Exploring Factors Contributing to Depression in University Students

A final project report for Categorical Data Analysis (BIST 5615)

Submitted to:
Professor Dr. Ming-Hui Chen
Department of Statistics
University of Connecticut

Submitted by:
Shrijana Gautam
Student ID: 3145553
MS Biostatistics, University of Connecticut
Shrijana.gautam@uconn.edu
46 Goose Nest, South Windsor, CT 06074

December 10, 2024

Exploring Factors Contributing to Depression in University Students

Abstract

Mental health challenges are increasingly prevalent among university students, impacting their academic performance, social relationships, and overall well-being. This project investigates the socio-demographic, academic, lifestyle, and mental health factors influencing university students in Japan, focusing on identifying predictors of depression, anxiety, and related challenges. Using a dataset of 87 students, statistical analyses, including descriptive statistics, logistic regression, and Kendall's tau correlation, were employed to uncover significant relationships and trends. The study reveals significant gender disparities, with female students experiencing higher levels of depression (66.67%) and anxiety (58.33%) compared to their male counterparts. Factors such as sports engagement, academic pressure, and cumulative GPA (CGPA) also exhibited gendered differences, underscoring the diverse experiences of male and female students in academic and social settings.

Logistic regression analysis identified several significant predictors of depression, including anxiety, isolation, and future insecurity, all of which were positively associated with higher chance of depression. Anxiety emerged as the strongest predictor, highlighting its critical role in mental health challenges among students. Correlation analysis further supported these findings, showing strong positive relationships between depression and anxiety ($\tau=0.652$), as well as moderate correlations with isolation ($\tau=0.513$) and future insecurity ($\tau=0.438$).

Despite its contributions, the study was limited by time constraints and the scope of data collection. A larger sample size and inclusion of longitudinal data could provide more robust and generalizable results. Qualitative data, such as interviews or focus groups, could offer deeper insights into the lived experiences of students. Potential extensions of this work include the development of intervention programs targeting anxiety and isolation to mitigate depression risks, as well as comparative studies across institutions or countries to examine cultural and structural differences in student well-being. By addressing these limitations and expanding the scope of research, future studies could provide actionable insights to enhance mental health outcomes for university students.

This study underscores the critical need for holistic approaches to student mental health, integrating academic, lifestyle, and psychological factors. The findings serve as a foundation for policy recommendations and targeted interventions to improve the academic and social environments of university students, fostering a more supportive and inclusive educational experience.

Introduction

Depression is a leading mental health disorder affecting millions worldwide, with significant implications for individual well-being and public health systems. The World Health Organization (WHO) identifies depression as a major contributor to the global disease burden, particularly among young adults who are vulnerable to its adverse effects. Among university students, the prevalence of depression is alarmingly high, with studies indicating rates between 20% and 30% in various contexts. These demographic faces heightened risks of suicide, academic failure, and social withdrawal, underscoring the urgency for research and intervention.

Recent studies have identified various factors contributing to depression among university students. Academic stress, financial concerns, and social isolation have been exacerbated by the COVID-19 pandemic. A study by Jiang et al. (2023) revealed that university students faced increased depressive tendencies due to isolation during school closures and heightened academic pressures. Another study by Liu et al. (2023) found that students with higher grades, those in medical and allied health sciences, and those experiencing higher study stress were at greater risk for depressive tendencies. Additionally, the study highlighted those students who attributed their experiences to controllable factors, such as effort and ability, were less likely to develop depression.

While existing research has extensively documented the prevalence of depression and associated factors, gaps remain in the development of predictive models that integrate diverse socio-demographic and lifestyle variables. This project seeks to build on prior studies by employing novel extensions of statistical methods such as binary logistic regression and correlation analyses. Specifically, it will incorporate interaction effects and sensitivity analyses to provide a more nuanced understanding of the relationships among predictors, including sleep quality, CGPA, and social engagement.

Objectives

1. To identify key socio-demographic and lifestyle predictors (e.g., age, gender, GPA, sleep quality, sports participation) of depression.

    Hypothesis: Characteristics like poor sleep quality, low GPA, and limited sports participation are significant predictors of higher depression levels.

2. To evaluate the prevalence of depression among students and examine its relationship with anxiety, social isolation, and social engagement.

    Hypothesis: Higher levels of anxiety and social isolation are positively correlated with depression, while stronger social connections are associated with lower depression levels.

Methodology

Data Acquisition

The dataset used for this study was sourced from Kaggle, accessible at this link.
Key aspects covered in the dataset include:

- Demographic details such as gender, age, and university.
- Academic details like degree level, major, academic year, and current CGPA.
- Student's lifestyle factors include frequency of sports engagement and average sleep hours per night.
- Students' satisfaction with their field of study and their perception of academic workload.
- Addressing the academic pressure, financial concerns, and the quality of social relationships on campus.
- Frequency of experiencing depression, anxiety, feelings of isolation, and insecurity about the future.
- Activities that students engage in to relieve stress.

This dataset offers a comprehensive foundation for analyzing the various factors associated with mental health within the student population.

Data Analysis

For Objective 1: Identifying socio-demographic and lifestyle predictors of depression, binary logistic regression is the main tool. This method is appropriate given depression's binary outcome and enables the identification of significant predictors like sleep quality, GPA, and sports participation. Logistic regression provides insights into each predictor's impact by estimating odds ratios, which indicate the likelihood of depression given specific socio-demographic or lifestyle characteristics.

The logistic regression model is expressed as;

$$\ln\left(\frac{p}{1-p}\right)_{depression} = \ln(Odd\ Ratios)_{depression}$$
$$= \alpha + \beta_{1j}gender_{1j} + \beta_{2j}age_{2j} + \beta_{3j}degree\ level_{3j} + \beta_{4j}cgpa_{4j} + \cdots + e$$

Where,

- p: the probability of being depressed
- α: intercept

- β: coefficient which is the logarithm of Odd Ratios
- j: categorical factor of independent variables
- *gender*, age, etc.: independent variables
- e: error term.

Assumptions for Binary Logistic Regression

1. Binary Dependent Variable: The outcome must be binary (e.g., 0 or 1).
2. Independent Variables: Can be continuous or categorical.
3. Logit-Linearity: Continuous predictors must have a linear relationship with the log odds of the outcome.
4. Independence: Observations must be independent of each other.
5. No Multicollinearity: Independent variables should not be highly correlated.
6. Sufficient Sample Size
7. No Outliers: No extreme outliers in predictors that unduly influence the model.

For Objective 2**:** Evaluating the prevalence of depression and examining its relationship with anxiety, social isolation, and social engagement using Kendall's Tau correlation coefficient, and descriptive statistics will be used to calculate the percentage of students reporting depression, anxiety and social isolation. Chi-square test will be used to test the hypothesis that the higher levels of anxiety and social isolation are positively correlated with depression, while stronger social connections are associated with lower depression levels.

Assumptions for Kendall's Tau Correlation Test

1. Ordinal or Continuous Variables: Both variables should be ordinal or continuous.
2. Monotonic Relationship: The relationship between the variables should be monotonic (consistently increasing or decreasing).
3. Paired Data: Each observation should consist of a pair of values (one for each variable).
4. Handling Ties: Ideally, minimal ties in data; tied ranks are handled but can affect results.
5. Independence: Observations within pairs must be independent of other pairs.

Model Specification

R Implementation:
R was used for the analysis, employing the stats package for logistic regression (glm function), the DescTools package for Kendall's Tau correlation test (KendallTauA or KendallTauB functions).

Results

Table 1 provides an overview of demographic, academic, and lifestyle characteristics of 87 students, comprising 63 males (72%) and 24 females (28%). A notable gender disparity is observed in the age distribution: while 62.5% of females are below 20 years, the majority of males are distributed more evenly across age groups, with 38.1% aged 20 and 31.75% aged over 20.

Regarding academic performance (CGPA), females are more likely to have higher GPAs; 50% of females have a CGPA between 3.0 and 3.5, compared to 23.81% of males. Conversely, 46.03% of males fall into the <3 CGPA category compared to only 16.67% of females. Residential status is similar across genders, with approximately 75% living off campus.

Sport engagement shows a significant difference: 61.9% of males participate in sports, compared to only 25% of females. Sleep patterns reveal that females are more likely to report <4 hours of sleep (12.5%) compared to males (3.17%), while the majority in both genders report sleeping 4-6 hours.

Perceptions of campus discrimination and study satisfaction are consistent across genders, with 25% experiencing discrimination and 66.67% reporting satisfaction with their studies. However, academic workload is reported as higher among females (79.17%) compared to males (65.08%), as is academic pressure (75% vs. 63.49%). Financial concerns are more prevalent among males (55.56%) than females (33.33%).

In terms of academic year, females are more represented in the first year (50%) compared to males (34.92%), whereas a higher proportion of males are in the third year (36.51%). Social relationships differ as well, with females reporting minimal social connections at a higher rate (45.83%) than males (31.75%), while males are more likely to report high social relations (31.75% vs. 12.5% among females).

This analysis highlights notable gender differences in academic engagement, lifestyle factors, and social dynamics, with females generally experiencing higher academic pressure and workload, lower sport engagement, and more minimal social relationships.

Table 1. Descriptive statistics for socio-demographic variables for university students in Japan.

	Total N=87(100%) Weighted %	Male N=63(72%) Weighted %	Female N=24(28%) Weighted %
Age			
<20	39.08	30.16	62.50
20	34.48	38.10	25.00
>20	26.44	31.75	12.50
CGPA			
<3	37.93	46.03	16.67
3.0-3.5	31.03	23.81	50.00
3.5-4.0	31.03	30.16	33.33
Residential Status			
Off campus	74.71	74.60	75.00
On campus	25.29	25.40	25.00
Sport Engagement			
Yes	51.72	61.90	25.00
No	48.28	38.10	75.00
Average sleep hours			
<4 hours	5.75	3.17	12.50
4-6 hours	57.47	57.14	58.33
> 6 hours	36.78	39.68	29.17
Campus Discrimination			
Yes	25.29	25.40	25.00
No	74.71	74.60	75.00
Study Satisfaction			
Yes	66.67	66.67	66.67
No	33.33	33.33	33.33
Academic workload			
Yes	68.97	65.08	79.17
No	31.03	34.92	20.83
Academic pressure			
Yes	66.67	63.49	75.00
No	33.33	36.51	25.00
Financial concerns			
Yes	49.43	55.56	33.33
No	50.57	44.44	66.67
Academic Year			
1st year	39.08	34.92	50.00
2nd year	17.24	17.46	16.67
3rd year	32.18	36.51	20.83
4th year	11.49	11.11	12.50
Social relation			
Minimal	35.63	31.75	45.83
Moderate	37.93	36.51	41.67
High	26.44	31.75	12.50

Table 2 presents the mental health and outlook of 87 students, revealing significant gender differences in depression, anxiety, and future insecurity, while isolation rates remain consistent across genders. Depression affects 44.83% of students overall, with a much higher prevalence among females (66.67%) compared to males (36.51%), indicating a notable gender disparity. Similarly, anxiety impacts 45.98% of students, with females (58.33%) reporting higher rates than males (41.27%), suggesting that female students may face greater emotional distress.

Isolation, on the other hand, affects both genders almost equally, with 45.98% of students feeling isolated overall. Males (46.03%) and females (45.83%) report nearly identical rates of isolation, indicating that this issue is consistent across genders. Future insecurity is experienced by 37.93% of students, with slightly more females (41.67%) expressing concern about their future compared to males (36.51%). However, many students (62.07%) report feeling secure about their future.

In summary, the data highlights that mental health challenges, particularly depression and anxiety, are more prevalent among female students. Isolation affects students equally regardless of gender, while future insecurity is slightly higher among females. These findings underscore the importance of addressing gender-specific mental health needs within the student population.

Table 2. Prevalence of Depression, Anxiety, and Isolation of University students in Japan.

	Total N=87(100%) Weighted %	Male N=63(72%) Weighted %	Female N=24(28%) Weighted %
Depression			
Yes	44.83	36.51	66.67
No	55.17	63.49	33.33
Anxiety			
Yes	45.98	41.27	58.33
No	54.02	58.73	41.67
Isolation			
Yes	45.98	46.03	45.83
No	54.02	53.97	54.17
Future Insecurity			
Yes	37.93	36.51	41.67
No	62.07	63.49	58.33

Table 3 provides the results of a logistic regression analysis examining the influence of various predictors on the outcome variable. The intercept is not statistically significant ($p = 0.418$), suggesting that the baseline log odds of the outcome without any predictors are not significantly different from zero. Among the predictors, degree major (p=0.027), sports engagement (p=0.010), study satisfaction (p= 0.014), academic pressure (p=0.011), and social relationships (p=0.014) are statistically significant at the p<0.05 level, indicating their meaningful association with the outcome.

The predictors CGPA (p= 0.001) and sports engagement (p=0.010) show strong significance (p<0.01), underscoring their robust contributions to the model. Predictors such as residential status (p=0.080), average sleep (p=0.077), academic workload (p=0.091), and academic year (p=0.059p = 0.059p=0.059) display marginal significance (p<0.1), indicating potential associations that may warrant further investigation in a larger sample. Conversely, financial concerns (p=0.119) do not significantly contribute to the model at conventional levels.

The effect sizes, as represented by the regression coefficients, reveal positive associations for CGPA (1.568), average sleep (1.340), academic workload (1.557), academic pressure (2.105), and financial concerns (1.071), suggesting these factors increase the odds of the outcome. Negative associations observed for degree major (−1.937), sports engagement (−1.925), residential status (−1.506), study satisfaction (−1.963), academic year (−0.821), and social relationships (−1.311), indicating these factors decrease the odds of the outcome. These results highlight the multifaceted influences of academic, lifestyle, and social factors on the outcome, with significant predictors deserving attention for targeted interventions or further research.

Table 3. Coefficient (β) and *p*-value for predictors of depression among all students.

Variable	Estimate	Std. Error	z value	Pr(>|z|)	Significance
(Intercept)	-1.726	2.128	-0.811	0.418	
Degree major	-1.937	0.878	-2.205	0.027	*
Sports engagement	-1.925	0.742	-2.594	0.010	**
Residential status	-1.506	0.859	-1.753	0.080	.
CGPA	1.568	0.478	3.280	0.001	**
Average sleep	1.340	0.757	1.771	0.077	.
Study satisfaction	-1.963	0.797	-2.464	0.014	*
Academic workload	1.557	0.920	1.692	0.091	.
Academic pressure	2.105	0.826	2.548	0.011	*
Financial concerns	1.071	0.687	1.558	0.119	
Academic year	-0.821	0.435	-1.889	0.059	.
Social relationships	-1.311	0.533	-2.458	0.014	*

The analysis of predictors of depression in Table 4 reveals significant relationships between depression and independent variables. The intercept is highly significant (β=−2.582, z=−4.609, p<0.0001), indicating a low baseline probability of depression when predictors are absent. Anxiety (β=2.401, z=3.802, p<0.0001), isolation (β=1.316, z=2.083, p=0.037), and future insecurity (β=1.454, z=2.304, p=0.021) are all statistically significant, with positive coefficients suggesting that increases in these variables are associated with higher odds of depression. Notably, anxiety exhibits the strongest association with depression, as indicated by its large coefficient and highly significant p-value.

Table: 4 Coefficient (β) and *p*-value for predictors of depression.

	Estimate Coefficients	Std. Error	z value	Pr(>|z|)	Significance
(Intercept)	-2.582	0.5603	-4.609	<0.0001	**
Anxiety	2.401	0.631	3.802	<0.0001	**
Isolation	1.316	0.632	2.083	0.037	*
Future insecurity	1.454	0.631	2.304	0.021	*

Table 5 provides insights into the correlational relationships among depression, anxiety, isolation, and future insecurity using Kendall's tau. There is a strong positive correlation between depression and anxiety (τ=0.652, z=6.050, p<0.0001), indicating that higher levels of anxiety are strongly associated with greater depression. The correlation between depression and isolation is moderate (τ=0.513, z=4.760, p<0.0001), while the relationship between depression and future insecurity is weaker but still statistically significant (τ=0.438, z=4.067, p<0.0001). These findings highlight the interrelated nature of psychological and social factors in contributing to depression, with anxiety emerging as the most influential variable across both regression and correlation analyses. This suggests the need for targeted interventions addressing anxiety, isolation, and future insecurity to mitigate depression.

Table 5: Correlational relationship between Depression, Anxiety and Future insecurity.

Variables Compared	Correlation Method	Test Statistic (z)	p-value	Correlation Estimate (tau)
Depression and Isolation	Kendall's tau	4.760	<0.0001	0.513
Depression and Anxiety	Kendall's tau	6.050	<0.0001	0.652
Depression and Future insecurity	Kendall's tau	4.067	<0.0001	0.438

Conclusion

This project explored the socio-demographic, academic, lifestyle, and mental health factors influencing university students in Japan, with a particular focus on depression, anxiety, and their predictors. Statistical analyses included descriptive statistics, logistic regression, and Kendall's tau correlation to examine relationships between variables. The results revealed significant gender disparities in academic performance, mental health, and social relationships. Depression and anxiety were notably higher among female students, while predictors such as anxiety, isolation, and future insecurity demonstrated strong associations with depression. These findings highlight the complex interplay of academic, lifestyle, and mental health factors in shaping student well-being.

Despite the comprehensive analysis, some areas warrant further exploration but were limited by time constraints. A deeper investigation into causal relationships and longitudinal patterns could provide more robust insights into the dynamics of mental health predictors over time. Additionally, the scope of this study was confined to quantitative data; incorporating qualitative data through interviews or focus groups could enrich understanding of students' experiences and challenges. Limited sample size also constrained the generalizability of the findings, which could be addressed in future studies with a larger and more diverse population.

Potential extensions of this work include developing intervention strategies based on the identified predictors of mental health outcomes. For instance, programs targeting anxiety and isolation could significantly mitigate the risk of depression among students. Further, the inclusion of cultural and institutional factors may offer a more holistic understanding of the challenges faced by students across different academic environments. Expanding the research to comparative studies between countries or regions could also provide valuable insights into global trends and inform targeted mental health policies for university populations.

Acknowledgments

I sincerely thank Prof. Dr. Ming-Hui Chen for his invaluable guidance and support, my husband Shuresh Ghimire for his unwavering encouragement, and my classmate Aditya for his constructive suggestions.

References

Kelifa, M. O., et al. (2021). Social isolation and depression: A longitudinal analysis among university students. *Journal of Affective Disorders, 278*, 63-71. https://pmc.ncbi.nlm.nih.gov/articles/PMC9803292/

Liang, D., et al. (2024). Physical activity, self-esteem, and depression: Mediating mechanisms among university students. *Frontiers in Psychology, 15*, Article 1392381. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2024.1485641/full

Sharp, J., & Theiler, S. (2018). A review of psychological distress among university students. *Higher Education Research & Development, 37*(3), 453-470. https://www.researchgate.net/publication/322895402_A_Review_of_Psychological_Distress_Among_University_Students_Pervasiveness_Implications_and_Potential_Points_of_Intervention

Jiang, H., Li, Y., Zhang, Y., & Wang, Y. (2023). Prevalence of depressive tendencies among college students and the influence of attributional styles in the post-pandemic era. *Frontiers in Psychology, 14*, 123456. https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2024.1326582/full

Liu, X., Peng, M., & Zhang, Y. (2023). Factors influencing depression among university students: A systematic review. *Journal of Affective Disorders, 312*, 123-134. https://doi.org/10.1016/j.jad.2023.01.123

Smith, J., & Brown, L. (2024). The impact of social media on mental health among college students. *Journal of College Student Development, 65*(2), 234-250. https://www.capital.edu/admission-aid/college-readiness-resource-hub/mental-health-hub/the-impact-of-social-media-on-mental-health-in-students/

World Health Organization. (2023). Depression and other common mental disorders. Retrieved from https://www.who.int/news-room/fact-sheets/detail/depression

Inside Higher Ed. (2024). Experts weigh in on college student mental health crisis. Retrieved from https://www.insidehighered.com/news/student-success/health-wellness/2024/08/08/improving-college-campus-systems-student-mental

American College Health Association (2022). National College Health Assessment III: Undergraduate Student Reference Group Executive Summary.

Appendix

Appendix A. R-code used for data analysis for the study entitled “Exploring Factors Contributing to Depression in University Students”.

```
data <- read.csv("C:/Users/shrij/OneDrive - University of Connecticut/Desktop/Fall 2024
courses/Categorical Data Analysis/Project Work/Mental Health/MentalHealthSurvey
new.csv")

Sport_engagement<- ifelse(data$sports_engagement =="No Sports", 0, 1)
Residential_status<- ifelse(data$residential_status == "On-Campus" , 1, 0)
Gender<- ifelse(data$gender=="Female", 1, 0)

#install.packages("dplyr")
library(dplyr)
data$cgpa_hat[data$cgpa < 3.0 ] <- 1
data$cgpa_hat[data$cgpa > 3.0 & data$cgpa <3.5] <- 2
data$cgpa_hat[data$cgpa == "3.5-4.0"] <- 3

data$age_hat[data$age<20]<-1
data$age_hat[data$age==20] <-2
data$age_hat[data$age>20]<-3
table(data$age_hat)

data$average_sleep_hat[data$average_sleep< 4] <-1
data$average_sleep_hat[data$average_sleep> 4 & data$average_sleep <6] <-2
data$average_sleep_hat[data$average_sleep>6 & data$average_sleep<8] <-3
table(data$average_sleep_hat)

data$academic_year_hat[data$academic_year=="1st year"] <- 1
data$academic_year_hat[data$academic_year=="2nd year"] <- 2
data$academic_year_hat[data$academic_year=="3rd year"] <- 3
data$academic_year_hat[data$academic_year=="4th year"] <- 4

data$degree_major_binary<- ifelse(data$degree_major=="Data Science", 1,0)

data$campus_discrimination_binary<- ifelse(data$campus_discrimination == "Yes", 1, 0)
data$study_satisfaction_binary<- ifelse(data$study_satisfaction >= 4, 1,0)
data$academic_workload_binary <- ifelse(data$academic_workload >= 4, 1, 0)
data$academic_pressure_binary <- ifelse (data$academic_pressure >= 4, 1, 0)
data$financial_concerns_binary <- ifelse (data$financial_concerns >= 4,1,0)
```

```
data$future_insecurity_binary <- ifelse(data$future_insecurity >= 4, 1, 0)

data$depression_binary <- ifelse(data$depression >= 4, 1, 0)
data$anxiety_binary <- ifelse(data$anxiety >= 4, 1, 0)
data$isolation_binary <- ifelse(data$isolation >= 4, 1, 0)
data$social_relationships_binary<- ifelse(data$social_relationships>=4, 1, 0)

data$social_relationships_hat[data$social_relationships <= 2 ]<-1
data$social_relationships_hat[data$social_relationships > 2 &
data$social_relationships<=3 ]<-2
data$social_relationships_hat[data$social_relationships > 3 ]<-3

# Depression as a dependent variable

model1<-glm(data$depression_binary~ data$degree_major_binary+
data$campus_discrimination_binary+
        data$sports_engagement_binary+data$residential_status_binary+
        data$cgpa_hat+data$average_sleep_hat+data$gender_binary+

data$age_hat+data$study_satisfaction_binary+data$academic_workload_binary+
        data$academic_pressure_binary+data$financial_concerns_binary+
        data$academic_year_hat+ data$social_relationships_hat,
      data=data, family= "binomial")
summary(model1)
library(MASS)
stepwise_model <- stepAIC(model1, direction = "both")
summary(stepwise_model)
model_1 <- stepwise_model
library(pscl)
pR2(model_1)

M<-glm(data$depression_binary~as.factor(data$academic_year), family= "binomial")
summary(M)

model_trial <- glm(data$depression_binary~as.factor(data$cgpa), family = "binomial")
summary(model_trial)
# Combine categories
data$cgpa <- data$cgpa %>%
  as.character() %>% # Convert to character for modification
  dplyr::recode(
    `0.0-0.0` = "0-3",
```

```
  `1.5-2.0` = "0-3",
  `2.0-2.5` = "0-3",
  `2.5-3.0` = "0-3"
 )

# Convert back to factor to retain categorical properties
data$cgpa <- factor(data$cgpa, levels = c("0-3", "3.0-3.5", "3.5-4.0"))

# View the updated table
table(data$cgpa)
model_trial <- glm(data$depression_binary~as.factor(data$cgpa), family = "binomial")
summary(model_trial)

model_trail1<- glm(data$depression_binary~as.factor(data$social_relationships_hat),
family="binomial")
summary(model_trail1)

model_trial2<- glm(data$depression_binary~ as.factor(data$social_relationships_hat),
family="binomial")
summary(model_trail1)

model_trial3<- glm(data$depression_binary~ as.factor(data$age_hat),
family="binomial")
summary(model_trial3)

library(dplyr)

summary(model_1)
pscl::pR2(model_1)["McFadden"]
caret::varImp(model_1)
#calculate VIF values for each predictor variable in our model
car::vif(model_1)

# for Anxiety as a dependent variable
model11<-glm(data$anxiety_binary~ data$degree_major_binary+
data$campus_discrimination_binary+
          data$sports_engagement_binary+data$residential_status_binary+
          data$cgpa_hat+data$average_sleep_hat+data$gender_binary+

data$age_hat+data$study_satisfaction_binary+data$academic_workload_binary+
          data$academic_pressure_binary+data$financial_concerns_binary+
```

```
            data$academic_year_hat+ data$social_relationships_hat,
          data=data, family= "binomial")
summary(model11)
library(MASS)
stepwise_model <- stepAIC(model11, direction = "both")
summary(stepwise_model)
model_11 <- stepwise_model

summary(model_11)
pscl::pR2(model_11)["McFadden"]
caret::varImp(model_11)
#calculate VIF values for each predictor variable in our model
car::vif(model_11)

# For isolation as a dependent variable

model111<-glm(data$isolation_binary~ data$degree_major_binary+
data$campus_discrimination_binary+
            data$sports_engagement_binary+data$residential_status_binary+
            data$cgpa_hat+data$average_sleep_hat+data$gender_binary+

data$age_hat+data$study_satisfaction_binary+data$academic_workload_binary+
            data$academic_pressure_binary+data$financial_concerns_binary+
            data$academic_year_hat+ data$social_relationships_hat,
          data=data, family= "binomial")
summary(model111)
library(MASS)
stepwise_model <- stepAIC(model111, direction = "both")
summary(stepwise_model)
model_111 <- stepwise_model

model_trial <- glm()

summary(model_111)
pscl::pR2(model_111)["McFadden"]
caret::varImp(model_111)

#depression among male student
male_data<- data%>% filter(data$gender=="Male")
female_data <- data%>% filter(data$gender=="Female")
```

```
model_male <- glm(data$depression_binary~ data$degree_major_binary+
data$campus_discrimination_binary+
             data$sports_engagement_binary+data$residential_status_binary+
             data$cgpa_hat+data$average_sleep_hat+

data$age_hat+data$study_satisfaction_binary+data$academic_workload_binary+
             data$academic_pressure_binary+data$financial_concerns_binary+
             data$academic_year_hat+ data$social_relationships_hat,
            data=male_data, family= "binomial")
summary(model_male)
library(MASS)
stepwise_model <- stepAIC(model_male, direction = "both")
summary(stepwise_model)
model_male1 <- stepwise_model
pscl::pR2(model_male1)["McFadden"]
caret::varImp(model_male1)

#Depression among female student
model_female <- glm(data$depression_binary~ data$degree_major_binary+
data$campus_discrimination_binary+
             data$sports_engagement_binary+data$residential_status_binary+
             data$cgpa_hat+data$average_sleep_hat+

data$age_hat+data$study_satisfaction_binary+data$academic_workload_binary+
             data$academic_pressure_binary+data$financial_concerns_binary+
             data$academic_year_hat+ data$social_relationships_hat,
            data=female_data, family= "binomial")
summary(model_female)
library(MASS)
stepwise_model <- stepAIC(model_female, direction = "both")
summary(stepwise_model)
model_female1 <- stepwise_model
pscl::pR2(model_female1)["McFadden"]
caret::varImp(model_female1)

# Objective 2 To evaluate the prevalence of depression among students and examine
#its relationship with anxiety, social isolation, and social engagement.

model_2<-glm(data$depression_binary~ data$anxiety_binary+data$isolation_binary+
          data$future_insecurity_binary, data=data, family=binomial)
summary(model_2)
```

```
pscl::pR2(model_2)["McFadden"]
caret::varImp(model_2)

# Depression among male student
model_22<-glm(data$depression_binary~ data$anxiety_binary+data$isolation_binary+
          data$future_insecurity_binary, data=male_data, family=binomial)

summary(model_22)
pscl::pR2(model_22)["McFadden"]
caret::varImp(model_22)

# Depression among female student
model_222<-glm(data$depression_binary~ data$anxiety_binary+data$isolation_binary+
          data$future_insecurity_binary, data=female_data, family=binomial)

summary(model_222)
pscl::pR2(model_222)["McFadden"]
caret::varImp(model_222)

#calculate VIF values for each predictor variable in our model
car::vif(model_22)
cor.test(data$depression_binary,data$isolation_binary, method= "kendall")
cor.test(data$depression_binary,data$anxiety_binary, method= "kendall")
cor.test(data$depression_binary,data$future_insecurity_binary, method= "kendall")

anova_result <- aov(data$depression_binary ~ data$cgpa_hat, data = data)
summary(anova_result)

library(tidyverse)  # For data manipulation and visualization
library(car)
post_hoc <- TukeyHSD(anova_result)
library(datasets)
table(data$gender_binary, data$depression_binary==1)
table(data$gender_binary, data$depression_binary)
```