

Predicting Depression in University Students: The Role of Anxiety

Parsa Rismanchi

2025-11-25

Abstract

Depression and anxiety are two of the most prevalent mental-health concerns among university students, often co-occurring and mutually reinforcing. Understanding how anxiety relates to depression can support early identification of at-risk students and guide preventative mental-health services on campus. In this mini project, I used a publicly available dataset of 101 university students to examine whether anxiety is associated with higher rates of depression. After excluding missing data, I compared depression prevalence between anxious and non-anxious students using descriptive statistics, a chi-square test of independence, and a logistic regression model. Results showed a clear pattern: anxious students had substantially higher depression rates (52.9%) compared with non-anxious peers (25.4%). This association was statistically significant, $\chi^2(1) = 6.40$, $p = .0114$, and a logistic model indicated that anxiety more than tripled the odds of depression (OR = 3.31, 95% CI [1.40, 8.03]). Findings replicate well-established comorbidity patterns and highlight anxiety screening as a practical strategy for early detection of depression among students.

1. Introduction

Depression and anxiety represent two of the most common psychological problems experienced by university students. Academic pressure, financial stress, social transition, and emerging adulthood all contribute to elevated vulnerability during this period. Numerous studies have documented that anxiety and depression frequently co-occur, with shared cognitive, emotional, and biological mechanisms. For example, both conditions involve negative affectivity, rumination, and difficulties regulating stress responses. In clinical settings, the presence of anxiety symptoms is often considered an early warning sign for depressive disorders.

Given this evidence, an important question is whether self-reported anxiety can meaningfully predict depression risk within non-clinical student populations. This question is relevant for mental-health services on university campuses, which often operate with limited resources and rely on brief screenings to identify students who may benefit from early intervention.

The present project aims to replicate the well-established association between anxiety and depression using a publicly available dataset from Kaggle (Islam, 2023). Although the dataset is simple and based on self-report, it provides a useful opportunity to examine whether anxious students display higher depression rates and to quantify the magnitude of this association using statistical tests commonly used in psychological research.

The study follows a straightforward analytic plan: (1) compare depression prevalence between anxious and non-anxious groups, (2) statistically test the difference using a chi-square test of

independence, and (3) estimate the odds ratio using logistic regression. Beyond replicating known patterns, the project also demonstrates reproducible research practices using R.

2. Methods

Data Source

Data were obtained from the *Student Mental Health* dataset published on Kaggle by Shariful Islam (2023). The dataset includes self-reported information about students' mental-health symptoms, including whether they experience anxiety and/or depression.

Variables

Two binary variables were used:

- **Anxiety:** Yes / No
- **Depression:** Yes / No

Both were recoded numerically (1 = Yes, 0 = No).

Sample and Exclusions

Rows with missing values in either variable were removed, resulting in a final analytic sample of **101 students**.

Analytic Approach

The analysis proceeded in three steps:

1. **Descriptive statistics** to compute depression prevalence in both groups.
2. **Chi-square test** to evaluate group differences.
3. **Logistic regression** to compute odds ratios and confidence intervals.

All analyses were conducted in R using the tidyverse ecosystem.

3. Results

Table 1: Table 1. Percentage of Depressed Students by Anxiety Status (n = 101)

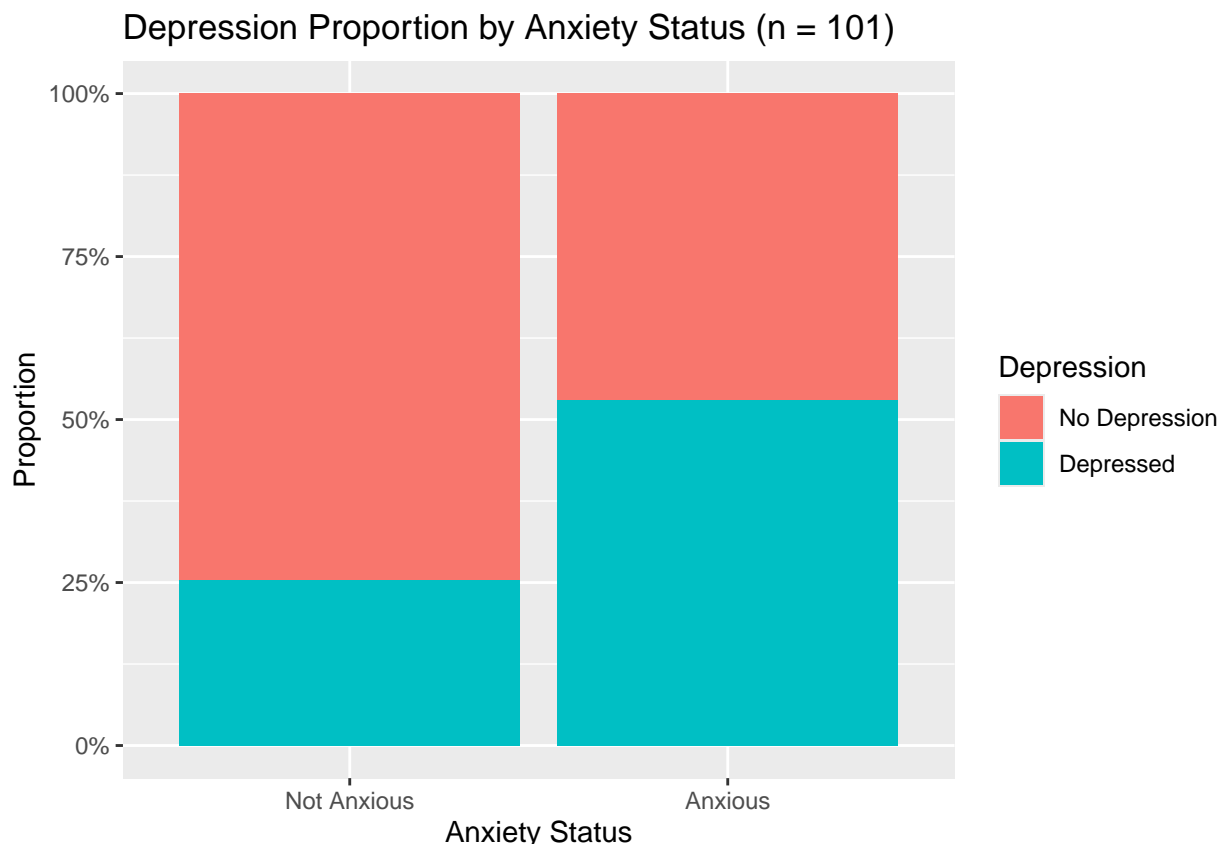
Anxiety Status	Depressed (n)	Group Size (n)	Percent Depressed
Not Anxious	17	67	25.4
Anxious	18	34	52.9

Descriptive results showed clear differences between the two groups. Among non-anxious students, 17 out of 67 (25.4%) reported depression. In contrast, among anxious students, 18 out of 34 (52.9%)

reported depression. These percentages indicate that symptoms of anxiety were associated with roughly double the rate of depression compared with peers without anxiety symptoms.

A chi-square test confirmed that this difference was statistically significant, $\chi^2(1) = 6.40$, $p = .0114$, indicating that anxiety and depression were not independent in this sample.

The logistic regression model indicated that anxious students were more than **three times** as likely to report depression (OR = 3.31, 95% CI [1.40, 8.03], $p = .007$). Taken together, the descriptive, inferential, and regression analyses consistently demonstrate a strong association between anxiety and depression in this dataset.



4. Discussion

The findings of this mini project replicate a well-documented pattern in clinical and health psychology: anxiety is strongly associated with elevated depression risk. The observed effect size aligns with prior literature illustrating substantial comorbidity between the two conditions. Models of negative affectivity and stress-diathesis frameworks explain how shared cognitive tendencies or heightened reactivity to stress can increase vulnerability to depressive symptoms among individuals with anxiety.

From a clinical perspective, these results highlight the importance of early identification. University counseling centers often operate under heavy demand, making efficient screening processes critical. Brief anxiety assessments could help identify students who may need further evaluation or preventative support. This is particularly valuable in university settings, where students may delay help-seeking or lack awareness of worsening symptoms.

The study has several limitations: cross-sectional design, reliance on self-report, and lack of relevant covariates such as stress levels, academic workload, personality traits, or social support. Still, the findings provide a clear, reproducible demonstration of how anxiety can serve as a marker of depression risk in student populations.

5. Conclusion

This mini project shows that university students reporting anxiety symptoms have substantially higher odds of experiencing depression. Although limited by its self-report, cross-sectional design, the analysis replicates established patterns of comorbidity and demonstrates how basic statistical tools can yield clinically relevant insights. Increasing access to anxiety screening and early intervention may help reduce depression rates among students.

References

Islam, S. (2023). *Student Mental Health Dataset*. Kaggle. (See dataset page for access and license details; accessed 25 November 2025.)

Reproducibility & Integrity

This is an independent replication study using a public dataset (Islam, 2023). All code, analysis, and reporting were conducted by **Parsa Rismanchi**. No part of this work was copied from existing notebooks. Full reproducibility is ensured via the provided R Markdown file and session info below.

Data Availability

- **Source:** Kaggle – Student Mental Health (Shariful Islam, 2023)
- **License:** See dataset page for details
- **Ethics:** Public anonymized dataset; no human-subjects involvement

Session Info

```
## R version 4.5.1 (2025-06-13 ucrt)
## Platform: x86_64-w64-mingw32/x64
## Running under: Windows 10 x64 (build 19045)
##
## Matrix products: default
##   LAPACK version 3.12.1
##
## locale:
## [1] LC_COLLATE=English_United States.utf8
## [2] LC_CTYPE=English_United States.utf8
## [3] LC_MONETARY=English_United States.utf8
```

```

## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.utf8
##
## time zone: Asia/Tehran
## tzcode source: internal
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## other attached packages:
## [1] broom_1.0.10    knitr_1.50      lubridate_1.9.4 forcats_1.0.1
## [5] stringr_1.6.0   dplyr_1.1.4     purrr_1.1.0     readr_2.1.5
## [9] tidyr_1.3.1     tibble_3.3.0    ggplot2_4.0.0   tidyverse_2.0.0
##
## loaded via a namespace (and not attached):
## [1] bit_4.6.0      gtable_0.3.6    crayon_1.5.3    compiler_4.5.1
## [5] tidyselect_1.2.1 parallel_4.5.1   scales_1.4.0    yaml_2.3.10
## [9] fastmap_1.2.0  R6_2.6.1        labeling_0.4.3   generics_0.1.4
## [13] backports_1.5.0 pillar_1.11.1    RColorBrewer_1.1-3 tzdb_0.5.0
## [17] rlang_1.1.6    stringi_1.8.7    xfun_0.54        S7_0.2.0
## [21] bit64_4.6.0-1 timechange_0.3.0 cli_3.6.5        withr_3.0.2
## [25] magrittr_2.0.4 digest_0.6.37    grid_4.5.1       vroom_1.6.6
## [29] rstudioapi_0.17.1 hms_1.1.4        lifecycle_1.0.4  vctrs_0.6.5
## [33] evaluate_1.0.5 glue_1.8.0       farver_2.1.2     rmarkdown_2.30
## [37] tools_4.5.1    pkgconfig_2.0.3  htmltools_0.5.8.1

```