दिल्ली विश्वविद्यालय
University of Delhi

COLLEGE OF VOCATIONAL STUDIES

BSC HONS COMPUTER SCIENCE

DATA ANALYSIS AND VISUALIZATION

PROJECT FILE

Nishant 2k21/CS/72
21013570064

TOPIC : - ***Understanding Factors Affecting College Student Mental Health: A Comprehensive Survey and Analysis***

ACKNOWLEDGEMENT

I extend my sincere thanks to Mrs. Geetika Vashist, my esteemed professor at the College of Vocational Studies, for her valuable guidance and unwavering support throughout the duration of this project. Her expertise and insights significantly influenced the direction and quality of this research. I also appreciate the College of Vocational Studies for providing an environment conducive to learning and research. The resources and opportunities offered by the college have greatly contributed to the successful completion of this project.

Lastly, I want to acknowledge the hard work and commitment that I put into this project. It has been a challenging yet rewarding journey, and I am proud to have been able to contribute to the field of Computer Science through this research endeavour.

CERTIFICATE

This is to certify that the project entitled " The “Understanding Factors Affecting College Student Mental Health: A Comprehensive Survey and Analysis” project aims to achieve the following objectives" has been successfully completed by

Nishant ,under the guidance of Mrs. Geetika Vashist. This project was undertaken as a partial fulfilment of the requirements for Data Analysis and Visualization. The work presented in this project is the original outcome of the research conducted by Nishant , and it has not been submitted elsewhere for any academic or professional purpose.

Signature:_____

Mrs. Geetika Vashist

PROJECT TITLE : - ***Understanding Factors Affecting College Student Mental Health: A Comprehensive Survey and Analysis***

Abstract :

This research study explores the multifaceted landscape of factors influencing the mental health of college students through a comprehensive survey and analysis. As the demands and challenges faced by college students continue to evolve, understanding the intricate interplay of various elements becomes imperative for fostering a supportive academic environment.

The study employs a mixed-methods approach, combining quantitative surveys and qualitative analyses to delve into the diverse aspects affecting mental well-being. Academic pressures, social dynamics, familial support structures, lifestyle choices, and accessibility to mental health resources are among the key variables scrutinized. The data, gathered from a diverse sample of college students, offers nuanced insights into the complex relationships between these factors and mental health outcomes

The findings reveal a spectrum of influences, highlighting the interconnected nature of academic, social, and personal factors. The study identifies key stressors and protective factors, providing a foundation for targeted interventions. Additionally, the research underscores the importance of a holistic approach, acknowledging the symbiotic relationship between academic success and mental well-being.

The implications of this study extend to educators, mental health professionals, policymakers, and all stakeholders invested in the welfare of college students. By unraveling the intricacies of these influences, this research contributes to the development of informed strategies aimed at cultivating a conducive and supportive environment for the mental health of college students.

This study not only adds to the growing body of literature on college student mental health but also serves as a catalyst for future research endeavors seeking to address the evolving challenges faced by this demographic.

Introduction:

In an era marked by increasing academic pressures and evolving societal expectations, the mental health of college students stands as a critical concern. This project undertakes a thorough exploration of the multifaceted influences impacting the mental well-being of students within Delhi University. Through a structured survey methodology, the study probes factors including course selection, family pressure, college environment, and other pertinent variables. Employing advanced data analysis techniques, the project seeks to unearth crucial insights, ultimately aiming to inform targeted interventions and foster a supportive academic milieu.

Objective:

The “Understanding Factors Affecting College Student Mental Health: A Comprehensive Survey and Analysis” project aims to achieve the following objectives:

1. **Identify Motivations for Course Selection:**
   - Investigate whether students choose their courses based on personal interest, family pressure, or other factors.
2. **Evaluate Family Pressure:**
   Gauge the level of familial influence on students' choice of academic pursuits and its potential impact on mental health.
3. **Assess College Environment:**
   Examine the perceived quality of the college environment and its correlation with mental well-being.
4. **Analyze Distance from Home**
   Explore how the distance between students' homes and the college affects their mental health.
5. **Determine Interest in Chosen Course:**
   Investigate the correlation between genuine interest in the chosen course and mental health outcomes.

Data collection and survey design:

1. **Demographic Information**:
   - Name:
   - Age:
   - Email:
   - Gender:
   - Academic Year:
   - Course of Study:
2. **Reason for Choosing the Current Course**:
   - Did you choose your current course out of personal interest, family pressure, or other reasons? (Multiple choice)

3. **Family Pressure**:
   - On a scale of 1 to 5, how much pressure do you feel from your family regarding your choice of course?
4. **College Environment**:
   - How would you rate the overall environment of your college? (Scale: Very Poor, Poor, Neutral, Good, Very Good)
5. **Distance from Home to College**:
   - How far is your college from your home? (Options: <5 km, 5-10 km, 1020 km, >20 km)
6. **Interest in Chosen Course**:
   - Are you genuinely interested in the course you've chosen? (Yes/No)
7. **Time Table and Workload**:
   - How do you feel about your class schedule and workload? (Options: Too Light, Just Right, Too Heavy)
8. **Stress Levels**:
   - On a scale of 1 to 5, how would you rate your current stress levels?
9. **Overall Mental Well-being**:
   - How would you describe your overall mental well-being? (Options: Very Poor, Poor, Neutral, Good, Very Good.
10. How would you describe your eating habits?
    . Very poor, poor, Good, Very Good.

1. Data processing.

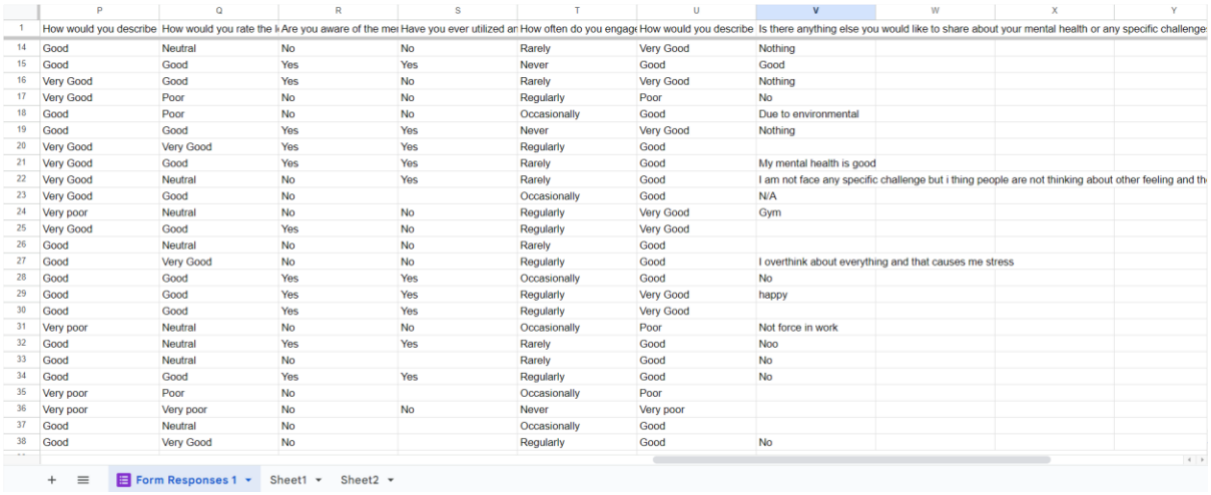
. import data set.
. Fill the missing values with mean of column in which null values is present.
. Removing any kind of space which cause any problem in getting data.
. Removing Duplicates.

Code:

```
import pandas as pd import
numpy as np import
matplotlib.pyplot as plt
import csv
 data =
pd.read_csv("data.csv3") df=
pd.DataFrame(data) print(df)

##df=df.drop(columns=['Timestamp'])
##df=df.drop(columns=['Is there is any gum and teeth related diseases have
been there in your family trade  '])
 columns_with_missing_values =
df.columns[df.isnull().any()].tolist()
 for column in
columns_with_missing_values:     if
df[column].dtype == 'float64':
        df[column].fillna(df[column].mean(), inplace=True)
         for column in
columns_with_missing_values:     if
df[column].dtype == 'object':
        df[column].fillna(df[column].mode().iloc[0], inplace=True)
         for column in
df.columns:
    if df[column].dtype == 'object':  # Check if the column
contains object/string data         df[column] =
df[column].str.upper()
 for column in
df.columns:
    if df[column].dtype == 'object':  # Check if the column contains
object/string data
        df[column] = df[column].str.strip()
     df.to_csv('Data4.csv',
index=False) print(df)
```

Data set before and after cleaning:

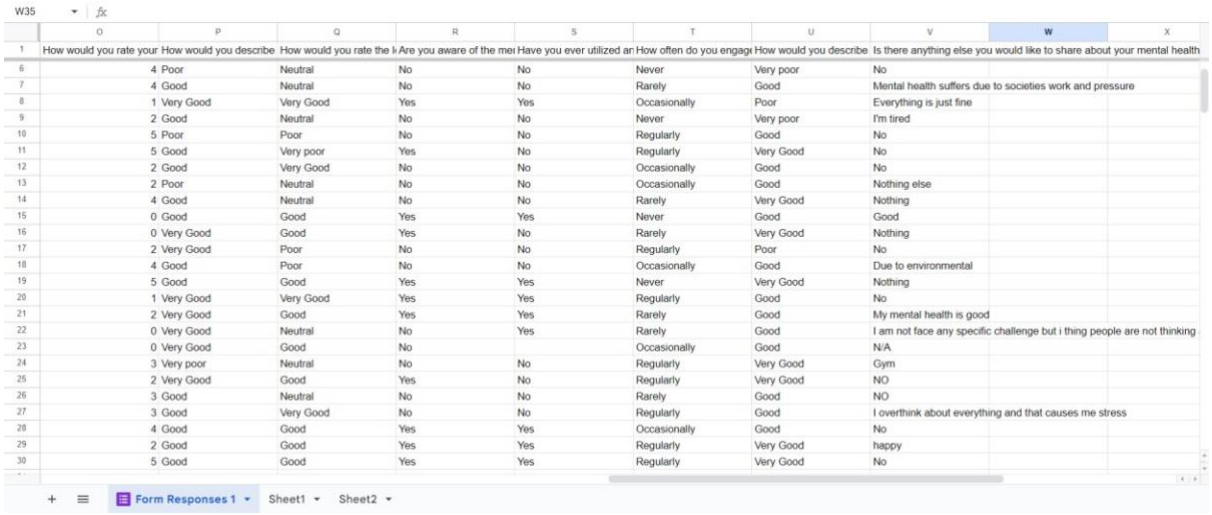
. Before :



	P	Q	R	S	T	U	V	W	X	Y
1	How would you describe	How would you rate the l	Are you aware of the mer	Have you ever utilized ar	How often do you engage	How would you describe	Is there anything else you would like to share about your mental health or any specific challenge			
14	Good	Neutral	No	No	Rarely	Very Good	Nothing			
15	Good	Good	Yes	Yes	Never	Good	Good			
16	Very Good	Good	Yes	No	Rarely	Very Good	Nothing			
17	Very Good	Poor	No	No	Regularly	Poor	No			
18	Good	Poor	No	No	Occasionally	Good	Due to environmental			
19	Good	Good	Yes	Yes	Never	Very Good	Nothing			
20	Very Good	Very Good	Yes	Yes	Regularly	Good				
21	Very Good	Good	Yes	Yes	Rarely	Good	My mental health is good			
22	Very Good	Neutral	No	Yes	Rarely	Good	I am not face any specific challenge but i thing people are not thinking about other feeling and th			
23	Very Good	Good	No		Occasionally	Good	N/A			
24	Very poor	Neutral	No	No	Regularly	Very Good	Gym			
25	Very Good	Good	Yes	No	Regularly	Very Good				
26	Good	Neutral	No	No	Rarely	Good				
27	Good	Very Good	No	No	Regularly	Good	I overthink about everything and that causes me stress			
28	Good	Good	Yes	Yes	Occasionally	Good	No			
29	Good	Good	Yes	Yes	Regularly	Very Good	happy			
30	Good	Good	Yes	Yes	Regularly	Very Good				
31	Very poor	Neutral	No	No	Occasionally	Poor	Not force in work			
32	Good	Neutral	Yes	Yes	Rarely	Good	Noo			
33	Good	Neutral	No		Rarely	Good	No			
34	Good	Good	Yes	Yes	Regularly	Good	No			
35	Very poor	Poor	No		Occasionally	Poor				
36	Very poor	Very poor	No	No	Never	Very poor				
37	Good	Neutral	No		Occasionally	Good				
38	Good	Very Good	No		Regularly	Good	No			

Form Responses 1 | Sheet1 | Sheet2

- After:



W35

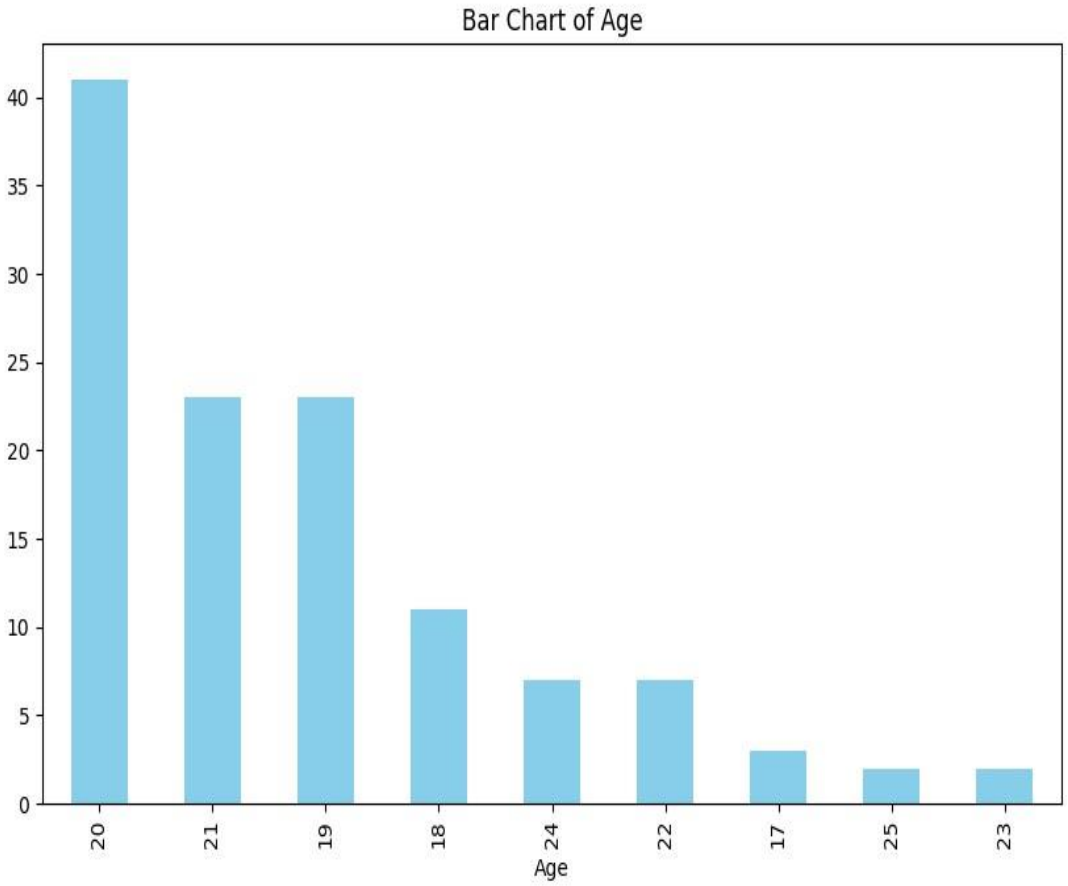
	O	P	Q	R	S	T	U	V	W	X
1	How would you rate your	How would you describe	How would you rate the l	Are you aware of the mer	Have you ever utilized ar	How often do you engage	How would you describe	Is there anything else you would like to share about your mental health		
6		4 Poor	Neutral	No	No	Never	Very poor	No		
7		4 Good	Neutral	No	No	Rarely	Good	Mental health suffers due to societies work and pressure		
8		1 Very Good	Very Good	Yes	Yes	Occasionally	Poor	Everything is just fine		
9		2 Good	Neutral	No	No	Never	Very poor	I'm tired		
10		5 Poor	Poor	No	No	Regularly	Good	No		
11		5 Good	Very poor	Yes	No	Regularly	Very Good	No		
12		2 Good	Very Good	No	No	Occasionally	Good	No		
13		2 Poor	Neutral	No	No	Occasionally	Good	Nothing else		
14		4 Good	Neutral	No	No	Rarely	Very Good	Nothing		
15		0 Good	Good	Yes	Yes	Never	Good	Good		
16		0 Very Good	Good	Yes	No	Rarely	Very Good	Nothing		
17		2 Very Good	Poor	No	No	Regularly	Poor	No		
18		4 Good	Poor	No	No	Occasionally	Good	Due to environmental		
19		5 Good	Good	Yes	Yes	Never	Very Good	Nothing		
20		1 Very Good	Very Good	Yes	Yes	Regularly	Good	No		
21		2 Very Good	Good	Yes	Yes	Rarely	Good	My mental health is good		
22		0 Very Good	Neutral	No	Yes	Rarely	Good	I am not face any specific challenge but i thing people are not thinking		
23		0 Very Good	Good	No		Occasionally	Good	N/A		
24		3 Very poor	Neutral	No	No	Regularly	Very Good	Gym		
25		2 Very Good	Good	Yes	No	Regularly	Very Good	NO		
26		3 Good	Neutral	No	No	Rarely	Good	NO		
27		3 Good	Very Good	No	No	Regularly	Good	I overthink about everything and that causes me stress		
28		4 Good	Good	Yes	Yes	Occasionally	Good	No		
29		2 Good	Good	Yes	Yes	Regularly	Very Good	happy		
30		5 Good	Good	Yes	Yes	Regularly	Very Good	No		

Form Responses 1 | Sheet1 | Sheet2

2.Analysis and visualization of data set using pandas :

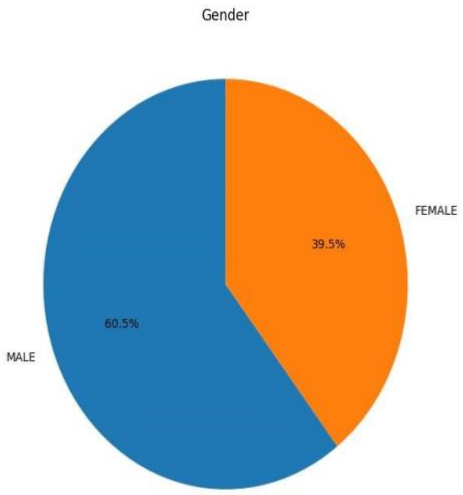
Data set 1: Age

```
import pandas as pd import
numpy as np import
matplotlib.pyplot as plt
import csv

#Available columns are: ['Timestamp', 'Name ', 'Age', 'Phone number ',
# 'Gender', 'Email id', 'Course of study.', 'Academic year', 'Unnamed: 8',
# 'How much pressure do you feel from your family regarding your choice of
course?',
# 'How would you rate the overall environment of your college?', 'Distance
from home to college:',
# "Are you genuinely interested in the course you've chosen?", 'How do you
feel about your class schedule and workload?', 'How would you rate your
current stress levels?', 'How would you describe your overall mental
wellbeing? ', 'How would you rate the level of social support you receive from
family, friends, and peers?', 'Are you aware of the mental health resources
available on campus?', 'Have you ever utilized any of the mental health
resources provided by the college? ', 'How often do you engage in physical
activity or exercise per week?', 'How would you describe your eating habits?',
'Is there anything else you would like to share about your mental health or
any specific challenges you face?']


df = pd.read_csv('Data4.csv')
column_name = 'Age'
##
# Check if the specified column exists in the DataFrame if
column_name not in df.columns:     print(f"Error: Column
'{column_name}' not found in the DataFrame. Available columns are:
{df.columns.tolist()}") else:
    # Count the occurrences of each unique value in the specified column
value_counts = df[column_name].value_counts()


value_counts = df['Age'].value_counts()
plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar',
color='skyblue') plt.title(f'Bar Chart of
Age') plt.xlabel('Age') plt.ylabel(' ')
plt.show()



 age_mean=df['Age'].mean() print("Average age is:") print(age_mean)
```



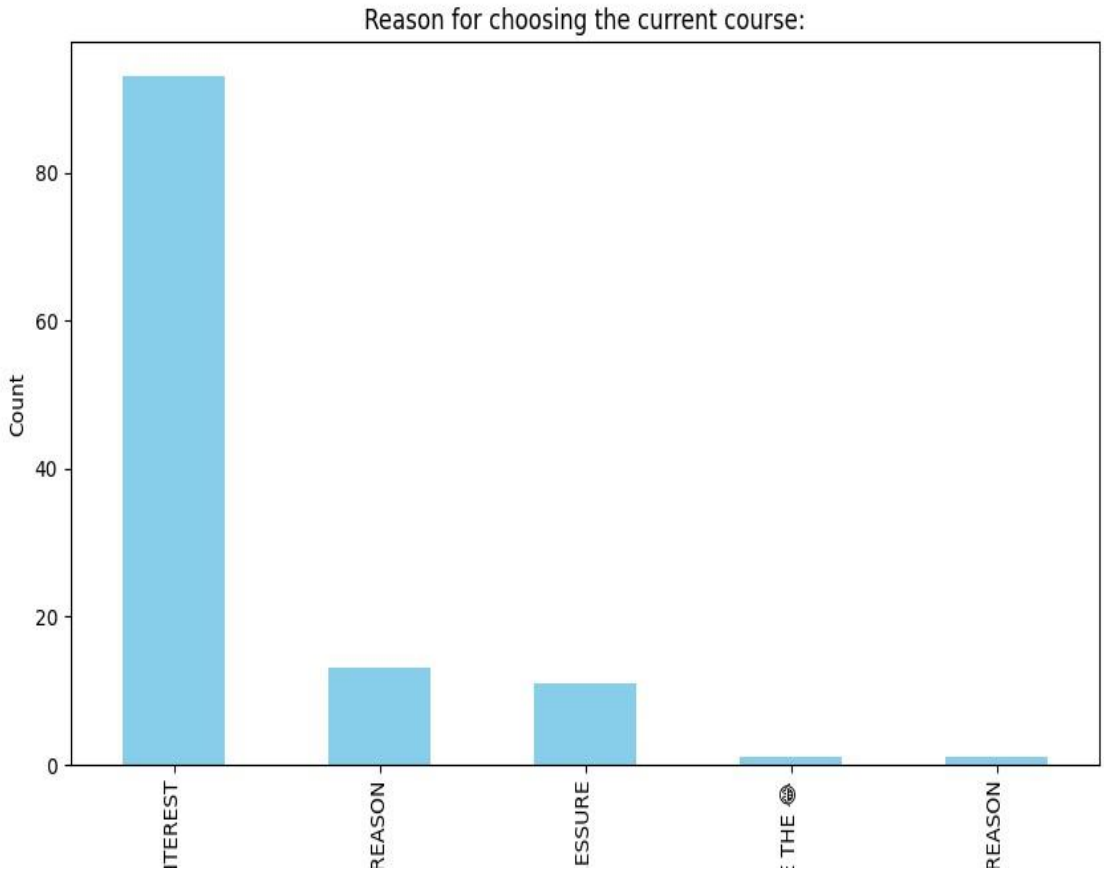
Data set 2: Gender

```
value_counts = df['Gender'].value_counts() plt.figure(figsize=(8,
8)) plt.pie(value_counts, labels=value_counts.index,
autopct='%1.1f%%', startangle=90) plt.title(f'Gender') plt.show()


mode_value = df['Gender'].mode()[0]  # Get the first mode (in case of
multiple modes) mode_count = df['Gender'].value_counts()[mode_value]
 total_values = len(df['Gender']) mode_percentage
= (mode_count / total_values) * 100
 print(mode_value,"
",mode_count,mode_percentage)
```



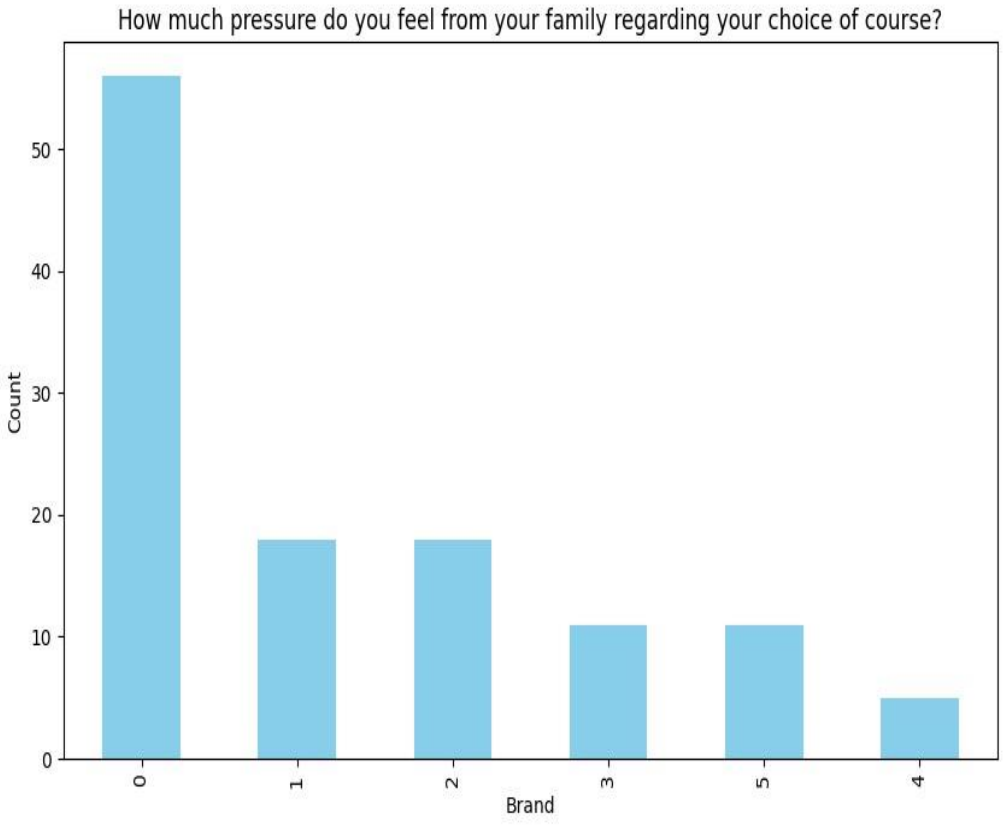
Data set 3: Reason for choosing the current course:

```
value_counts = df['Reason for choosing the current course:'].value_counts()
plt.figure(figsize=(10, 6)) value_counts.plot(kind='bar', color='skyblue')
plt.title(f'Reason for choosing the current course:') plt.xlabel('Reason
for choosing the current course:') plt.ylabel('Count') plt.show()
```



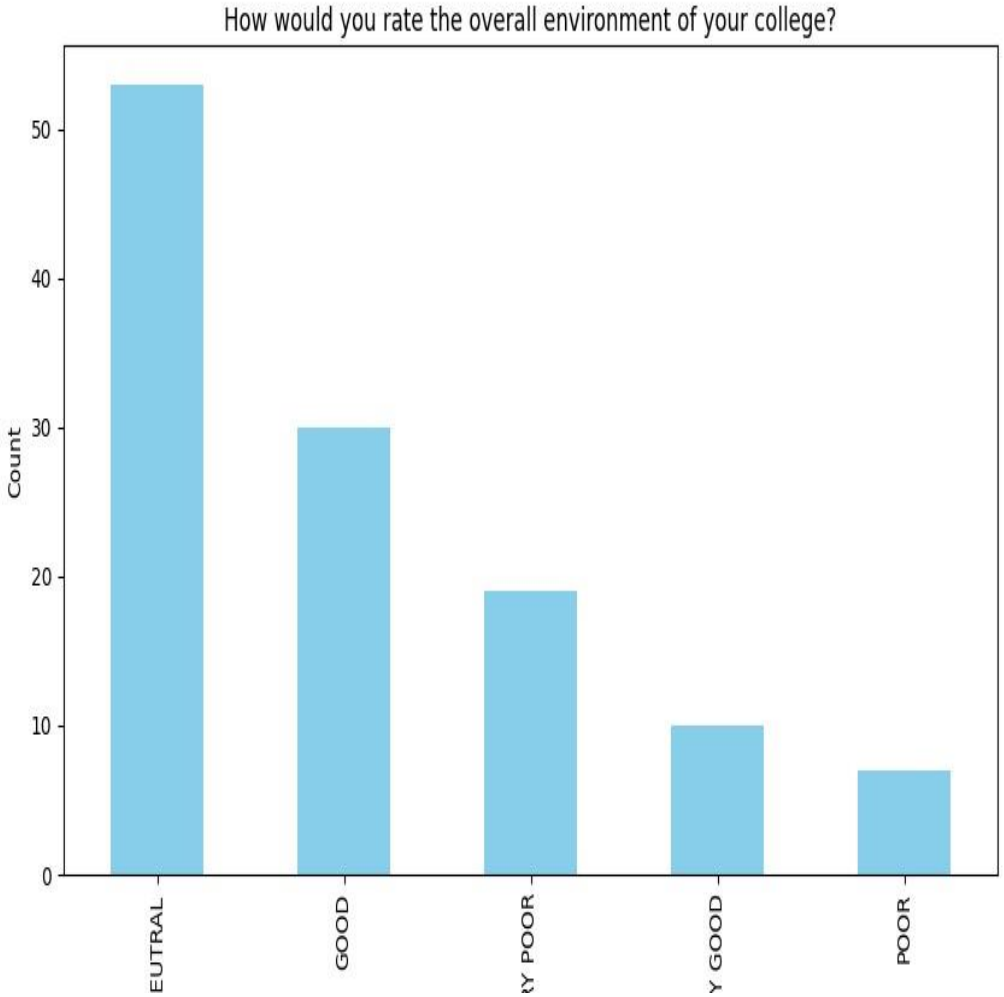
Data set 4. How much pressure do you feel from your family regarding your choice of course?

```
value_counts = df['How much pressure do you feel from your family regarding your choice of course?'].value_counts() plt.figure(figsize=(10, 6)) value_counts.plot(kind='bar', color='skyblue') plt.title(f'How much pressure do you feel from your family regarding your choice of course?') plt.xlabel('Brand') plt.ylabel('Count') plt.show() mode_value1= df['How much pressure do you feel from your family regarding your choice of course?'].mode()[0]  # Get the first mode (in case of multiple modes) mode_count1= df['How much pressure do you feel from your family regarding your choice of course?'].value_counts()[mode_value1] total_values1 = len(df['How much pressure do you feel from your family regarding your choice of course?']) mode_percentage1 = (mode_count1 / total_values1) * 100 print(mode_value1,"  ",mode_count1,mode_percentage1)
```



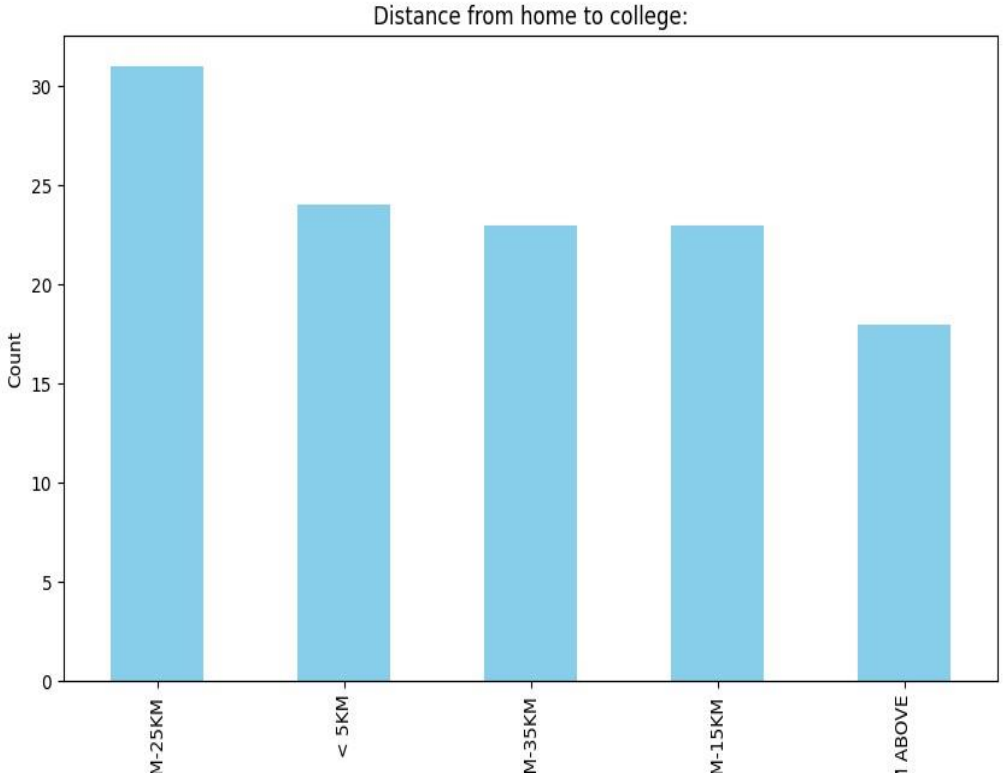
Data set 5.How would you rate the overall environment of your college?

```
value_counts = df["How would you rate the overall environment of your
college?"].value_counts() plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue') plt.title(f"How would you
rate the overall environment of your college?") plt.xlabel('Brand')
plt.ylabel('Count') plt.show()
```



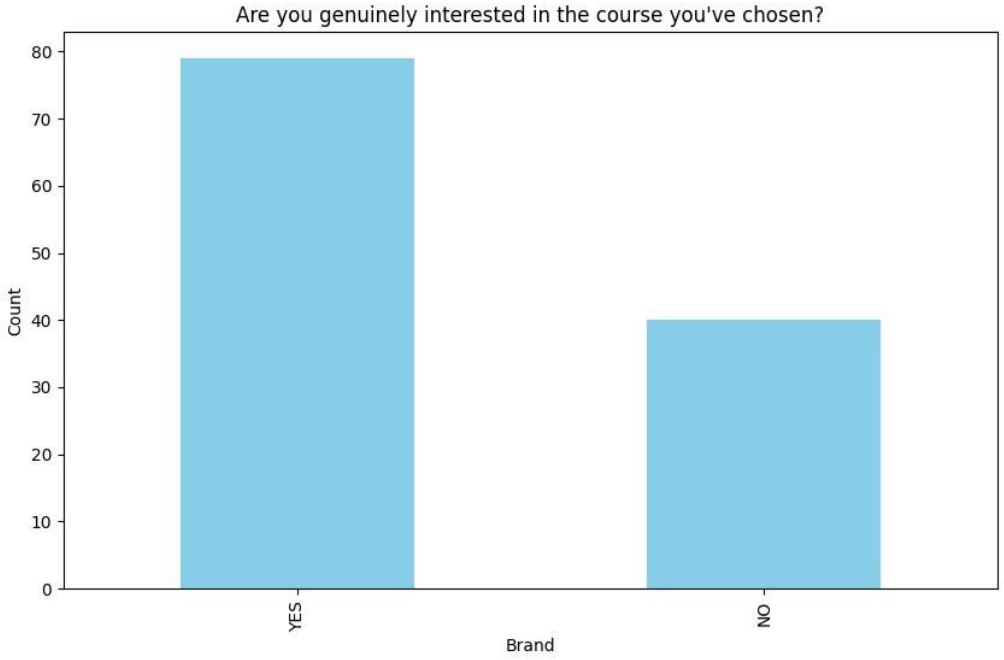
Data set 6. Distance from home to college:

```
value_counts = df['Distance from home to
college:'].value_counts() plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue')
plt.title(f'Distance from home to college:') plt.xlabel('Distance
from college') plt.ylabel('Count') plt.show()
 mode_value1= df['Distance from home to college:'].mode()[0]  # Get the
first mode (in case of multiple modes) mode_count1= df['Distance from home
to college:'].value_counts()[mode_value1]
 total_values1 = len(df['Distance from home to
college:']) mode_percentage1 = (mode_count1 /
total_values1) * 100
 print(mode_value1,"
",mode_count1,mode_percentage1)
```



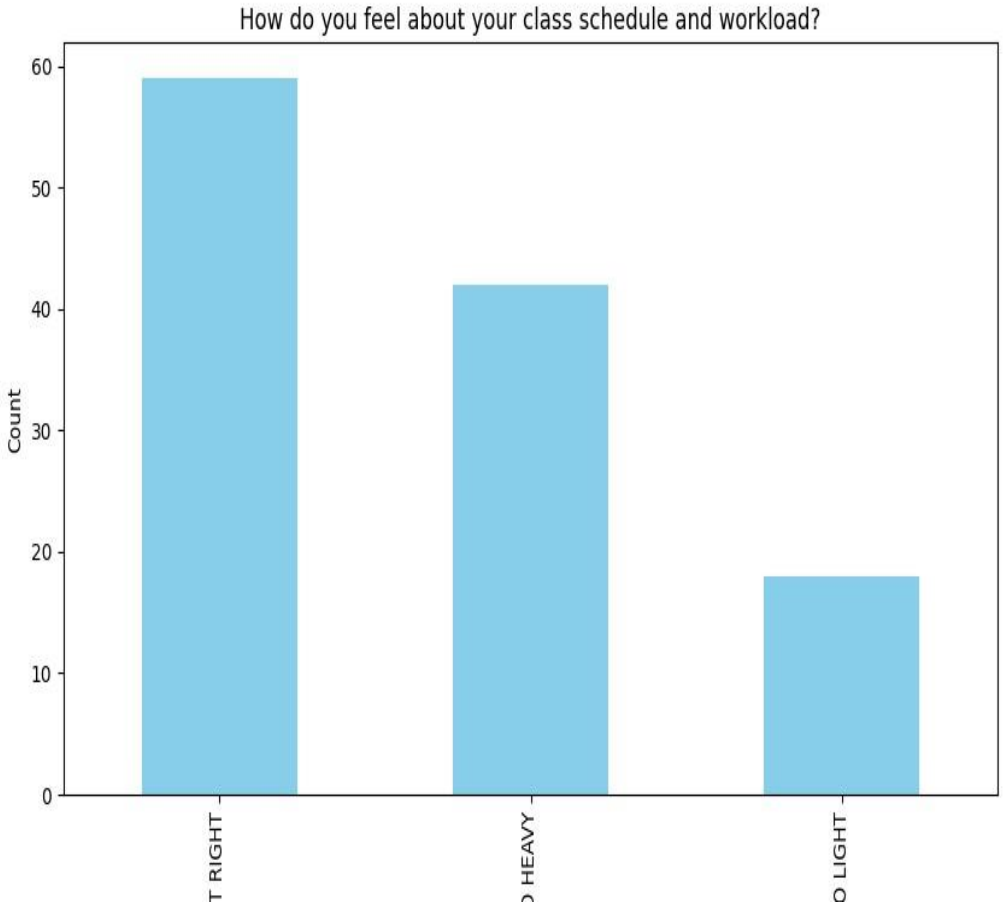
Data set 7. Are you genuinely interested in the course you`ve chosen?

```
value_counts = df["Are you genuinely interested in the course you've
chosen?"].value_counts() plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue') plt.title(f"Are you
genuinely interested in the course you've chosen?")
plt.xlabel('Brand') plt.ylabel('Count') plt.show()
 mode_value1= df["Are you genuinely interested in the course you've
chosen?"].mode()[0]  # Get the first mode (in case of multiple
modes) mode_count1= df["Are you genuinely interested in the course
you've chosen?"].value_counts()[mode_value1]
 total_values1 = len(df["Are you genuinely interested in the course
you've chosen?"]) mode_percentage1 = (mode_count1 / total_values1) * 100
 print(mode_value1,"
",mode_count1,mode_percentage1)
```



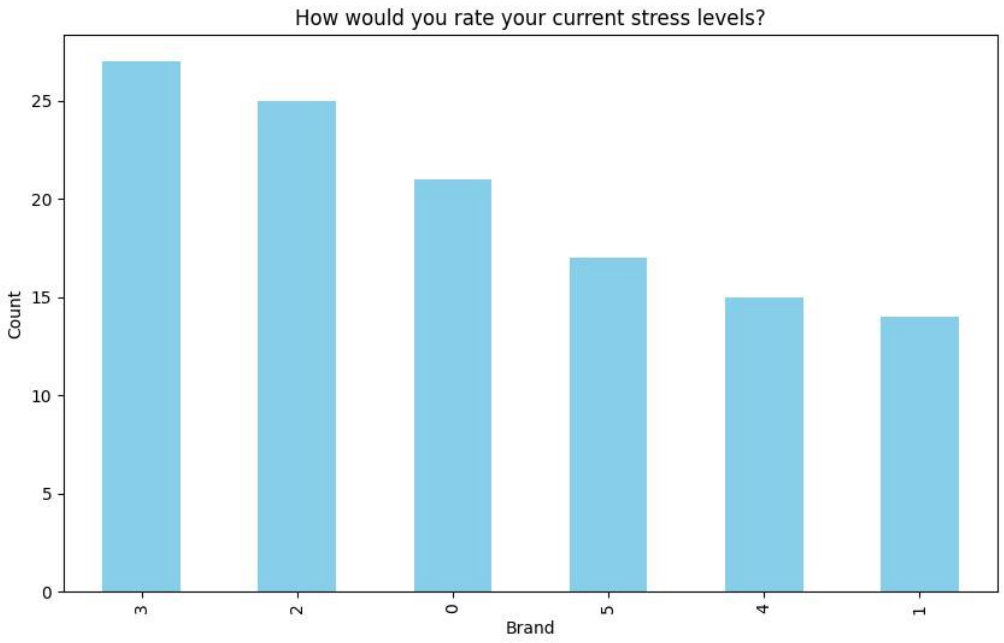
Data set 8. How do you feel about your class schedule and workload?

```
value_counts = df['How do you feel about your class schedule and
workload?'].value_counts() plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue') plt.title(f"How do you
feel about your class schedule and workload?")  plt.xlabel('Brand')
plt.ylabel('Count') plt.show()
```



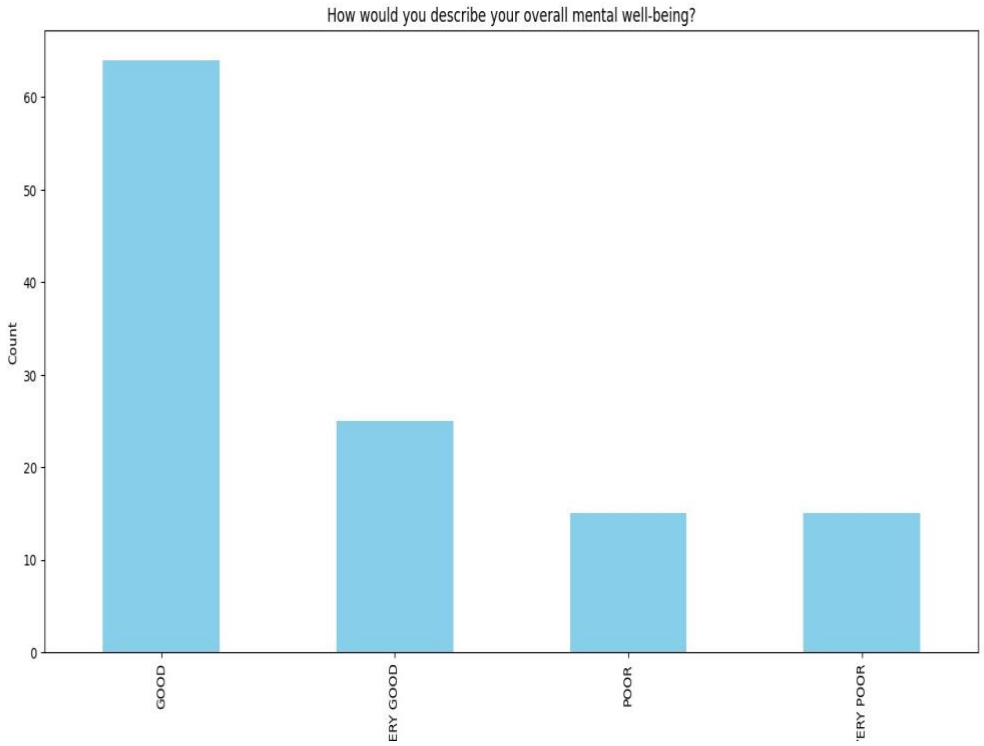
Data set 9. How would you rate your current stress levels?

```
value_counts = df["How would you rate your current stress levels?"].value_counts()
plt.figure(figsize=(10, 6)) value_counts.plot(kind='bar', color='skyblue')
plt.title(f"How would you rate your current stress levels?") plt.xlabel('Brand')
plt.ylabel('Count') plt.show()
```



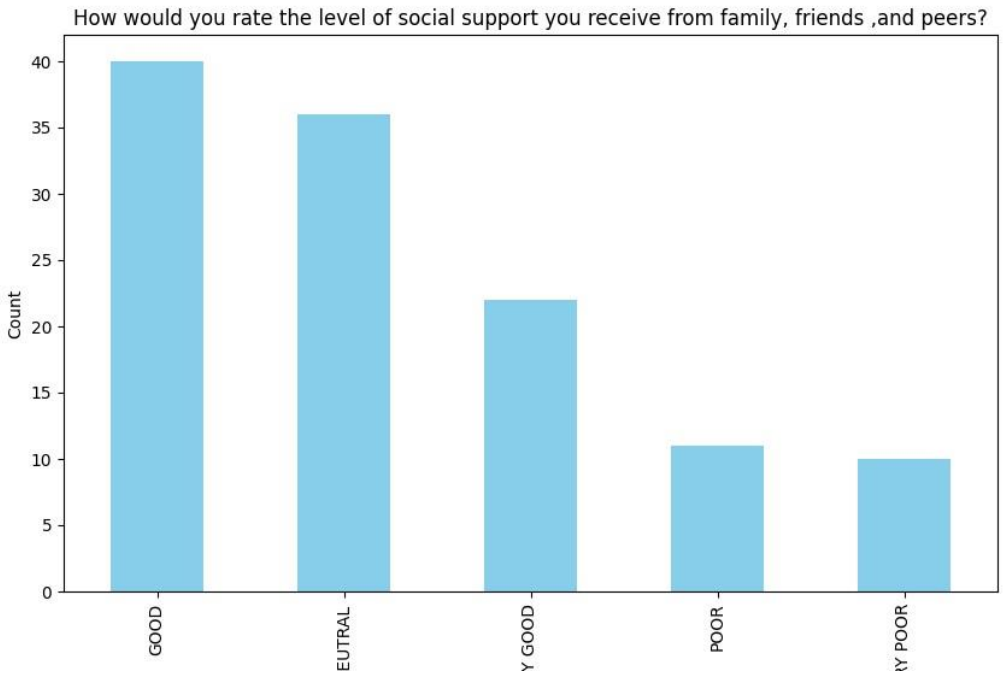
Data set 10.How would you describe your overall mental well-being?

```
value_counts = df["How would you describe your overall mental
wellbeing?"].value_counts() plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue') plt.title(f"How would
you describe your overall mental well-being?") plt.xlabel('Brand')
plt.ylabel('Count') plt.show()
```



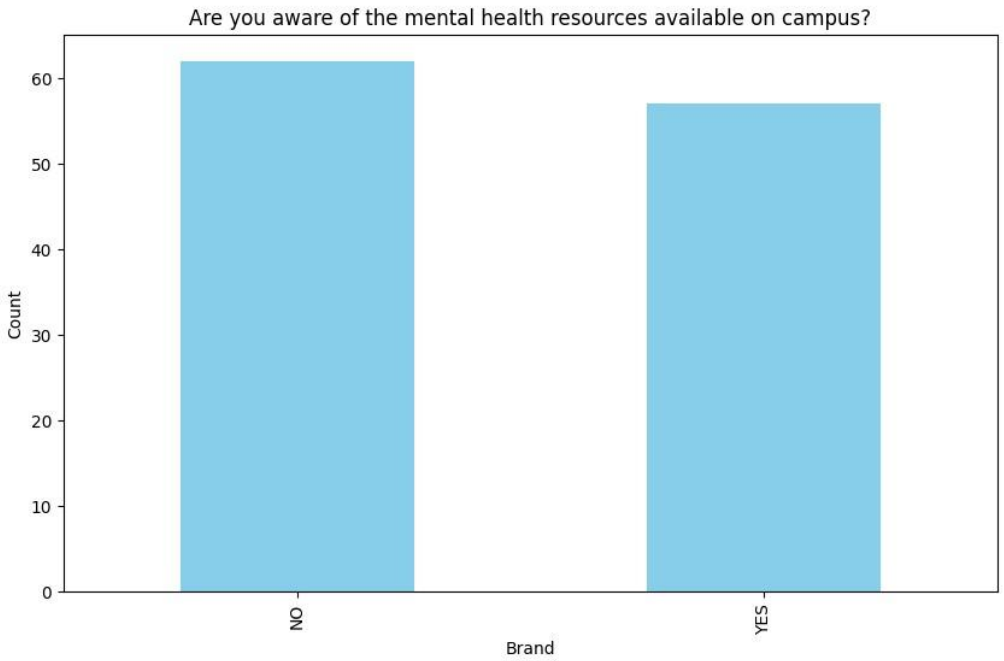
Data set 11.How would you rate the level of social support you receive from family, friends and peers?

```
value_counts = df["How would you rate the level of social support you receive from family, friends, and peers?"].value_counts() plt.figure(figsize=(10, 6)) value_counts.plot(kind='bar', color='skyblue') plt.title(f"How would you rate the level of social support you receive from family, friends ,and peers?") plt.xlabel('Brand') plt.ylabel('Count') plt.show()
```



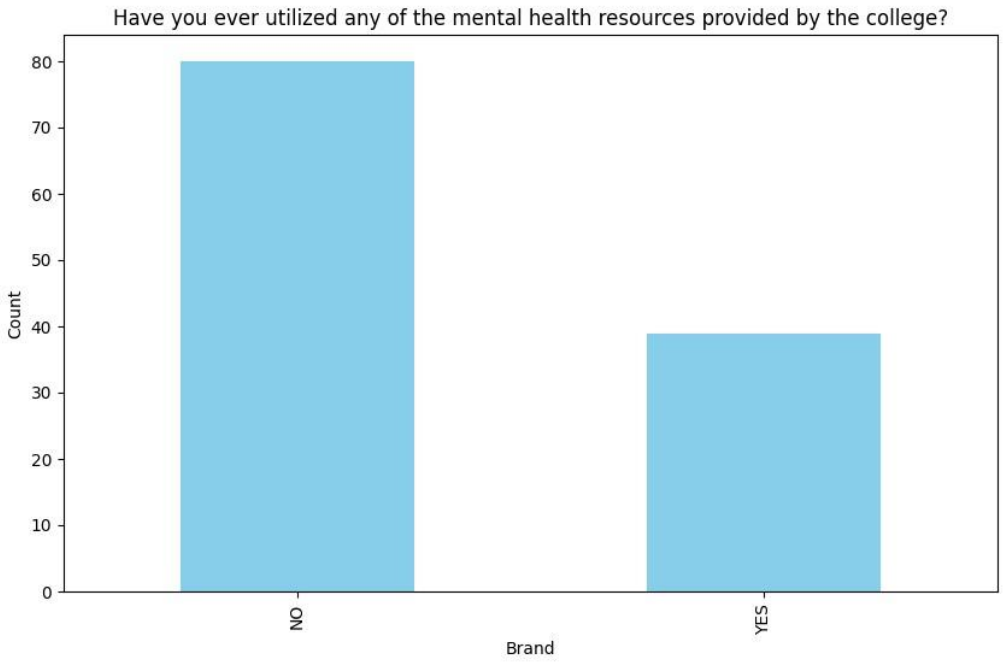
Data set 12.Are you aware of the mental health resources available on campus?

```
value_counts = df["Are you aware of the mental health resources available on campus?"].value_counts() plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue') plt.title(f"Are you aware of the mental health resources available on campus?") plt.xlabel('Brand')
plt.ylabel('Count') plt.show()
```



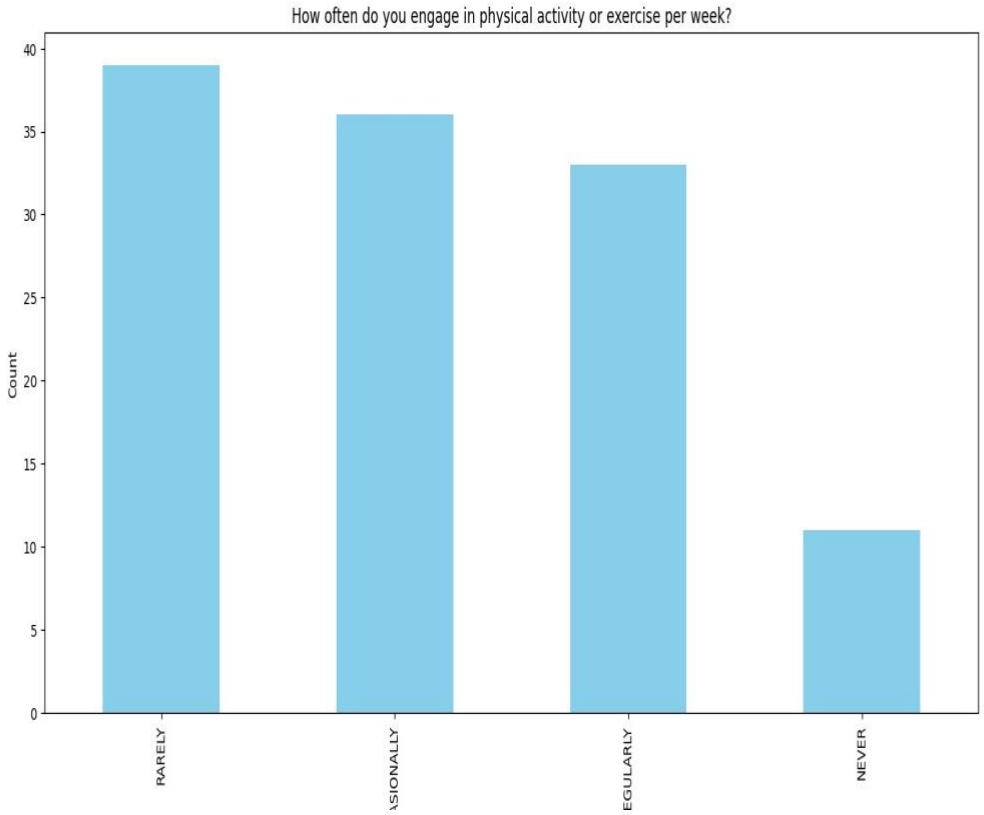
Data set 13. Have you ever utilized any of the mental health resources provided by the college?

```
value_counts = df["Have you ever utilized any of the mental health resources
provided by the college?"].value_counts() plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue') plt.title(f"Have you ever
utilized any of the mental health resources provided by the college?")
plt.xlabel('Brand') plt.ylabel('Count') plt.show()
```



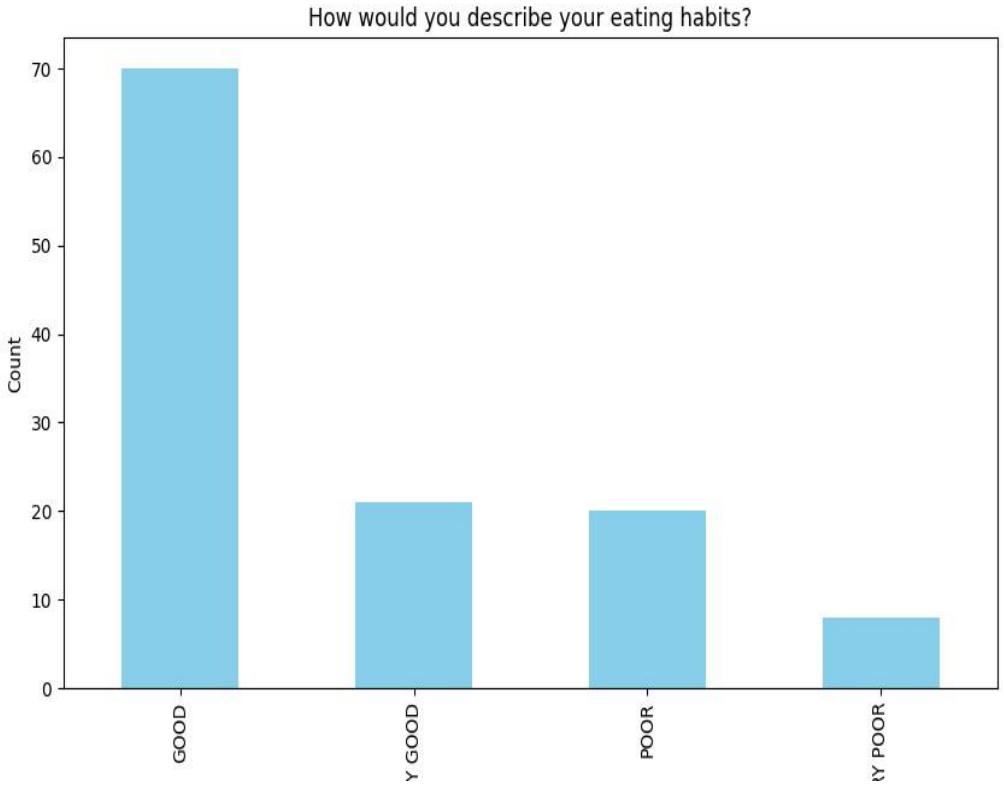
Data set 14. How often do you engage in physical activity or exercise per week?

```
value_counts = df["How often do you engage in physical activity or exercise per week?"].value_counts() plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue') plt.title(f"How often do you engage in physical activity or exercise per week?")
plt.xlabel('Brand') plt.ylabel('Count') plt.show()
```



Data set 15. How would you describe your eating habits?

```
value_counts = df["How would you describe your eating habits?"].value_counts()
plt.figure(figsize=(10, 6)) value_counts.plot(kind='bar', color='skyblue')
plt.title(f"How would you describe your eating habits?") plt.xlabel('Brand')
plt.ylabel('Count') plt.show()
```



Outcomes:

■ The project anticipates uncovering valuable insights into the factors influencing college student mental health. It is expected that the findings will provide a nuanced understanding of the challenges faced by students, enabling institutions to implement targeted interventions and support systems.

Reference

• McKiney, W. (2017). Python for Data Analysis: Data Wrangling with Pandas, NumPy and IPython. 2nd edition. O'Reilly Media. • O'Neil, C., & Schutt, R. (2013). Doing Data Science: Straight Talk from the Frontline O'Reilly Media https://www.w3schools.com/python/

https://www.geeksforgeeks.org/pandas-tutorial/

https://www.javatpoint.com/python-pandas