotation=45, ha='right')

# Show the plot
plt.tight_layout()
plt.show()
```

6. RESULTS

To Determine which age groups have moderate, severe, and mild anxiety and depression. For the

With a very low p-value of 9.025023e-37 and a high F-statistic of 106.589 for the component "Age," it is clear that the factor "Age" significantly affects the dependent variable. The factor "Gender" has a p-value of 0.036 and a lower (4.426) but significant F-statistic, indicating that it may have a less dramatic effect on the dependent variable than the factor "Age."

The interaction term "Age: Gender" has a reasonably high p-value (0.324) and a low F-statistic (1.131), suggesting that it is not significant.

The residual is the amount of variability in the dependent variable that cannot be explained after taking into account the impact of the factors "Gender" and "Age." Since its F-statistic is residual and there is no more variability to explain While their interaction may not be significant, overall it seems that "Age" and "Gender" have a considerable impact on the dependent variable.

	sum_sq	df	F	PR(>F)
C(Age)	44464.392573	2.0	106.589195	9.025023e-37
C(Gender)	923.218580	1.0	4.426244	3.611369e-02
C(Age):C(Gender)	471.599158	2.0	1.130509	3.240621e-01
Residual	71959.524142	345.0	NaN	NaN
	sum_sq	df	F	PR(>F)
C(Age)	44464.392573	2.0	106.589195	9.025023e-37
C(Gender)	923.218580	1.0	4.426244	3.611369e-02
C(Age):C(Gender)	471.599158	2.0	1.130509	3.240621e-01
Residual	71959.524142	345.0	NaN	NaN

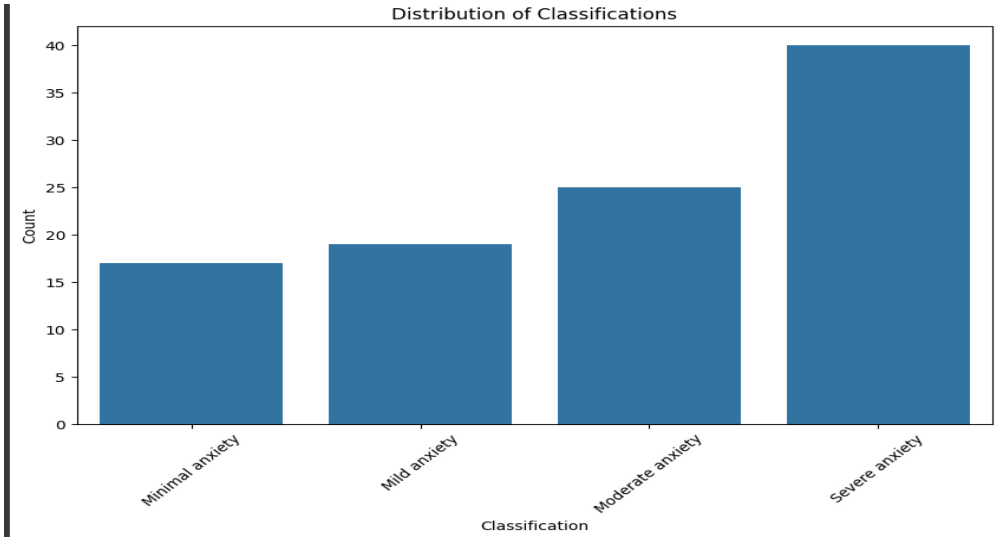
Total depression score and total anxiety score

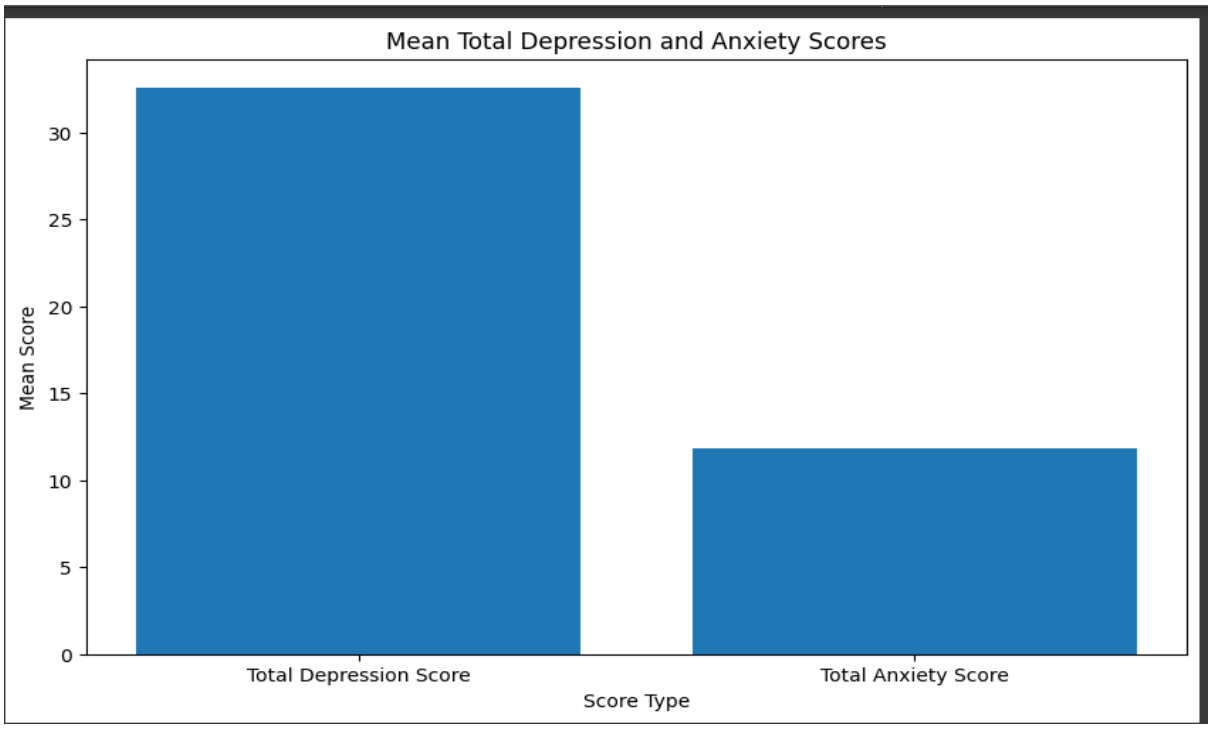
	sum_sq	df	F	PR(>F)
C(Age)	44464.392573	2.0	106.589195	9.025023e-37
C(Gender)	923.218580	1.0	4.426244	3.611369e-02
C(Age):C(Gender)	471.599158	2.0	1.130509	3.240621e-01
Residual	71959.524142	345.0	NaN	NaN

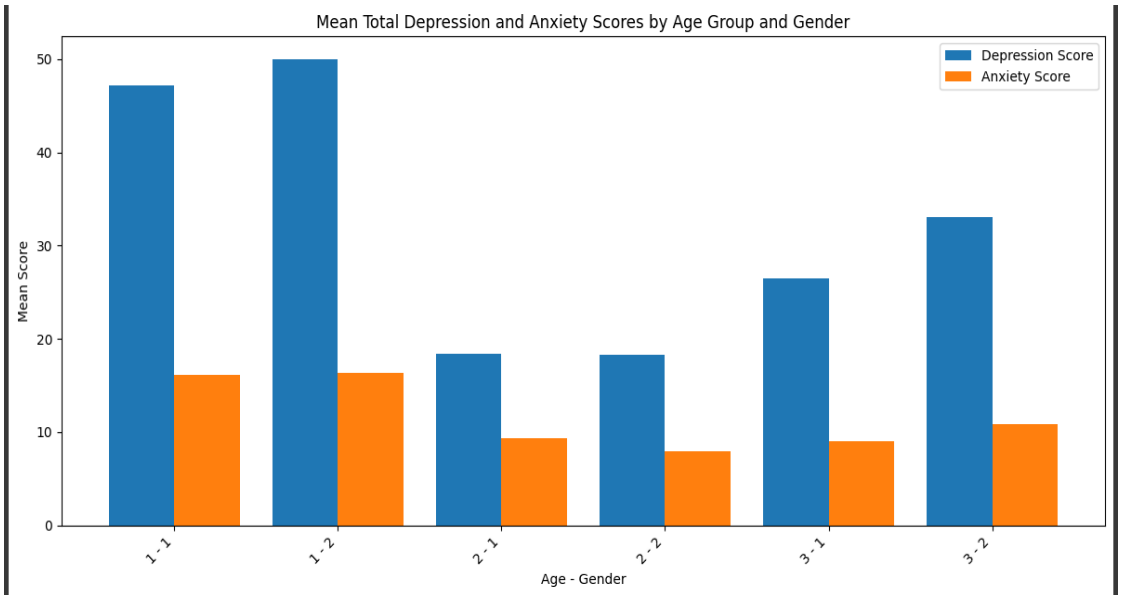
The null hypothesis (H0) was set as there will be no association between the age group and depression, or anxiety. The Chi-square measures the discrepancy between observed and expected frequencies in a contingency table. A high Chi-square value indicates a stronger association. The p-value represents the probability of observing an extreme Chi-square statistic, indicating the significance of the relationship. Since we got a p-value < 0.05, we can reject the null hypothesis. and conclude that variables have a significant relationship. In our analysis, age group has a significant impact on depression and anxiety.

```
Chi-Square: 17.770811355261017
P-value: 0.00013839402639050712
Degrees of Freedom: 2
Expected Frequencies:
[[31.63291139 10.36708861]
 [42.93037975 14.06962025]
 [44.43670886 14.56329114]]
```

Observed frequencies of outcome







The inference of my study is that: High school students have extreme depression and anxiety.

Undergraduate Students have mild depression and anxiety. While postgraduate students have moderate depression and anxiety.

7. CONCLUSION

It is clear from the study's final results, which compare three age groups (15–17, 18–22, and 23–25), that people in the 15–17 age group had considerably higher levels of anxiety and depression than people in the older age groups.
In particular, the results show that as compared to older people, those in the 15–17 age bracket have far greater levels of anxiety and despair. This implies that adolescence—more specifically, the middle to late adolescent years—represents a crucial time when mental health issues are more noticeable.
On the other hand, there seems to be a decline in the reported levels of anxiety and sadness as people enter the young adult (18–22 and 23–25 age groups). This could be due to a number of things, such as growing older, developing more coping strategies, and becoming more resilient in the face of adversity.

To effectively treat mental health difficulties in teenagers, particularly those in the 15–17 age range, our findings highlight the significance of early intervention and support systems. At this developmental stage, anxiety and depression can be lessened with the help of programs that promote mental health, such as mental health education, counseling services, and friendly surroundings.

Additionally, continued study is necessary to expand our knowledge of the fundamental causes of anxiety and depression in a variety of age groups and to guide the creation of specialized interventions that cater to the particular requirements of each group.

In the end, we can work to build healthier and more resilient communities by elevating mental health awareness and offering complete support throughout the lifespan.

8. REFERENCES

1. Nabarun Karmakar, Jayati Saha, Anjan Datta, Kaushik Nag, Kaushik Tripura, Partha Bhattacharjee (2021). A Comparative Study on Depression, Anxiety, and Stress among Medical and Engineering Students in North-East India https://journals.lww.com/chri/fulltext/2021/08010/a_comparative_study_on_depression,_anxiety,_and.5.aspx

2. Dian Asih, Andi Nurhafidzah, Hidayah Askin, and Hoirun Nisa (2023). Factors associated with anxiety and depression in young adults in metropolitan city https://ejournal.poltekkes-smg.ac.id/ojs/index.php/jrk/article/view/9446

3. Girum Nakie, Tesfaye Segon, Mamaru Melkam, Getachew Tesfaw Desalegn, Tadele Amare Zeleke (2021). Prevalence and associated factors of depression, anxiety, and stress among high school students in, Northwest Ethiopia https://link.springer.com/article/10.1186/s12888-022-04393-1

4. Saba Asif, Azka Mudassar, Talala Zainab Shahzad, Mobeen Raouf, and Tehmina Pervaiz (2017). Frequency of depression, anxiety, and stress among university students. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7372668/

5. Muhammad Faris Fauzi, Tengku Shahrul Anuar, Lay Kek The, Wai Feng Lim, Richard Johari James, Rohana Ahmad, Mawarni Mohamed, Sahol Hamid Abu Bakar, Farida Zuraina Mohd Yusof and Mohd Zaki Salleh (2021). Stress, Anxiety, and Depression among a Cohort of health sciences undergraduate students: The prevalence and risk factors. https://www.mdpi.com/1660-4601/18/6/3269

6. Kevin WC Lun, CK Chan, Patricia KY Ip, Samantha YK Ma, WW Tsai, CS Wong, Christie HT Wong, TW Wong, D Yan (2018). Depression and Anxiety among university students in Hong Kong. https://pdfs.semanticscholar.org/3d4e/f2677042bf7f1575a477d09440295ff0939a.pdf

7. Ahmad A Mirza, Mukhtiar Baig, Ghada M Beyari, Mohammed Aiman Halawani & Abdulrahim A Mirza (2021). Depression and Anxiety Among Medical Students: A Brief Overview. https://www.tandfonline.com/doi/pdf/10.2147/AMEP.S302897

8. Mohammad Mofatteh (2020).Risk factors associated with stress, anxiety and depression among university undergraduate students. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7870388/.