**Decoding Depression in College Students: A Data-Driven Approach**

**AUTHORS-KRISHNA LODHA AND ARYANEEL BHADURI**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**ABSTRACT**

This research project focuses on a comprehensive analysis of depression among college students, utilizing a data-driven methodology to identify and understand the various factors contributing to this pressing issue. Our study gathers extensive data from university students regarding their personal lives, encompassing a range of variables such as academic performance, social relationships, lifestyle habits, and mental health history. By employing statistical analysis and machine learning techniques, we aim to identify the key risk factors that correlate with higher instances of depression within this demographic.

In addition to pinpointing those students who are most at risk, our research will delve into the specific triggers that may precipitate or exacerbate depressive episodes. Understanding these triggers is crucial for developing effective intervention strategies that can support students who may be struggling with their mental health. Furthermore, we will analyse how various stressors, such as academic pressure, financial concerns, and social isolation, can impact students' mental well-being.

At the culmination of our research, we aim to offer actionable solutions and recommendations that can be implemented by educational institutions, mental health professionals, and the students themselves. These solutions will focus on promoting awareness, enhancing support systems, and fostering resilience among students. By addressing the complex interplay of factors influencing depression, this project seeks to contribute valuable insights and tools that can help create a healthier, more supportive academic environment for all college students.

**INTRODUCTION**

Depression among college students has emerged as a significant public health concern over the past few decades, reflecting broader societal shifts and increasing pressures faced by young adults. Historically, mental health issues in academic settings were often overlooked, with a stigma surrounding mental illness that discouraged open discussion and support. As enrolment in higher education has expanded, so too has the awareness of mental health challenges. Studies indicate that college students are at a higher risk for depression compared to their non-college peers, exacerbated by factors such as academic stress, social isolation, and financial burdens. Despite growing recognition of the issue, gaps remain in understanding the complex interplay of factors contributing to depression among this demographic. Existing research often focuses on specific populations or singular aspects of mental health, leaving a need for comprehensive, data-driven studies that consider the multitude of influences—academic, social, and personal—that affect student well-being. Additionally, many studies rely on self-reported data, which can introduce biases, highlighting the necessity for robust methodologies that include diverse data sources and analytical techniques.

Our research aims to address these gaps by conducting a comprehensive analysis of depression among college students. We will utilize a data-driven approach, gathering extensive information on students’ academic performance, social relationships, lifestyle habits, and mental health histories. Employing advanced statistical analysis and machine learning techniques, we seek to identify key risk factors and triggers that may precipitate or exacerbate depressive episodes.

By understanding these contributing factors, we aspire to develop actionable interventions that educational institutions, mental health professionals, and students themselves can implement. Our goal is to foster awareness, enhance support systems, and promote resilience within the student population. Through this research, we hope to provide valuable insights that can lead to a healthier, more supportive academic environment, ultimately contributing to the well-being and success of all college students.

**LITERATURE**

The prevalence of depression among college students has been the subject of extensive research in recent years. Studies indicate that approximately 30% of college students report experiencing significant depressive symptoms, with factors such as academic pressure, social isolation, and financial strain often cited as key contributors (Eisenberg et al., 2009; Kitzrow, 2003). Research has also explored the impact of lifestyle choices, such as sleep patterns and substance use, on mental health outcomes (Beiter et al., 2015).

Moreover, several studies have utilized quantitative methodologies, including surveys and statistical analyses, to identify correlations between academic performance and mental health (Cameron & Burch, 2017). For example, a study by Conley et al. (2013) found that students with lower GPAs were more likely to report depressive symptoms, highlighting the link between academic success and mental well-being. Additionally, qualitative research has provided insights into the lived experiences of students, revealing how societal pressures and stigma influence their willingness to seek help (Wang et al., 2019).

Despite the wealth of existing studies, notable gaps remain in the literature. Many studies tend to focus on specific populations, such as first-year students or those involved in specific academic programs, which limits the generalizability of findings. Furthermore, much of the existing research relies heavily on self-reported data, which can introduce biases and fail to capture the complexity of students' experiences.

Additionally, previous studies often investigate individual factors in isolation rather than examining the interplay between academic, social, and personal influences on mental health. This presents a need for comprehensive, data-driven analyses that consider the multifaceted nature of depression among college students. Our research seeks to fill these gaps by employing a holistic approach that integrates various data sources and analytical techniques.

To guide our research, we will employ the Biopsychosocial Model of Health, which posits that health outcomes are the result of an interplay between biological, psychological, and social factors. This framework is particularly relevant for understanding depression in college students, as it allows for an examination of how academic pressures (psychological), social relationships (social), and lifestyle habits (biological) collectively influence mental health.

By utilizing this theoretical framework, our study will systematically analyse the various contributing factors to depression among college students. This comprehensive approach not only addresses the gaps identified in previous research but also provides a nuanced understanding of how these factors interact, ultimately guiding the development of effective interventions.

In summary, our research aims to build on the existing literature by providing a thorough analysis of the multifaceted nature of depression among college students, addressing critical gaps, and applying a robust theoretical framework to inform our findings and recommendations.

**METHODOLOGY**

The methodology for analyzing depression risk among college students is structured around the Biopsychosocial Model of Health, which posits that health outcomes arise from the complex interplay of biological, psychological, and social factors. This holistic approach is particularly relevant in the context of college students, who face unique stressors that can significantly impact their mental health. The increasing prevalence of depression within this demographic has garnered attention in recent years, highlighting the need for thorough investigations that encompass various influences on well-being.

Initially, participants are recruited from diverse academic backgrounds through university channels and social media, ensuring a representative sample. This recruitment strategy aims to capture a wide range of experiences and perspectives, reflecting the multifaceted nature of student life. An online survey is designed to gather extensive information, covering demographics such as age, gender, and year of study, as well as biological factors (e.g., sleep patterns, physical activity, and substance use). Psychological factors include academic performance metrics, coping strategies, and personal mental health histories, while social factors address the quality of relationships, social support systems, and feelings of isolation.

Informed consent is obtained to ensure ethical standards are met, allowing participants to engage with the study while understanding their rights and the use of their data. Once data is collected, the next step involves creating a binary variable, Depression\_Risk, based on specific indicators related to lifestyle and mental health. This variable is crucial for defining the target in our analysis and allows for the identification of students who may be at higher risk for depressive symptoms.

The class distribution of the Depression\_Risk variable is then examined to understand the balance between the groups, as imbalances could skew model results. Features and the target variable are separated, with categorical variables encoded using LabelEncoder to convert them into numerical formats suitable for machine learning algorithms. This preprocessing step is essential, as many machine learning models require numerical input to function effectively.

Following this, the dataset is split into training and testing sets, maintaining stratification to preserve the class distribution. This ensures that both sets reflect the original sample characteristics, which is critical for building robust models. StandardScaler is applied to normalize the features, addressing potential disparities in data scales that could impact model performance, particularly in distance-based algorithms like KNN.

Logistic Regression is employed as one of the primary analytical techniques, allowing us to explore the relationship between predictors and the likelihood of depression risk. This model is particularly effective for binary outcomes and provides interpretable coefficients that can inform on the significance of each predictor. Predictions are made, and the model's performance is evaluated using accuracy, precision, recall, and F1-score, with the results presented in a classification report. This thorough evaluation enables us to assess how well the model identifies students at risk for depression, providing insights into potential intervention points.

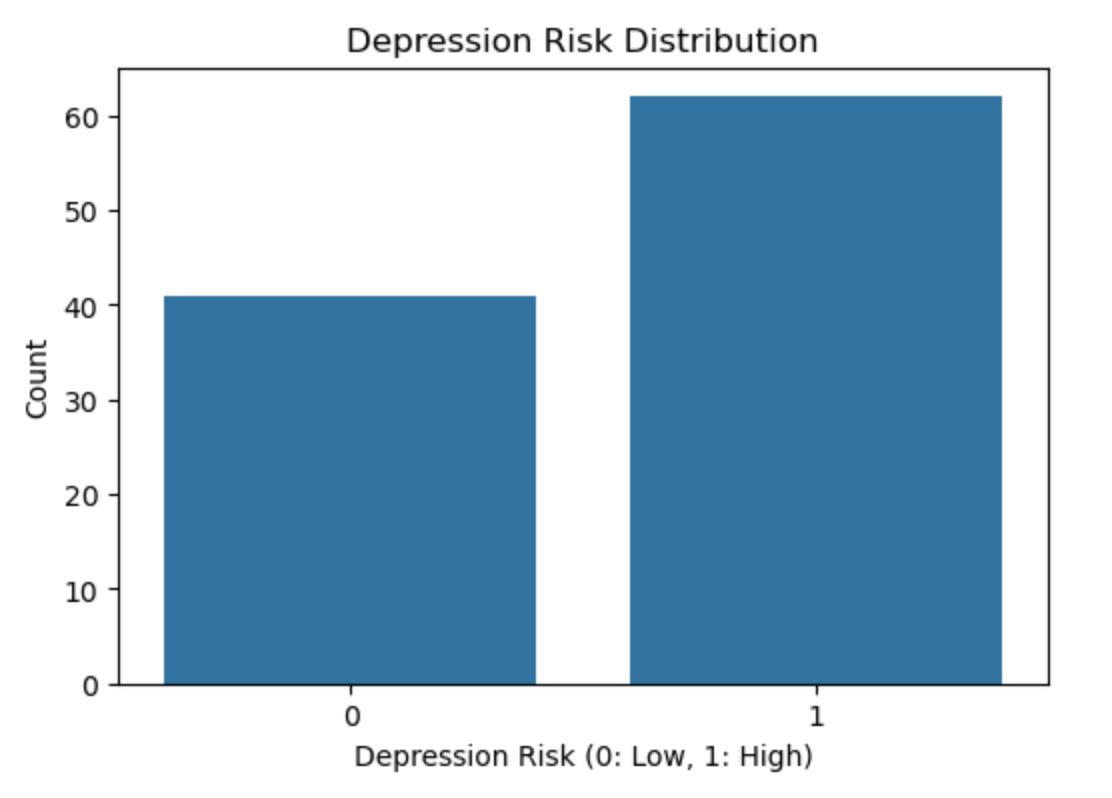
In addition to Logistic Regression, a K-Nearest Neighbors (KNN) classifier is implemented to compare results and further understand the data patterns. KNN is a non-parametric method that classifies data points based on their proximity to labeled instances in the feature space, making it particularly useful for detecting local structures within the data. After training the KNN model, its predictions are similarly assessed, offering an alternative perspective on the factors contributing to depression risk.

To provide additional insights into model performance, confusion matrices are generated for both models, visualizing true versus predicted classifications. This allows for a more nuanced understanding of where each model excels or struggles, facilitating targeted improvements in predictive accuracy. Moreover, visualizations such as heatmaps can be employed to identify relationships between different features and the Depression Risk, enriching the interpretative depth of the findings.

**FINDINGS**

In understanding depression, several psychological, emotional, and behavioral indicators are essential for assessing how closely an individual aligns with depressive symptoms. Depression extends beyond fleeting emotional states and represents a complex mental health condition that can severely impair an individual’s well-being and daily functioning. One of the most critical challenges in clinical practice is distinguishing between clinical depression and temporary sadness. Misdiagnosing either condition can have serious consequences—overdiagnosis of depression might result in unnecessary medical treatments, while underdiagnosis could delay essential interventions for those in need. Therefore, a nuanced and thorough approach to identifying depression is necessary to ensure individuals receive appropriate care and support.

From a clinical standpoint, duration plays a vital role in distinguishing depression from sadness. Depression does not emerge suddenly or pass within a few hours or days; instead, it is characterized by a persistent state of emotional and physical decline. Individuals experiencing depression may report symptoms such as hopelessness, extreme fatigue, a loss of interest in activities they once enjoyed, significant changes in appetite and sleep patterns, and difficulties in concentrating. For these symptoms to meet the diagnostic threshold, they need to persist consistently for at least two weeks. This extended duration ensures that mental health professionals differentiate depression from temporary emotional reactions that may result from life events, such as the loss of a loved one, stress at work, or personal conflicts. Temporary sadness, while often intense, tends to dissipate without clinical intervention, whereas depression typically requires therapeutic, medical, or psychosocial support to resolve.



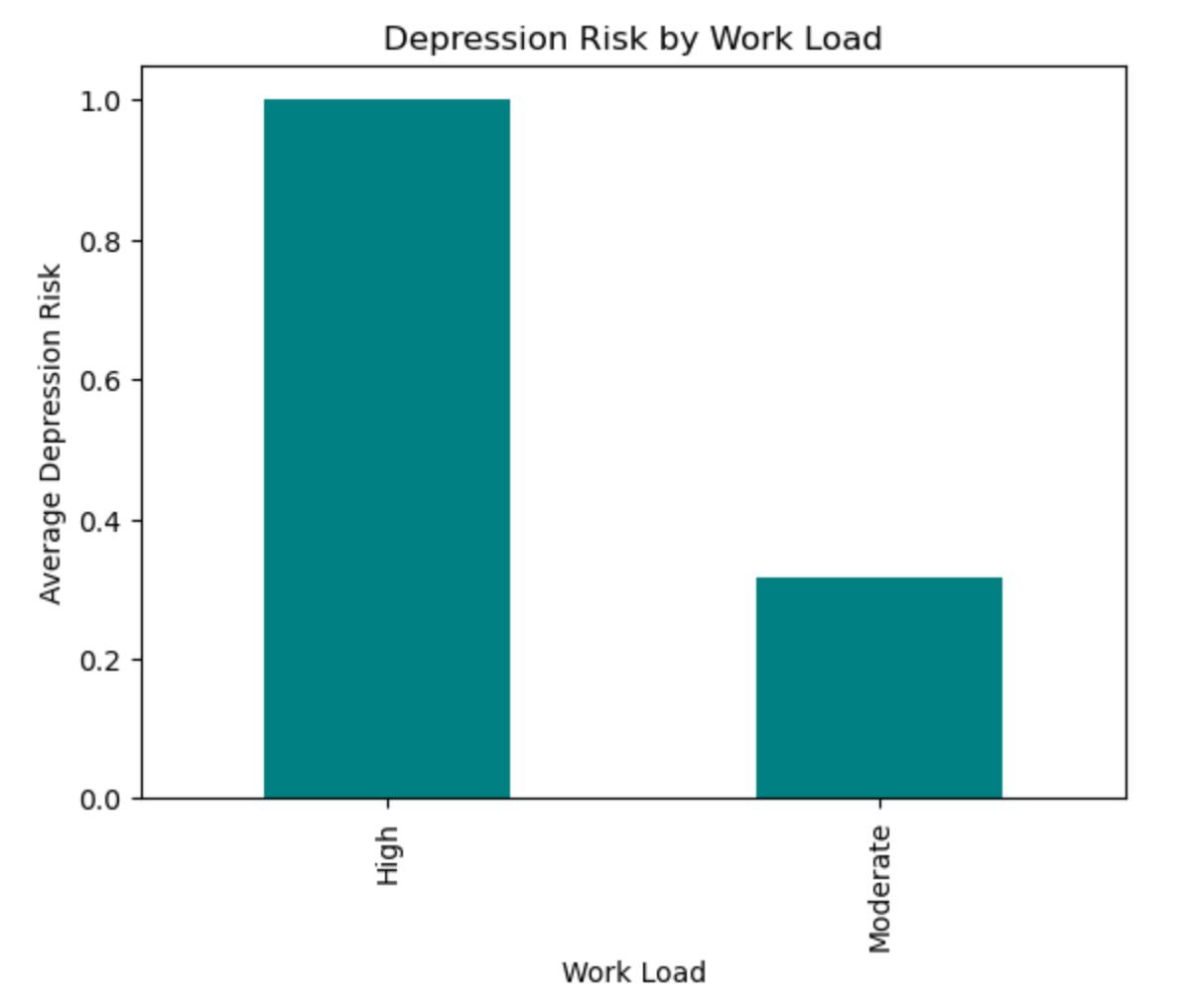
**FINDINGS FROM PSYCHIATRIST**

According to the psychiatrist consulted, the duration of symptoms is a key criterion for distinguishing depression from sadness. Depression is characterized by emotional and behavioral patterns that last for a minimum of two weeks or more. If the experience of sadness or distress does not persist for this duration, it is unlikely to be diagnosed as clinical depression. Instead, such shorter periods of emotional disturbance are more likely to reflect sadness or a stress-induced emotional reaction, which usually resolves naturally over time. This criterion is essential for mental health practitioners, as it prevents misdiagnosing temporary emotions as depression and ensures only those with sustained symptoms receive clinical treatment. Monitoring the individual over time also helps capture recurring depressive episodes, which may point to chronic conditions such as major depressive disorder or bipolar disorder.

In addition to the duration of symptoms, the psychiatrist highlighted the role of color preferences in mental health diagnostics. Contrary to popular beliefs, the psychiatrist emphasized that depression is not influenced by the colors a person prefers. This observation challenges the idea that certain color choices might directly indicate depressive states. However, while color preferences are not reliable indicators for diagnosing depression, the psychiatrist suggested they may provide subtle clues about other mental health conditions. For instance, individuals with heightened anxiety or mood instability might exhibit preferences for specific colors, which could reflect underlying emotional tendencies or psychological states. Although this information can be informative in broader psychological evaluations, it should be treated with caution and considered alongside other clinical assessments, such as interviews, behavioral observations, and self-reported symptoms.

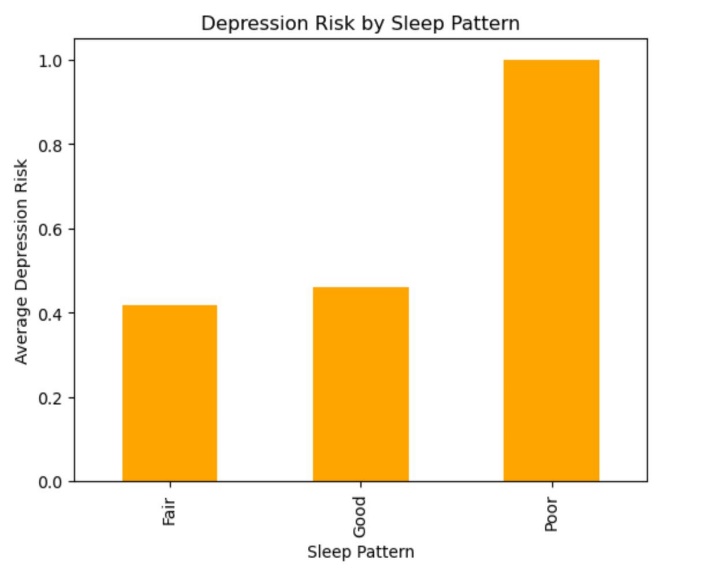
These insights from the psychiatrist highlight the importance of a multidimensional approach to mental health diagnosis. Relying on a single marker, such as color preferences or transient emotional states, could lead to diagnostic errors. Instead, clinicians must assess multiple dimensions of an individual’s experience, including symptom severity, duration, and behavioral patterns, over time. This comprehensive approach ensures that individuals struggling with genuine depressive symptoms receive the support they need, while those experiencing temporary sadness are not subjected to unnecessary medical interventions. It also reflects the broader challenges in mental health diagnostics, where complex conditions like depression require careful observation, professional judgment, and personalized care to promote recovery and well-being. literature but also contributes to a deeper understanding of the factors influencing depression among college students. By utilizing a robust theoretical framework and employing a range of analytical techniques, the research aims to guide effective intervention strategies that educational institutions and mental health professionals can implement. Ultimately, the goal is to foster a more supportive academic environment that prioritizes mental well-being, promoting resilience and success among college students.

**WORKLOAD AMONG STUDENTS**

Workload among students has become a significant factor influencing mental health, with mounting academic pressures contributing to increased stress and the risk of depression. In the context of our research, excessive workload—manifesting through demanding assignments, tight deadlines, exams, and extracurricular commitments—plays a critical role in the onset of depressive symptoms. Prolonged exposure to academic stress can disrupt sleep patterns, diminish motivation, and impair cognitive functions such as concentration and decision-making, which are also key markers of depression. Students experiencing a heavy workload often struggle to maintain a healthy work-life balance, leading to feelings of isolation, fatigue, and emotional exhaustion, which, if sustained for two weeks or more, align with clinical depression as described by the psychiatrist. While short-term stress may evoke temporary sadness, chronic academic stress has the potential to trigger long-term depressive episodes, reinforcing the need for early detection and intervention. Recognizing the connection between workload and mental health is essential for developing supportive environments in educational settings that promote student well-being and resilience against depression.

**SLEEP PATTERN**

Sleep patterns among students play a critical role in both their mental well-being and academic performance, with significant implications for depression. Irregular or insufficient sleep is common among students due to academic stress, social commitments, and screen time habits, often leading to chronic sleep deprivation. Research indicates that poor sleep quality—such as delayed sleep onset, fragmented sleep, or reduced total sleep duration—can increase the risk of developing depressive symptoms. Inconsistent sleep patterns disrupt the body’s circadian rhythms, affecting emotional regulation and cognitive functioning, which are essential for maintaining mental health. Depression, in turn, can further impair sleep, creating a vicious cycle where students experiencing depressive symptoms struggle to achieve restorative sleep, exacerbating their emotional difficulties. This bidirectional relationship between sleep disturbances and depression underscores the need for early intervention, including promoting sleep hygiene practices and mental health awareness among students. Identifying and addressing sleep-related issues can serve as a preventive measure, reducing the likelihood of depression and fostering healthier psychological outcomes.

From the graph we can infer that poor sleep often tends to higher depression risks in comparison to good and fair sleep pattern.

In the graph beside –

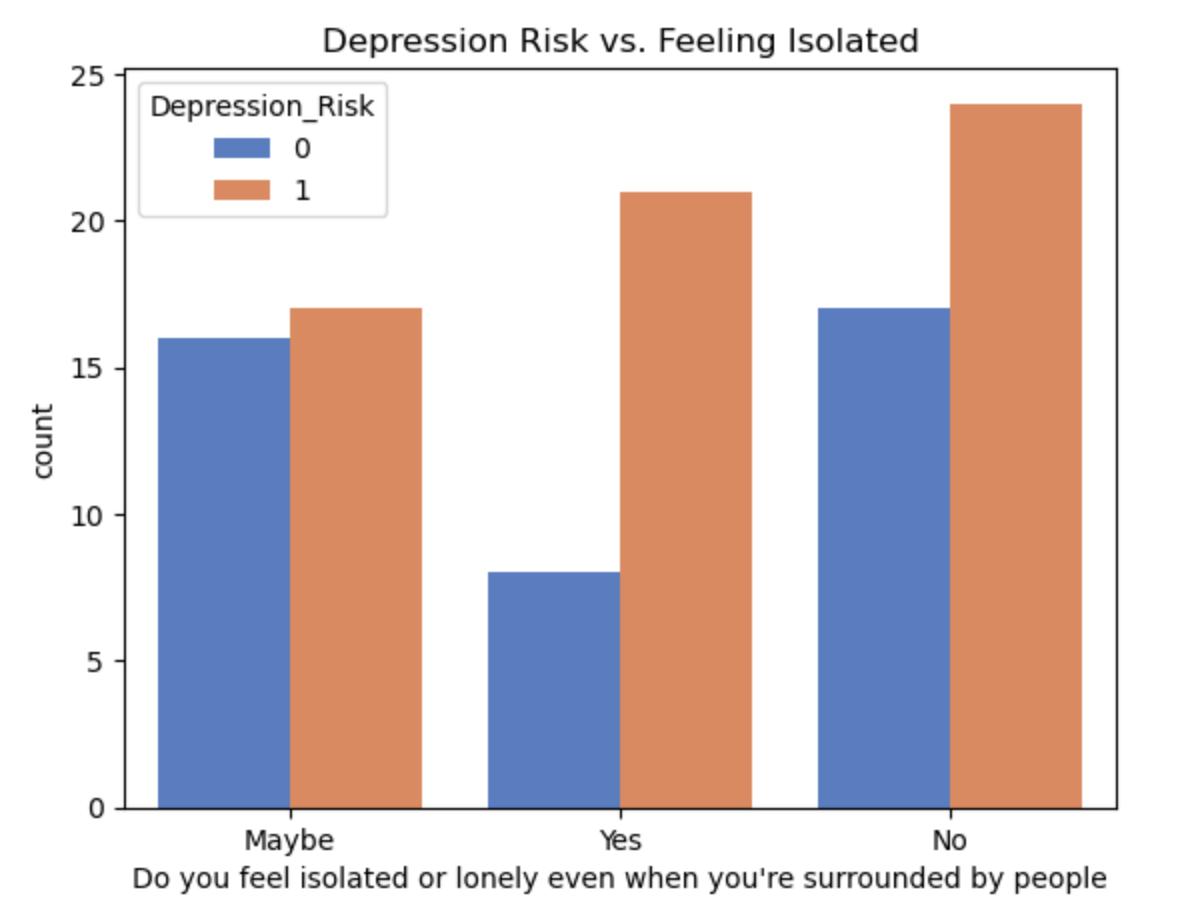
Scale of y axis- 20 percent

X axis-denotes the sleep pattern

The relationship between workload and sleep patterns is particularly significant, especially among students and working professionals, as increased workload often leads to disrupted sleep. When academic or professional demands escalate, individuals may sacrifice sleep to meet deadlines, prepare for exams, or manage responsibilities. This leads to shorter sleep duration, irregular sleep schedules, and poor sleep quality, all of which negatively impact cognitive performance, concentration, and emotional well-being. Over time, this imbalance between workload and sleep can contribute to mental health issues, including anxiety and depression, as well as reduced productivity and motivation. Establishing time management strategies and promoting work-life balance is essential to mitigate these effects, ensuring that adequate sleep remains a priority even during periods of high workload.

**ISOLATION**

Isolation and depression are closely interconnected, yet the relationship between them is more complex than simply the number of people surrounding an individual. While social isolation—characterized by limited social interactions and feelings of loneliness—can contribute to the onset or worsening of depression, the presence of people around an individual does not necessarily protect them from experiencing depressive symptoms. Depression is an internal condition driven by emotional, psychological, and neurochemical factors, and a person can feel isolated or disconnected even in a crowded environment or within their social circles. This highlights that the feeling of isolation is more subjective, rooted in an individual’s sense of emotional connection and support rather than physical proximity to others. Conversely, someone who spends time alone may not feel isolated if they have meaningful relationships and emotional stability. This demonstrates that depression’s causes are complex and multifaceted, and it cannot be solely attributed to external factors like social interaction or isolation, but rather how one perceives and processes these experiences.



Isolation can significantly contribute to the development of depression among students who refrain from engaging in social interactions and fail to communicate their feelings with those around them. This phenomenon can be understood through several psychological and social frameworks.

Firstly, social isolation disrupts the essential human need for connection and support. Humans are inherently social beings, and meaningful relationships are crucial for emotional well-being. When students withdraw from their peers, they deprive themselves of the emotional support that can mitigate stress and foster resilience. The absence of social interaction can exacerbate feelings of loneliness, leading to a cycle of withdrawal and increased emotional distress.

Secondly, the lack of open communication can hinder the processing of emotions. When students do not express their thoughts and feelings, they may internalize negative emotions, which can lead to a distorted self-perception and an overwhelming sense of hopelessness. This internalization can contribute to the onset of depressive symptoms, as unaddressed feelings of sadness and frustration accumulate over time.

Additionally, the cognitive patterns associated with isolation can reinforce depressive symptoms. Students may develop negative thought patterns, characterized by pessimism and self-criticism, which can be compounded by their isolation. The absence of positive social interactions prevents the challenge of these cognitive distortions, making it more difficult for students to cultivate a positive outlook on life.

In summary, the interplay of social isolation, lack of communication, stigma surrounding mental health, and cognitive patterns can significantly contribute to the development of depression among students. Addressing these factors through increased awareness, support systems, and open dialogue is essential in mitigating the adverse effects of isolation on mental health.

