

A Deep Dive into the Interplay of Depression, Sleep Deprivation, and Self-Esteem: An Application of an In-depth Advanced Statistical Analysis

Muslihat Adejoke Gaffari and Mostafa Zahed, Mathematics and Statistics Department, East Tennessee State University, TN, USA.

ABSTRACT

Depression, intertwined with sleep deprivation and self-esteem, presents a significant challenge to mental health worldwide. The research shown in this paper employs advanced statistical methodologies to unravel the complex interactions among these factors. Through log-linear homogeneous association, multinomial logistic regression, and generalized linear models, the study scrutinizes large datasets to uncover nuanced patterns and relationships. By elucidating how depression, sleep disturbances, and self-esteem intersect, the research aims to deepen understanding of mental health phenomena. The study clarifies the relationship between these variables and explores reasons for prioritizing depression research. It evaluates how statistical models, such as log-linear, multinomial logistic regression, and generalized linear models, shed light on their intricate dynamics. Findings offer insights into risk and protective factors associated with these variables, guiding tailored interventions for individuals in psychological distress. Additionally, policymakers can utilize these insights to develop comprehensive strategies promoting mental health and well-being at a societal level.

Keywords: Depression, Sleep Deprivation, Self-Esteem, Mental Health, Log-Linear Model, Multinomial Logistic Regression Model, Generalized Linear Model.

INTRODUCTION

Sadness is a main part of the human condition. Feeling down and a little blue is sometimes a temporary experience connected to specific events. A sense of sadness or hopelessness can be steadier, which is well-known as depression. Depression is a severe problem that can affect people's lives. What they think and feel will affect their ability to sleep [1].

Depression, also known as Major Depressive Disorder (MDD) or clinical depression, is a common and severe mental health condition characterized by persistent and pervasive feelings of sadness, hopelessness, and a lack of interest or pleasure in most activities. It goes beyond the regular ups and downs that people experience in life. Depression can affect a person's thoughts, feelings, and physical well-being, often leading to a range of emotional and physical symptoms [2].

There are different types of depression, some of which develop due to specific circumstances, which include symptoms of depressed mood or loss of interest, most of the time for at least two weeks. For example, *Persistent Depressive Disorder*, also called Dysthymia or Dysthymic disorder, consists of less severe symptoms than other Major Depressive Disorders. Still, it is longer lasting, usually up to 2 years. *Perinatal Depression* occurs during or after pregnancy. *Seasonal Affective Disorder* comes and goes with the seasons. Depression with Symptoms of Psychosis is a severe form of depression in which a person experiences psychosis symptoms such as delusions or hallucinations [2].

There are some common causes of depression, such as a brain chemistry imbalance, drug abuse, stress, female sex hormones, and physical health problems [3]. Depression symptoms can include suicidal thoughts or attempts, trouble concentrating, a loss of interest in things that previously brought joy, insomnia, and irritability [4]. Depression also affects the body physically. For instance, an increased risk of heart disease, a weakened immune system, impaired digestion and the nervous system, a lowered libido, and a lowered tolerance for pain are possible signs of depression [5].

The degrees of depression can vary from moderate to severe. Whether it is mild or severe, it can be treated with either medication, psychotherapy, or a combination of the two. Antidepressants are used to treat depression, which works by changing how the brain produces or processes certain chemicals involved in mood or stress. Psychotherapy or talk therapy is sometimes used alone for the treatment of mild depression. For moderate to severe depression, psychotherapy is often implored along with antidepressant medications. If medication or psychotherapy does not reduce the symptoms of depression, brain stimulation therapies can be tried. There are different types of brain stimulation therapies, such as Electro-Convulsive Therapy (ECT), Repetitive Transcranial Magnetic Stimulation (RTMS), Vagus Nerve Stimulation (VNS), Magnetic Seizure Therapy (MST), and Deep Brain Stimulation (DBS). ECT and RTMS are the most widely used brain stimulation therapies [2].

This research examines two potential causes of depression, which are trouble sleeping and feeling bad about oneself. Trouble sleeping, also known as insomnia, is most often described as a subjective complaint of poor sleep quality or quantity despite adequate time for sleep, resulting in daytime fatigue, irritability, and decreased concentration.

Insomnia is classified as idiopathic or comorbid. Comorbid insomnia is associated with psychiatric disorders, medical disorders, substance abuse, and specific sleep disorders. Idiopathic insomnia is essentially a diagnosis of exclusion. There are some causes of insomnia, such as poor sleep habits, stress, and anxiety [6].

On the other hand, sleeping too much or hypersomnia is the inability to stay awake and alert during the day despite having more than adequate nighttime sleep [7]. The treatment of sleep disorders (insomnia or hypersomnia) varies from moderate to severe. Medications, lifestyle changes, and addressing underlying medical conditions can be used to treat mild sleep disorders. At the same time, Cognitive Behavioral Therapy (CBT) is a therapy used to treat severe sleep disorders [8].

Feeling bad about oneself, commonly called low self-esteem or poor self-image, is a psychological state characterized by negative perceptions and beliefs about one's worth, abilities, and overall value. This state can manifest in various ways and profoundly impact an individual's emotions, thoughts, behaviors, and relationships. Early life experiences such as negative criticism and abuse or bullying, traumatic events such as loss, failure, or rejection, social comparison, and interpersonal relationships can cause low self-esteem. Low self-esteem can be treated by having healthy relationships, setting realistic goals, engaging in positive affirmations, seeking support, and practicing self-compassion [9].

Depression, sleep deprivation, and self-esteem are integral components of human experience, each exerting a profound influence on mental well-being and quality of life. These psychological constructs, individually and collectively, play pivotal roles in shaping individuals' emotional states, cognitive functioning, and interpersonal relationships. Understanding the complex interplay among these variables is essential for elucidating the mechanisms underlying mental health disorders and developing targeted interventions to mitigate their impact.

The relationship between trouble sleeping, depression, and feeling bad about yourself is complicated. Some people find they cannot sleep at all, while others find they cannot stop sleeping. Sleep problems and depression may also be innate biological factors that cause people to feel bad about themselves. Sleep problems are also associated with more severe depressive illness [10].

The primary purpose of this research is to clarify the relationship between depression, trouble sleeping or sleeping too much, and feeling bad about yourself. We will address that by answering the following questions:

RQ 1: What are the critical reasons for prioritizing the study of depression, and how does a deeper understanding of this mental health condition contribute to improved prevention, intervention, and overall well-being in individuals and society?

RQ 2: How can Log-Linear Models, Multinomial Logistic Regression, and Generalized Linear Models (GLM) be employed to analyze the association between depression, sleep disturbances, and self-esteem, shedding light on the intricate relationships within mental health?

RQ 3: How can the performance of proposed statistical models be effectively compared using statistical measures such as the Likelihood ratio test, Pearson chi-square, Mean Square Error, Bayesian Information Criterion (BIC), and Akaike Information Criterion (AIC)?

LITERATURE REVIEW

Depression, sleep disturbances, and self-esteem are interconnected aspects of mental health that significantly impact individuals' well-being. Understanding the complex relationships among these factors is crucial for developing effective interventions and promoting mental health. In this literature review, we synthesize findings from various studies to explore the bidirectional relationships between depression, sleep disturbances, and self-esteem. The review encompasses research spanning different populations and methodologies, providing insights into the nuanced dynamics of these psychological phenomena.

A systematic review of the relationship between sleep disturbances and mental health assessed the bidirectionality between sleep disturbances, anxiety, and depression, indicating a one-way relationship where anxiety predicted excessive daytime sleepiness, highlighting the intricate interplay between sleep disturbances and mental health problems [11]. Postpartum sleep disturbance and postpartum depression revealed a strong relationship between the two among predominantly educated, middle-class, older, white participants [12].

The study of depression is carried out on specific populations, such as the quality of sleep and depression in college students, shedding light on the unique challenges faced by a demographic group that highlighted the prevalence of depression among college students and its association with sleep quality, emphasizing the importance of addressing mental health issues in educational settings and the effect of anxiety and depression on sleep quality in individuals at high risk for insomnia revealed the complex relationship between psychiatric comorbidity and sleep disturbance [13, 14].

The meta-analysis examination of the relationship between sleep disorders and suicidal behavior in patients with depression indicated a significant association between sleep disorders and various manifestations of suicidal behavior and estimated more specific categories, including insomnia, nightmares, hypersomnia, suicidal ideation, suicide attempt, and completed suicide underscoring the importance of addressing sleep disturbances in suicide prevention efforts and highlighting the critical role of sleep in the context of depression and suicidal thoughts [15].

The investigation of the relationship between sex hormones, sleep problems, and depression emphasized the complex interplay between biological factors, sleep disturbances, and depression, which provides insights into potential mechanisms underlying these relationships and the examination of the relationship between depression and sleep quality, diseases, and general characteristics elucidated the multifaceted nature of depression and its associations with various demographic and clinical factors [16, 18]. The association between depression and sleep quality

in patients with Systemic Lupus Erythematosus (SLE) reveals the significant impact of chronic illness on mental health outcomes. It highlights the need for targeted interventions addressing chronic conditions' physical and psychological aspects [17].

In conclusion, the findings provide valuable insights into the complex dynamics of depression, sleep disturbances, and self-esteem, underscoring the importance of holistic approaches to mental health care guided by advanced statistical methodologies to tailor interventions and promote well-being across diverse populations.

RESEARCH METHODOLOGY

The techniques used to address the research question included data description and analysis techniques.

DATA DESCRIPTION

The National Center for Health Statistics (NCHS) obtains the dataset. The data was downloaded from the National Health and Nutrition Examination Survey (NHANES) website [22]. In this data set, a nine-item depression screening instrument, also called the Patient Health Questionnaire, was administered to determine the frequency of depression symptoms over the past two weeks with a follow-up question to assess the overall impairment of the symptoms [19]. The responses for the nine-item instrument were categorized as “not at all,” “several days,” “more than half the days,” and “nearly every day” and were given a point ranging from 0 to 3. In this data, mental health is measured as “little interest in doing things,” “depression,” “trouble sleeping” or “sleeping too much,” feeling tired” or “having little energy,” poor appetite” or “overeating,” feeling bad about yourself,” lack of concentration,” “moving or speaking slowly or too fast,” and suicidal thought. The questions are asked at the Mobile Examination Center (MEC) by trained interviewers using the Computer-Assisted Personal Interview system as part of the MEC interview. Both male and female participants aged twelve and older are eligible for the screening, but only data from participants aged eighteen and older are studied in this research. Participants requiring a proxy were not eligible because of the sensitive nature of the questions. The data for youth aged twelve to seventeen is accessible through the NCHS Research Data Center [23]. The dataset consists of 10 variables with 5162 observations. In this study, three categories, “depression,” “trouble sleeping,” and “feeling bad about yourself,” are considered.

ANALYSIS TECHNIQUES

The data analysis is done by checking the descriptive statistics such as mosaic plots and contingency tables, and the relationship of these variables and their effect on mental health is analyzed using the Log-linear Model, Multinomial Logistic Regression Model, and Generalized Linear Models. Each variable will be examined as a dependent variable, and the other two will be independent. The Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Deviance Statistic, and Mean Square Error (MSE) of these models are used to obtain the best model for the analysis.

a. Mosaic Plot

To explain multidimensional categorical data, it would be better to visually examine the structure of the variables, as each variable is analyzed separately.

A mosaic plot is an area-proportional visualization of frequencies composed of cells created by recursive vertical and horizontal splits of a rectangle [27]. It is a graphical display that allows the examination of relationships among two or more categorical variables. A particular type of stacked bar chart provides a visual representation of a contingency table.

b. Contingency Table

A contingency table, also known as a cross-tabulation or two-way table, is a statistical table used to display the frequency distribution of two or more categorical variables. Each cell in the table represents the frequency count or percentage of cases that fall into a particular combination of categories for the studied variables. The primary purpose of cross-tabulation is to uncover patterns, associations, and dependencies between categorical variables in a dataset [24].

For more than two variables, controlling for a third variable, Z , is often useful in assessing the relationship between two other variables, X and Y . There are two options:

- A marginal contingency table provides the marginal distributions of the variables X and Y , showing the total frequencies or percentages for each variable category. In other words, it displays each row and column total in the contingency table [24].
- A conditional contingency table examines the relationship between two variables, X and Y while holding one variable constant. It provides the frequency distribution of one variable for each category of the other variable. It is also known as cell percentages or proportions, and it is calculated by dividing the frequency in each contingency table cell by the corresponding marginal total. Conditional frequencies express the proportion of observations in each cell relative to the total number of observations for that row or column [24].

c. Linear Trend

The linear trend test determines whether there is an increasing or decreasing trend in the levels observed of one variable as the levels of the other “increase.” The test assumes both variables X and Y are ordinal. When the variables are ordinal, a trend association is common. As the level of X increases, responses on Y tend to increase toward higher levels, or responses on Y tend to decrease toward lower levels.

The test statistic, sensitive to positive or negative linear trends, utilizes correlation information in the data. The test statistic is calculated as follows:

$$Q = \sqrt{(n_{++} - 1)r^2}$$

where n_{++} is the total number of observations in the dataset, r^2 is the Pearson correlation coefficient [28].

d. The Cochran-Mantel-Haenszel (C M H) test

The Cochran-Mantel-Haenszel (C M H) test is an alternative test of XY conditional independence in $2 \times 2 \times K$ contingency tables. This test conditions the row and column totals in each partial table. In the partial table k, the row totals are n_{1+k} , n_{2+k} , and the column totals are n_{+1k} , n_{+2k} .

The associated hypotheses for the Cochran-Mantel-Haenszel (C M H) Test for Conditional Independence are stated as follows:

Null Hypothesis (H_0): X and Y are conditionally independent.

Alternative Hypothesis (H_a): X and Y are not conditionally independent [28].

e. Log-Linear Models

Log-linear Models are a class of statistical models used to analyze the relationship between two or more categorical variables. Log-linear models model cell counts regarding the row and column variables. The counts in the cells are treated as a third variable, as the response. These models evaluate associations that can be visualized using contingency tables [28].

A saturated statistical model perfectly fits the observed data, achieving zero discrepancy between the observed and expected values. It represents the maximum possible degree of fit to the data, capturing all the variability present in the observed data without any loss of information.

The saturated model is given by:

$$\ln(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY}$$

where μ_{ij} is the expected count or mean response for the combination of values (i, j) for X, and Y, λ is the intercept term, which is the expected count when all predictor variables are at their

reference levels, λ_i^X is the effect of level i of variable X on the expected count relative to the reference level of X, λ_j^Y is the effect of level j of variable Y on the expected count relative to the reference level of Y and λ_{ij}^{XY} is the interaction effect between levels i of variable X and j of variable Y, indicating how the joint presence of X and Y affects the expected count beyond the effects of the individual variables.

f. Multinomial Logistic Regression Model

The multinomial logistic regression model is an extension of binary logistic regression that allows the prediction of outcomes with more than two categories. It is used when the dependent variable is categorical with more than two levels, but the independent variables are still continuous or categorical. The Multinomial Logistic Regression Model uses three standard link functions: cumulative logit, adjacent categories logit, and continuation ratio logit.

The cumulative logit model assumes the cumulative log - odds of being in the category j or below. It is precisely for situations where the outcome variable has ordered responses, splits categories into half, and models the lower and upper half. It can be expressed as a linear combination of the predictor variables:

$$\log \left(\frac{P(Y \leq j)}{1 - P(Y \leq j)} \right) = \ln \left(\frac{\sum_{k=1}^j \pi_k}{\sum_{k=j+1}^J \pi_k} \right) = \beta_{0j} + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p$$

where $P(Y \leq j)$ is the cumulative probability of being in category j or below, $\beta_{0j}, \beta_1 X_1, \beta_2 X_2 \dots \beta_p$ are the model coefficients.

g. Generalized Linear Models (GLMs)

Generalized linear models are statistical models that generalize linear regression to handle response variables with error distributions other than the normal distribution. Log-linear, binary, and multinomial logistic regression models all involve the function of response parameter(s) equal to the linear combination of predictors and model parameters.

This approach can be formulated for any response from the exponential family of distributions, such as Normal, Poisson, and Binary distributions. The right-hand side of the equation looks like a standard linear model, but the left-hand side involves transforming the mean [28].

Since the data's response variables are count variables, we will use the Poisson and negative binomial regression. Using a log-linear model, Poisson regression models the relationship between predictor variables and the expected value of a count variable (response variable). The general form of the model is given by:

$$\ln(\mu_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i}$$

where μ_i is the mean count, β_0 is the intercept term of the model, β_1 and β_2 are the coefficients of the predictor variables X_{1i} and X_{2i} respectively.

The assumptions and fit of the Poisson model, which is independent of response, there is no collinearity among predictors, Poisson sampling is good, and variance is equal to the mean, are checked to determine whether the model is a good fit. The model is a Poisson sampling in which time is fixed, and the total is unknown. The data was collected over two weeks.

If all the assumptions are not met, the model will be changed to a negative binomial model with the same general form as the Poisson model.

RESULT AND DISCUSSION

a. Mosaic Plot

The plot in Figure 1 is drawn to highlight if there are too many depressed people who have trouble sleeping. The plot depicts relationships between “Depression”, “Trouble Sleeping,” and “Feeling Bad About Yourself.” “Depression” levels are shown on the vertical axis, and “Trouble Sleeping” levels are shown on the horizontal axis, ranging from 0 to 3. On the right side of the plot, there is a vertical bar labeled “Pearson residuals” with a scale from – 12 to 69, which provides a visual presentation of the deviations of observed cell frequencies from the frequencies expected under the assumption of independence between “Depression” and “Trouble Sleeping.” Large positive residuals greater than 4 can be found for “Not at All” levels of “Depression” and “Trouble Sleeping” and are colored in green. These positive residuals (4 to 69) indicate that the observed frequency is higher than expected under the independence assumption. On the other hand, large negative residuals are less than 4 for “Several Days of Depression” or “Several Days of Trouble Sleeping,” colored orange. These negative residuals indicate that the observed frequency is lower than expected under the independence assumption. Residuals between – 4 and 4, shaded in gray, suggest that the observed frequency is close to the expected frequency under the independence assumption. The p-value less than 2.22e–16 represents the probability of observing deviations from independence computed from a chi-square distribution with a degree of freedom 9. The degree of freedom of the chi-square (χ^2) The test of independence is calculated as follows:

$$\begin{aligned} df &= (I - 1) (J - 1) \\ &= (4 - 1) (4 - 1) \\ &= 9, \end{aligned}$$

where df is the degree of freedom, I is the number of levels of “Depression,” and J is the number of levels of “Trouble Sleeping.”

The association between “Depression” and “Trouble Sleeping” will be evaluated statistically with the p-value from the chi-square distribution in the mosaic plot. The associated hypotheses for the chi-square distribution are as follows:

Null Hypothesis (H_0): There is no association between depression and trouble sleeping.

Alternative Hypothesis (H_a): There is an association between depression and trouble sleeping.

The p-value is small enough to reject the null hypothesis of independence and conclude that there is an association between depression and trouble sleeping.

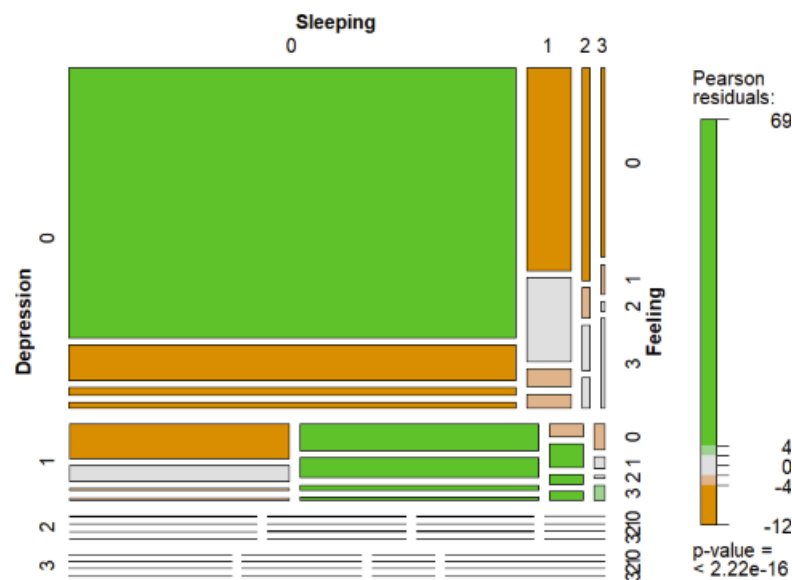


Figure 1: The mosaic plot of feeling bad about yourself

The mosaic plot in Figure 2 is drawn to highlight if too many trouble-sleeping people feel bad about themselves. The plot depicts relationships between “Feeling Bad About Yourself,” “Trouble Sleeping,” and “Depression.” “Trouble Sleeping” levels are shown on the vertical axis, and “Feeling Bad About Yourself” levels on the horizontal axis, ranging from 0 to 3. On the right side of the plot, there is a vertical bar labeled “Pearson residuals” with a scale from -12 to 69 , which provides a visual presentation of the deviations of observed cell frequencies from the frequencies expected under the assumption of independence between “Feeling Bad About Yourself” and “Trouble Sleeping.” Large positive residuals greater than 4 can be found for “Not at All” levels of “Trouble Sleeping” and “Feeling Bad About Yourself” and are colored in green. These positive residuals (4 to 69) indicate that the observed frequency is higher than expected under the assumption of independence. On the other hand, there are large negative residuals, which are less than 4 for “Several Days of “Feeling Bad About Yourself” or “Several Days of Trouble

Sleeping,” colored in orange. These negative residuals indicate that the observed frequency is lower than the expected frequency under the independence assumption. Residuals between – 4 and 4, shaded in gray, suggest that the observed frequency is close to the expected frequency under the independence assumption. The p-value less than $2.22\text{e-}16$ represents the probability of observing deviations from independence computed from a chi-square (χ^2) distribution with a degree of freedom 9. The test of independence is calculated as follows:

$$\begin{aligned} \text{df} &= (I - 1) (J - 1) \\ &= (4 - 1) (4 - 1) \\ &= 9, \end{aligned}$$

where df is the degree of freedom, I is the number of levels of “Feeling Bad About Yourself,” and J is the number of levels of “Trouble Sleeping.”

The association between “Feeling Bad About Yourself” and “Trouble Sleeping” will be evaluated statistically with the p-value from the chi-square distribution in the mosaic plot. The associated hypotheses for the chi-square distribution are:

Null Hypothesis (H_0): There is no association between feeling bad about yourself and trouble sleeping.

Alternative Hypothesis (H_a): There is an association between feeling bad about yourself and trouble sleeping.

The p-value is small enough to reject the null hypothesis of independence and conclude that there is an association between feeling bad about yourself and having trouble sleeping.

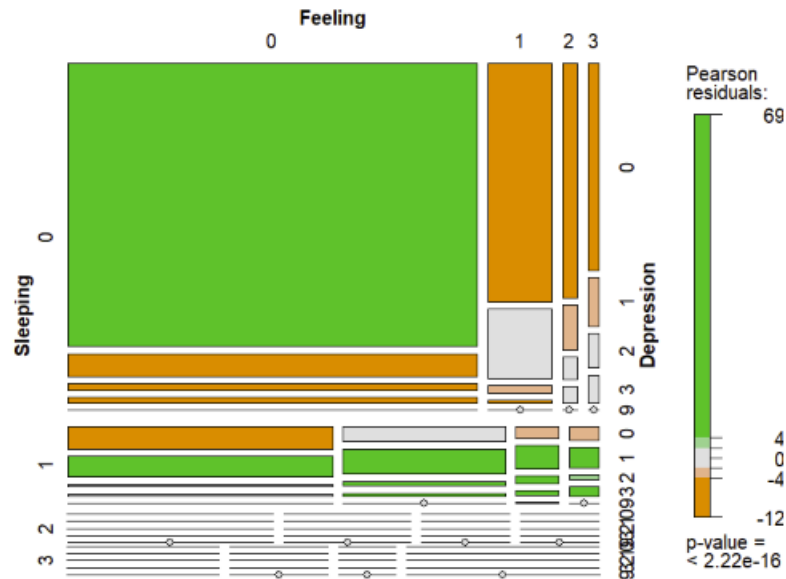


Figure 2: The mosaic plot of Depression.

The mosaic plot in Figure 3 is drawn for this data, and we would like to highlight if there are too many depressed people who feel bad about themselves. The plot depicts relationships between “Feeling Bad About Yourself,” “Depression,” and “Trouble Sleeping.” “Depression” levels are shown on the vertical axis, and “Feeling Bad About Yourself” levels on the horizontal axis, ranging from 0 to 3. On the right side of the plot, there is a vertical bar labeled “Pearson residuals” with a scale from – 12 to 69, which provides a visual presentation of the deviations of observed cell frequencies from the frequencies expected under the assumption of independence between “Feeling Bad About Yourself” and “Depression.” Large positive residuals greater than 4 can be found for “Not at All” levels of “Feeling Bad About Yourself” and “Depression” and are colored in green. These positive residuals (4 to 69) indicate that the observed frequency is higher than expected under the independence assumption. On the other hand, there are large negative residuals, which are less than 4 for “Several Days of “Feeling Bad About Yourself” or “Several Days of Depression,” colored in orange. These negative residuals indicate that the observed frequency is lower than expected under the independence assumption. Residuals between – 4 and 4, shaded in gray, suggest that the observed frequency is close to the expected frequency under the independence assumption. The p-value, less than 2.22e–16, represents the probability of observing deviations from independence computed from a chi-square distribution with a degree of freedom 9. The degree of freedom of the chi-square (χ^2) The test of independence is calculated as follows:

$$\begin{aligned}
 df &= (I - 1)(J - 1) \\
 &= (4 - 1)(4 - 1) \\
 &= 9,
 \end{aligned}$$

where df is the degree of freedom, I is the number of levels of “Feeling Bad About Yourself,” and J is the number of levels of “Depression.” The association between “Feeling Bad About Yourself” and “Depression” will be evaluated statistically with the p-value from the chi-square distribution in the mosaic plot. The associated hypotheses of a chi-square distribution are stated:

Null Hypothesis (H_0): There is no association between feeling bad about yourself and depression.

Alternative Hypothesis (H_a): There is an association between feeling bad about yourself and depression.

The p-value is small enough to reject the null hypothesis of independence and conclude that there is an association between feeling bad about yourself and depression.

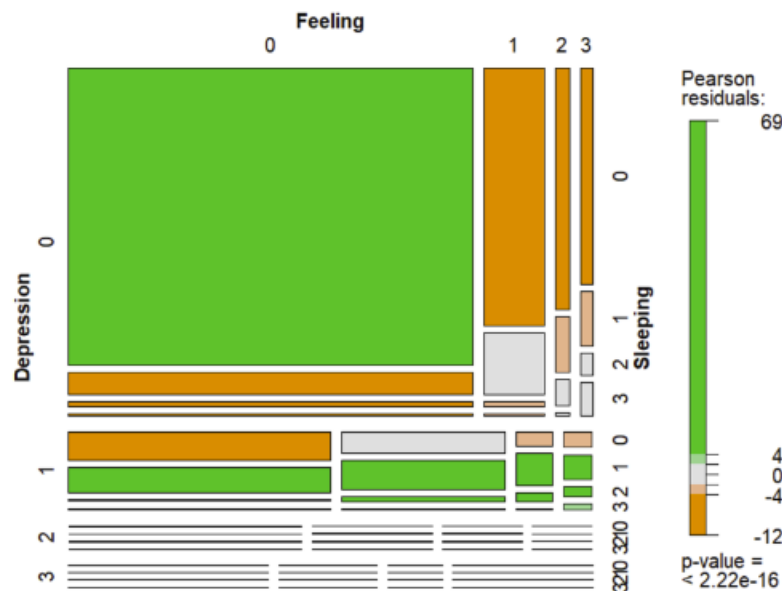


Figure 3: The mosaic plot of Trouble Sleeping.

b. Contingency Table

There are two ways to account for a third variable, which are conditional and marginal contingency tables.

Tables 1, 2, 3, and 4 display the conditional contingency table between “Depression” and “Trouble sleeping” for each level of “Feeling bad about yourself.” Table 1 displays the conditional

contingency table between “Depression” and “Trouble sleeping” for the “Not At All” level of “Feeling bad about yourself.”

Table 1: Conditional contingency table for “Not At All” level of “Feeling Bad About Yourself.”

	Trouble Sleeping			
Depression	Not at all	Several	More than half	Nearly
Not at all	2851	216	42	19
Several	233	198	14	9
More than half	65	26	22	8
Nearly	57	20	8	11

Table 2 displays the conditional contingency table between “Depression” and “Trouble sleeping” for “Several Days” level of “Feeling bad about yourself.”

Table 2: Conditional contingency table for “Several Days” level of “Feeling Bad About Yourself.”

	Trouble Sleeping			
Depression	Not at all	Several	More than half	Nearly
Not at all	374	90	6	3
Several	109	146	25	4
More than half	12	28	14	8
Nearly	5	18	8	16

Table 3 displays the conditional contingency table between “Depression” and “Trouble sleeping” for the “More Than Half the Days” level of “Feeling bad about yourself.”

Table 3: Conditional contingency table for “More Than Half the Days” level of “Feeling Bad About Yourself.”

	Trouble Sleeping			
Depression	Not at all	Several	More than half	Nearly
Not at all	83	19	9	1
Several	16	36	10	1
More than half	8	11	14	8
Nearly	6	8	4	8

Table 4 displays the conditional contingency table between “Depression” and “Trouble sleeping” for the “Nearly Everyday” level of “Feeling bad about yourself.”

Table 4: Conditional contingency table for “Nearly Everyday” level of “Feeling.

	Trouble Sleeping			
Depression	Not at all	Several	More than half	Nearly
Not at all	59	15	6	9
Several	14	22	9	5
More than half	10	5	9	7
Nearly	8	10	9	40

Since each category has four levels, we can calculate several odds ratios. For simplicity, one level of the “Feeling Bad About Yourself” variable, which is more meaningful, is picked to compute the odds ratio. The calculation of the odds ratio between “Several Days” and “More Than Half the Days” levels of “Feeling Down, Depressed, or Hopeless” and “Trouble Sleeping” for “More Than Half the Days” level of “Feeling Bad About Yourself” is done with the circled values in the conditional contingency table 5.

Table 5: Conditional contingency table for “More Than Half the Days” level of “Feeling Bad About Yourself” for odds ratio calculation.

	Trouble Sleeping			
Depression	Not at all	Several	More than half	Nearly
Not at all	83	19	9	1
Several	16	(36)	(10)	1
More than half	8	(11)	(14)	8
Nearly	6	8	4	8

The conditional probability, $\theta_{XY|Z}$ can be calculated as follows:

$$\begin{aligned}
 \theta_{XY|Z} &= \theta_{XY|k} = \frac{n_{ijk}n_{i'j'k}}{n_{ij'k}n_{i'j'k}} \\
 &= \theta_{XY|4} = \frac{n_{222}n_{322}}{n_{232}n_{332}} \\
 &= \frac{36 \times 14}{11 \times 10} = \frac{504}{110} \\
 &= 4.582
 \end{aligned}$$

It means that for those who have had depression for several days, the odds of “several days of trouble sleeping” will be 4.58 times more than those who have had depression for more than half the days due to more than half the days of feeling bad about themselves.

Also, the calculation of the odds ratio between “Several Days” and “Nearly Every Day of “Feeling Down” and having “Trouble Sleeping” for “More Than Half the Days “of “Feeling Bad About Yourself” is done with the circled figures in the conditional contingency table 6.

Table 6: Conditional contingency table for “More Than Half the Days” level of “Feeling Bad About Yourself” for second odds ratio calculation.

	Trouble Sleeping			
Depression	Not at all	Several	More than half	Nearly
Not at all	83	19	9	1
Several	16	(36)	10	(1)
More than half	8	11	14	8
Nearly	6	(8)	4	(8)

The conditional probability, $\theta_{XY|Z}$ can be calculated as follows:

$$\begin{aligned}
 \theta_{XY|Z} &= \theta_{XY|k} = \frac{n_{ijk}n_{i'j'k}}{n_{ij'k}n_{i'jk}} \\
 &= \theta_{XY|4} = \frac{n_{222}n_{322}}{n_{232}n_{332}} \\
 &= \frac{36 \times 8}{8 \times 1} = \frac{288}{8} \\
 &= 36
 \end{aligned}$$

It means that for those who have depression for several days, the odds of “several days of trouble sleeping” will be 36 times more than those who have depression nearly every day due to “more than half the days feeling bad about themselves.” If the rest of the odds ratio is calculated, it is found that all odd ratios are greater than one. In other words, $\theta_{XY|Z} = c > 1$, for all X and Y, where c is a constant. We conclude that depression has a greater effect on mental health than trouble sleeping, which are independent of each other condition on “more than half the days feeling bad about yourself.”

Table 7 shows the marginal contingency table, which presents the total counts across all levels of “Feeling Bad About Yourself. ”

Table 7: Marginal contingency table between “Depression” and “Trouble Sleeping.”

	Trouble Sleeping			
Depression	Not at all	Several	More than half	Nearly
Not at all	3367	340	63	32
Several	372	(402)	(58)	19
More than half	95	(70)	(59)	31
Nearly	76	56	29	75

The odds ratio between “several days” and “more than half the days” of feeling down and having trouble sleeping would be calculated. Therefore, the joint probability (θ_{XY}) between “Depression” (X) and “Trouble Sleeping” (Y) can be calculated as follows:

$$\begin{aligned}
\theta_{XY|Z} &= \frac{n_{11+} \cdot n_{22+}}{n_{12+} \cdot n_{21+}} \\
&= \frac{402 \times 59}{70 \times 58} \\
&= \frac{23718}{4060} \\
&= 5.84
\end{aligned}$$

It means that for those who have depression for several days, the odds of several days of trouble sleeping will be 5.84 times more than those who have depression for more than half the days. If we calculate all odds ratios for the marginal table, we will find that all are greater than 1. Therefore, we would conclude the odds of success are more significant for those who have depression.

Tables 8, 9, 10, and 11 display the conditional contingency table between “Feeling Bad About Yourself” and “Trouble Sleeping” for each level of “Depression.”

Table 8 displays the conditional contingency table between “Feeling Bad About Yourself” and “Trouble Sleeping” for the “Not at All” level of “Depression.”

Table 8: Conditional contingency table for “Not at All” level of “Feeling Depressed.”

	Feeling Bad About Yourself			
Trouble Sleeping	Not at all	Several	More than half	Nearly
Not at all	2851	374	83	59
Several	216	90	19	15
More than half	42	6	9	6
Nearly	19	3	1	9

Table 9 displays the conditional contingency table for the “Several Days” level of depression between “Feeling Bad About Yourself” and “Trouble Sleeping.”

Table 9: Conditional contingency table for “Several Days” level of “Feeling Depressed.”

	Feeling Bad About Yourself			
Trouble Sleeping	Not at all	Several	More than half	Nearly
Not at all	233	109	16	14
Several	198	146	36	22
More than half	14	25	10	9
Nearly	9	4	1	5

Table 10 displays the conditional contingency table between “Feeling Bad About Yourself” and “Trouble Sleeping” for the “More Than Half the Days” level of “Depression.”

Table 10: Conditional contingency table for “More Than Half the Days” level of “Feeling Depressed.”

	Feeling Bad About Yourself			
Trouble Sleeping	Not at all	Several	More than half	Nearly
Not at all	65	12	8	10
Several	26	28	11	5
More than half	22	14	14	9
Nearly	8	8	8	7

Table 11 displays the conditional contingency table for the “Nearly Everyday” level of depression between “Feeling Bad About Yourself” and “Trouble Sleeping.”

Table 11: Conditional contingency table for “Nearly Every Day level of “Feeling Depressed.”

	Feeling Bad About Yourself			
Trouble Sleeping	Not at all	Several	More than half	Nearly
Not at all	57	5	6	8
Several	20	18	8	10
More than half	8	8	4	9
Nearly	11	16	8	40

Since each category has four levels, we can calculate several odds ratios. For simplicity, one level of the “Depression” variable, which is more meaningful, is picked to compute the odds ratio. The calculation of the odds ratio between “Several Days” and “More Than Half the Days” levels of “Feeling Bad About Yourself” and “Trouble Sleeping” for “More Than Half the Days” level of “Feeling Down, Depressed and Hopeless” is done with the circled values in the conditional contingency table 12.

Table 12: Conditional contingency table for “More Than Half the Days” level of “Feeling Depressed” for odds ratio calculation.

	Feeling Bad About Yourself			
Trouble Sleeping	Not at all	Several	More than half	Nearly
Not at all	65	12	8	1
Several	26	28	11	5
More than half	22	14	14	9
Nearly	8	8	8	7

The conditional probability, $\theta_{XY|Z}$ can be calculated as follows:

$$\begin{aligned}\theta_{XY|Z} &= \theta_{XY|k} = \frac{n_{ijk}n_{i'j'k}}{n_{ij'k}n_{i'j'k}} \\ &= \theta_{XY|4} = \frac{n_{222}n_{322}}{n_{232}n_{332}} \\ &= \frac{28 \times 14}{14 \times 11} = \frac{28}{11} \\ &= 2.545\end{aligned}$$

It means that for those who have trouble sleeping for several days, the odds of “Several Days” of Feeling Bad About Yourself” will be 2.545 times more than those who have trouble sleeping for more than half the days due to more than half the days of feeling down, depressed, and hopeless. Also, the calculation of the odds ratio between “Several Days” and “Nearly Everyday” levels of “Trouble Sleeping” and “Feeling Bad About Yourself” for “More Than Half The Days” level of “Feeling Down, Depressed, and Hopeless” is done with the circled values in the conditional contingency table 13.

Table 13: Conditional contingency table for “More Than Half the Days” level of “Feeling Depressed” for second odds ratio calculation.

	Feeling Bad About Yourself			
Trouble Sleeping	Not at all	Several	More than half	Nearly
Not at all	65	12	8	1
Several	26	28	11	5
More than half	22	14	14	9
Nearly	8	8	8	7

The conditional probability, $\theta_{XY|Z}$ can be calculated as follows:

$$\begin{aligned}\theta_{XY|Z} &= \theta_{XY|k} = \frac{n_{ijk}n_{i'j'k}}{n_{ij'k}n_{i'j'k}} \\ &= \theta_{XY|4} = \frac{n_{222}n_{322}}{n_{232}n_{332}} \\ &= \frac{28 \times 7}{8 \times 5} = \frac{196}{40} \\ &= 4.90\end{aligned}$$

It means that for those who have trouble sleeping for several days, the odds of several Days of feeling bad about yourself” will be 4.90 times more than those who have trouble sleeping nearly every day due to more than half the days of feeling down, depressed, and hopeless. If the rest of the odds ratio is calculated, it is found that all odds ratios are greater than one. In other words, $\theta_{XY|2} = c > 1, \forall X, Y$, where c is a constant. We conclude that trouble sleeping has a greater effect on mental health than feeling bad about yourself, which are independent of each other condition on “more than half the days feeling down, depressed, and hopeless.”

Then, the marginal contingency table that presents the total counts across all levels of “Depression” is shown below in Table 14.

Table 14: Marginal contingency table between “Feeling Bad About Yourself” and “Trouble Sleeping”.

	Feeling Bad About Yourself			
Trouble Sleeping	Not at all	Several	More than half	Nearly
Not at all	3209	500	113	91
Several	461	282	75	52
More than half	86	53	37	33
Nearly	48	31	18	61

The odds ratio between “Several Days” and “More Than Half the Days” levels of “Trouble Sleeping” and “Feeling Bad About Yourself” would be calculated. The joint probability ($\theta_{XY|Z}$) between “Feeling Bad About Yourself” (X) and “Trouble Sleeping” (Y) can be calculated as follows:

$$\begin{aligned}
 \theta_{XY|Z} &= \frac{n_{11+} \cdot n_{22+}}{n_{12+} \cdot n_{21+}} \\
 &= \frac{282 \times 37}{53 \times 75} \\
 &= \frac{10434}{3975} \\
 &= 2.625
 \end{aligned}$$

It means that for those who have trouble sleeping for several days, the odds of several days of feeling bad about yourself will be 2.63 times more than those who have trouble sleeping for more than half the days of feeling down, depressed, and hopeless. If we calculate the odds for the marginal table, we will find that all are greater than 1. Therefore, we would conclude that the odds of success are greater for those with trouble sleeping.

Tables 15,16,17 and 18 display the conditional contingency table between “Feeling Bad About Yourself” and “Depression” for each level of “Trouble Sleeping”. Table 23 displays the conditional contingency table between “Feeling Bad About Yourself” and “Depression” for “Not at All” level of “Trouble Sleeping”.

Table 15: Conditional contingency table for “Not at All” level of “Trouble Sleeping” .

	Feeling Bad About Yourself			
Depression	Not at all	Several	More than half	Nearly
Not at all	2851	374	83	59
Several	233	109	16	14
More than half	65	12	8	10
Nearly	57	5	6	8

Table 16 displays the conditional contingency table between “Feeling Bad About Yourself” and “Depression” for the “Several Days” level of “Trouble Sleeping.”

Table 16: Conditional contingency table for “Several Days” level of “Trouble Sleeping”.

	Feeling Bad About Yourself			
Depression	Not at all	Several	More than half	Nearly
Not at all	216	90	19	15
Several	198	146	36	22
More than half	26	28	11	5
Nearly	20	18	8	10

Table 17 displays the conditional contingency table between “Feeling Bad About Yourself” and “Depression” for the “More Than Half The Days” level of “Trouble Sleeping.”

Table 17: Conditional contingency table for “More Than Half Days” level of “Trouble Sleeping”.

	Feeling Bad About Yourself			
Depression	Not at all	Several	More than half	Nearly
Not at all	42	6	9	6
Several	14	25	10	9
More than half	22	14	14	9
Nearly	8	8	4	9

Table 18 displays the conditional contingency table between “Feeling Bad About Yourself” and “Depression” for the “Nearly Everyday” level of “Trouble Sleeping.”

Table 18: Conditional contingency table for “Nearly Everyday” level of “Trouble Sleeping” .

	Feeling Bad About Yourself			
Depression	Not at all	Several	More than half	Nearly
Not at all	19	3	1	9
Several	9	4	1	5
More than half	8	8	8	7
Nearly	11	16	8	40

Since each category has four levels, we can calculate several odds ratios. For simplicity, one level of the “Sleeping” variable, which is more meaningful, is picked to compute the odds ratio. The calculation of the odds ratio between “Several Days” and “More Than Half the Days” levels of “Depression” and “Feeling Bad About Yourself” for “More Than Half the Days” of “Trouble Sleeping” is done with the circled values in the conditional contingency table 19.

Table 19: Conditional contingency table for “More Than Half the Days” level of “Trouble Sleeping” for odds ratio calculation.

	Feeling Bad About Yourself			
Trouble Sleeping	Not at all	Several	More than half	Nearly
Not at all	42	6	9	6
Several	14	25	10	9
More than half	22	14	14	9
Nearly	8	8	4	9

The conditional probability $\theta_{XY|Z}$ between the “Several Days” and the “More Than Half the Days” levels of “Feeling Bad About Yourself” and “Depression” can be calculated as follows:

The conditional probability, $\theta_{XY|Z}$ can be calculated as follows:

$$\begin{aligned}
 \theta_{XY|Z} &= \theta_{XY|k} = \frac{n_{ijk}n_{i'j'k}}{n_{ij'k}n_{i'j'k}} \\
 &= \theta_{XY|4} = \frac{n_{222}n_{322}}{n_{232}n_{332}} \\
 &= \frac{25 \times 14}{14 \times 10} = \frac{25}{10} \\
 &= 2.50
 \end{aligned}$$

It means that for those who have had depression for several days, the odds of several days of feeling bad about yourself will be 2.50 times more than those who have had depression for more than half the days due to more than half the days of trouble sleeping.

Also, the circled values in conditional contingency table 20 are used to calculate the odds ratio between the “Several Days” and “Nearly Every day “levels of “Depression” and “Feeling Bad About Yourself” for the “More Than Half Days “level of “Trouble Sleeping.”

Table 20: Conditional contingency table for “More Than Half the Days” level of “Trouble Sleeping” for second odds ratio calculation.

	Feeling Bad About Yourself			
Depression	Not at all	Several	More than half	Nearly
Not at all	42	6	9	6
Several	14	25	10	9
More than half	22	14	14	9
Nearly	8	8	4	9

The Conditional Probability $\theta_{XY|Z}$ between “Several Days” and “Nearly Everyday “levels of “Feeling Bad About Yourself” and “Trouble Sleeping” are calculated as follows:

The conditional probability, $\theta_{XY|Z}$ can be calculated as follows:

$$\begin{aligned}
\theta_{XY|Z} &= \theta_{XY|k} = \frac{n_{ijk}n_{i'j'k}}{n_{ij'k}n_{i'j'k}} \\
&= \theta_{XY|4} = \frac{n_{222}n_{442}}{n_{242}n_{422}} \\
&= \frac{25 \times 9}{8 \times 9} = \frac{25}{8} \\
&= 3.125
\end{aligned}$$

It means that for those who have depression for several days, the odds of several days of feeling bad about themselves will be 3.125 times more than those who have depression nearly every day due to more than half the days of trouble sleeping. If the rest of the odds ratio is calculated, it is found that all odd ratios are greater than one. In other words, $\theta_{XY|4} = c > 1$, $\forall X, Y$, where c is a constant. We conclude that depression has a greater effect on mental health than feeling bad about yourself which are independent of each other condition on “more than half the days trouble sleeping”.

Table 21 shows the marginal contingency table, which presents the total counts across all levels of “Trouble Sleeping. ”

Table 21: Marginal contingency table for “More Than Half the Days” of “Trouble Sleeping” .

	Feeling Bad About Yourself			
Depression	Not at all	Several	More than half	Nearly
Not at all	3128	473	112	89
Several	454	284	63	50
More than half	121	62	41	31
Nearly	96	47	26	67

The odds ratio between “Several Days” and “More Than Half the Days” levels of “Feeling Bad About Yourself” and “Depression” would be calculated.

The joint probability ($\theta_{XY|Z}$) between “Feeling Bad About Yourself” (X) and “Depression” (Y) can be calculated as follows:

The conditional probability, $\theta_{XY|Z}$ can be calculated as follows:

$$\begin{aligned}
\theta_{XY|Z} &= \frac{n_{11+} \cdot n_{22+}}{n_{12+} \cdot n_{21+}} \\
&= \frac{284 \times 41}{62 \times 63} \\
&= \frac{11644}{3906} \\
&= 2.98
\end{aligned}$$

It means that for those who have depression for several days, the odds of several days of feeling bad about yourself will be 2.98 times more than those who have depression for more than half the days of trouble sleeping. If we calculate the odds for the marginal table, we will find that all

are greater than 1. Therefore, we would conclude the odds of success are greater for those who have trouble sleeping.

c. Linear Trend

For Feeling Bad About Yourself, we will determine whether there is an increasing or decreasing trend between the levels of depression and trouble sleeping. The associated hypotheses for the linear trend are given by:

Null Hypothesis (H_0): There is no trend between depression and trouble sleeping.

Alternative Hypothesis (H_a): A positive linear trend exists between depression and trouble sleeping.

The linear trend test statistic is calculated as follows:

$$\begin{aligned} Q &= \sqrt{(n_{++} - 1)r^2} \\ &= \sqrt{(242 - 1)(1)^2} \\ &\approx 15.524, \end{aligned}$$

where n_{++} is the total number of counts in the conditional contingency table for the “More Than the Days” level of “Feeling Bad About Yourself” and r is the weighted correlation between “Depression” and “Trouble Sleeping” .

The linear trend test statistic value is compared with the critical value of the chi-square with a degree of freedom 3 ($\chi^2(3) = 7.815$). The degree of freedom of chi-square is calculated as follows:

$$\begin{aligned} df &= k - 1 \\ &= 4 - 1 \\ &= 3, \end{aligned}$$

where df is the degree of freedom and k is the number of levels of “Feeling Bad About Yourself”.

This comparison suggests the null hypothesis (H_0) should be rejected since the test statistic value is greater than the critical value. The data provide sufficient evidence to conclude there is a positive linear association between trouble sleeping and depression due to more than half the days feeling bad about yourself.

For Depression, we will determine whether there is an increasing or decreasing trend between the levels of depression and trouble sleeping. The associated hypotheses for the linear trend are stated as follows:

Null Hypothesis (H_0): There is no trend between feeling bad about yourself and trouble sleeping.

Alternative Hypothesis (H_a): A positive linear trend exists between feeling bad about yourself and trouble sleeping.

The test statistic of the linear trend is calculated as follows:

$$\begin{aligned} Q &= \sqrt{(n_{++} - 1)r^2} \\ &= \sqrt{(255 - 1)(1)^2} \\ &\approx 15.937, \end{aligned}$$

where n_{++} is the total number of counts in the conditional contingency table for the “More Than the Days” level of “Feeling Down, Depressed, and Hopeless” and r is the weighted correlation between “Feeling Bad About Yourself” and “Trouble Sleeping.”

The linear trend test statistic value is compared with the critical value of the chi-square with degree of freedom 3 ($\chi^2(3) = 7.815$). The degree of freedom of chi-square is calculated as follows:

$$\begin{aligned} df &= k - 1 \\ &= 4 - 1 \\ &= 3, \end{aligned}$$

where df is the degree of freedom and k is the number of levels of “Feeling Down, Depressed, and Hopeless.”

The comparison of the test statistic value to the critical value $\chi^2(3)$ suggests the null hypothesis (H_0) should be rejected since the test statistic value is greater than the critical value. The data provide sufficient evidence to conclude there is a positive linear association between feeling bad about yourself and having trouble sleeping due to more than half the days feeling down, depressed, and hopeless.

For trouble sleeping, we will determine whether there is an increasing or decreasing trend between the levels of depression and trouble sleeping. The associated hypotheses of the linear trend are stated as follows:

Null Hypothesis (H_0): There is no trend between feeling bad about yourself and depression.

Alternative Hypothesis (H_a): There is a positive linear trend between feeling bad about yourself and depression.

The test statistic of the linear trend is calculated as follows:

$$\begin{aligned} Q &= \sqrt{(n_{++} - 1)r^2} \\ &= \sqrt{(209 - 1)(1)^2} \\ &\approx 14.422, \end{aligned}$$

where n_{++} is the total number of counts in the conditional contingency table for “More Than the Days” level of “Trouble Sleeping” and r is the weighted correlation between “Feeling Bad About Yourself” and “Depression” .

The linear trend test statistic value is compared with the critical value of chi-square with degree of freedom 3 ($\chi^2(3) = 7.815$). The degree of freedom of chi-square is calculated as follows:

$$\begin{aligned} df &= k - 1 \\ &= 4 - 1 \\ &= 3, \end{aligned}$$

where df is the degree of freedom and k is the number of levels of “Trouble Sleeping” .

The comparison of the test statistic value to the critical value $\chi^2(3)$, suggests that the null hypothesis (H_0) should be rejected since the test statistic value is greater than the critical value. The data provide sufficient evidence to conclude there is a positive linear association between depression and feeling bad about yourself due to more than half the days of feeling down, depressed, and hopeless.

d. Cochran- Mantel- Haenszel (C M H) Test

The Cochran-Mantel-Haenszel Test is used to assess the conditional independence of categorical predictors associated with categorical outcomes.

The associated hypotheses of the Cochran-Mantel-Haenszel Test for Feeling Bad About Yourself are stated as:

Null Hypothesis (H_0): Depression and Trouble sleeping are conditionally independent.

Alternative Hypothesis (H_a): Depression and Trouble sleeping are not conditionally Independent.

The Cochran-Mantel-Haenszel (C M H) χ^2 statistic is approximately 1178, which is greater than the chi-square critical value of the degree of freedom 9 ($\chi^2(9) = 16.919$). The degree of freedom of the chi-square test of conditional independence is calculated as follows:

$$\begin{aligned}
df &= (I - 1) (J - 1) \\
&= (4 - 1) (4 - 1) \\
&= 9,
\end{aligned}$$

where df is the degree of freedom, I is the number of levels of “Depression” and J is the number of levels of “Trouble Sleeping.”

We would reject H_0 since the test statistic value is greater than the critical value and conclude that “Depression” and “Trouble Sleeping” are not conditionally independent conditional on each “Feeling Bad About Yourself” level. For at least one feeling level, there is a significant association between depression and trouble sleeping.

The associated hypotheses of the Cochran-Mantel-Haenszel Test for Depression are stated as follows:

Null Hypothesis (H_0): Feeling bad about yourself and having Trouble sleeping are conditionally independent.

Alternative Hypothesis (H_a): Feeling bad about yourself and having Trouble sleeping are not conditionally independent.

The Cochran-Mantel-Haenszel (C M H) χ^2 statistic is approximately 280.35, which is greater than any $\chi^2(9)$ critical value of the degree of freedom 9 ($\chi^2(9) = 16.919$).

The degree of freedom of the chi-square test of conditional independence is calculated as follows:

$$\begin{aligned}
df &= (I - 1) (J - 1) \\
&= (4 - 1) (4 - 1) \\
&= 9,
\end{aligned}$$

where df is the degree of freedom, I is the number of levels of “Feeling Bad About Yourself” and J is the number of levels of “Trouble Sleeping.”

We would reject the null hypothesis (H_0) since the test statistic value is greater than the critical value and concludes that “Feeling Bad About Yourself” and “Trouble Sleeping” are not conditionally independent of “Depression” levels. For at least one “Depression” level, there is a significant association between trouble sleeping and feeling bad about yourself.

The associated hypotheses for the Cochran-Mantel-Haenszel Test for Trouble Sleeping are stated as follows:

Null Hypothesis (H_0): Feeling bad about yourself and depression are conditionally independent.

Alternative Hypothesis (H_a): Feeling bad about yourself and depression are not conditionally independent.

The Cochran-Mantel-Haenszel (C M H) χ^2 statistic is approximately 222.61, which is greater than the chi-square critical value of the degree of freedom 9 ($\chi^2(9) = 16.919$). The degree of freedom of the chi-square test of conditional independence is calculated as follows:

$$\begin{aligned} df &= (I - 1)(J - 1) \\ &= (4 - 1)(4 - 1) \\ &= 9, \end{aligned}$$

where df is the degree of freedom and k is the number of levels of “Trouble Sleeping” .

We would reject the null hypothesis (H_0) since the test statistic value is greater than the critical value and concludes depression and feeling bad about yourself are not conditionally independent of trouble sleeping levels. For at least one trouble sleeping level, there is a significant association between feeling bad about yourself and depression.

e. Log-Linear Model

For Feeling Bad About Yourself, we will investigate if there is any homogeneous association between all two-factor interaction terms, i.e., XY (Depression $\times$ Sleeping). The appropriate model would be the Loglinear Homogeneous Association Model (M_0). The general form of the Loglinear Homogeneous Association Model for this data is given by:

$$\ln(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY} \quad \text{for } i = 1, 2, 3, 4, j = 1, 2, 3, 4$$

The fitted model can be represented as follows:

$$\ln(\mu_{ij}) = -1.348789 + 0.313508_i^{Depression} + 0.525302_j^{Sleeping} - 0.062089_{ij}^{Depression \times Sleeping}$$

To perform a test to assess the overall fit of this model, the deviance statistic, $G^2 M_0$, is calculated by finding the difference between the null deviance, which is the saturated or full model, and the residual deviance which is the reduced model.

The associated hypotheses for the deviance statistic are stated as follows:

Null Hypothesis (H_0): Extra model parameters are zero (not significant).

Alternative Hypothesis (H_a): Extra model parameters are non-zero (at least one is significant).

The test statistic of the deviance statistic is calculated as follows:

$$\begin{aligned} G^2 M_0 | M_1 &= G^2(M_0) - G^2(M_1) \\ &= 6177.6 - 5242.2 \\ &= 935.35 \end{aligned}$$

The degree of freedom of the chi-square test of the Log-linear Homogeneous Association Model is calculated as follows:

$$\begin{aligned} df &= null_{df} - residual_{df} \\ &= 5161 - 5158 \\ &= 3, \end{aligned}$$

where df is the degree of freedom, $null_{df}$ is the degree of freedom of the null deviance and $residual_{df}$ is the degree of freedom of the residual deviance.

By comparing the test value to the critical value ($\chi^2_{0.05(3)} = 7.815$), we would reject the null hypothesis. Therefore, the data provide sufficient evidence to conclude the homogeneous association model fits well. It means that there is an association between any pair of variables. Additionally, it implies that we have no three-way interaction between depression, sleeping, and feeling bad about ourselves.

For Depression, we will investigate if there is any homogeneous association between all two-factor interaction terms, i.e., XY (Sleeping $\times$ Feeling).

The general form of the Log-linear Homogeneous Association Model for this data is given by:

$$\ln(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY} \text{ for } i = 1,2,3,4, j = 1,2,3,4$$

The fitted model can be represented as follows:

$$\ln(\mu_{ij}) = -1.42978 + 0.63156_i^{Sleeping} + 0.36246_j^{Feeling} - 0.07570_{ij}^{Sleeping \times Feeling}$$

To perform a test to assess the overall fit of this model, the deviance statistic, $G^2 M_0$, is calculated by finding the difference between the null deviance, which is the saturated or full model, and the residual deviance, which is the reduced model.

The associated hypotheses of the deviance statistics are stated as follows:

Null Hypothesis (H_0): Extra model parameters are zero (not significant)

Alternative Hypothesis (H_a): Extra model parameters are non-zero (at least one is significant)

The test statistic of the deviance statistic is calculated as follows:

$$\begin{aligned} G^2 M_0 | M_1 &= G^2(M_0) - G^2(M_1) \\ &= 6442.9 - 5062.3 \\ &= 1380.6 \end{aligned}$$

The degree of freedom of the chi-square test of the Log-linear Homogeneous Association Model is calculated as follows:

$$\begin{aligned} df &= null_{df} - residual_{df} \\ &= 5161 - 5158 \\ &= 3, \end{aligned}$$

where df is the degree of freedom, $null_{df}$ is the degree of freedom of the null deviance and $residual_{df}$ is the degree of freedom of the residual deviance.

By comparing the test value to the critical value ($\chi^2_{0.05(3)} = 7.815$), we would reject the null hypothesis since the test value is greater than the critical value. Therefore, the data provide sufficient evidence to conclude the homogeneous association model fits well. It means that there is an association between any pair of variables. Additionally, it implies that we have no three-way interaction between depression, sleeping, and feeling bad about ourselves.

For Trouble Sleeping, we will investigate whether a homogeneous association exists between all two-factor interaction terms, i.e., XY (Depression $\times$ Feeling).

The general form of the Log-linear Homogeneous Association Model for this data is given by:

$$\ln(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY} \text{ for } i = 1,2,3,4, j = 1,2,3,4$$

The fitted model can be represented as follows:

$$\ln(\mu_{ij}) = -1.621580 + 0.434453_i^{Depression} + 0.428898_j^{Feeling} - 0.049198_{ij}^{Depression \times Feeling}$$

To perform a test to assess the overall fit of this model, the deviance statistic, $G^2(M_0)$, is calculated by finding the difference between the null deviance, the saturated or full model, and the residual deviance, the reduced model.

The associated hypotheses for the deviance statistics are stated as follows:

Null Hypothesis (H_0): Extra model parameters are zero (not significant).

Alternative Hypothesis (H_a): Extra model parameters are non-zero (at least one is significant).

The test statistic of the deviance statistic is calculated as follows:

$$\begin{aligned} G^2 M_0 | M_1 &= G^2(M_0) - G^2(M_1) \\ &= 5503.7 - 4291.8 \\ &= 1211.9 \end{aligned}$$

The degree of freedom of the chi-square test of the Log-linear Homogeneous Association Model is calculated as follows:

$$\begin{aligned} df &= null_{df} - residual_{df} \\ &= 5161 - 5158 \\ &= 3, \end{aligned}$$

where df is the degree of freedom, $null_{df}$ is the degree of freedom of the null deviance and $residual_{df}$ is the degree of freedom of the residual deviance.

By comparing the test value to the critical value $\chi^2_{0.05(9)}$, we would reject the null hypothesis (H_0). Therefore, the data provide sufficient evidence to conclude the homogeneous association model fits well. It means that there is an association between any pair of variables. Additionally, it implies that we have no three-way interaction between depression, sleeping, and feeling bad about ourselves.

f. Multinomial Logistic Regression Model

We will model a categorical response, "Feeling Bad About Yourself, using any predictor and assume the total number of responses is fixed. The appropriate model would be the Multinomial Logistic Regression model. The Multinomial Logistic Regression Model uses three standard link functions: cumulative logit, adjacent categories logit, and continuation ratio logit.

The Cumulative Logit Model is used in this analysis that split categories in half model lower half versus upper half. The model has different intercepts because it models the cumulative probabilities of observing an outcome falling into a specific category or below. Each intercept represents the log odds of being in or below a specific category relative to a reference category.

The model can be written mathematically as follows:

$$\text{logit} \left(\frac{P \leq j}{P < j} \right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} \quad , \quad j = 0, 1, 2$$

So, the model for Feeling Bad About Yourself is written as follows:

$$\text{logit}\left(\frac{P \leq j}{P < j}\right) = \beta_0 + \beta_1 \text{Depression}_i + \beta_2 \text{Sleeping}_i$$

More specifically, the model can be expanded as follows:

$$\log\left(\frac{P_0}{P_1 + P_2 + P_3}\right) = 1.5523 + 0.4206 \text{Depression}_i + 0.7741 \text{Sleeping}_i,$$

$$\log\left(\frac{P_0 + P_1}{P_2 + P_3}\right) = 2.9644 + 0.4206 \text{Depression}_i + 0.7741 \text{Sleeping}_i, \text{ and}$$

$$\log\left(\frac{P_0 + P_1 + P_2}{P_3}\right) = 3.8131 + 0.4206 \text{Depression}_i + 0.7741 \text{Sleeping}_i$$

To perform a test to assess the overall fit of this model, the deviance statistics, $G^2(M_0)$ is calculated and compared to the chi-square critical value. The associated hypotheses of the deviance statistic are stated as follows:

Null Hypothesis (H_0): Model fit is good.

Alternative Hypothesis (H_a): Model fit is not good.

The test statistic of the deviance statistic is calculated as follows:

$$\begin{aligned} G^2 M_0 | M_1 &= G^2(M_0) - G^2(M_1) \\ &= 8414.769 - 7699.3 \\ &= 715.44 \end{aligned}$$

The degree of freedom of the chi-square test of the Cumulative Logit Model is calculated as follows:

$$\begin{aligned} \text{df} &= K - 1 \\ &= 4 - 1 \\ &= 3, \end{aligned}$$

where df is the degree of freedom, and K is the number of levels of “Feeling Bad About Yourself”.

By comparing the test value to the critical value ($\chi^2_{0.05}(3) = 7.815$), we would reject the null hypothesis. Therefore, the data provide sufficient evidence to conclude the cumulative logit model fits well. It means that there is an association between any pair of variables. Additionally, it implies that we have no three-way interaction between depression, sleeping, and feeling bad about ourselves.

For the Multinomial Logistic Regression Model, the model can be written mathematically as follows:

$$\text{logit} \left(\frac{P \leq j}{P < j} \right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} \quad , \quad j = 0,1,2$$

So, the model for Depression is written as follows:

$$\text{logit} \left(\frac{P \leq j}{P < j} \right) = \beta_0 + \beta_1 \text{Sleeping}_i + \beta_2 \text{Feeling}_i$$

More specifically, the model can be expanded as follows:

$$\log \left(\frac{P_0}{P_1 + P_2 + P_3} \right) = 1.7994 + 1.264 \text{Sleeping}_i + 0.495 \text{Feeling}_i,$$

$$\log \left(\frac{P_0 + P_1}{P_2 + P_3} \right) = 3.3841 + 1.264 \text{Sleeping}_i + 0.495 \text{Feeling}_i, \text{ and}$$

$$\log \left(\frac{P_0 + P_1 + P_2}{P_3} \right) = 4.3852 + 1.264 \text{Sleeping}_i + 0.495 \text{Feeling}_i$$

To perform a test to assess the overall fit of this model, the deviance statistics, $G^2(M_0)$ is calculated and compared to the chi-square critical value. The associated hypotheses of the deviance statistic are stated as follows:

Null Hypothesis (H_0): Model fit is good.

Alternative Hypothesis (H_a): Model fit is not good.

The test statistic of the deviance statistic is calculated as follows:

$$\begin{aligned} G^2 M_0 | M_1 &= G^2(M_0) - G^2(M_1) \\ &= 8557.117 - 7214.218 \\ &= 1342.899 \end{aligned}$$

The degree of freedom of the chi-square test of the Cumulative Logit Model is calculated as follows:

$$\begin{aligned} \text{df} &= k - 1 \\ &= 4 - 1 \\ &= 3, \end{aligned}$$

where df is the degree of freedom and k is the number of levels of “Feeling Down, Depressed, and Hopeless” .

By comparing the test value to the critical value ($\chi^2_{0.05}(3) = 7.815$), we would reject the null hypothesis (H_0) since the test value is greater than the critical value. Therefore, the data provide sufficient evidence to conclude the cumulative logit model fits well. It means that there is an association between any pair of variables. Additionally, it implies that we have no three-way interaction between depression, sleeping, and feeling bad about ourselves.

The Multinomial Logistic Regression model can be written mathematically as:

$$\text{logit} \left(\frac{P \leq j}{P < j} \right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} \quad , \quad j = 0,1,2$$

So, the model for Trouble Sleeping is written as follows:

$$\text{logit} \left(\frac{P \leq j}{P < j} \right) = \beta_0 + \beta_1 \text{Depression}_i + \beta_2 \text{Feeling}_i$$

More specifically, the model can be expanded as follows:

$$\log \left(\frac{P_0}{P_1 + P_2 + P_3} \right) = 1.9811 + 1.0219 \text{Depression}_i + 0.6357 \text{Feeling}_i,$$

$$\log \left(\frac{P_0 + P_1}{P_2 + P_3} \right) = 3.8039 + 1.0219 \text{Depression}_i + 0.6357 \text{Feeling}_i, \text{ and}$$

$$\log \left(\frac{P_0 + P_1 + P_2}{P_3} \right) = 4.9173 + 1.0219 \text{Depression}_i + 0.6357 \text{Feeling}_i.$$

To perform a test to assess the overall fit of the Cumulative Logit Model, the deviance statistic $G^2(M_0)$ is calculated and compared to the chi-square critical value. The associated hypotheses of the deviance statistic are stated as follows:

Null Hypothesis (H_0): Model fit is good.

Alternative Hypothesis (H_a): Model fit is not good.

The test statistic of the deviance statistic is calculated as follows:

$$\begin{aligned} G^2 M_0 | M_1 &= G^2(M_0) - G^2(M_1) \\ &= 7758.285 - 6502.324 \\ &= 1255.961 \end{aligned}$$

The degree of freedom of the chi-square test of the Cumulative Logit Model is calculated as follows:

$$\begin{aligned} df &= K - 1 \\ &= 4 - 1 \\ &= 3, \end{aligned}$$

where df is the degree of freedom, and K is the number of levels of “Trouble Sleeping” .

By comparing the test value to the chi-square critical value ($\chi^2_{0.05}(3) = 7.815$), we would reject the null hypothesis. Therefore, the data provide sufficient evidence to conclude the cumulative logit model fits well. It means that there is an association between any pair of variables. Additionally, it implies that we have no three-way interaction between depression, trouble sleeping, and feeling bad about ourselves.

g. Generalized Linear Model

We will model the response variable, “Feeling Bad About Yourself,” with the Generalized Linear Models (GLMs). The Poisson Regression and the Negative Binomial Regression are used since the response variable of the data is a count variable.

The general form of the Poisson model is given by:

$$\ln(\mu_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} \quad ,$$

where μ_i is the mean count, β_0 is the intercept term of the model, β_1 and β_2 are the coefficients of the predictor variables X_{1i} (Depression) and X_{2i} (Trouble Sleeping) respectively.

So, the model will be represented as follows:

$$\ln(\hat{\mu}) = -1.27599 + 0.23081(\text{Depression}_i) + 0.41123(\text{Sleeping}_i)$$

The general form of the Negative Binomial Regression model is given as follows:

$$\ln(\mu_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} \quad ,$$

where μ_i is the mean count, β_0 is the intercept term of the model, β_1 and β_2 are the coefficients of the predictor variables X_{1i} (Depression) and X_{2i} (Trouble Sleeping) respectively.

Therefore, the fitted model would be:

$$\ln(\hat{\mu}) = -1.37535 + 0.31645(\text{Depression}_i) + 0.47472(\text{Sleeping}_i)$$

We will model the response variable, “Depression,” with generalized linear models (GLMs).

The general form of the Poisson model is given by:

$$\ln(\mu_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} \quad ,$$

where μ_i is the mean count, β_0 is the intercept term of the model, β_1 and β_2 are the coefficients of the predictor variables X_{1i} (Trouble Sleeping) and X_{2i} (Feeling Bad About Yourself) respectively.

So, the fitted model will be represented as follows:

$$\ln(\hat{\mu}) = -1.36237 + 0.54893(\text{Sleeping}_i) + 0.25197(\text{Feeling}_i)$$

The general form of the Negative Binomial Regression model is:

$$\ln(\mu_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} \quad ,$$

where μ_i is the mean count, β_0 is the intercept term of the model, β_1 and β_2 are the coefficients of the predictor variables X_{1i} (Trouble Sleeping) and X_{2i} (Feeling Bad About Yourself) respectively.

$$\ln(\hat{\mu}) = -1.47999 + 0.69791(\text{Sleeping}_i) + 0.27767(\text{Feeling}_i)$$

h. Statistical Measures

The statistical measures obtained from the analysis of “Feeling Bad About Yourself” are shown in Table 22 below. These measures are visualized graphically below, and colored vertical lines represent the best model.

Table 22: Statistical Measures of Feeling Bad About Yourself.

Models	Statistical Measures			
	AIC	BIC	G^2	MSE
Log-Linear	8347.642	8373.839	935.4	0.615654
Multinomial	7711.33	7750.624	715.439	0.07756587
GLM(Poisson)	8400.514	8420.161	880.5	1.384698
GLM(NB)	8113.765	8139.961	655.3	1.069236

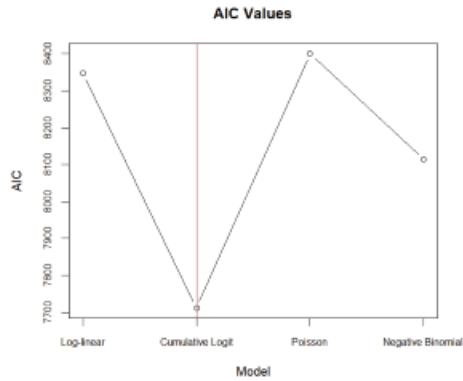


Figure 4: Graph of AIC values for Feeling Bad About Yourself

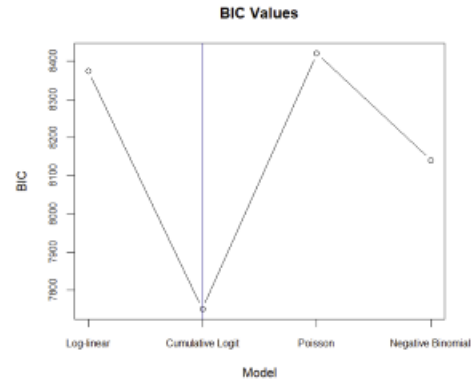


Figure 5: Graph of BIC values for Feeling Bad About Yourself

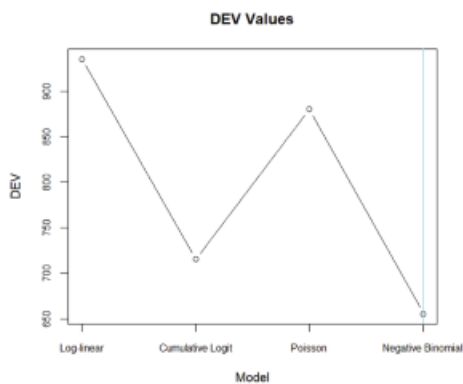


Figure 6: Graph of Deviance values for Feeling Bad About Yourself

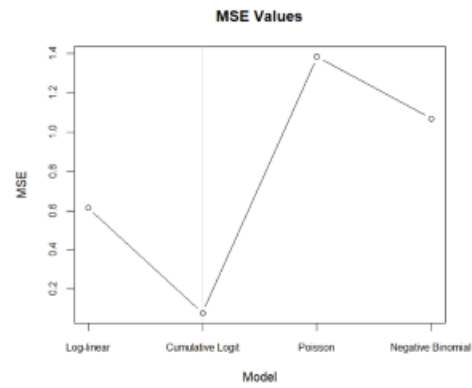


Figure 7: Graph of MSE values for Feeling Bad About Yourself

The graphs of AIC, BIC, and MSE values for Feeling Bad About Yourself depict that the cumulative logit model consistently has the lowest value among other models. This finding suggests that the cumulative logit model is the best model to describe the relationship between feeling bad about yourself, depression, and trouble sleeping. On the other hand, the graph of Deviance Statistic value shows that the negative binomial model has the lowest value among other models. This finding indicates that the negative binomial model is the best model to describe the relationship between the variables.

The statistical measures obtained from the analysis of “Depression” are shown in Table 23. These measures are visualized graphically below, and colored vertical lines represent the best model.

Table 23: Statistical measures of “Feeling Down, Depressed, and Hopeless” .

Models	Statistical Measures			
	AIC	BIC	G^2	MSE
Log-Linear	8152.288	8178.484	1431.3	0.7365462
Multinomial	7228.218	7274.062	1342.899	0.05878926
GLM(Poisson)	8201.003	8220.65	1380.6	1.542902
GLM(NB)	7947.697	7973.894	1064.6	1.324138

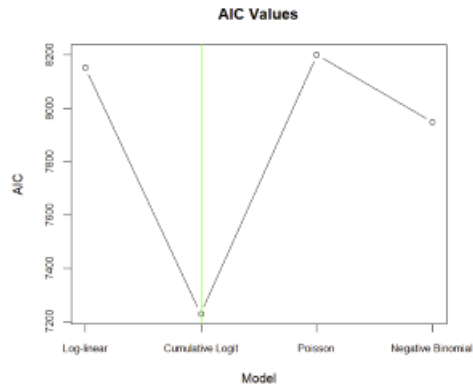


Figure 8: Graph of AIC values for Depression

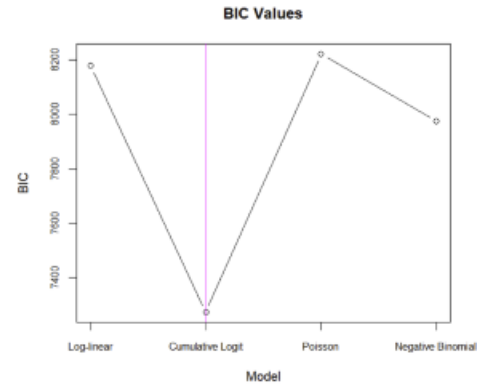


Figure 9: Graph of BIC values for Depression

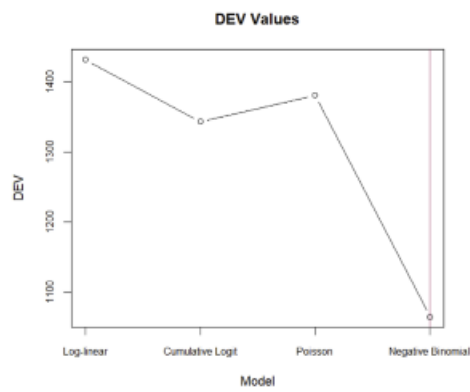


Figure 10: Graph of Deviance Values for Depression

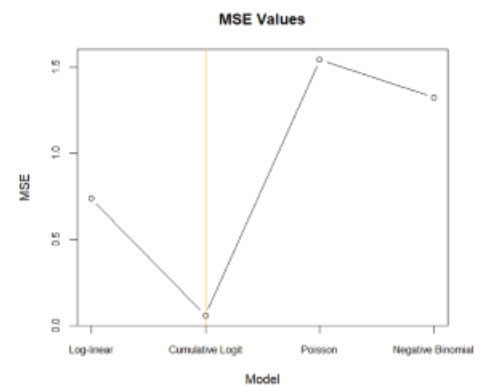


Figure 11: Graph of MSE Values for Depression

The graphs of AIC, BIC, and MSE values for Depression depict that the cumulative logit model consistently has the lowest value among other models. This finding suggests that the cumulative logit model is the best model to describe the relationship between feeling bad about yourself, depression, and trouble sleeping. On the other hand, the graph of Deviance Statistic value shows that the negative binomial model has the lowest value among other models. This finding indicates that the negative binomial model is the best model to describe the relationship between the variables.

The statistical measures obtained from the analysis of “Trouble Sleeping” are shown in Table 24. These measures are visualized graphically below, and colored vertical lines represent the best model.

Table 24: Statistical measures of Trouble Sleeping.

Models	Statistical Measures			
	AIC	BIC	G^2	MSE
Log-Linear	7053.431	7079.627	1238.3	0.5409994
Multinomial	6514.324	6553.619	1255.961	0.06607545
GLM(Poisson)	7077.584	7097.405	1211.9	1.109951
GLM(NB)	6907.558	6933.754	1120.3	0.9982505

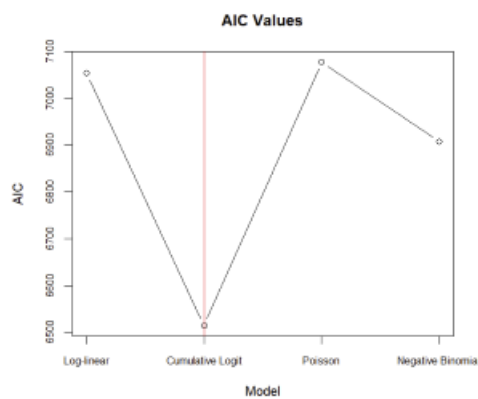


Figure 12: Graph of AIC Values for Trouble Sleeping

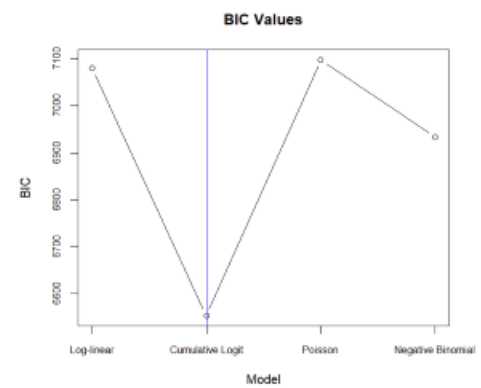


Figure 13: Graph of BIC Values for Trouble Sleeping

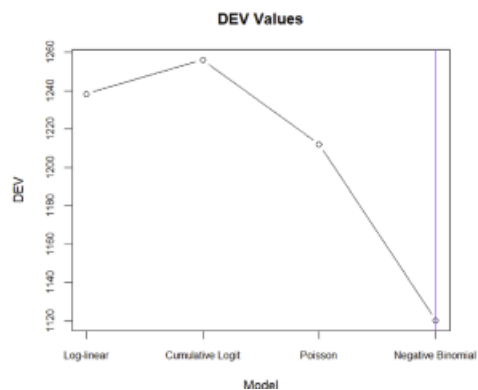


Figure 14: Graph of Deviance Values for Trouble Sleeping

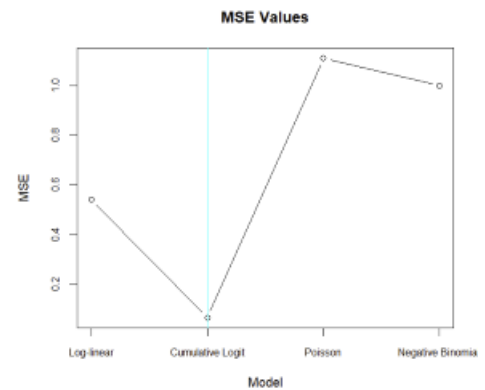


Figure 15: Graph of MSE Values for Trouble Sleeping

The graphs of AIC, BIC, and MSE values for Trouble Sleeping depict that the cumulative logit model consistently has the lowest value among other models. This finding suggests that the cumulative logit model is the best model to describe the relationship between feeling bad about yourself, depression, and trouble sleeping. On the other hand, the graph of Deviance Statistic value shows that the negative binomial model has the lowest value among other models. This finding indicates that the negative binomial model is the best model to describe the relationship between the variables.

CONCLUSION AND FUTURE WORK

In this research, we wanted to investigate if there is any association between depression, trouble sleeping, and feeling bad about yourself. To do this, we employed three models. The first model was the log-linear model which showed We concluded that there are relationships between any pair of variables but no relationship between three variables at the same time. In other words, people with depression commonly experience disturbed sleep patterns, people with trouble sleeping will feel bad about themselves, and those who have depression will feel bad about themselves. But there is not a relationship between three variables together at the same time.

In future studies, exploring the dynamics of the relationship between depression, trouble sleeping, and feeling bad about yourself, non-parametric models such as those that do not assume any assumptions could be employed. Additionally, supervised techniques in machine learning, such as random forest, could assist researchers in seeing if there is any improvement in the relationship of the variables. Furthermore, since there was no association between depression, trouble sleeping, and feeling bad about yourself, so adding more variables as predictors could potentially show more information regarding mental health.

REFERENCES

- [1] Kaneita, Y., Ohida, T., Uchiyama, M., Takemura, S., Kawahara, K., Yokoyama, E., and Fujita, T.(2006). The relationship between depression and sleep disturbances: a Japanese nationwide general population survey. The Journal of Clinical Psychiatry.
- [2] National Institute of Mental Health. (2024). Depression.
<https://www.nimh.nih.gov/health/topics/depression>
- [3] Freepik. (2020). 8 common causes of depression [Infographic].
<https://image.freepik.com/free-vector/8-common-causes-depression-infographics115990-222.jpg>
- [4] Dreamstime. (2020). Depression signs symptom infographic [Image of people with mental health problems: sad woman, despair, stress, loneliness].
<https://thumbs.dreamstime.com/b/depression-signs-symptominfographic-people-mental-health-problems-sad-woman-despair-stress-lonelinessflat-vector-135097853.jpg>
- [5] Bel Marra Health. (2019). Effects of Depression on the Human Body [Infographic]<https://www.belmarrahealth.com/infographics/infographic-effectsof-depression-on-the-human-body/>
- [6] Passarella, S., Duong, M. T. (2008). Diagnosis and treatment of insomnia. American Journal of Health-System Pharmacy, 65(10), 927-934. Cleveland Clinic (2017). Hypersomnia:

Symptoms, Causes and Treatment.<https://my.clevelandclinic.org/health/diseases/21591-hypersomnia>

[8] Healthline. (2018). Insomnia vs. Hypersomnia: Understanding the Differences. <https://www.healthline.com/health/insomnia/insomnia-vs-hypersomnia#treatment>

[9] Healthline. (2022). Low Self-Esteem: Signs, Causes, and How to Improve. <https://www.healthline.com/health/low-self-esteem>

[10] Tsuno, N., Besset, A., and Ritchie, K. (2005). Sleep and depression. *The Journal of Clinical Psychiatry*.

[11] Alvaro, P. K., Roberts, R. M., and Harris, J. K. (2013). A systematic review assessing bidirectionality between sleep disturbances, anxiety, and depression. *Sleep*, 36(7), 1059-1068.

[12] Bhati, S., Richards, K. (2015). A systematic review of the relationship between postpartum sleep disturbance and postpartum depression. *Journal of Obstetrics, Gynecologic and Neonatal Nursing*, 44(3), 350-357.

[13] Dinis, J., Bragança, M. (2018). Quality of sleep and depression in college students: a systematic review. *Sleep Science*, 11(04), 290-301.

[14] Oh, C. M., Kim, H. Y., Na, H. K., Cho, K. H., Chu, M. K. (2019). The effect of anxiety and Depression on Sleep Quality of Individuals with high risk for Insomnia: a population-based study. *Front Neurol*. 2019; 10: 849.

[15] Wang, X., Cheng, S., Xu, H. (2019). Systematic review and meta-analysis of the relationship between sleep disorders and suicidal behavior in patients with depression. *BMC Psychiatry*, 19, 1-13.

[16] Morssinkhof, M. W. L., Van Wylick, D. W., Priester-Vink, S., van der Werf, Y. D., den Heijer, M., van den Heuvel, O. A., Broekman, B. F. P. (2020). Associations between sex hormones, sleep problems, and depression: A systematic review. *Neuroscience Biobehavioral Reviews*, 118, 669-680.

[17] Yin, R., Li, L., Xu, L., Sui, W., Niu, M. E., Xu, R., Srirat, C. (2022). Association between depression and sleep quality in patients with systemic lupus erythematosus: a systematic review and meta-analysis. *Sleep and Breathing*, 26(1), 429-441.

[18] Jiang, Y., Jiang, T., Xu, L. T., Ding, L. (2022). Relationship of depression and sleep quality, diseases and general characteristics. *World journal of psychiatry*, 12(5), 722.

[19] Kroenke, K., Spitzer, R. L. (2002). The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric Annals*, 32(9), 509-515.

[20] Walck, C. (1996). Handbook on statistical distributions for experimentalists (pp. 54-56). Stockholms universitet.

- [21] Krishnamoorthy, K. (2006). Handbook of statistical distributions with applications. Chapman and Hall/CRC.
- [22] NHANES 2015-2016: Mental Health - Depression Screener Data Documentation, Codebook, and Frequencies (cdc.gov)
- [23] Centers for Disease Control and Prevention. (2015). NHANES Questionnaire.
- [24] Lauritzen, S. L. (1979). Lectures on contingency tables. University of Copenhagen.
- [25] Fabozzi, F. J., Focardi, S. M., Rachev, S. T., Arshanapalli, B. G., Hoechstetter, M. (2014). Appendix E: model selection criterion: AIC and BIC. The basics of financial econometrics, 41(1979), 399-403.
- [26] Sullivan, J. (2024). Mean Squared Error (MSE). Statistics. <https://statisticsbyjim.com/regression/mean-squared-error>
- [27] Hornik, K., Zeileis, A., and Meyer, D. (2006). The strucplot framework: visualizing multi-way contingency tables with vcd. Journal of Statistical Software, 17(3), 1 - 48
- [28] Agresti, A. (1996). An Introduction to Categorical Data Analysis (Vol.135). New York: Wiley.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Please feel free to contact the authors at:

- Muslihat Adejoke Gaffari: Gaffari@etsu.edu
- Mostafa Zahed: zahedm@etsu.edu

APPENDIX

Abbreviations

For easy understanding and readability, these are the terms used in this research:

- MDD - Major Depressive Disorder
- ECT - Electroconvulsive Therapy
- RTMS - Repetitive Transcranial Magnetic Stimulation
- VNS - Vagus Nerve Stimulation
- MST - Magnetic Seizure Therapy
- DBS - Deep Brain Stimulation
- GLM - Generalized Linear Models
- AIC - Akaike Information Criterion

- BIC - Bayesian Information Criterion
- MSE - Mean Square Error
- SLE - Systemic Lupus Erythematosus
- MEC - Mobile Examination Center
- CMH - Cochran-Mantel-Haenszel
- GEE - Generalized Estimation Equations
- MLE - Maximum Likelihood Estimation
- NCHS - National Center for Health Statistics
- MEC - Mobile Examination Center
- NHANES - National Health and Nutrition Survey