NOVA IMS Information Management School

MDSAA

Master Degree Program in

Data Science and Advanced Analytics

Data-Driven Exploration of Student Mental Health and Well-being in Prestigious University Settings: A Case Study

Yasmine Boubezari

Master Thesis

presented as partial requirement for obtaining a Master's Degree in Data Science and Advanced Analytics

NOVA Information Management School

Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

Data Driven Exploration of Student Mental Health and Well-Being in Prestigious University Settings: A Case Study

by

Yasmine Boubezari

Master Thesis presented as partial requirement for obtaining the Master's degree in Data Science and Advanced Analytics, with a specialization in Data Science

Supervised by

Maria Helena Baptista, PhD, Nova Information Management School

July, 2025

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism, any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledged the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Lisbon, July 14th 2025

Yasmine Boubezari

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Dr. Maria Helena Baptista, for her guidance and support throughout this thesis.

I am also thankful to the health professional, Dr. Ana Margarida Saraiva who kindly helped me develop the questionnaire, and to everyone who contributed to the data collection step.

My deepest thanks go to my parents for all the times they encouraged me to not drop from university when I had the thought of it. Mom, dad, you were right.

Finally, I would like to thank my family and friends for their constant support, kindness, and attentiveness, which consistently sustained me in all my pursuits.

ABSTRACT

The escalating prevalence of mental health challenges among university students has become a paramount issue in contemporary higher education, particularly within prestigious academic institutions where students face distinctive pressures associated with academic rigor and future career expectations. This study explores the well-being and mental health landscape of students at Institution XYZ, with a dual aim: to identify the main factors associated with student well-being, and to develop Machine Learning models capable of stratifying students by risk level, thereby supporting proactive and data-informed intervention strategies. Adopting a quantitative approach, a comprehensive survey was administered to 258 students, capturing psychological, academic, and contextual dimensions, including measures of self-concept, institutional attachment, academic performance, and perceived satisfaction from the student journey. The research addresses the question: "Which factors impact student well-being and mental health the most at Institution XYZ?" The dataset underwent thorough preprocessing, exploratory data analysis, feature engineering, and model configuration and optimization. Exploratory analyses highlighted meaningful patterns linking student well-being to satisfaction with the university experience, self-concept, and academic performance. Four machine learning classification algorithms -Logistic Regression, Random Forest, Support Vector Machines, and XGBoost- were implemented using cross-validated pipelines to predict risk levels (Low, Moderate, High), with special attention to address class imbalance challenges. Model performance is assessed through macro-averaged F1-scores, recall and precision metrics to ensure fair performance across all risk groups. Random Forest emerged as the most effective model for early risk detection, achieving a recall of 0.60 and macro-F1 of 0.55 on the test set for the High Risk class, correctly identifying three out of five students in this category on the test while maintaining balanced performance overall. Combined with Logistic Regression for interpretability, a hybrid approach is chosen. Feature importance analysis consistently identified student satisfaction, self-concept, social interactions and academic performance as the most influencing factors associated with lower risk across all models. The findings support the initial hypothesis (H1), confirming that student well-being and mental health at Institution XYZ are significantly influenced by at least one of the factors mentioned, as four factors out of six had an impact. These results offer actionable insights for targeted support measures and contribute to the emerging field of data-driven student mental health analytics, promoting evidence-based strategies to foster well-being within academically demanding environments.

KEYWORDS

Machine Learning; Multiclass Classification; Student Mental Health; Well-being; Risk Detection.

Sustainable Development Goals (SDG):

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

EDA	Exploratory Data Analysis
LR	Logistic regression
MDI	Mean Decrease Impurity
MHR	Mental Health Research
ML	Machine Learning
NLP	Natural Language Processing
RBF	Radial Basis Function
RF	Random Forest
SMH	Student Mental Health
SMOTE	Synthetic Minority Oversampling Technique
SVC	Support Vector Classifier
SVM	Support Vector Machine
XGBOOST	Extreme Gradient Boosting

1. INTRODUCTION

The golden era of personal development is widely recognized as a period of transformation, occurring roughly between late adolescence and early adulthood (Arnett, 2000). During these years, individuals stand at the crossroads of challenges and opportunities, navigating academic, social, and emotional experiences that converge to shape their identity (Galambos et al., 2003). It is a time of exploring talents, refining aspirations, and facing limitations—yet also of discovering the resilience necessary to overcome them (Ryff, 1989).

This pivotal phase often coincides with university life, drawing increasing scholarly attention to mental health and well-being among university students (Eisenberg et al., 2007; Stallman, 2010). For instance, Brandão et al. (2011) examined how academic exposure affects students' health status, concluding that the university environment exerts a high influence on both short- and long-term well-being. Similarly, Acosta-Gonzaga (2023) investigated the effects of self-esteem and academic engagement on student performance, which revealed the critical interplay between mental health variables and academic success. Social behavior and interactions with peers can profoundly impact not only academic outcomes but the entire process of personal growth during these formative years (Hefner & Eisenberg, 2009).

Indeed, while universities present vast opportunities—both social and professional—the associated pressures (e.g., academic workload, competitive atmosphere) can exacerbate stress and anxiety, sometimes leading to dropouts (Eisenberg et al., 2009).

Prestigious universities, often perceived as incubators of excellence, offer unique prospects for development, yet their competitive nature intensifies these pressures (Lipson et al., 2019). Top-performing students from diverse backgrounds often encounter heightened competition, where academic excellence becomes the norm, potentially leading to phenomena like impostor syndrome (Pákozdy et al., 2024) and increased pressure to align with the institution's esteemed reputation (Mael & Ashforth, 1992).

At Institution XYZ, a prestigious higher education institution renowned for its data-centric ethos, the intersection of student well-being, institutional dynamics, and analytical innovation presents an interesting environment to explore in the field of mental health and well-being. For confidentiality reasons, the real name of the institution has been anonymized and is referred to as Institution XYZ throughout this document.

This research explores the interplay of well-being and mental health among Institution XYZ students, by employing to psychological theory a data-driven methodology rooted in machine learning to uncover patterns, segment profiles, and predict vulnerabilities among its student. The study aims to not only identify factors influencing student mental health but also propose actionable institutional strategies to strengthen resilience for its students and provide support for the adequate population on the right time. (Lipson et al., 2019).

2. LITERATURE REVIEW

This chapter establishes the theoretical and methodological foundations for investigating mental health risk and well-being among students of the institution XYZ, and adopts an interdisciplinary lens that integrates psychology and data science. It seeks to outline the transformative role of data science in mental health research, contextualize the study's variables, and explore the analytical approaches usually employed in Mental Health Research (MHR).

2.1. Student Mental Health in Prestigious Universities

Mounting evidence indicate a significant rise in psychological challenges for university students such as anxiety, depression, and stress (Drira et al., 2024). Recent studies report that over 30% of university students experience moderate to severe levels of anxiety and depression, a trend that has intensified over the past decade (Auerbach et al., 2018; Lipson et al., 2019).

This escalation is particularly noticeable in high-ranked academic institutions, where the convergence of academic rigor, competitive environments, and institutional prestige creates unique psychological pressures (Pákozdy et al., 2024). Students in such settings often face intense academic competition, which can exacerbate stress and foster low self-esteem, where individuals doubt their accomplishments despite evident success (Clance & Imes, 1978; Pákozdy et al., 2024).

Additionally, the expectations tied to the prestigious reputation of these institutions can amplify pressure to conform to high standards of achievement, further straining mental well-being (Mael & Ashforth, 1992; Wilkins et al., 2012).

These challenges are compounded by social and financial stressors, such as the need to maintain social status or manage high tuition costs, which are particularly salient in elite academic contexts (Adams et al., 2016).

2.2. Application of Data Science in Mental Health Research

Data science has emerged as a transformative force in mental health research, offering tools to analyze complex datasets, uncover hidden patterns, and predict outcomes with unprecedented precision (Drira et al., 2024). This section explores its applications in psychology and medicine –with a focus on student mental health- and examines the methodological approaches in different papers that cover similar themes.

2.2.1. DATA COLLECTION APPROACHES

Data collection is a critical step in our research for capturing mental health risk factors and the characteristics associated to these risk factors. From previously documented data collection approaches, we have retained some points that aligns with our research.

Self-reported questionnaires, such as the ones used in the Healthy Minds Study (Lipson et al., 2019), are widely employed due to their scalability and ability to standardize psychological constructs like stress and well-being. These studies often use 5-point Likert scales to measure subjective well-being (Diener et al., 2010), which are suitable for this research to assess students' mental health states reliably.

Categorical measures are also valuable for capturing help-seeking behaviors, which are critical for identifying at-risk students (Rickwood et al., 2005). Demographic variables, like age and gender, are essential for contextualizing risk factors and ensuring equitable representation, as mentioned by Drira et al. (2024), who note that only 18% of Student Mental Health (SMH) studies adequately address demographic diversity.

To ensure the reliability of survey instruments, Cronbach's alpha is used to assess the internal consistency of scales with three or more items, such as those measuring well-being or anxiety (Eisinga et al., 2013). For constructs captured by two-item scales, inter-item correlations like Pearson's (for approximately normal data) or Spearman's rank correlation (for non-normal or ordinal data) are applied to validate reliability, as recommended by Eisinga et al. (Eisinga et al., 2013).

A further research argues that parametric methods, including Pearson's r, are sufficiently robust for Likert-type data -especially when using 5-point scales- as violations of normality tend to have minimal impact on results (Norman, 2010).

Nonetheless, given the ordinal and potentially skewed nature of Likert data -where responses often concentrate toward high values, prior research suggests to use a dual approach with both pearson and spearman for solid results (Sullivan & Artino, 2013).

2.2.2. MODELLING AND MACHINE LEARNING APPROACHES

Supervised Machine Learning (ML) is widely used in MHR to classify students into risk categories, based on symptom severity or composite scores (Auerbach et al., 2018; Ding et al., 2025; Lipson et al., 2019).

From a selection of papers in MHR with similar classification and pattern learning objectives, we notice the frequent usage of tree based models, logistic regression and Neural networks models. Logistic Regression, valued for its interpretability and suitable to our research, effectively models multiclass probabilities in SMH contexts (Bhavani & Naveen, 2024; Graham et al., 2019). Random Forests, an ensemble method, provide robustness to noisy data and

feature importance to identify key risk factors and which is one of this research's objective (Breiman, 2001; Shatte et al., 2019). Gradient boosting models like XGBoost excel in multiclass prediction by capturing intricate variable interactions (Chen & Guestrin, 2016; Ding et al., 2025). Another model commonly used in MHR and which is adapted to our task is Support Vector Machines (SVM), as it capture non-linear patterns with radial basis function kernels (Bzdok & Meyer-Lindenberg, 2018; Madububambachu et al., 2024).

As neural networks is commonly used with large data, it is not suitable for our research since the collected data from questionnaire is estimated to less and we focus on logistic regression, tree based models and support vector machines models.

According to Chawla et al. (2002), class imbalance is a common issue in SMH as minority populations are underrepresented, and addressed it by generating synthetic minority samples with techniques like SMOTE. Another alternative found for addressing imbalance is the use of class weighting in Logistic Regression and Support Vector machines (SVM) as discussed by Hastie et al. (2017).

2.3. THEORETICAL FOUNDATIONS AND CONCEPTS RELATED TO STUDENTS MENTAL HEALTH AND WELL-BEING

Psychological and higher education research has long sought to understand the factors shaping students' mental health, revealing a complex interplay of variables that influence well-being, a multidimensional construct encompassing emotional health, life satisfaction, and purpose (Diener et al., 2018; Ryff, 1989). Well-being, as a central outcome, is shaped by psychological, social, academic, and economic factors, each grounded in robust theoretical frameworks. For instance, self-concept—how students perceive their identity and abilities—bolsters resilience but can falter under competitive pressures (Shavelson et al., 1976). Social interactions, rooted in social support theory, act as a buffer against stress by fostering emotional and structural support (Thoits, 2011). Academic performance, tied to self-efficacy and pressure, serves as both a motivator and a stressor, influencing psychological health (Richardson et al., 2012). Financial stress, driven by economic pressures like debt, consistently undermines well-being (Adams et al., 2016). Satisfaction with the university experience, reflecting alignment between expectations and reality, enhances mental health and retention (Schreiner & Nelson, 2013). Reliance on institutional image, tied to university prestige, could inspire pride but also heighten stress (Wilkins et al., 2012).

Crucially, the proactive seeking of mental health resources emerges as a pivotal coping mechanism. Grounded in the Theory of Planned Behavior, this behavior mitigates mental health risks by enabling students to manage distress, particularly in high-pressure academic settings (Ajzen, 1991; Gulliver et al., 2018). By integrating well-being with help-seeking, researchers stratify students into risk categories (low, moderate, high), as informed by stress and coping frameworks (Biggs et al., 2017; Eisenberg et al., 2016).

Demographic variables -age, gender, level of study, area of study, and resideny status- further contextualize these dynamics, shaping stress experiences and help-seeking tendencies (Auerbach et al., 2018). Exploratory factors, such as rewarding aspects of university life (e.g., belonging) and potential institutional initiatives (e.g., support programs), enhance well-being by fostering resilience and engagement (Seligman, 2018; Tinto, 2017). However, psychological research acknowledges that questionnaires cannot capture the full complexity of mental health. Latent factors, such as resilience, past trauma, or external stressors, often explain residual variance in well-being, introducing bias in predictive models (Keyes & Martin, 2017). This study leverages these validated constructs to explore mental health risks, recognizing the interplay of measurable and unmeasurable factors in shaping student outcomes.

3. CONCEPTUAL MODEL

This chapter aims to understand the study's objective and core construct. It starts with the research hypothesis, followed by the detailed definitions of the variables to use -mentioned on literature review, point 2.3-, and finishes with a summary of the relation and links between those variables.

3.1. HYPOTHESIS BUILDING

Regardless of the growing body of research on student mental health (SMH), several gaps remain that this study seeks to address, particularly in the context of Institution XYZ.

First, despite the existence of numerous studies on student mental health, particularly in prestigious universities, none of them have specifically examined the factors influencing mental health at Institution XYZ. This gap is particularly significant given the rising global interest in student well-being, as institutions increasingly recognize its impact on academic success and retention (Lipson et al., 2019). Investigating factors like academic performance, financial stress, and social interactions at Institution XYZ is thus interesting to understanding and supporting its student population.

Second, while individual influencing factors such as academic performance, financial stress, and social interactions have each been identified in prior studies as impactors on student mental health, no single study has combined this particular selection of factors into a one single and comprehensive analysis. Pascoe et al. (2020) highlight academic performance as a key stressor affecting well-being, Adams et al. (2016) emphasize the role of financial stress in mental health challenges, and Thoits (2011) underscores the protective effect of social interactions. However, these factors are studied in isolation or with other variables, limiting a holistic understanding of their combined impact on student mental health. Additionally, the inclusion of 'Unmeasured Factors' in this study addresses a critical limitation noted by Keyes et al. (2017), as latent factors like resilience or hidden stressors often remain unmeasured.

To address these gaps, this study proposes a quantitative research approach to identify students of institute most likely at risk of developing depression and other mental health issues, explore the factors contributing to their mental health challenges, and understand the solutions they expect.

The **research question** guiding this study is:
"Which factors impact the most students ‘well-being and mental health at Institution XYZ?"
This question aims to uncover the key determinants of mental health at institution XYZ, and leads to the following hypothesis:

H0: The factors “self-concept”, “social interactions”, “reliance on institutional image”, “academic performance”, “financial stress”, and “satisfaction from the student experience” do not hold a significant influence on students’ well-being and mental health at Institution XYZ.

H1: At least one of the factors “self-concept”, “social interactions”, “reliance on institutional image”, “academic performance”, “financial stress”, and “satisfaction from the student experience” significantly influence the well-being and mental health of students at Institution XYZ.

To answer the hypothesis, we structure the research objectives into two parts:

- Explore the factors that influence student well-being and their engagement with mental health resources, with a focus on the high risk students category.
- Apply Machine Learning models to predict the target variable "Student Risk Status", using the impacting factors as predictive variables.

- **Expected theoretical and practical outputs**

The research is expected to yield on the theoretical side a deeper understanding of how academic, social, psychological and financial factors interact to influence student mental health in this specific academic environment, and further validation of the role of data-driven methods in psychological research.

On the practical side, the expected output is the production of a segmentation model to identify at-risk student profiles and tailored recommendations for administrators at Institution XYZ to enhance student well-being through targeted support initiatives.

- **Feasibility and coherence of the approach**

Institution XYZ exemplifies the environment where the student mental health dynamics are at play. As a high-ranked university, it attracts top-performing students from diverse backgrounds, who navigate both the opportunities and pressures of an academically demanding and competitive setting. This makes it a compelling environment for investigating student mental health, particularly through a lens that leverages the institution’s strength in data-driven methodologies.

An interdisciplinary approach that combines psychological theory with advanced data science techniques is essential to uncover latent patterns, predict vulnerabilities, and propose targeted interventions (Drira et al., 2024).

The literature review confirms that the proposed research topic examining the mental health of students at Institution XYZ through a quantitative, data-driven approach is both feasible

and coherent. The use of machine learning models to identify students at risk of developing mental health issues, the recognized importance of student feedback in tailoring interventions (Tinto, 2017), and the potential of data science to bridge psychological constructs with actionable insights (Drira et al., 2024) provide a solid and coherent foundation for this study.

3.2. CORE CONSTRUCT

The conceptual model hinges on a carefully curated set of variables designed to elucidate the mental health dynamics of students at Institution XYZ, with "Well-being" positioned as the central construct, and six influencing factors -reliance on institutional image, self-concept, satisfaction from the student experience, academic performance, financial stress, and social interactions-, selected as its common influent factors. This selection is theoretically grounded in psychological and educational research, reflecting the multifaceted nature of well-being within a high-ranking university context.

Each variable's inclusion is justified by its established relevance to student mental health, supported by contemporary empirical evidence, and which is developed on this chapter gradually in order to highlight the model's applicability to the study population. The logical flow found between our variables is illustrated on this following concept model:

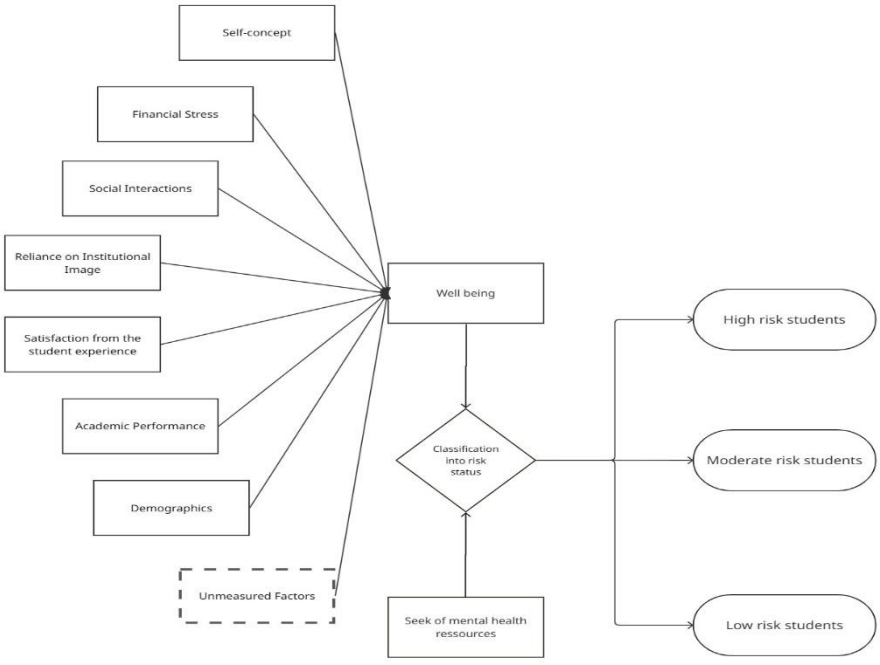


Figure 1: Proposed Conceptual Model of Student Well-being Determinants and Risk Classification

3.2.1. CENTRAL VARIABLE: WELL-BEING

Well-being serves as the principal variable of interest, which combines the emotional and psychological health of students as the cornerstone of the model. It is chosen for its comprehensive scope, aligning with validated frameworks like the subjective well-being model, which integrates affective and cognitive dimensions of mental health (Diener et al., 2018).
Its centrality is substantiated by extensive research that identified well-being as a critical indicator of student resilience and vulnerability in higher education settings, where academic and social pressures converge (Lipson et al., 2019).

3.2.2. WELL-BEING'S INFLUENCING FACTORS VARIABLES

Six factors are selected based on their documented impact on well-being, and tailored to the unique environment of a high-ranked institution like XYZ.

Self-concept refers to an individual's overall perception of themselves, encompassing both descriptive beliefs about their identity and evaluative judgments of their abilities (Shavelson et al., 1976). In competitive academic environments, a positive self-concept enhances well-being by fostering resilience and satisfaction, while a challenged self-concept can amplify stress and undermine psychological health, which subsequently leads to the risk of depression and student failure (Marsh & Craven, 2006). In other terms, self-concept positively influence well-being.

Social Interactions denote the quality, frequency, and extent of interpersonal relationships with peers and the broader community, framed within social support theory as a buffer against psychological distress (Thoits, 2011). Cohen & Wills (1985) distinguish between structural support (e.g., network size) and functional support (e.g., emotional support), with recent studies affirming its protective effect on student mental health. In university settings, where social integration is vital, this construct shapes emotional resilience, warranting its role as a determinant of well-being (Thoits, 2011).

Academic Performance is defined as the level of success in educational tasks, typically measured through grades or self-reported achievement (Pascoe et al., 2020). It is highlighted by Educational Psychology as a source of validation that becomes a stressor, with meta-analyses linking it to well-being through self-efficacy and pressure (Richardson et al., 2012). At Institution XYZ, where academic excellence is a hallmark, its impact on psychological health is particularly salient, supporting its inclusion as an influencing factor positively associated to well-being.

Financial stress is introduced to account for the economic pressures that students may face, a factor consistently identified as a major contributor to diminished well-being in higher education (Adams et al., 2016). This variable is especially relevant for students at a prestigious

institution, where costs and expectations may amplify financial burdens, which introduce this factor as negatively influent to well-being.

Satisfaction from the Student Experience is the degree of contentment with the academic and social facets of university life (Schreiner & Nelson, 2013). Building on student engagement theories (Astin, 1984), it reflects the alignment between expectations and reality, with studies linking higher satisfaction to improved mental health and retention (Schreiner & Nelson, 2013). In a high-performing context like Institution XYZ, this variable captures the psychological balance critical to well-being. The capacity to balance between the academic and the personal life positively influence students' well-being.

Reliance on Institutional Image reflects students' dependence on the prestige and reputation of their university (Wilkins et al., 2012). Emerging from higher education branding literature, this construct influences well-being through pride and belonging as it can lead to a negative impact via heightened expectations and fear of not being enough (Stensaker, 2015; Wilkins et al., 2012). Its significance in prestigious academic settings justifies its inclusion as a negative well-being influencer in this study, as the institution's status can significantly shape students' psychological health.

3.2.3. DEMOGRAPHICS

Demographic variables are individual characteristics that contextualize mental health outcomes, by offering essential initial insights into a student's background (Arnett, 2000; Smith & Khawaja, 2011). Although these metrics are straightforward to gather, they do not directly capture the student's mental health state. As prior research has noted, demographic segments can differ in how they experience stress (Eisenberg et al., 2007) or in their likelihood of seeking help (Hefner & Eisenberg, 2009).

Age

Age can significantly influence students' academic experiences and mental health outcomes. Young adults (often those aged 18–24) may be in the midst of critical developmental transitions -such as leaving the parental home or starting university for the first time- which can heighten stress and vulnerability (Arnett, 2000). Mature students (25 or older) may juggle additional responsibilities such as full-time employment or family obligations, potentially leading to different patterns of stress and well-being (Tilley, 2014).

Including age categories allows the study can capture these varying life-stage challenges and better understand how mental health correlates with distinct age brackets.

Gender

Gender is another key demographic factor that can shape mental health experiences. Research indicates that women and men often differ in how they perceive stress, seek help, and manifest symptoms of anxiety or depression (Eisenberg et al., 2007; Piccinelli & Wilkinson,

2000). Additionally, non-binary or transgender students may face unique stressors related to identity and discrimination (Hendricks & Testa, 2012).

Including gender thus allows a more nuanced analysis of mental health trends and the identification of potential disparities or specific needs among diverse gender groups.

Level of Study

Students at different academic levels frequently report varied stressors and concerns. Undergraduate students may grapple with transitioning to university life and managing foundational coursework, while Master's or PhD candidates often face pressures tied to research requirements, thesis or dissertation deadlines, and career transitions (Offstein et al., 2004). Doctoral students, in particular, can experience heightened stress due to longer program durations and the expectation of producing original research (Pyhältö et al., 2009). Clustering respondents by their level of study might provide clarity into how academic demands correlate with mental health and well-being.

Area of Study

According to multiple studies, academic demands can vary widely depending on the discipline, and different fields may foster distinct campus cultures (Levecque et al., 2017; Salanova et al., 2010). In this study, it corresponds to the different area programs offered by Intitution XYZ in order to investigate if some programs affect students well being more than others, or require more efforts and attention.

By distinguishing participants' fields, it becomes possible to investigate whether certain academic areas align with higher or lower stress and if specific resources might be tailored to each group.

Residency Status

International students typically face unique stressors including language barriers, cultural adaptation, homesickness, and visa-related pressures(Sawir et al., 2008; Smith & Khawaja, 2011). These challenges can exacerbate feelings of isolation and psychological distress, and which potentially impact academic performance and the overall well-being (Rosenthal et al., 2008).
Differentiating between local and international students thus illuminates how cultural and contextual factors intersect with mental health experiences in the university context.

Collectively, these demographic variables connect to well-being by shaping the context in which psychological health is experienced, which might facilitate exploratory segmentation of well-being and risk profiles, enhancing the study's interpretability.

3.2.4. BEHAVIORAL MODERATORS

Seek of Mental Health Resources is defined as the proactive inclination to engage with psychological support services, encompassing attitudes and behaviors toward help-seeking (Rickwood et al., 2005). This variable directly relates to well-being by acting as a behavioral moderator, reflecting the extent to which students actively manage their psychological health.

The construct has evolved as a critical coping mechanism in academic settings according to health psychology and the Theory of Planned Behavior (Ajzen, 1991), with research demonstrating its protective role against mental health decline (Gulliver et al., 2018). In this study, it serves as a behavioral moderator of well-being and a protective factor against mental health decline in university settings, informing risk stratification and reflecting resource utilization at Institution XYZ.

Beyond its bi-directional influence with well-being, this variable's behavioral and exploratory constructs enrich the model by capturing students' actions, preferences, and contextual diversity.

3.2.5. DERIVED VARIABLE

The formulation of **Student Risk Status** lies on a structured combination of "Well-being" and "seek of mental health resources". The logic underpinning this derivation rests on the premise that risk arises from the interaction between a student's psychological state and their engagement with support mechanisms, a perspective informed by the transactional model of stress and coping (Biggs et al., 2017). "Well-being" serves as the foundational measure of emotional health, while "Seek of Mental Health Resources" moderates this state by reflecting the degree of proactive behavior -or absence thereof- in managing distress. This approach is bolstered by recent studies demonstrating that help-seeking significantly influences mental health trajectories, particularly in environments where resources are accessible yet variably utilized (Gulliver et al., 2018).

Based on Eisenberg's research on help-seeking behaviors among college students(Eisenberg et al., 2016), Rickwood's framework on professional mental health service utilization (Rickwood et al., 2005) and Gulliver's exploration of barriers to mental health care in youth (Gulliver et al., 2018)**,** the output classes of the variable "Student Risk Status" are defined as *Low Risk, Moderate Risk*, and *High Risk.*

This classification scheme (detailed in chapter five, point 5.1.2.) is informed by Lazarus and Folkman's transactional model of stress and coping (Biggs et al., 2017), and reflects the interaction between a student's current psychological state (Well-being) and their behavioral response to distress (Seek of Mental Health Resources). High levels of well-being are assumed to confer resilience, regardless of help-seeking activity. Moderate well-being is interpreted contextually, with help-seeking behavior serving as a moderating factor—those who express

openness to seeking help are categorized as Low Risk, while those who remain disengaged are classified as Moderate Risk. Low well-being in combination with a lack of help-seeking indicates High Risk, as consistent with previous literature emphasizing that untreated psychological distress increases vulnerability (Eisenberg et al., 2016; Rickwood et al., 2005).

3.2.6. LATENT INFLUENCES: UNMEASURED FACTORS

The variable "Unmeasured Factors" is integrated in this study to estimate the impact of latent influences on students' well-being that are not accounted for by the explicitly measured factors. According to literature, Well-being is a complex and multifaceted construct that cannot be fully explained by a limited set of factors circumstances (Keyes et al., 2010; Thoits, 2011). Numerous **unmeasured influences** may also contribute to students' mental health, including:

- **Family Background**: Parental support, family conflicts, or financial dependency may impact students' stress levels (Tilley, 2014).
- **Past Traumas**: Students who have experienced traumatic events may exhibit lower well-being regardless of their academic success (Keyes, 2002).
- **Eating Disorders and Health Issues**: Psychological distress can stem from physical health concerns, dietary habits, or sleep disorders (Diener, 2000).
- **Unspecified External Stressors**: Work commitments, cultural pressures, and other contextual factors may also shape student well-being in ways that are difficult to measure through a survey (Ryff, 1989).

To consider other factors that are out of our focus on this study but that might still influence the well-being, we create the variable "unmeasured factors", derived from the variable "well-being" and the six explicitly measured influencing factors: self-concept, reliance on institutional image, social interactions, financial stress, academic performance, and satisfaction from the student experience. The inclusion of this variable is theoretically grounded in psychological and statistical literature, which highlights that observable predictors often fail to fully explain complex constructs such as well-being, leaving room for unobservable factors such as personal resilience, prior traumatic experiences, or external life circumstances (Keyes et al., 2010; Thoits, 2011).

Based on the study of Keyes & Martin (Keyes & Martin, 2017) that demonstrate the directional distinction of unmeasured factors -unobserved variables can either exacerbate or ameliorate mental health outcomes-, and the statistical methodologies of Wooldrige (2020) to assess unexplained variance where the extent of deviation from expected patterns reflects the strength of omitted variables, "Unmeasured Factors" is operationalized into six distinct categories:

- Low unmeasured positive influent factors: low presence or significance of unmeasured factors, which largely affect the well-being in a positive sense.

- Low unmeasured negative influent factors: low presence/impact of unmeasured factors, which largely affect the well-being in a negative sense.
- Medium unmeasured positive influent factors: moderate presence of unmeasured factors, which largely affect well-being in a positive sense.
- Medium unmeasured negative influent factors: moderate presence of unmeasured factors, which largely affect well-being in a negative sense.
- High unmeasured positive influent factors: high presence of unmeasured factors, which largely affect well-being in a positive sense.
- High unmeasured negative influent factors: high presence of unmeasured factors, which largely affect well-being in a negative sense.

The methodological derivation of this variable from our collected variables is detailed in chapter five, point 5.1.2.

3.2.7. EXPLORATORY VARIABLES

Rewarding Aspects refer to the elements of the university experience perceived as fulfilling by students, such as career development, personal growth, or institutional reputation (Kinzie & Kuh, 2017). This variable connects to well-being by highlighting positive experiences that enhance students' emotional and psychological health, acting as resilience-building factors. Studies in positive psychology and student engagement demonstrate that rewarding aspects—such as recognition of achievement or a sense of belonging—bolster well-being by fostering satisfaction and motivation (Kinzie & Kuh, 2017; Seligman, 2018). Drawing from the National Survey of Student Engagement and positive psychology (Seligman, 2018), these aspects enhance well-being by fostering satisfaction and motivation. The purpose of including "Rewarding Aspects" in this study is exploratory, as we want to better understand the protective elements specific to the student experience at Institution XYZ, and explore the different satisfaction points of each student risk class.

Potential Initiatives are prospective institutional actions or programs desired by students to improve their university experience, such as enhanced support services or wellness programs (Tinto, 2017). This variable links to well-being by identifying interventions that could proactively support psychological health and engagement, addressing gaps in current offerings. Research on student persistence and institutional effectiveness underscores that targeted initiatives—such as mental health resources or community-building programs—can significantly improve well-being and reduce attrition (Eisenberg et al., 2016; Tinto, 2017). For Institution XYZ students, these initiatives are particularly relevant given the high-stakes context, where tailored support could mitigate stress and enhance resilience. The purpose of "Potential Initiatives" in this study is exploratory, aiming to uncover student-preferred strategies that could strengthen well-being, offering actionable insights for institutional policy and complementing the diagnostic focus of "Student Risk Status".

Demographics (detailed in point 3.2.3) are further explored across risk profiles and identify tailored interventions, leveraging their role in EDA to uncover patterns in well-being and risk tendencies.

3.3. SUMMARIZATION OF THE DYNAMICS BETWEEN VARIABLES

The conceptual model integrates well-being, its influencing factors, behavioral moderators, derived variables, and exploratory/contextual elements to explain mental health outcomes among Institution XYZ students. Well-being, as the central construct, is influenced by positive and negative factors. Positive factors -self-concept, social interactions, academic performance, and satisfaction from the student experience- enhance well-being by fostering resilience, emotional support, self-efficacy, and contentment. Conversely, the negative factors -financial stress and reliance on institutional image- undermine well-being by amplifying anxiety and pressure in Institution XYZ high-stakes environment.

Demographic variables (age, gender, degree studied, area of study, residency status) indirectly influence well-being by shaping experiences of stress or support, often mediated through primary factors like financial stress or social interactions. For instance, international students may face heightened financial or social challenges, affecting their psychological health. Help-seeking behavior moderates the relationship between well-being and mental health outcomes, with proactive engagement reducing vulnerability, particularly for students with moderate or low well-being.

The interaction of well-being and help seeking informs the derived Student Risk Status (Low, Moderate, High Risk), which identifies students' vulnerability to mental health decline and academic failure. High well-being typically indicates Low Risk, while low well-being without help seeking signals High Risk, guiding targeted interventions. Unmeasured factors capture latent influences (e.g., personal resilience or external stressors), ensuring the model accounts for unobserved heterogeneity.

Exploratory variables -rewarding aspects and potential initiatives- complement this framework by identifying protective factors and student-suggested interventions to enhance well-being. Demographics further enrich risk profiling by segmenting students (e.g., by degree or age) to tailor strategies like counseling or financial aid. Together, these dynamics form a comprehensive framework that links well-being to risk assessment and actionable institutional policies, addressing the unique needs of Institution XYZ students.

Scope and Limitations of Variable Interactions in the Conceptual Model

While the conceptual model is built upon a robust selection of variables, it is imperative to acknowledge the complex interplay inherent among these variables within psychological frameworks.

In the broader literature, the factors selected are not isolated one from another; rather, they often exert reciprocal influences on each other, creating a web of interactions that could, if fully explored, lead to infinite loops of causality (Sowislo & Orth, 2013; Thoits, 2011). For instance, self-concept may shape satisfaction from the student experience, while academic performance could simultaneously influence financial stress and social interactions. Such multidirectional relationships are well documented in psychological research, reflecting the dynamic nature of mental health constructs (Pascoe et al., 2020).

However, to maintain focus and ensure the model's analytical tractability, this study deliberately prioritizes the most robust and direct links pertinent to the research objective and which are the unidirectional links with our main variable "Well-being".

By focusing on these core and unidirectional connections, the model avoids diverging into tangential explorations that could dilute its focus or complicate its interpretability. This approach is consistent with methodological recommendations in psychological modeling, where clarity and specificity are favored over exhaustive mapping of all possible interactions when addressing a targeted research question (Funder & Ozer, 2019).

3.4. TOOLS

To elaborate this study, we are using the following tools for their advantageous and easy usage:

Table 1 : Overview of the research tools

Tools	Purpose	Application in this study
Jupyter Notebook (Via Anaconda) and google collab	Coding environment	Usage of the interactive environment for coding.
Python liberaries	Data analysis, preprocessing and modelling	Usage of Python and its libraries for EDA, preprocessing, and modeling.

Qualtrics survey	Data collection	Online survey design and distribution.
Miro	Diagrams	Diagram creation.
Litmaps	Literature mapping	Usage in mapping connections between articles.
ResearchGate	Literature review	Access to full text articles.
PubMed	Literature review	Sources peer-reviewed psychology and health articles.
Zotero	Reference management	Citation and management of references.
Grok (third version)	Text fluency support	Enhances text clarity and coherence and ensures an academic writing tone.

4. DATA COLLECTION INSTRUMENTS

The quality and reliability of data collection instruments are critical to ensure the validity of research findings, particularly when we explore sensitive topics like well-being and mental health. In this chapter we mention in detail the systematic process undertaken to design, validate, and distribute the questionnaire used in this study, based on the methodology found on literature review for data collection (Boateng et al., 2018; Crameri et al., 2016; Sallehuddin Md Yusof, 2023; Thomas et al., 2024).

The undertaken steps of the data collection are summarized on the following diagram:

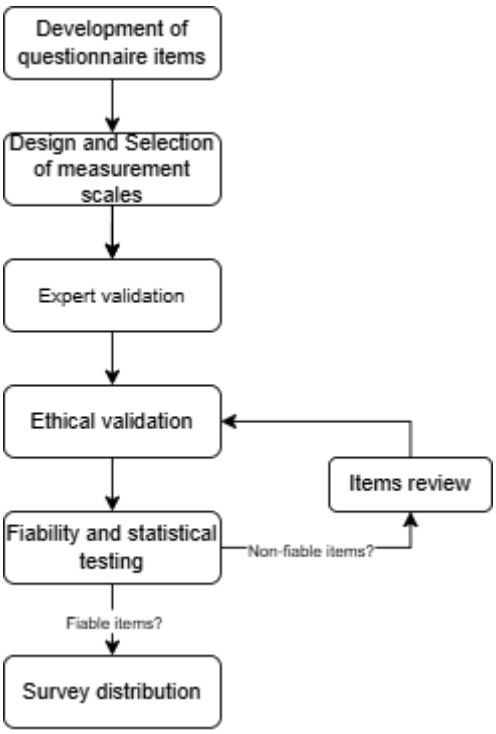


Figure 2: Flowchart of the Data Collection Process

Each step of this schema is carefully executed to capture the desired variables while adhering to methodological and ethical standards, and which are detailed on the following points.

4.1. DEVELOPMENT OF THE ITEMS QUESTIONS

The development of the questionnaire begins with the identification of key variables derived from the literature review and detailed in *conceptual model's core constr*uct point 3.2, and which are selected for their theoretical relevance and empirical grounding in prior research. Their formulation is guided by definitions from the literature and inspired by validated instruments, including the *Rosenberg Self-Esteem Scale* (Rosenberg, 2011), the *WHO-5 Well-*

Being Index (*The World Health Organization-Five Well-Being Index (WHO-5)*, 2024)**,** and the *Mental Help Seeking Attitudes Scale* (Hammer et al., 2018)**.**

Each item is carefully built to reflect the conceptual and operational dimensions of its respective variable. Particular attention is given to the clarity, tone, and sensitivity of the wording, due the personal nature of the topics explored. The phrasing is refined to minimize discomfort and encourage openness, thus maintaining ethical standards in dealing with sensitive self-reported data.

4.2. SCALES AND MEASUREMENT SELECTION

The design of the questionnaire involves the selection of appropriate measurement scales tailored to the variables under investigation, drawing from established practices in the literature, and takes practical considerations for respondent experience in order to make it easy to use and not long.

This process is facilitated by Qualtrics, a data collection platform, which supports the construction and structuring of the instrument. The following measurements are used for our questions:

- **Multiple-choice questions**

We used five multiple choice questions for the demographic variables -age, gender, area of study, level of study, and residency status-, where every student of Institution XYZ could find his corresponding profile with the available options.

The formulation type of questions (e.g., “What is your gender?”) are automatically suggested by Qualtrics during the form-building process, which reflects their widespread usability and acceptance in survey design.

- **Likert matrix**

To measure well-being and its associated factors, we chose the Likert matrix, a method widely recognized in psychological research as the self-esteem test scaling of Rosenberg demonstrate (Rosenberg, 1965).

Typically, a single matrix captures one variable. However, to simplify the questionnaire for respondents -given the sensitive nature of the topic- and to shorten its length to maximize completion rates and minimize dropouts, a single matrix is used for multiple related items. All responses correspond to the same range (from “1 = Strongly Disagree” to “5 = Strongly Agree”), and are compatible with these scale labels.

While this decision might increase preprocessing time (e.g., separating items into distinct variables, pilot testing, etc), we consider it as a deemed a worthwhile time and effort sacrifice in order to create a user-friendly instrument for participants, and adhere to the ethics committee's guideline of approximately 15 questions per questionnaire.

The selection of a 5-point Likert scale ("1 = Strongly Disagree" to "5 = Strongly Agree") is chosen over a 7-point scale after consideration of similar psychological studies, such as Rosenberg's Self-Esteem Scale referred previously, which commonly employ 5-point scales. We consider that this decision balances granularity with simplicity, and facilitates respondent comprehension and consistency across items.

To reduce response biases and encourage more authentic answers, items to capture well-being and its influencing factors, since they follow the same scaling (from "1=strongly disagree" to "5= strongly agree"), are intentionally shuffled throughout the Likert matrix. This approach helps prevent artificial consistency and minimizes the influence of question (Schwarz, 1999). Additionally, it draws on **the incubation effect**, where cognitive distancing between related items encourages more thoughtful and independent responses (Sio & Ormerod, 2009). This design choice aligns with our limited number of question authorized, in order to obtain the most accurate answers and guarantee the best practices in this psychological test construction.

- **Multiple-choice questions with open-ended options**

For the items "Rewarding Aspects" and "Potential Initiatives", requiring a variety of infinite and creative possibilities according to each participants optic, we use a multiple-choice format allowing multiple selections across the most common options, and which is supplemented by an open-ended option ("Other: please specify") for a larger choice.

This approach provides participants with flexibility to express additional perspectives beyond predefined categories.

- **Trichonomous questions**

Three additional questions related to the questionnaire's assessment and evaluation are included with a "Yes/No/Neutral" scale.

First, we use a question on the comprehension of terms (Q10. Were the questions and words used on this survey clear and understandable?), to identify potential bias from jargon. Second, a question on the ease of understanding and navigating the questionnaire (Q11. Were the scales utilized to answer easy to use and appropriate?), to evaluate its user-friendliness as intended.
Finally, a question on whether the questionnaire caused discomfort (Q12. Did you experience

any sort of discomfort or anxiety while answering this survey?), to verify if the goal of minimizing respondent unease was achieved.

- **Open-ended feedback question**

A final open-ended question is added to gather qualitative feedback, in order to allow participants to share any additional thoughts on the questionnaire or the broader research topic.

This scaling approach is based on the methods taken for survey development in existing psychological and health surveys, and that were referred to all along the steps. The measurements are adapted and tailored for each of question, in order to capture the maximum of insights while facilitating the answering process for participants and minimizing discomfort for participants.

4.3. EXPERT VALIDATION

To validate the initial draft of the questionnaire, a review of the questionnaire's items and scaling measurements is conducted by Professor Ana Margarida Faro Craveiro Saraiva, psychologist with an expertise with students in the adolescence phase, working in "Agrupamento de Escolas de Parede". This step is taken to validate and approve the two first steps of our data collection (Development of Questionnaire Items, and Design and Selection of Measurement Scales) by assessing the suitability of the questions in capturing the targeted variables effectively, particularly given the sensitive nature of the research topic. Professor Saraiva evaluated the clarity, relevance, and alignment of the items with the study's objectives, and provided a feedback that informed subsequent revisions.

The feedback approved our questionnaire (Appendix A), after suggesting the following adjustments:

- Clarity and simplicity in wording: questions should employ clear, straightforward and understandable language by the target population.
- Focus on single constructs per item: each question should address a single concept to maintain precision. For example, an item such as "I feel good about my life thanks to my finances" reflects general well-being AND financial influence, and requires a separation into two distinct questions like "I am satisfied with my life" and "My financial situation supports my well-being".
- Polarity of items: the questionnaire should incorporate a mix of positively and negatively worded items to mitigate response biases.
- Length consideration to reduce respondent fatigue: lengthy questionnaires are to avoid, as they may lead to rushed or unreflective responses motivated by a desire to complete the task quickly.

- Sensitivity to the topic's delicate nature: given the sensitive subject matter, it is very important to use a careful tone and words. Overly direct or intrusive questions (e.g., "I feel poor") should be rephrased in neutral terms (e.g., "I have sufficient financial resources") to remain informative while avoiding discomfort or social desirability bias among participants.

Based on this feedback, the required adjustments are implemented to the wording and organization of the questions before going to the following steps.

4.4. Ethical Approval

To ensure compliance with research ethics standards, we have submitted a formal request to the Ethics Committee of Nova IMS, detailing on the thesis topic and objectives, the characteristics of the data collection process joined with the survey. The committee reviewed the request and granted approval (Appendix B), accompanied by the following suggestions for improvements to align with the highest ethical research standards:

- Include a trigger warning to alert participants about sensitive topics.
- Offer access to mental health resources or counselors for potential distress.
- Detail in the consent form the research nature, sensitive topics, and possible discomfort.
- Explain data usage, storage, and confidentiality measures in the consent process.
- Justify alumni participation if they are included, and address potential recall bias in their consent.
- Specify data security, anonymization processes, and access restrictions.
- Provide participants with a findings summary or resources as reciprocity.

Based on this feedback, necessary revisions are made to the survey. The revisions include adjustments to the consent form –trigger warning, access to mental health councelors, details on the research topic-, adjustments on the participant population –by including only current students to avoid memory bias-, and the necessary explanation on data usage and security in order to guarantee participant protection and ethical integrity.

After sending the updated version to the ethics committee for verification and obtaining their approval, we started the pilot testing distribution with the questions and measurements built (Appendix C).

4.5. Pilot Testing

This phase's objective is to validate the reliability of the Likert matrix questions on the questionnaire that tend to capture the "well-being" variables and its influencing factors variables (Q6).

The trichonomous questions regarding the participants' feedback on the usability and understandability of the survey (Q10, Q11, Q12) are reviewed as well.

For demographics questions (Q1, Q2, Q3, Q4, Q5), mental health resources seeking (Q7), rewarding aspects (Q8), and potential initiatives (Q9), as there is no wrong answer or a specific pattern to follow on the available answers suggested, we did not apply on them any reliability validation method.

The use of a single Likert matrix to capture multiple variables, combined with efforts to minimize the number of questions, necessitate a specific approach to data processing and statistical validation, as noted in the " Scaling and measurement Selection" section, point 4.2.

The pilot test is conducted on a sample of 30 participants, representing 10-15% of the expected final response range (250–300 total responses). Three coefficients are employed to assess the reliability and consistency of the items:

- **Cronbach's Alpha Coefficient**

Cronbach's Alpha is a widely used statistic that measures the **internal consistency** of a set of items meant to capture the same underlying construct. It reflects how closely related a set of items are as a group, and is considered a measure of scale reliability (Cronbach, 1951). The formula for Cronbach's Alpha is the following

$$\alpha = (\frac{k}{k-1})\left(1 - \frac{\sum_{i=0}^{k} \sigma_{yi}^2}{\sigma_x^2}\right)$$

Where:

K = the number of scale items

σ_{yi}^2 = the variance associated with item I

σ_x^2 = the variance associated with the observed total score.

Figure 3: Alpha Cronbach Formula

This coefficient is computed using Python in Jupyter Notebook for the variable "well-being", which was measured through three Likert-scale items. Since Cronbach's Alpha is most appropriate for constructs measured by three or more items, it is not applied to the variables captured by two items. Therefore, we opt for Pearson and spearman correlation coefficients for this case, as it is more suitable (Eisinga et al., 2013).

- **Pearson Correlation Coefficient**

For constructs captured by two items, we apply the Pearson correlation coefficient, which quantifies the strength and direction of the linear relationship between two continuous variables (Cohen, 1988). The formula for Pearson's *r* is :

$$\mathrm{r} = \frac{\sum(x_i - \bar{\mathrm{x}})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

Where:

r = pearson correlation coefficient

x_i = x variable samples y_i= y variable sample

$\bar{\mathrm{x}}$ = mean of values in x variable $\bar{y}$= mean of values in y variable

Figure 4: Pearson Correlation Coefficient Formula

This method is applied to assess the internal coherence of the following two-item constructs: self-concept, social interactions, academic performance, reliance on institutional image**,** and satisfaction with student experience. Pearson's *r* is suitable for two-item reliability estimation and has been recommended over Cronbach's Alpha in such cases (Eisinga et al., 2013).

- **Spearman rank correlation coefficient**

As a non-parametric alternative to Pearson's *r*, the Spearman rank correlation coefficient is computed. It assesses the monotonic relationship between two ordinal variables without assuming a normal distribution (Field, 2018). The coefficient is based on ranked values, and is defined by the following theorem :

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)}$$

Where:

ρ = Spearman's rank correlation coefficient

d_i = difference between the two ranks of each observation

n = number of observations.

Figure 5: Spearman Rank Correlation Formula

This additional analysis helped ensure that internal consistency remained acceptable even if the data does not meet parametric assumptions.

- **The percentage**

To evaluate the clarity and usability of the questionnaire, three additional questions are included at the end of the pilot survey, each with Yes/No/Neutral responses.

The responses are analyzed using simple percentage calculations across the 30 pilot responses to identify potential usability issues and assess participant comfort and engagement. This step helps to ensure that the questionnaire remains accessible, user-friendly, and non-intrusive, minimizing dropout risk due to confusion or emotional discomfort.

4.6. QUESTIONNAIRE DISTRIBUTION

Following the successful completion of the previous steps -question development, scaling development, psychologist expert validation, ethical approval, and the copilot testing-, the distribution of the questionnaire continued. The questionnaire link is shared with Institution XYZ students through the institutional emails, the WhatsApp student groups, and with the QR code shared by teachers who volunteered to help in this research during their classes.

These methods are selected to maximize participation rates while maintaining convenience for respondents, aligning with the goal of minimizing dropouts and ensuring a robust sample for analysis in the most ethical approach.

Notably, the "financial stress" variable encountered a technical issue which lead us to reconsider its usage in order to avoid bias. Given its critical relevance to this study, and the time constraints preventing the addition of a replacement item, an alternative approach is taken, based on the existing responses from the questions Q6_9 and Q8 (see appendix C) to infer this construct, and which is detailed in Empirical methodology chapter, point 5.1.2.

With this solid data collection framework, along with the data of the current population demographics at Institution XYZ for the year 2024/2025 -obtained from their administration-, we now turn to the technical part of this research. The next chapter, *Empirical Methodology*, details the analytical techniques applied to process and interpret the 258 collected responses, and advance the study's objectives.

5. EMPIRICAL METHODOLOGY

Grounded in a quantitative framework, this chapter details the technical processes employed to explore, preprocess, engineer, model, and optimize prediction for "Student Risk Status".

5.1. RAW DATA PREPARATION

As the collected data is not understandable in its raw form through Exploratory Data Analysis (EDA), a primary preparation is required to extract clear and interpretable insights.

5.1.1. INITIAL DATA EXPLORATION

The analysis begins with a first look at the dataset, comprising 258 observations. We examine the dataset's structure, identify data types, missing values, and assess data quality. The presence of missing values in collected data is due to the sensitive nature of mental health research, as questions considered sensitive and/or personal by the psychologist were optional on the questionnaire in order to respect participants comfort.
We implement quality control mechanisms to ensure data integrity:

- Attention Check Filtering:

  One attention check question ("For attention check, please Select Agree") was embedded in the survey. The rows of the respondents that failed this check are dropped to ensure participant engagement and avoid biased answers from blind answering.

- reCAPTCHA Score Validation:

  reCAPTCHA scores are analyzed, and responses below a 0.5 threshold (which might indicate potential bot activity) are excluded to ensure having only human-generated data.

- Missing Values Assessment:

  We check missing values across the dataset, particularly in Likert scale responses for sensitive constructs (e.g., well-being, financial stress). With approximately 5% of responses missing, we apply the mean as they are numerical variables.

- Metadata Cleaning:

  Unnecessary metadata (e.g., response start times, participant IDs) are removed to streamline the dataset and eliminate noise on the future steps.

5.1.2. DATA ENGINEERING AND VARIABLE CONSTRUCTION

- Variables extraction from Likert scale:

  The Likert matrix scale (1 = strongly disagree, 5 = strongly agree) measured seven psychological constructs: student satisfaction, social interactions, self-concept, academic performance, well-being, financial stress, and reliance on institutional image. Each

construct is operationalized through 2-3 questions, and the final composite scores are calculated by averaging responses (Hair, 2010). We create new columns to keep the new average scores with the names that reflect their captured variable, while original question-level columns are dropped to reduce dimensionality, which aligns with psychological measurement standards.

- Categorical Variable Encoding:

Variables collected through multiple choice questions that have more than one answer and that aim for data exploration (such as “potential initiatives” and “rewarding aspects”) are one-hot encoded to facilitate statistical analysis of preferences across risk classes, and to better explore the target variable classes.

- “Financial stress” technical issue assessment:

To complement the single-item "Financial Stress" measure, which encountered a technical issue during data collection, we took the methodology explained on the following table:

Table 2: Illustrative table of "Financial Stress" variable assessment steps

Aspects	Description
Input Variables	- Financial stress: Categorical scale from one to five (five classes). - Potential Initiatives : multiple categorical entries, containing "scholarships and financial aid" among its options.
Financial Aid option Extraction Method	- We create a copied dataset to not affect our main dataset - We apply One vs rest encoding to the variable “potential initiatives”, where “1” means the "scholarships / financial aid" option is selected, and “0” means the option is not selected.
“Financial stress” and “financial aid option” comparison	- We check the five categorical classes of ‘financial stress” and compare them to the encoded variable "Potential Initiatives". - The overall percentage of respondents selecting "scholarships/ financial aid" option is calculated.

	- The distribution of this selection is table crossed with the five "Financial Stress" classes to identify potential patterns. - For each class, the counts and percentages of presence (1) and absence (0) of "scholarships/ financial_aid" option are computed to provide a detailed profile of response trends.
Result analysis and comparison to the literature review	We compare the final distribution to literature: - If the distribution of "1" is more important on the high financially stressed class and less important on the low financially stressed class, we maintain the variable "financial stress". - Otherwise, we drop the variable "financial stress" from our research due to its lack of fiability.

This approach draws on the principle of data triangulation, utilizing "potential initiative" options as an indirect indicator to enrich the understanding of "Financial Stress" (Denzin & Lincoln, 2005). The strategy aligns with the theoretical insights which suggests that higher financial stress levels correlate with increased interest in financial aid or scholarships (Selenko & Batinic, 2011). The analysis is designed to uncover trends that may support the single-item measure's relevance, with detailed findings presented in Chapter 6, point 6.1.1. For future iterations, incorporating multiple items for "Financial Stress" is recommended to enable formal reliability testing (Hair et al., 2010).

- Variable construction: from the collected data, we have built based on theoretical concepts two variables, "Unmeasured factors" and "Student risk status".

  - Unmeasured Factors

  The derivation of "Unmeasured_Factors" the variable is based on the premise that significant discrepancies between a student's well-being score and the scores of the six measured factors indicate the presence of additional explanatory variables not directly measured in this study. Specifically, for each student, the difference between their well-being score and each factor's score is calculated on a 5-point scale (1 to 5, integers). The methodology accounts for the expected relationships between well-being and the factors:

Table 3: Process summary for deriving the *Unmeasured Factors* variable

Aspects	Description
Input Variables	- Well-being: Continuous scale from 1.0 to 5.0 (Likert, with possible decimals). - Influencing Factors (six variables): "Reliance on Institutional Image", "Self-concept", "Satisfaction from the Student Experience", "Academic Performance", "Financial Stress", "Social Interactions" (each measured separately from 1.0 to 5.0).
Detection of significant difference between "well being" and its influencing factors variables	For each student, we compute differences: well_being - factor_score. - For positive factors (self_concept, social_interactions, academic_performance, student_satisfaction) and well_being - inverted_factor_score - for negative factors (reliance_on_institution, Financial Stress, inverted as 6 - score). - A difference is significant if the difference between "well being" and the absolute value surpass the threshold of 0.5.
Detection of "low", "high" and "moderate" unmeasured factors	- "low unmeasured factors": if 1 or less variables are not aligned with "well-being" (≤ 1 significant difference). -"High unmeasured factors": if more than 3 variables are not aligned with "well-being" (: ≥ 4 significant differences). - "moderate unmeasured factors": if 2 or 3 variables are aligned with "well-being" (2 or 3 significant differences).
Detection of "Positive" and "Negative" unmeasured factors	-"Positive influent factors" : More positive significant differences (well_being > factor_score) than negative ones, suggesting unmeasured factors are boosting well-being. - "Negative influent factors": More negative significant differences (well_being < factor_score), suggesting unmeasured factors reducing well-being.

Final Observations	- **Low** unmeasured **positive** influent factors - **Low** unmeasured **negative** influent factors - **Medium** unmeasured **positive** influent factors - **Medium** unmeasured **negative** influent factors - **High** unmeasured **positive** influent factors - **High** unmeasured **negative** influent factors

This directional distinction is supported by recent studies on well-being, which demonstrate that contextual or unobserved variables can either exacerbate or ameliorate mental health outcomes (Keyes et al., 2017).

- Student risk status

The categorization of "Student Risk Status" into "Low Risk", "Moderate Risk", and "High Risk" is achieved through a precise classification delineated in the accompanying table.

Table 4: Summary of thresholds and procedures applied to derive the "student risk status" variable

Well-being	The seek of mental health ressources	Student risk status
[3.5 ; 5[	All options	Low risk
[2.5 ;3.49[	“yes”or “I’ve considered it”	Low risk
	“neutral”	Moderate risk
	“I don’t feel the need for now”, “No, never”	Moderate risk
[1.0 ; 2.49[	“yes”or “I’ve considered it”	Moderate risk
	“Neutral”	High risk
	“I don’t feel the need for now”, “No, never”	High risk

The decision to base "Student Risk Status" solely on "Well-being" and "Seek of Mental Health Resources" is driven by their direct relevance to mental health risk assessment. “Well-being" encapsulates a holistic evaluation of psychological health, consistent with its established use in student mental health research (Lipson et al., 2019), while "Seek of Mental Health Resources" provides a behavioral indicator of risk mitigation, aligning with prevention-oriented paradigms (Eisenberg et al., 2016). This tiered structure is supported by empirical evidence that proactive help-seeking can temper the effects of poor well-being, whereas inaction exacerbates risk (Rickwood et al., 2005).

5.2. EXPLORATORY DATA ANALYSIS

Once the raw data is prepared for exploration, we look for pertinent insights through the following steps:

5.2.1. DESCRIPTIVE STATISTICS

To explore the distribution of the numerical variables, we calculate the mean, the standard deviation, the minimum and maximum values, and the skewness using the numpy library (Harris et al., 2020). This step, inspired by standard EDA practices, is employed to identify potential ceiling effects or asymmetries in well-being measures (Hernández-Torrano et al., 2020; Tukey, 1977) which gives a better idea of the distribution of the predictive variables.

5.2.2. TARGET VARIABLE ANALYSIS

We cross the target variable “student risk status” with demographics and exploratory variables using the matplotlib (Hunter, 2007) and seaborn (Waskom, 2021) libraries to examine their distribution across the three classes “High risk”, “Low risk” and “Moderate risk”. The well-being influencing factors (self-concept, academic performance, among others) are analyzed by risk class to explore their relationships and correlations, a common practice in student mental health studies (Eisenberg et al., 2016).

5.2.2.1. DEMOGRAPHICS COMPARISON ANALYSIS

We compare the demographics of the valid answers from our collected sample with the real student population of Institution XYZ on the school year 2024/2025. The difference is calculated between the percentages of the sample and the true population for each category, and is then visualized in barplot.

5.2.2.2. EXPLORATORY VARIABLES AND TEXT ANALYSIS

The variables “Potential Initiatives” and “Rewarding Aspects” are examined by risk class. For each student risk class, we plot the most selected answers for each variable to obtain a better

understanding of students' satisfaction point by class, as well as the potential initiatives that they consider as effective for an optimized mental health.

Next, we explore the textual data and insights from open-ended questions "Potential Initiatives Text", "Rewarding Aspects Text", and "Additional Feedback", designed to capture students' perspectives on their academic environment, suggestions for improvement, and positive aspects of their experience at institution XYZ. The initial plan consisted in employing Natural Language Processing (NLP) techniques (sentiment analysis) to process the textual data. However, a preliminary assessment of the responses sowed that the three questions have more than 90% missing values, wish is a limited volume of textual data to treat (figure21). This low response rate rendered advanced NLP techniques impractical, as they typically require larger datasets for reliable outcomes. Consequently, a manual qualitative approach is adopted to evaluate the available responses.

The text analysis process comprise three key steps: data cleaning, thematic categorization, and association with the "Student risk status" variable. First, responses are manually reviewed to eliminate irrelevant or non-informative entries, such as vague or off-topic comments, and keep only in-theme responses. Second, the cleaned responses are qualitatively examined to identify recurring themes relevant to the research objectives, such as academic support, campus facilities, or student well-being. Third, each response is linked to the respondent's "Student risk status" to examine potential patterns between students' risk profiles and their qualitative feedback in subsequent analyses.

5.3. MODELS

To focus on the modelling task, we drop all columns from the EDA that are not necessary for the model and might result in noise, and keep only the predictive variables along the target variable renamed in "y_student_risk_status" (in order to differentiate from the independent variables).

5.3.1. MODELS PREPROCESSING

Four models -logistic regression, random forest, support vector machines and XGboost with SMOTE- are selected for the classification task according to their relevance in the literature. To configure them to our dataset, we undertook a proper preprocessing, implemented using scikit-learn and imbalanced-learn libraries (Pedregosa et al., 2012), which consists in the following steps:

5.3.1.1. DATA SPLITTING

We divide the dataset into training (75%) and test (25%) sets using stratified sampling to maintain the correct proportion of risk classes in "Student Risk Status" (Low, Moderate, High Risk), especially for the minority "High Risk" class.

Features (X) include the six primary influencing factors (self-concept, social interactions, academic performance, satisfaction from the student experience, financial stress, reliance on institutional image) and the demographic variables.

5.3.1.2. FEATURE ENCODING AND SCALING

Two preprocessing pipelines are defined using ColumnTransformer:

For Logistic Regression, categorical features are encoded with One Hot Encoder to avoid multicollinearity, as linear models cannot handle perfectly correlated dummy variables (Hastie et al., 2017). Numerical features are standardized using StandardScaler to zero mean and unit variance, which ensures a stable optimization and aligns with one-hot encoded dummies' variance (Field, 2018).

For Tree-Based Models and SVM, the categorical features are one-hot encoded without dropping levels to retain all categories, as tree-based models (Random Forest, XGBoost) and SVM handle collinearity effectively (Breiman, 2001; Cortes & Vapnik, 1995). Numerical features are still scaled to maintain pipeline consistency, though scaling is unnecessary for tree-based models.

The dual preprocessing approach balances the needs of linear and non-linear models while simplifying pipeline management.

5.3.2. MODELS CONFIGURATION

We configure our four models within preprocessing pipelines, after fixing a random seed at 42 for reproducibility.

5.3.2.1. LOGISTIC REGRESSION

The model is configured with a regularization strength of 0.3, which imposes stronger regularization compared to the default 1.0, to prevent overfitting by constraining the magnitude of the coefficients. The elastic net mixing parameter, set to 0.5, balances between L1 (Lasso) and L2 (ridge) penalties, encouraging both sparsity in the coefficients and stability of the model. The 'saga' solver is chosen as it supports elastic net with multinomial logistic regression and handles imbalanced data robustly. The maximum number of iteration is set to 300 to ensure convergence. To address the imbalanced dataset, we apply balanced class weights, which penalizes misclassifications of minority classes more heavily (Hastie et al., 2017).

5.3.2.2. RANDOM FOREST

We implement a Random Forest classifier composed of 600 decision trees. Bootstrap sampling is retained, following the procedure introduced by Breiman (Breiman, 2001), while the option "balanced" for class weights is activated to compensate for class-imbalance. The learning algorithm automatically assigns larger mis-classification costs to the rarer classes, encouraging

each tree to model their patterns without discarding any observations. All categorical predictors are one-hot encoded and left in the design matrix so that individual category levels can serve as split candidates; tree-based methods are able to ignore uninformative dummies and to choose the most discriminative levels when constructing their decision rules.

5.3.2.3. SUPPORT VECTOR MACHINE (SVM)

We employ a SVM with a radial basis function (RBF) kernel and balanced class weights to perform multiclass classification, as the RBF kernel effectively captures non-linear relationships (Cortes & Vapnik, 1995).
The model uses a preprocessing pipeline retaining all dummy variables to allow direct modeling of every categorical level, which is suitable given SVM's robustness to collinearity. Probability estimates are disabled to focus on hard labels.

5.3.2.4. XGBOOST

We utilize an XGBoost Classifier combined with SMOTE oversampling within a pipeline for multiclass classification, as XGBoost's gradient boosting excels at capturing complex patterns in multi-class tasks (Chen & Guestrin, 2016).

The model is configured with the 'multi:softprob' objective, multi-class logarithmic loss evaluation, and a preprocessing step retaining all dummy variables to enable direct splits on every categorical level, which is advantageous for tree-based models. SMOTE generates synthetic samples to address the small "High Risk" class, enhancing pattern learning, while regularization parameters prevent overfitting on synthetic data (Chawla et al., 2002).

5.3.3. HYPERPARAMETER TUNING

Randomised search is performed for 25 iterations, which is sufficient to explore the predefined hyper-parameter spaces while remaining computationally affordable. Model quality is assessed with the macro-averaged F1 score, which gives each class equal influence on the objective function regardless of its frequency in the sample (Sokolova & Lapalme, 2009).

For Logistic Regression, the weighted Random Forest and the Support Vector Machines, we employ five-fold stratified cross-validation so that every fold preserves the original class proportions.

In contrast, the XGBoost model combined with SMOTE is evaluated with three-fold stratified cross-validation, leaving a larger share of minority examples in each training fold improves the neighborhood structure that SMOTE needs to synthesize realistic new instances (Chawla et al., 2002).

The hyper-parameter domains searched are as follows. Logistic Regression is tuned solely on the inverse regularization strength C, examined over the interval 0.1 – 2.1. For Random forest

we explore tree depths between three and eight levels, a minimum leaf size ranging from two to ten observations, and the fraction of predictors considered at each split between 0.30 and 1.00. The SVM search varies C from 0.1 to 2.1 together with the radial-basis kernel scale (γ) from 0.001 to 0.01. Finally, XGBoost is tuned over tree depths of two to five levels, learning-rate values between 0.01 and 0.20, 50 – 200 boosting rounds, L2-regularisation (λ) from 1 to 10, L1-regularisation (α) from 0.1 to 1.0, and minimum child-weight values of 1, 3, and 5.

5.3.4. Evaluation Metrics

We evaluate the models performance using two metrics to assess performance across training, cross-validation, and test sets:

- F1-macro
  The primary metric, averaging F1-scores across classes without weighting by class size, ensures balanced performance for the imbalanced dataset (Powers & Ailab, 2011).
- Accuracy
  A secondary metric provides overall correctness, which in this particular study is less reliable due to class imbalance.
- Detailed performance
  More precise details on performance using classification reports and confusion matrices, are employed to capture precision, recall, and F1-score for the three classes, as good performance for specific classes (high risk and moderate risk) is relevant for targeted interventions (Pedregosa et al., 2012).

5.3.5. Feature Importance

To understand which factors significantly influence each class of the target value, we drive a feature importance analysis across the four predictive models. Each model employs its own adapted method for quantifying feature contributions, tailored to its specific algorithmic structure (Pedregosa et al., 2012).

- Logistic Regression
  We start by extracting the raw coefficients and intercepts from the model to obtain the raw data needed to compute odds ratios and feature importance while understanding the baseline probabilities for each class.
  These **coefficients** reveal how each feature influences the predicted probability for each risk category in our multi-class classification setting (Hastie et al., 2017). To simplify the interpretability, the coefficients are transformed into odds ratios using the exponential function, generating a table of odds ratios for each feature across the three classes (High Risk, Low Risk, Moderate Risk). We plot the results in a bar chart that illustrate the odds of features for each risk class.

- Random Forest
  The Random Forest model calculates feature importance using the **mean decrease in impurity (MDI)** method, which measures the average reduction in Gini impurity across all decision trees, a standard approach for tree-based models (Breiman, 2001). We focus on the top seven features based on their importance scores, with feature names obtained from our preprocessing pipeline. To ensure clarity in visualization, we wrap longer feature names to a maximum of 25 characters per line. The results are presented in a horizontal bar plot, and are ordered by the most to the least important.

- Support Vector Machine (SVM)
  Given the complexity of interpreting SVM models directly, we employ **permutation importance,** a model-agnostic approach that evaluates feature contributions by shuffling each feature's values and measuring the resulting decrease in F1-macro score (Altmann et al., 2010). To obtain robust results, we repeat this process ten times and utilize parallel processing for efficiency. We compile the mean and standard deviation of importance scores across repetitions into a DataFrame, using feature names from our preprocessed dataset. The visualization includes horizontal bar plots with error bars representing standard deviations, where the critical features to risk prediction are clearly readable.

- XGBoost
  For the XGBoost model, we determine feature importance using **the gain** metric, which quantifies the average improvement in model performance (reduction in multi-class log-loss) attributed to each feature across all splits in the boosted trees (Chen & Guestrin, 2016). The results are organized in descending order and are visualized through a plot showing the top features ranked by their gain values.

6. RESULTS, CONCLUSIONS AND FUTURE RESEARCH

6.1. RESULTS

This section will detail the pertinent results obtained from pilot testing, EDA, and modelling.

6.1.1. PILOT TESTING RESULTS

The pilot testing driven on 30 participants show the following results:

Participant feedback reinforced the questionnaire's clarity, ease of use, and sensitivity, with 100% affirming comprehension, 96% confirming usability, and 100% reporting the absence of discomfort.

These findings confirm that the questionnaire achieved its design goals of being clear, accessible, and ethically sensitive while minimizing respondent burden.

- **Items fiability**

The following table summarizes the items reliability scores for each variables, using Cronbach, Pearson and spearmans coefficients, and their respective interpretation:

Table 5: Reliability scores of items in pilot tests

Variable	Pearson	spearman	Cronbach	Score Interpretation
Well-being	/	/	0.75	Fair to good reliability.
Self-concept	0.49	0.50	/	Moderate correlation.
Social interactions	0.55	0.53	/	Moderate to strong correlation.
Academic performance	0.51	0.54	/	Moderate to strong correlation.
Reliance on institutional image	0.24	0.21	/	Low correlation.
Satisfaction rom the student experience	0.59	0.60	/	Moderate to strong correlation.

- Cronbach coefficient: according to established interpretation guidelines, the obtained score of 0.751 represents an adequate internal consistency stage of the study, as values higher than 0.7 are considered reliable, particularly in exploratory research (Hair, 2010).
- Pearson correlation: According to the score interpretation table of Pearson, for values between 0 and 1, the obtained scores are acceptable while the value of 0.24 for the variable "Reliance on institutional image" is low.
- Spearman rank correlation: according to the spearman rank correlation interpretation, the obtained scores are acceptable except for reliance on institutional image, which showed a low value (0.21).

Once the full sample collected in the end of data collection, we calculated again the Pearson coefficient for the variable with the low score "Reliance on Institutional Image", and obtained a higher score (0.43).

With the item review from the expert reinforcing the questions item, the relevance of this variable from literature review, and time constraint; we decided to keep the variable "Reliance on Institutional Image" and did not apply any further adjustments.

- **"Financial stress" variable assessment**

The results obtained from the triangulation approach, undertaken to consider keeping the single item variable "financial stress", indicate a clear gradient in the selection of "scholarships_financial_aid" across the "Financial Stress" spectrum.

Table 6: Breakdown of "Financial aids" options across levels of "Financial Stress"

Reported Financial stress level, from "1= very low" to "5=very high"	Number of students by financial stress level	Number of times "scholarships/financial aid" is selected	Number of times "scholarships/financial aid" is not selected	Percentage (%) of selected option "scholarships/financial aid" by financial stress level	Percentage (%) of selected option "scholarships/financial aid" by financial stress level
1	21	**2**	19	**9.52**	90.48
2	55	**17**	38	**30.91**	69.09
3	42	**18**	24	**42.86**	57.14
4	79	**46**	33	**58.23**	41.77
5	42	**25**	17	**59.52**	40.48

The observed trend supports the theoretical expectation that individuals experiencing higher financial strain are more likely to seek financial assistance, as posited by Selenko and Batinic (2011), who highlighted a link between perceived financial strain and adaptive behaviors such as seeking aid. The gradual increase in the percentage of presence from 9.52% at the lowest stress level to 59.52% at the highest provides preliminary evidence that the single-item measure may capture a meaningful aspect of financial stress, despite the absence of a multi-item reliability check (Nunnally & Bernstein, 1994).

Due to the theoretical importance of this variable in this study, we keep 'financial stress", as it aligns with the exploratory triangulation strategy outlined in Empirical methodology.

However, this correlation may be subject to bias due to the single-item measure's potential lack of sensitivity to nuanced stress levels, which could influence the respondents' selection of "scholarships/financial aid" option (Hair, 2010).

6.1.2. Exploratory Data Analysis

This section highlights the most pertinent findings from the EDA. We start with sample characteristics, where we visualize and compare the demographics from the collected data to the real current student population of Institution XYZ. Then, we explore and compare the demographics and characteristics of the three classes of our target variable. Finally, we review the pertinent open-ended questions.

6.1.2.1. Sample Characteristics

The final dataset comprised 239 valid responses. After applying attention check and reCAPTCHA filtering (originally 258 responses, with 19 dropped for quality concerns).

By comparing the demographics of the sample with the real population, we obtained the following results:

Gender Distribution: Gender distribution is relatively balanced across the sample (Figure 14):

- Male: 52.30%
- Female: 46.90%
- Non-binary/third gender: 0.40%
- Prefer not to say: 0.40%

When compared to the overall XYZ institution population, the sample gender distribution closely aligns with the institutional average, where the gender splits into 54.46% for males and 45.54% for females (Table 13). This near equal gender representation enhances the generalizability of findings and allow meaningful gender-based comparisons.

Age

The sample predominantly consists of young adults, with the majority (59.80%, n=143) falling within the 18-24 age range, followed by 34.30% (n=82) in the 25-34 range. Older age groups are poorly represented (figure 15).

This age distribution aligns with typical higher education demographics and reflects the target population for mental health interventions in academic settings. The sample's concentration in the 18-34 age range (94.1%) aligns well with the population's average ages per cycle (bachelor: 19.87, master's: 27.77), confirming representativeness for the target demographic.

Residency Status

The sample shows a balanced representation between Portuguese and international students, with 61.1% being Portuguese students and 38.9% being international students. This distribution reflects the international character of Institution XYZ programs (Figure 16).

Study area (academic program)

Analysis of study areas revealed the following distribution:

- Study area C : 42.29%
- Study area D : 20.16%
- Study area B : 17.39%
- Study area A : 16.60%
- Study area E : 1.58%
- Study area F: 1.98%

When compared to the real Institution XYZ population, the sample shows notable representation differences: Study area C is overrepresented (42.29% vs. 30.03% in population), while the study area A is significantly underrepresented (16.60% vs. 36.30% in population). Study area B shows higher representation (17.39.% vs. 7.39% in population). Additionally, some students took more than one program, with the most common combination being Study areas A and B. This overlap reflects the interest of students in complementing these two programs in bachelor and master, or in master and PhD (figure 17).

Academic Level

The sample is heavily weighted toward graduate-level students:

- Master's degree students: 55.64%
- Bachelor's degree students: 32.30%
- Post-graduation program: 8.17%
- PhD students: 3.89%

Compared to the real population (table 14), the sample shows overrepresentation in bachelor's programs (32.30% vs. 18.86% in population) and underrepresentation in post-graduation programs (8.17% vs. 18.47% in population). Combinations in degrees are observed as well, with 15 students pursued more than one program at Institution XYZ, primarily Master's with other degrees (seven Master's-PhD combinations, six Bachelor's-Master's combinations).

6.1.2.2. TARGET CLASS ANALYSIS

- **Student risk status distribution**

The mental health risk classification results in the following distribution:

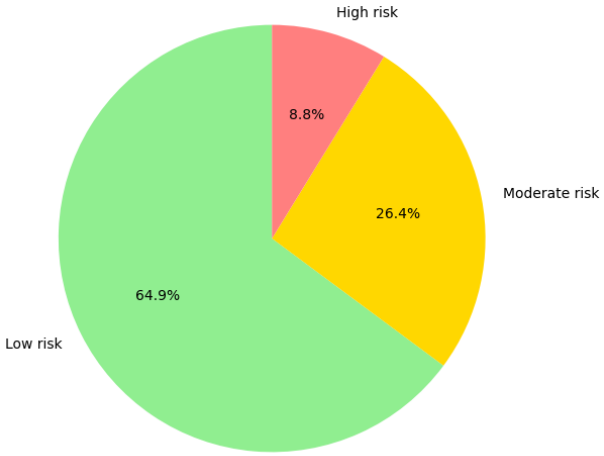


Figure 6: Distribution of students by risk status (Pie Chart)

The well-being risk classification estimates 8.8% of students at high risk of struggling with mental health issues, 26.4% at moderate risk, and 64.9% at low risk. Compared to existing literature, the proportion of students at moderate risk aligns with global estimates, which typically range between 20–35% (Auerbach et al., 2016; Ibrahim et al., 2013). However, the high-risk group appears underrepresented, as studies often report 15–30% of students experiencing severe mental health concerns (Beiter et al., 2015).

- **Patterns and insights by risk status**

Age

The results reveal distinct risk patterns across age groups:

Table 7: Percentage Distribution of Student Risk Classes Across Age Categories

Age intervals	High risk (%)	Low risk (%)	Moderate risk (%)
18-24	9.79	65.03	25.17
25-34	7.32	64.63	28.05

35-44	20.00	20.00	60.00

Older groups (45-54 and 55-64) are predominantly in Low Risk (85.71% and 100.00%, respectively), suggesting greater stability, though sample sizes are limited.

Further analysis within the "High Risk" category indicate that the 18-24 age group is the most frequent, accounting for 66.67% of High Risk students (14 students aged from 18 to 24 at high risk, out of a total of 21 students at high risk), with 25-34 at 28.57% and 35-44 at 4.76%.

These findings partially align with the literature review, which suggested that younger individuals, particularly those aged 18-24, are generally more affected by risk factors due to transitional life stages and academic pressures (Selenko & Batinic, 2011).

Gender

Table 8: Percentage Distribution of Student Risk Classes Across Gender Categories

Gender	High risk (%)	Low risk (%)	Moderate risk (%)
Male	12.00	62.40	25.60
Female	5.36	66.96	27.68
Non-binary / Third Gender	00.00	100.00	00.00

Contrary to a potential hypothesis that females might be more affected the data shows males have a higher perentage in High Risk (15 males at high risk out of 125, vs. six females at high risk out of 112). This suggests males may be more vulnerable to high-risk status.

Study Area (academic program)

The analysis of study areas reveals that the combination of area A and C (20.00% High Risk) and the study area B (19.51% High Risk) are the most affected (**figure 18**) suggesting heightened risk from dual or specialized demands. Conversely, the combination of the study area B and C (0.00% High Risk) emerge as the most tranquil, indicating stability for these students.

Other combinations, such as study area C and E show 0.00% High Risk but are entirely in Moderate Risk, limiting their classification as tranquil due to small samples. The higher risk in combined programs supports the literature's implication that academic specialization can heighten stress (Selenko & Batinic, 2011).

Residency Status

Despite prior research suggesting that international students are generally more vulnerable to adaptation challenges and mental health risks (Smith & Khawaja, 2011), the current data reveals no significant difference in risk levels between international and Portuguese students.

Table 9: Percentage Distribution of Student Risk Classes Across Portuguese and International students

Residency status	High risk (%)	Low risk (%)	Moderate risk (%)
International	8.80	65.90	25.30
Portuguese	8.80	64.20	27.00

Degree Level

The results reveal distinct patterns across risk categories.

Table 10: Percentage distribution of Student Risk Classes Across Degree Levels

Degree level	High risk (%)	Low risk (%)	Moderate risk (%)
Bachelor	12.05	61.45	26.51
Master	6.29	70.63	23.08
PhD	10.00	70.00	20.00
Post-graduation	14.29	42.85	42.86

Contrary to expectations from the literature review, which anticipated a higher representation of Master's and PhD students in the High Risk category, the data indicate that these groups are predominantly in Low Risk, with Post-graduation Program students showing a High Risk presence surpassing 10% and all degree levels high risk population.

Potential initiatives

The results in figure 19 indicate that Moderate Risk students show the highest preference for "Enhanced Career Support" (58.7%) and "Scholarships/Financial Aid" (52.4%), suggesting a focus on career and financial stability. High Risk students exhibit the highest selection for "Mental Health Support" (0.381), reflecting potential prioritization of well-

being under stress. Low Risk students show moderate interest across initiatives, with "Larger Infrastructures" (43.2%) and "Wellness Programs" (34.8%) standing out.

Rewarding aspects

According to figure 20, Low risk class report the highest valuation of "Personal Growth" (68.4%) and "Career Development" (61.9%), indicating a strong focus on personal and professional gains. Moderate risk class prioritize "University Reputation" (65.1%), suggesting an emphasis on institutional prestige. High risk class show a more balanced distribution, with "University Reputation" (38.1%) and "Personal Growth" (33.3%) as key aspects, alongside "Career Development" (33.3%). These findings align with adaptive behavior theories under varying stress levels (Selenko & Batinic, 2011).

Unmeasured Factors

The unmeasured factors distribution across the three risk classes show a considerable presence of high negative unmeasured factors which are not captured by this study (Figure 21). According to literature review, this bias is a persistent and common issue in psychology research and SMH research.

6.1.2.3. TEXT ANALYSIS

The text analysis examined responses to three open-ended survey questions ("Additional Feedback," "Potential Initiatives Text," and "Rewarding Aspects Text") to capture qualitative insights into student experiences at Institution XYZ. After manual cleaning to remove irrelevant entries, seven responses are analyzed and categorized by "Student at Risk Status" (Low risk, Moderate risk, High risk). Low risk students dominate, contributing 71.43% (5/7) of responses, while Moderate and High risk students each provided one response (14.29% each). This distribution suggests greater engagement from Low risk students, though the small sample size limits broader conclusions.

Low risk responses primarily expressed appreciation (e.g., a participant showed appreciation for The Bridges program for mental health support, and considers that as an amazing experience changing) and proposed initiatives (e.g., a participant suggested on-campus gym), with some complaints about inadequate facilities or faculty engagement (e.g., a participant complained about not having enough space in study rooms”). Moderate risk students balance appreciation (e.g., a participant expressed his appreciation for project-based classes) with proposals for amelioration (e.g., for more teaching assistance support). The high risk response focused solely on complaints, highlighting the lack of international student representation (e.g., a respondent complained that international students often feels left out of decision making”). These patterns align with research suggesting that student well-being influences feedback tone (Tinto, 2017).

The findings indicate diverse needs across risk profiles, with Low risk students valuing existing resources but seeking improvements, and high risk students emphasizing systemic issues. However, the low response rate precludes generalizability and advanced NLP techniques, as noted by (Braun & Clarke, 2006). Future studies with larger samples could validate these insights.

6.1.3. MODELS RESULTS

6.1.3.1. MODELS PERFORMANCE

All models demonstrate moderate performance in predicting student risk categories, with F1-macro scores on the test set ranging from 0.48 to 0.55, as presented on the following table:

Model	Macro F1 (Train)	Macro F1 (CV)	Macro F1 (Test)	Accuracy (Test)
Logistic regression	0.56	0.44± 0.10	**0.55**	0.63
Random forest	0.72	0.48	**0.55**	0.60
SVM	0.56	0.48	0.49	0.63
XGBoost	0.78	0.49± 0.05	0.48	0.58

Table 11: Model performance scores (macro F1 and accuracy)

- **Logistic Regression** emerges as the most consistent performer, achieving the highest test F1-macro score of 0.55 and the best test accuracy of 0.63. Importantly, this model shows good generalization capability, with minimal performance degradation between training (0.56) and testing (0.55), suggesting robust learning without significant overfitting.

Dataset	F1-macro	Accuracy
Train (174 rows)	0.56	0.64
CV validation (5-fold)	0.44 ±0.10	0.57 ±0.08
Hold-out test set (59 rows)	0.55	0.63

Detailed test-set report

	precision	recall	f1-score	support
High risk	0.30	0.60	0.40	5
Low risk	0.84	0.67	0.74	39
Moderate risk	0.47	0.56	0.51	16
accuracy			0.63	60
macro avg	0.54	0.61	0.55	60
weighted avg	0.70	0.63	0.65	60

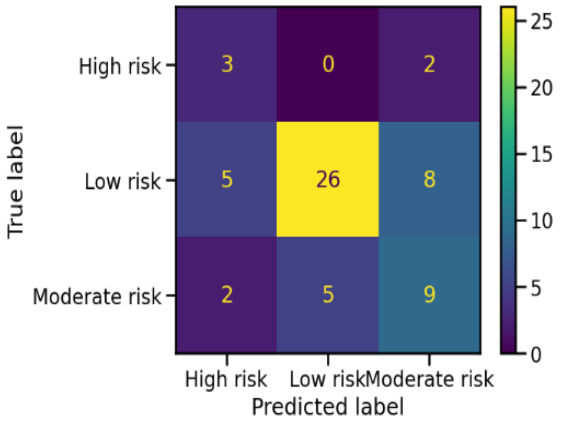


Figure 7: Detailed Test-set Performance and Confusion Matrix for Logistic Regression

- **Random Forest** displays concerning signs of overfitting, with a substantial drop from training F1-macro (0.72) to test performance (0.55). This 17-point decrease suggests the model may be learning noise in the training data rather than generalizable patterns. The cross-validation score of 0.48 provides a more realistic estimate of its true performance.

Dataset	F1-macro	Accuracy
Random Forest (train)	0.72	0.77
Random Forest (CV mean)	0.48	0.59
Random Forest (test)	0.55	0.6

Detailed test-set report

	precision	recall	f1-score	support
High risk	0.43	0.60	0.50	5
Low risk	0.78	0.64	0.70	39
Moderate risk	0.38	0.50	0.43	16
accuracy			0.60	60
macro avg	0.53	0.58	0.55	60
weighted avg	0.65	0.60	0.61	60

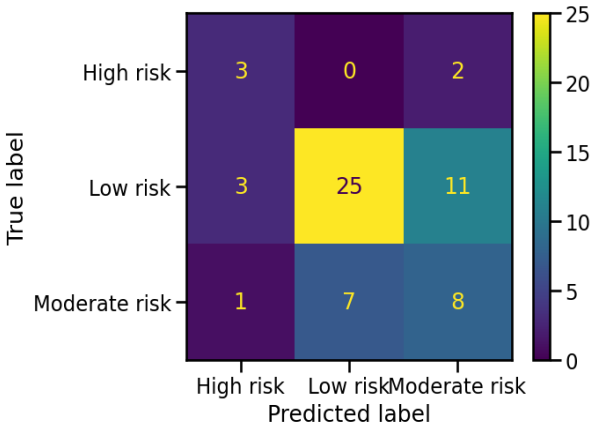


Figure 8: Detailed Test-set Performance and Confusion Matrix for Random Forest

- **Support Vector Machine** shows stable but modest performance across all metrics. With consistent F1-macro scores around 0.48-0.49 and competitive test accuracy (0.63), the SVM demonstrates reliable generalization but limited predictive power for this particular task.

SVM Results:

Dataset	F1-macro	Accuracy
SVM (train)	0.56	0.66
SVM (CV mean)	0.48	0.59
SVM (test)	0.49	0.63

SVM Detailed test-set report

	precision	recall	f1-score	support
High risk	0.20	0.20	0.20	5
Low risk	0.84	0.69	0.76	39
Moderate risk	0.43	0.62	0.51	16
accuracy			0.63	60
macro avg	0.49	0.51	0.49	60
weighted avg	0.68	0.63	0.65	60

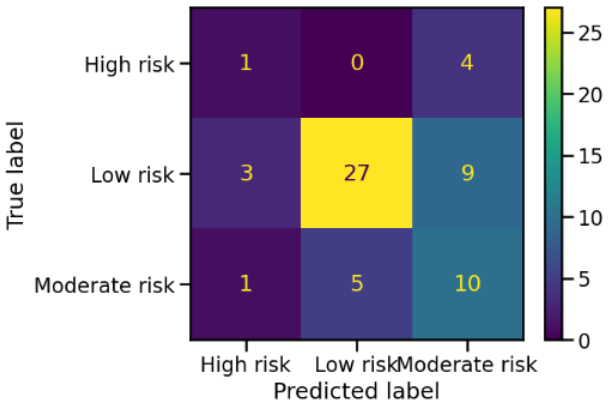


Figure 9: Detailed Test-set Performance and Confusion Matrix for Support Vector Machines

- **XGBoost with Smote** presents a classic case of overfitting despite regularization techniques. The dramatic drop from training F1-macro (0.78) to test performance (0.48) indicates the model is struggling to generalize to unseen data, even with the SMOTE technique applied to address class imbalance.

XGBoost + SMOTE Results

Dataset	F1-macro	Accuracy
XGBoost + SMOTE (train)	0.78	0.81
XGBoost + SMOTE (CV mean)	0.49 ±0.05	0.59 ±0.04
XGBoost + SMOTE (test)	0.48	0.58

Detailed test-set report

	precision	recall	f1-score	support
High risk	0.29	0.40	0.33	5
Low risk	0.76	0.67	0.71	39
Moderate risk	0.37	0.44	0.40	16
accuracy			0.58	60
macro avg	0.47	0.50	0.48	60
weighted avg	0.62	0.58	0.60	60

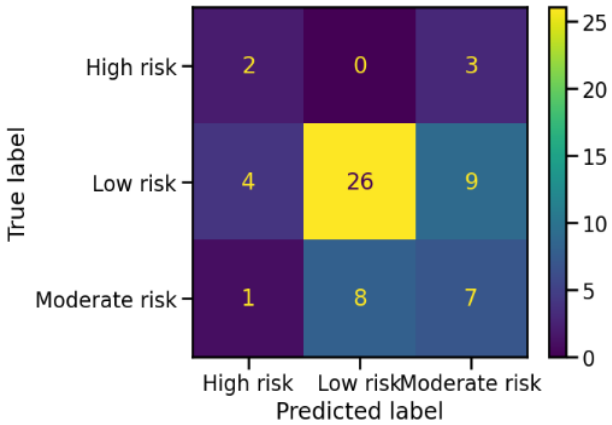


Figure 10: Detailed Test-set Report and Confusion Matrix for XGBoost

6.1.3.2. CLASS SPECIFIC PERFORMANCE

The test-set classification reports provide detailed performance for each class (High risk, Low risk, Moderate risk), with support sizes of 5, 39, and 16, respectively.

- **"High Risk" class (Support: 5)**

- Logistic Regression: Precision 0.30, Recall 0.60, F1-score 0.40. Moderate recall indicates some ability to identify High risk cases, but low precision suggests many false positives, likely due to the small sample size.
- Random Forest: Precision 0.43, Recall 0.60, F1-score 0.50. Similar recall to Logistic Regression but with better precision, which shows a better handling of the minority class.

- SVM: Precision 0.20, Recall 0.20, F1-score 0.20. Poor performance, with low recall and precision, indicating difficulty detecting High risk cases.
- XGBoost: Precision 0.29, Recall 0.40, F1-score 0.33. Improved recall compared to SVM, thanks to SMOTE, but precision remains low, suggesting synthetic data helps but doesn't fully resolve the imbalance.

Analysis: High risk classification is challenging across all models due to the small support (5 instances), leading to low precision and variable recall. RF and Logistic Regression achieve the highest recall (0.60), critical for early-warning systems, while SMOTE in XGBoost improves recall over SVM.

- **"Low Risk" class (Support: 39)**

- Logistic Regression: Precision 0.84, Recall 0.67, F1-score 0.74. High precision and good recall reflect strong performance, leveraging the larger sample size.
- Random Forest: Precision 0.78, Recall 0.64, F1-score 0.70. Slightly lower recall than Logistic Regression but still robust, benefiting from class balancing.
- SVM: Precision 0.84, Recall 0.69, F1-score 0.76. Best F1-score, with high precision and recall, indicating excellent Low risk detection.
- XGBoost: Precision 0.76, Recall 0.67, F1-score 0.71. Good performance, though slightly lower precision than Logistic Regression and SVM.

Analysis: Low risk classification is consistently strong (F1-scores: 0.70–0.76) due to the larger support, with SVM performing best. High precision across models suggests reliable identification of low-risk students.

- **"Moderate Risk" class (Support: 16):**

- Logistic Regression: Precision 0.47, Recall 0.56, F1-score 0.51. Moderate performance, with balanced precision and recall, but limited by sample size.
- Random Forest: Precision 0.38, Recall 0.50, F1-score 0.43. Weaker performance compared to other classes, with lower precision and recall, indicating difficulty distinguishing Moderate risk.
- SVM: Precision 0.43, Recall 0.62, F1-score 0.51. Highest recall, suggesting better detection of Moderate risk cases, with F1-score matching Logistic Regression.
- XGBoost: Precision 0.37, Recall 0.44, F1-score 0.40. Lower F1-score, with weak precision, possibly due to SMOTE's impact on class boundaries.

Analysis: Moderate risk performance is moderate (F1-scores: 0.40 - 0.51), with SVM and Logistic Regression performing best. The smaller support (16) compared to Low risk limits precision, and class overlap may confuse models like RF and XGBoost.

6.1.3.3. FEATURE IMPORTANCE

The feature importance results from the four models highlight the factors most predictive of student mental health risk status (High risk, Low risk, Moderate risk). Below, we summarize the top features for each model and identify common patterns:

- **Logistic Regression (Coefficients, Class-Specific):**

The raw coefficients (log-odds) and intercepts were extracted for the three classes: High Risk (Class 0), Low Risk (Class 1), and Moderate Risk (Class 2).
Many coefficients are zero (e.g., age groups), which is expected due to the ElasticNet strong regularization applying an L1 penalty to eliminate less relevant features. Non-zero coefficients like "Gender_Male" with a value of 0.87 for High Risk, for example, indicate a strong positive effect on the log-odds of being at high risk.

The intercepts are -0.36 for high risk, 0.22 for low risk, and 0.14 for moderate risk, which translates into baseline odds of 0.70, 1.24 and 1.15 respectively. These scores mean that in the absence of feature influences, low risk is the most likely outcome, followed by moderate risk, and high risk being the least likely.

For an enhanced interpretability, we calculate the odds ratios of the coefficients for the risk classes. Key results include:

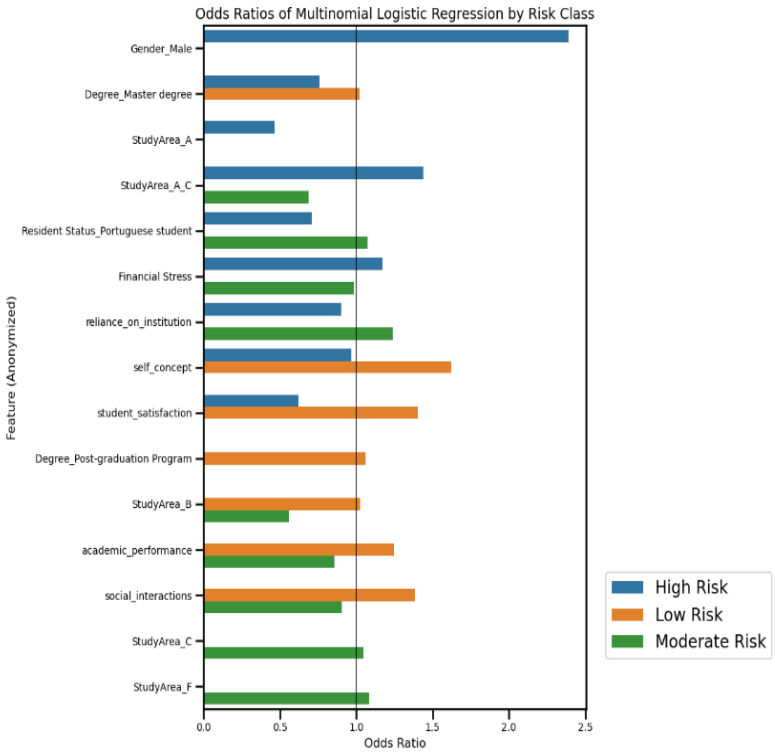


Figure 11: Feature Importance Coefficients Barplot from Logistic Regression

- High Risk:
  Gender_Male: Odds ratio = 2.39 means that being male increases the odds of being at high risk by 139%, compared to females.
  Study Area_A: Odds ratio = 0.46, means that belonging to this program reduces the odds of being at high risk by 54%.
  student_satisfaction: Odds ratio = 0.62, means that a satisfied student is less likely to be at risk by 38% than a student who does not feel satisfied from his student journey.
  Resident Status_Portuguese student: Odds ratio = 0.71, means that being Portuguese reduces the odds of being at high risk by 29% compared to international students.

- Low Risk:
  Study Area_B: Odds ratio = 0.56, meaning that belonging to the study area B reduces the odds of being at low risk by 44%.
  reliance_on_institution: Odds ratio = 1.24, increasing the odds of being at low risk by 24%, suggesting that the feeling of belonging to Institution XYZ lowers the risk.
  self_concept: Odds ratio = 1.619, meaning that self-concept increases the odds of being at low risk by 62%.

- Moderate Risk:
  student_satisfaction: Odds ratio = 1.40, increasing the odds of being at moderate risk by 40%, meaning that even satisfied students can still experience moderate risk.
  social_interactions: Odds ratio = 1.38, meaning that social interactions increase the odds of being at moderate risk by 38%.

- **Random Forest (Mean Decrease Impurity):**

Top features: student_satisfaction (0.27), self_concept (0.18), social_interactions (0.11), academic_performance (0.08), reliance_on_institution (0.07), gender_Female (0.04), Financial_stress (0.04).

Key Insight: Student satisfaction, self-concept, and social interactions are the dominant predictors of risk levels, followed by academic performance and reliance on institutional support. Gender and specific study areas play secondary roles, with financial stress also contributing modestly (Figure 25).

- **SVM (Permutation Importance):**

Top features: student_satisfaction (0.10), self_concept (0.07), academic_performance (0.07), reliance_on_institution (0.07), social_interactions (0.06), Resident Status (0.04), Gender_Male (0.03), Financial Stress (0.03).

Key Insight : similar to RF, satisfaction from the student experience, self-concept, and academic performance are the most critical features for SVM model. Financial stress is less prominent on this models best features (Figure 26).

- **XGBoost (Feature Importance by Gain):**

Top features: self_concept (0.19), student_satisfaction (0.16), academic_performance (0.10), Study Area_C (0.09), social_interactions (0.08), Age_25-34 (0.07), reliance_on_institution (0.07), Gender_Male (0.06).

Key Insight : Self-concept and satisfaction from the student experience are the most important for this model, followed by Academic performance and the study area (the C one, more precisely). Social interactions, Age and Gender follow, with less impact.

6.2. DISCUSSION

6.2.1. MODEL SELECTION RATIONALE

The selection of an appropriate model for predicting student mental health risk status balances performance, interpretability, and the critical need to identify high risk students for early intervention.

We keep Logistic Regression as a transparent backup for interpreting risk factors, given its clear coefficients and competitive overall performance. The coefficients provided the most impacting factors to our sample, facilitating explanations to administrators and counselors. However, is should not be the main model for risk detection as its test-set performance indicates limitations in consistently identifying High risk students, potentially missing critical cases.

Random Forest model is preferred for early warning systems for its “High risk” recall, reasonable macro-F1, and balanced performance across classes. But its feature importance is less interpretable than Logistic Regression.

Despite its precision for Low risk and Moderate risk, SVM’s poor High-risk detection limits its suitability for early intervention, where missing at-risk students is a significant concern.

XGBoost, despite SMOTE’s benefits, is less optimal due to lower macro-F1 and Moderate risk performance, while SVM’s low High risk recall makes it unsuitable for primary screening.

Thus, we opt for a hybrid approach: Random forest for students at risk detection, and Logistic regression for explainability and interpretation.

6.2.2. PREDICTOR VARIABLES ALIGNMENT WITH LITERATURE REVIEW:

- Core Features: student satisfaction, self-concept, social interactions and academic performance are consistently among the top features across all models, indicating they are

the primary drivers of risk prediction. It aligns with literature review that consider these factors as protective factors (Hefner & Eisenberg, 2009; Marsh & Martin, 2011) which in this study are reducing the likelihood of High or Moderate risk.

- Secondary Features: Gender and degree program are relevant but rank lower in most models, except for gender in Logistic Regression. As the literature revealed that males tend to be more resistant to stress and depression, others have showed that this gender is less likely seeking mental health resources and help, which leads to a higher risk (Auerbach et al., 2016).

- Supporting Features: Financial Stress and reliance on institution appear in multiple models but with varying importance, suggesting context-specific influence. Financial stress role as a risk factor for High risk supports studies linking financial strain to anxiety and depression (Andrews & Wilding, 2004). While reliance on Institution's low link to low risk aligns with the literature as over-reliance may indicate underlying stress or lack of personal coping mechanisms (Conley et al., 2014).

- Non-Dominant Features: Age is less predictive in our research as our sample includes predominantly a young age categories, which is balanced with the Institution XYZ population, but does not add relevancy to the model and thus can not be compared to the literature findings regarding age.

6.2.3. RESPONSE TO HYPOTHESIS

The empirical results confirm and support the hypothesis (H1), which suggested that at least one of the factors from "Self-Concept", "Social Interactions", "Reliance on Institutional Image", "Academic Performance", "Financial Stress", and "Satisfaction from the Student Experience" influence the mental health and well-being of Institution XYZ students.

"Self-Concept", "Social Interactions", "Satisfaction from the Student Experience" and "Academic Performance" emerged as top predictors across all models, with high feature importance and strong associations with Low Risk, confirming their significant protective influence (Hefner & Eisenberg, 2009; Marsh & Martin, 2011).

"Financial Stress" showed moderate impact, particularly for High Risk in Logistic Regression, aligning with its role as a risk factor (Andrews & Wilding, 2004) though its lower prominence in XGBoost suggests contextual variability.

"Reliance on Institutional Image" had a limited effect, linked to Moderate Risk, indicating over-reliance may reflect stress rather than protection (Conley et al., 2014).

6.3. CONCLUSION

From the results and discussion, we highlight the following conclusions:

The factors "Student Satisfaction", "Self-Concept", “Social interactions” and "Academic Performance" are the most influential measured factors to our sample, consistently identified across Logistic Regression, Random Forest, SVM, and XGBoost models, supporting their protective roles against mental health risks (Lipson et al., 2019; Pascoe et al., 2020).

Male students, and students completing the two study areas A and C are more likely to show higher vulnerability.

High Risk students prioritize "Mental Health Support", Moderate Risk students value "Enhanced Career Support" and "Scholarships/Financial Aid", and Low Risk students prefer "Personal Growth" and "Career Development".

The Random Forest model, with a High Risk recall of 0.60 and macro-F1 of 0.55 on the test, effectively predicts the student at risk category for early intervention, complementing Logistic Regression’s interpretability. These findings, aligned with the exploratory framework (Denzin & Lincoln, 2005), inform targeted interventions, though the single-item "Well-being" measure and small High Risk sample suggest caution and future multi-item validation (Nunnally & Bernstein, 1994). Thus, a hybrid model approach consisting in Balanced Random Forest for target intervention, and Logistic Regression for interpretability is suitable for our risk detection task.

6.4. LIMITATIONS

Despite the pertinent findings of this study, we encountered several limitations.

First, the sensitive nature of the topic -mental health- posed challenges in collecting a sufficiently large dataset, as it proved difficult to encourage students to participate despite assurances of anonymity. This may have limited both the sample size and its representativeness.

Second, the questionnaire was constrained to an average of 15 questions in order to minimize participant dropout. Consequently, it was not possible to meticulously capture all relevant dimensions and influencing factors. In psychological research, fully capturing the complexity of mental health dynamics is inherently challenging, and future work will need to address potential biases arising from unmeasured factors (Keyes & Martin, 2017).

Finally, the imbalance in the target variable -an issue commonly encountered in psychological research involving risk and minority classes- made accurate prediction particularly difficult given the small dataset.

6.5. FUTURE RESEARCH DIRECTIONS

Based on the sample characteristics and findings, several avenues for future research emerge to address current limitations and expand understanding of student mental health dynamics.

6.5.1. METHODOLOGICAL ENHANCEMENTS

Expanded Measurement Instruments: Future research should address the current limitation of two to three questions per construct by expanding questionnaires to include at least four to five questions per factor. This enhancement would allow to capture a larger dimension of mental health and improve measurement reliability and validity.

Additional Psychological Constructs: Incorporating supplementary factors identified in student mental health literature would provide deeper insights. Notably, impostor syndrome is highlighted in several studies as a significant contributor to student mental health challenges and should be systematically included in future investigations.

Thresholds precision: the variables built on measured factors might benefit better from statistically defined thresholds, as methods like k-means might easily correct the imbalance data problem.

6.5.2. INSTITUTIONAL AND COMPARATIVE STUDIES

Cross-School Analysis: Extending the methodology to the other schools of the University and comparing findings across all nine schools would offer broader understanding of how different academic environments and study areas influence mental health dynamics. This comparative approach could reveal field-specific risk factors and protective elements.

6.5.3. ADVANCED ANALYTICAL APPROACHES

Longitudinal Design: Implementing longitudinal data collection would enable analysis of long-term trends and the dynamic nature of mental health throughout students' academic trajectories.

Advanced Data Science Techniques: Future studies with larger datasets could benefit from applying sophisticated analytical methods with larger datasets, like deep learning algorithms and natural language processing to analyze student feedback and open-ended responses, potentially uncovering latent patterns not captured by the approaches used.

6.5.4. ALTERNATIVE DATA INTEGRATION

Multi-Modal Data Sources: Integrating alternative data sources such as physiological measures (like sleep patterns) or anonymized social media activity patterns could provide objective indicators complementing self-reported measures and enhance the identification of latent mental health factors.

Digital Behavioral Indicators: Incorporating digital footprints from learning management systems, library usage patterns, or campus facility utilization could offer unobtrusive indicators of student engagement and well-being trajectories.

6.5.5. INTERVENTION DEVELOPMENT

Personalized Intervention Protocols: Future research should focus on developing and testing targeted interventions based on the identified risk factors, particularly addressing the primary predictors: student satisfaction, self-concept, and academic performance integration.

BIBLIOGRAPHICAL REFERENCES

Author, A. A., Author, B. B., & Author, C. C. (Year). Title of article. *Title of Periodical, volume number* (issue number), pages. DOI or permanent link.

Acosta-Gonzaga, E. (2023). The Effects of Self-Esteem and Academic Engagement on University Students' Performance. *Behavioral Sciences (Basel, Switzerland)*, *13*(4), 348. https://doi.org/10.3390/bs13040348

Adams, D. R., Meyers, S. A., & Beidas, R. S. (2016). The relationship between financial strain, perceived stress, psychological symptoms, and academic and social integration in undergraduate students. *Journal of American College Health*, *64*(5), 362–370. https://doi.org/10.1080/07448481.2016.1154559

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Altmann, A., Toloşi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: A corrected feature importance measure. *Bioinformatics*, *26*(10), 1340–1347. https://doi.org/10.1093/bioinformatics/btq134

Andrews, B., & Wilding, J. M. (2004). The relation of depression and anxiety to life-stress and achievement in students. *British Journal of Psychology*, *95*(4), 509–521. https://doi.org/10.1348/0007126042369802

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, *55*(5), 469–480. https://doi.org/10.1037/0003-066X.55.5.469

Auerbach, R. P., Alonso, J., Axinn, W. G., Cuijpers, P., Ebert, D. D., Green, J. G., Hwang, I., Kessler, R. C., Liu, H., Mortier, P., Nock, M. K., Pinder-Amaker, S., Sampson, N. A., Aguilar-Gaxiola, S., Al-Hamzawi, A., Andrade, L. H., Benjet, C., Caldas-de-Almeida, J. M.,

Demyttenaere, K., … Bruffaerts, R. (2016). Mental disorders among college students in the World Health Organization World Mental Health Surveys. *Psychological Medicine*, *46*(14), 2955–2970. https://doi.org/10.1017/S0033291716001665

Auerbach, R. P., Mortier, P., Bruffaerts, R., Alonso, J., Benjet, C., Cuijpers, P., Demyttenaere, K., Ebert, D. D., Green, J. G., Hasking, P., Murray, E., Nock, M. K., Pinder-Amaker, S., Sampson, N. A., Stein, D. J., Vilagut, G., Zaslavsky, A. M., Kessler, R. C., & WHO WMH-ICS Collaborators. (2018). WHO World Mental Health Surveys International College Student Project: Prevalence and distribution of mental disorders. *Journal of Abnormal Psychology*, *127*(7), 623–638. https://doi.org/10.1037/abn0000362

Beiter, R., Nash, R., McCrady, M., Rhoades, D., Linscomb, M., Clarahan, M., & Sammut, S. (2015). The prevalence and correlates of depression, anxiety, and stress in a sample of college students. *Journal of Affective Disorders*, *173*, 90–96. https://doi.org/10.1016/j.jad.2014.10.054

Bhavani, B. H., & Naveen, N. C. (2024). An Approach to Determine and Categorize Mental Health Condition using Machine Learning and Deep Learning Models. *Engineering, Technology & Applied Science Research*.

Biggs, A., Brough, P., & Drummond, S. (2017). Lazarus and Folkman's Psychological Stress and Coping Theory. In C. L. Cooper & J. C. Quick (Eds.), *The Handbook of Stress and Health* (1st ed., pp. 349–364). Wiley. https://doi.org/10.1002/9781118993811.ch21

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in Public Health*, *6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Brandão, M. P., Pimentel, F. L., & Cardoso, M. F. (2011). Impact of academic exposure on health status of university students. *Revista de Saúde Pública*, *45*, 49–58. https://doi.org/10.1590/S0034-89102011000100006

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Bzdok, D., & Meyer-Lindenberg, A. (2018). Machine Learning for Precision Psychiatry: Opportunities and Challenges. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(3), 223–230. https://doi.org/10.1016/j.bpsc.2017.11.007

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. https://doi.org/10.1613/jair.953

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Clance, P. R., & Imes, S. A. (1978). The imposter phenomenon in high achieving women: Dynamics and therapeutic intervention. *Psychotherapy: Theory, Research & Practice*, *15*(3), 241–247. https://doi.org/10.1037/h0086006

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, *98*(2), 310–357. https://doi.org/10.1037/0033-2909.98.2.310

Conley, C. S., Kirsch, A. C., Dickson, D. A., & Bryant, F. B. (2014). Negotiating the Transition to College: Developmental Trajectories and Gender Differences in Psychological

Functioning, Cognitive-Affective Strategies, and Social Well-Being. *Emerging Adulthood*, *2*(3), 195–210. https://doi.org/10.1177/2167696814521808

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018

Crameri, A., Schuetz, C., Andreae, A., Koemeda, M., Schulthess, P., Tschuschke, V., & Von Wyl, A. (2016). The Brief Symptom Inventory and the Outcome Questionnaire-45 in the Assessment of the Outcome Quality of Mental Health Interventions. *Psychiatry Journal*, *2016*, 1–14. https://doi.org/10.1155/2016/7830785

Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist*, *55*(1), 34–43. https://doi.org/10.1037/0003-066X.55.1.34

Diener, E., Oishi, S., & Tay, L. (2018). Advances in subjective well-being research. *Nature Human Behaviour*, *2*(4), 253–260. https://doi.org/10.1038/s41562-018-0307-6

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., & Biswas-Diener, R. (2010). New Well-being Measures: Short Scales to Assess Flourishing and Positive and Negative Feelings. *Social Indicators Research*, *97*(2), 143–156. https://doi.org/10.1007/s11205-009-9493-y

Ding, H., Li, N., Li, L., Xu, Z., & Xia, W. (2025). Machine learning-enabled mental health risk prediction for youths with stressful life events: A modelling study. *Journal of Affective Disorders*, *368*, 537–546. https://doi.org/10.1016/j.jad.2024.09.111

Drira, M., Ben Hassine, S., Zhang, M., & Smith, S. (2024). Machine Learning Methods in Student Mental Health Research: An Ethics-Centered Systematic Literature Review. *Applied Sciences*, *14*(24), 11738. https://doi.org/10.3390/app142411738

Eisenberg, D., Golberstein, E., & Hunt, J. B. (2009). Mental Health and Academic Success in College. *The B.E. Journal of Economic Analysis & Policy*, *9*(1). https://doi.org/10.2202/1935-1682.2191

Eisenberg, D., Gollust, S. E., Golberstein, E., & Hefner, J. L. (2007). Prevalence and correlates of depression, anxiety, and suicidality among university students. *American Journal of Orthopsychiatry*, *77*(4), 534–542. https://doi.org/10.1037/0002-9432.77.4.534

Eisenberg, D., Lipson, S. K., & Posselt, J. (2016). Promoting Resilience, Retention, and Mental Health. *New Directions for Student Services*, *2016*(156), 87–95. https://doi.org/10.1002/ss.20194

Eisinga, R., Grotenhuis, M. T., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health*, *58*(4), 637–642. https://doi.org/10.1007/s00038-012-0416-3

Funder, D. C., & Ozer, D. J. (2019). Evaluating Effect Size in Psychological Research: Sense and Nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202

Galambos, N. L., Barker, E. T., & Tilton-Weaver, L. C. (2003). Who gets caught at maturity gap? A study of pseudomature, immature, and mature adolescents. *International Journal of Behavioral Development*, *27*(3), 253–263. https://doi.org/10.1080/01650250244000326

Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H.-C., & Jeste, D. V. (2019). Artificial Intelligence for Mental Health and Mental Illnesses: An Overview. *Current Psychiatry Reports*, *21*(11). https://doi.org/10.1007/s11920-019-1094-0

Gulliver, A., Farrer, L., Bennett, K., Ali, K., Hellsing, A., Katruss, N., & Griffiths, K. M. (2018). University staff experiences of students with mental health problems and their

perceptions of staff training needs. *Journal of Mental Health*, *27*(3), 247–256. https://doi.org/10.1080/09638237.2018.1466042

Hair, J. F. (Ed.). (2010). *Multivariate data analysis: A global perspective* (7. ed., global ed). Pearson.

Hammer, J. H., Parent, M. C., & Spiker, D. A. (2018). Mental Help Seeking Attitudes Scale (MHSAS): Development, reliability, validity, and comparison with the ATSPPH-SF and IASMHS-PO. *Journal of Counseling Psychology*, *65*(1), 74–85. https://doi.org/10.1037/cou0000248

Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Van Kerkwijk, M. H., Brett, M., Haldane, A., Del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hastie, T., Tibshirani, R., & Friedman, J. H. (2017). *The elements of statistical learning: Data mining, inference, and prediction* (Second edition). Springer.

Hefner, J., & Eisenberg, D. (2009). Social support and mental health among college students. *American Journal of Orthopsychiatry*, *79*(4), 491–499. https://doi.org/10.1037/a0016918

Hendricks, M. L., & Testa, R. J. (2012). A conceptual framework for clinical work with transgender and gender nonconforming clients: An adaptation of the Minority Stress Model. *Professional Psychology: Research and Practice*, *43*(5), 460–467. https://doi.org/10.1037/a0029597

Hernández-Torrano, D., Ibrayeva, L., Sparks, J., Lim, N., Clementi, A., Almukhambetova, A., Nurtayev, Y., & Muratkyzy, A. (2020). Mental Health and Well-Being of University

Students: A Bibliometric Mapping of the Literature. *Frontiers in Psychology*, *11*, 1226. https://doi.org/10.3389/fpsyg.2020.01226

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/mcse.2007.55

Ibrahim, A. K., Kelly, S. J., Adams, C. E., & Glazebrook, C. (2013). A systematic review of studies of depression prevalence in university students. *Journal of Psychiatric Research*, *47*(3), 391–400. https://doi.org/10.1016/j.jpsychires.2012.11.015

Keyes, C. L. M. (2002). The Mental Health Continuum: From Languishing to Flourishing in Life. *Journal of Health and Social Behavior*, *43*(2), 207. https://doi.org/10.2307/3090197

Keyes, C. L. M., Dhingra, S. S., & Simoes, E. J. (2010). Change in Level of Positive Mental Health as a Predictor of Future Risk of Mental Illness. *American Journal of Public Health*, *100*(12), 2366–2371. https://doi.org/10.2105/AJPH.2010.192245

Keyes, C. L. M., & Martin, C. C. (2017). The Complete State Model of Mental Health. In M. Slade, L. Oades, & A. Jarden (Eds.), *Wellbeing, Recovery and Mental Health* (1st ed., pp. 86–98). Cambridge University Press. https://doi.org/10.1017/9781316339275.009

Kinzie, J., & Kuh, G. (2017). Reframing Student Success in College: Advancing Know-What and Know-How. *Change: The Magazine of Higher Learning*, *49*(3), 19–27. https://doi.org/10.1080/00091383.2017.1321429

Levecque, K., Anseel, F., De Beuckelaer, A., Van Der Heyden, J., & Gisle, L. (2017). Work organization and mental health problems in PhD students. *Research Policy*, *46*(4), 868–879. https://doi.org/10.1016/j.respol.2017.02.008

Lipson, S. K., Lattie, E. G., & Eisenberg, D. (2019). Increased Rates of Mental Health Service Utilization by U.S. College Students: 10-Year Population-Level Trends (2007–2017). *Psychiatric Services*, *70*(1), 60–63. https://doi.org/10.1176/appi.ps.201800332

Madububambachu, U., Ukpebor, A., & Ihezue, U. (2024). Machine Learning Techniques to Predict Mental Health Diagnoses: A Systematic Literature Review. *Clinical Practice & Epidemiology in Mental Health*, *20*(1), e17450179315688. https://doi.org/10.2174/0117450179315688240607052117

Mael, F., & Ashforth, B. E. (1992). Alumni and their alma mater: A partial test of the reformulated model of organizational identification. *Journal of Organizational Behavior*, *13*(2), 103–123. https://doi.org/10.1002/job.4030130202

Marsh, H. W., & Craven, R. G. (2006). Reciprocal Effects of Self-Concept and Performance From a Multidimensional Perspective: Beyond Seductive Pleasure and Unidimensional Perspectives. *Perspectives on Psychological Science*, *1*(2), 133–163. https://doi.org/10.1111/j.1745-6916.2006.00010.x

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering: Academic self-concept. *British Journal of Educational Psychology*, *81*(1), 59–77. https://doi.org/10.1348/000709910X503501

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, *15*(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3. ed., [Nachdr.]). McGraw-Hill.

Offstein, E. H., Larson, M. B., McNeill, A. L., & Mjoni Mwale, H. (2004). Are we doing enough for today's graduate student? *International Journal of Educational Management*, *18*(7), 396–407. https://doi.org/10.1108/09513540410563103

Pákozdy, C., Askew, J., Dyer, J., Gately, P., Martin, L., Mavor, K. I., & Brown, G. R. (2024). The imposter phenomenon and its relationship with self-efficacy, perfectionism and

happiness in university students. *Current Psychology*, *43*(6), 5153–5162. https://doi.org/10.1007/s12144-023-04672-4

Pascoe, M. C., Hetrick, S. E., & Parker, A. G. (2020). The impact of stress on students in secondary school and higher education. *International Journal of Adolescence and Youth*, *25*(1), 104–112. https://doi.org/10.1080/02673843.2019.1596823

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2012). *Scikit-learn: Machine Learning in Python*. https://doi.org/10.48550/ARXIV.1201.0490

Piccinelli, M., & Wilkinson, G. (2000). Gender differences in depression: Critical review. *British Journal of Psychiatry*, *177*(6), 486–492. https://doi.org/10.1192/bjp.177.6.486

Powers, D. M., & Ailab. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*.

Pyhältö, K., Stubb, J., & Lonka, K. (2009). Developing scholarly communities as learning environments for doctoral students. *International Journal for Academic Development*, *14*(3), 221–232. https://doi.org/10.1080/13601440903106551

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353–387. https://doi.org/10.1037/a0026838

Rickwood, D., Deane, F. P., Wilson, C. J., & Ciarrochi, J. (2005). Young people's help-seeking for mental health problems. *Australian E-Journal for the Advancement of Mental Health*, *4*(3), 218–251. https://doi.org/10.5172/jamh.4.3.218

Rosenberg, M. (2011). *Rosenberg Self-Esteem Scale* [Dataset]. https://doi.org/10.1037/t01038-000

Rosenthal, D. A., Russell, J., & Thomson, G. (2008). The health and wellbeing of international students at an Australian university. *Higher Education*, *55*(1), 51–67. https://doi.org/10.1007/s10734-006-9037-1

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, *57*(6), 1069–1081. https://doi.org/10.1037/0022-3514.57.6.1069

Salanova, M., Schaufeli, W., Martínez, I., & Bresó, E. (2010). How obstacles and facilitators predict academic performance: The mediating role of study burnout and engagement. *Anxiety, Stress & Coping*, *23*(1), 53–70. https://doi.org/10.1080/10615800802609965

Sallehuddin Md Yusof. (2023). *The Six Stages of Test Construction*. https://doi.org/10.13140/RG.2.2.14509.26081/1

Sawir, E., Marginson, S., Deumert, A., Nyland, C., & Ramia, G. (2008). Loneliness and International Students: An Australian Study. *Journal of Studies in International Education*, *12*(2), 148–180. https://doi.org/10.1177/1028315307299699

Schreiner, L. A., & Nelson, D. D. (2013). The Contribution of Student Satisfaction to Persistence. *Journal of College Student Retention: Research, Theory & Practice*, *15*(1), 73–111. https://doi.org/10.2190/CS.15.1.f

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, *54*(2), 93–105. https://doi.org/10.1037/0003-066x.54.2.93

Selenko, E., & Batinic, B. (2011). Beyond debt. A moderator analysis of the relationship between perceived financial strain and mental health. *Social Science & Medicine*, *73*(12), 1725–1732. https://doi.org/10.1016/j.socscimed.2011.09.022

Seligman, M. (2018). PERMA and the building blocks of well-being. *The Journal of Positive Psychology*, *13*(4), 333–335. https://doi.org/10.1080/17439760.2018.1437466

Shatte, A. B. R., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: A scoping review of methods and applications. *Psychological Medicine*, *49*(09), 1426–1448. https://doi.org/10.1017/S0033291719000151

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-Concept: Validation of Construct Interpretations. *Review of Educational Research*, *46*(3), 407–441. https://doi.org/10.3102/00346543046003407

Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, *135*(1), 94–120. https://doi.org/10.1037/a0014212

Smith, R. A., & Khawaja, N. G. (2011). A review of the acculturation experiences of international students. *International Journal of Intercultural Relations*, *35*(6), 699–713. https://doi.org/10.1016/j.ijintrel.2011.08.004

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, *45*(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Sowislo, J. F., & Orth, U. (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychological Bulletin*, *139*(1), 213–240. https://doi.org/10.1037/a0028931

Stallman, H. M. (2010). Psychological distress in university students: A comparison with general population data. *Australian Psychologist*, *45*(4), 249–257. https://doi.org/10.1080/00050067.2010.482109

Stensaker, B. (2015). Organizational identity as a concept for understanding university dynamics. *Higher Education*, *69*(1), 103–115. https://doi.org/10.1007/s10734-014-9763-8

Sullivan, G. M., & Artino, A. R. (2013). Analyzing and Interpreting Data From Likert-Type Scales. *Journal of Graduate Medical Education*, *5*(4), 541–542. https://doi.org/10.4300/jgme-5-4-18

The World Health Organization-Five Well-Being Index (WHO-5). (2024). https://www.who.int/publications/m/item/WHO-UCN-MSD-MHE-2024.01

Thoits, P. A. (2011). Mechanisms Linking Social Ties and Support to Physical and Mental Health. *Journal of Health and Social Behavior*, *52*(2), 145–161. https://doi.org/10.1177/0022146510395592

Thomas, N., Schneider, J., & Zhou, S. (2024). Qualitative research on language learning strategies and self-regulation. *AILA Review*, *37*(2), 177–187. https://doi.org/10.1075/aila.00059.tho

Tilley, B. P. (2014). What Makes a Student Non-traditional? A Comparison of Students Over and Under Age 25 in Online, Accelerated Psychology Courses. *Psychology Learning & Teaching*, *13*(2), 95–106. https://doi.org/10.2304/plat.2014.13.2.95

Tinto, V. (2017). Reflections on Student Persistence. *Student Success*, *8*(2), 1–8. https://doi.org/10.5204/ssj.v8i2.376

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley publ.

Waskom, M. (2021). seaborn: Statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021. https://doi.org/10.21105/joss.03021

Wilkins, S., Balakrishnan, M. S., & Huisman, J. (2012). Student Choice in Higher Education: Motivations for Choosing to Study at an International Branch Campus. *Journal of*

Studies in International Education, *16*(5), 413–433. https://doi.org/10.1177/1028315311429002

Wooldridge, J. M. (2020). *Introductory econometrics: A modern approach* (Seventh edition). Cengage Learning.

APPENDIX A : PSYCHOLOGIST QUESTIONNAIRE APPROVAL

Ana Margarida Faro Craveiro Saraiva

Licensed Psychologist, OPP 9210

Agrupamento de Escolas de Parede

Email: ana.saraiva65@gmail.com

05 May 2025

To Whom It May Concern,

This letter confirms my participation in the review of the questionnaire used in the Master's dissertation entitled "The Impact of High Ranked Universities on Students' Mental Health and Well-Being: Study Case of [redacted], prepared by Yasmine Boubezari.

As a licensed psychologist, I evaluated the coherence and relevance of the questionnaire[1], ensuring it adequately assesses students' mental health and well-being. The final version incorporates my suggestions and complies with psychological research standards.

For any further inquiries, I am available via the contact details provided above.

Yours sincerely,

Ana Margarida Saraiva

Ana Margarida Faro Craveiro Saraiva

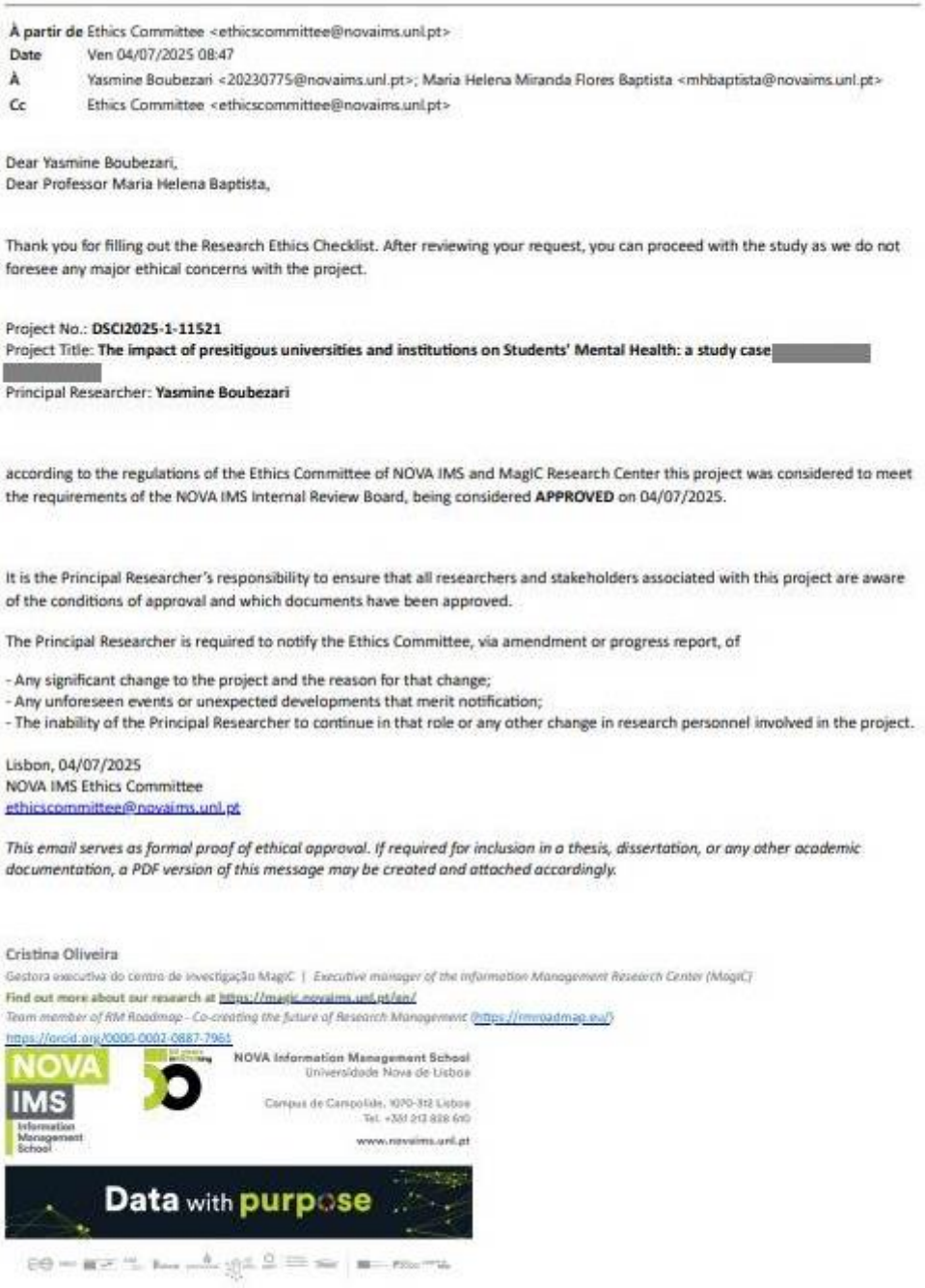
Licensed Psychologist, OPP 9210

[1] The questionnaire is available at: https://novaims.eu.qualtrics.com/jfe/form/SV_4Z6zSW9Nv6CSctg

Note: The name of the institution has been intentionally anonymized in all supporting documents and materials to ensure confidentiality for publication purposes.

Figure 12: Psychologist's Approval for Questionnaire Development

APPENDIX B: ETHICS COMMITTEE APPROVAL



À partir de Ethics Committee <ethicscommittee@novaims.unl.pt>
Date Ven 04/07/2025 08:47
À Yasmine Boubezari <20230775@novaims.unl.pt>; Maria Helena Miranda Flores Baptista <mhbaptista@novaims.unl.pt>
Cc Ethics Committee <ethicscommittee@novaims.unl.pt>

Dear Yasmine Boubezari,
Dear Professor Maria Helena Baptista,

Thank you for filling out the Research Ethics Checklist. After reviewing your request, you can proceed with the study as we do not foresee any major ethical concerns with the project.

Project No.: **DSCI2025-1-11521**
Project Title: **The impact of presitigous universities and institutions on Students' Mental Health: a study case**
Principal Researcher: **Yasmine Boubezari**

according to the regulations of the Ethics Committee of NOVA IMS and MagIC Research Center this project was considered to meet the requirements of the NOVA IMS Internal Review Board, being considered **APPROVED** on 04/07/2025.

It is the Principal Researcher's responsibility to ensure that all researchers and stakeholders associated with this project are aware of the conditions of approval and which documents have been approved.

The Principal Researcher is required to notify the Ethics Committee, via amendment or progress report, of

- Any significant change to the project and the reason for that change;
- Any unforeseen events or unexpected developments that merit notification;
- The inability of the Principal Researcher to continue in that role or any other change in research personnel involved in the project.

Lisbon, 04/07/2025
NOVA IMS Ethics Committee
ethicscommittee@novaims.unl.pt

This email serves as formal proof of ethical approval. If required for inclusion in a thesis, dissertation, or any other academic documentation, a PDF version of this message may be created and attached accordingly.

Cristina Oliveira
Gestora executiva do centro de investigação MagIC | Executive manager of the Information Management Research Center (MagIC)
Find out more about our research at https://magic.novaims.unl.pt/en/
Team member of RM Roadmap - Co-creating the future of Research Management (https://rmroadmap.eu/)
https://orcid.org/0000-0002-0887-7961
NOVA IMS Information Management School
NOVA Information Management School
Universidade Nova de Lisboa
Campus de Campolide, 1070-312 Lisboa
Tel. +351 213 828 610
www.novaims.unl.pt
Data with purpose

Note: The name of the institution has been intentionally anonymized in all supporting documents and materials to ensure confidentiality for publication purposes.

Figure 13: Ethics committee Approval of the Research Topic

APPENDIX C: QUESTIONNAIRE

Table 12: Overview of Questionnaire Items, Corresponding Questions, and Measurement Types

Variables	Scaling	Associated questions
Age	Multiple choice question limited to one choice.	Q1. What is your age range? - 18-24 - 25-34 - 35-44 - 45-54 - 54-65 - +65
Gender	Multiple choice question, limited to one choice.	Q2. What is your gender? - Male. - Female. - Non-binary/ third gender. - Prefer not to say.
Degree level	Multiple choice question, as one student can have more than one degree from Institution XYZ.	Q3. What degree are you or have you studied at institution XYZ? - Bachelor degree. - Master's degree. - Post-graduation program. - PhD and above.
Area of study (program)	Multiple choice question, as a student can have different study areas in his academic journey.	Q4. Please select your study area. - Study area A - Study area B - Study area C - Study area D - Study area E - Study area F - Study area G
Residency Status	Multiple choice question,	

	limited to one choice.	Q5. Are you a local or an international student? - Portuguese student. - International student.
Well-being	From "1. strongly Disagree to 5. Strongly Agree", in Likert matrix	- My overall well-being increased since I have joined this institution. - I generally feel physically healthy and emotionally balanced in my daily life. - I manage daily stress effectively and maintain a positive outlook.
Self-concept	From "1. strongly Disagree to 5. Strongly Agree", in Likert matrix	- My self-image is shaped more by internal values than external recognition. - I feel confident in my ability to overcome challenges in my academic, personal, and professional life.
Academic performance	From "1. strongly Disagree to 5. Strongly Agree", in Likert matrix	- I feel satisfied with my overall academic performance (e.g., grades, progress, etc.). - My academic results accurately reflect my capabilities.
Financial stress	From "1. strongly Disagree to 5. Strongly Agree", in Likert matrix	- Financial concerns (e.g., tuition fees, living expenses) cause me significant anxiety.
social interactions	From "1. strongly Disagree to 5. Strongly Agree", in Likert matrix	- Interacting with peers on and off campus significantly enriches my student experience. - I enjoy and actively seek out social activities where I can meet new people.
Reliance on institutional image	From "1. strongly Disagree to 5.	- I often highlight my academic achievements and university. affiliation in social settings, even without intending to.

	Strongly Agree", in Likert matrix	- When I introduce myself, I make sure to mention my affiliation with Institution XYZ early in the conversation.
Satisfaction from the student experience	From "1. strongly Disagree to 5. Strongly Agree", in Likert matrix	- I would recommend this institution to a friend or family member without hesitation. - I feel satisfied with my overall student journey.
Help seeking for mental health	Scaled multiple choice question, with a single selection.	Q7. Have you sought any mental health support service since you have joined this institution? (e.g. stress management workshops, psychologist, support hotlines, etc) - Yes, I've already did. - I've considered it. - Neutral. - I don't feel the need for now. - No, I never thought about it.
Rewarding Aspects	Hybrid question (Multiple-choice with Open-ended Option)	Q8 What aspects of your university experience do you find the most rewarding? (Please select all that apply) - Career development. - Personal growth. - Access to high-quality teaching, libraries, and academic tools. - Extracurricular activities. - Diversity and Inclusion. - Opportunities for research. - The reputation of the university and its benefits for personal and professional opportunities. - None of them. - Other (please mention).
Potential Initiatives	Hybrid question (Multiple-choice with	Q9. Which initiatives and/or changes would you like to see? (please select all that apply)

	Open-ended Option)	- Enhanced career support services (e.g., internships, job affairs, networking events, etc). - Larger infrastructures and facilities (e.g., open study room 24/7, more tables in open spaces, etc). - Expansion of scholarships or financial aid opportunities. - Stronger inclusion and diversity initiatives within the university. - Workshops on personal development (e.g., time management, critical thinking, etc) - Opportunities to participate in volunteering or social responsibility projects. - Intern anonymous support hotline or online mental health tools (an intern hotline that you could call anonymously for issues related to the campus or studies). - Frequent wellness programs incorporating physical and emotional health activities (e.g. yoga, fitness classes, scheduled running sessions). - None. - Other (please mention).
Understability of the terms and vocabulary	Trichotomous question	Q10. Were the questions and words used on this survey clear and understandable? - Yes. - No. - Neutral.
Easiness and user-friendly	Trichotomous question	Q11. Were the scales utilized to answer easy to use and appropriate? (from "Strongly Agree" to "Strongly Disagree", multiple choice, open questions) - Yes. - No. - Neutral.

Discomfort checking	Trichotomous question	Q12. Did you experience any sort of discomfort or anxiety while answering this survey? - No. - Yes. - Neutral.
Optional for additional feedback.	Open-ended question	Q13. Any comment or suggestion that you would like to add?

Note: The name of the institution and its study areas and programs have been intentionally anonymized in all supporting documents and materials to ensure confidentiality for publication purposes.

APPENDIX D: EXPLORATORY DATA ANALYSIS

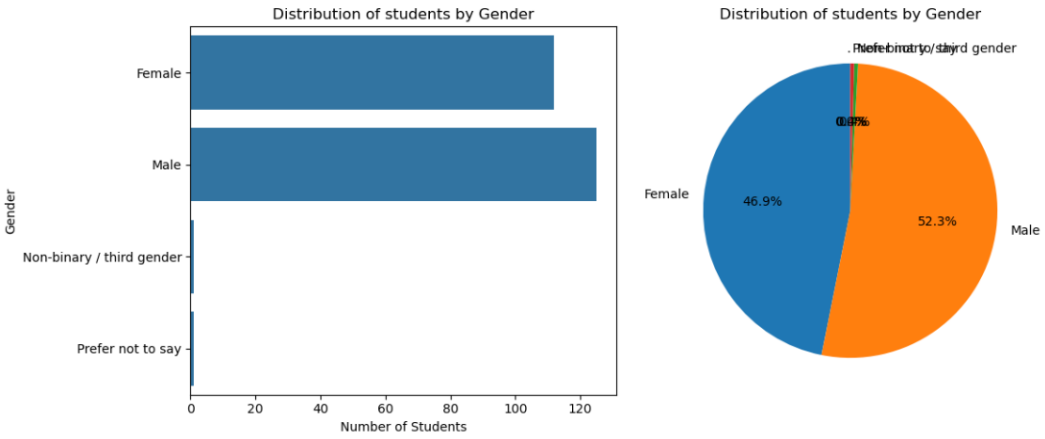


Figure 14: Gender Distribution Illustrated by Barplot and Pie Chart

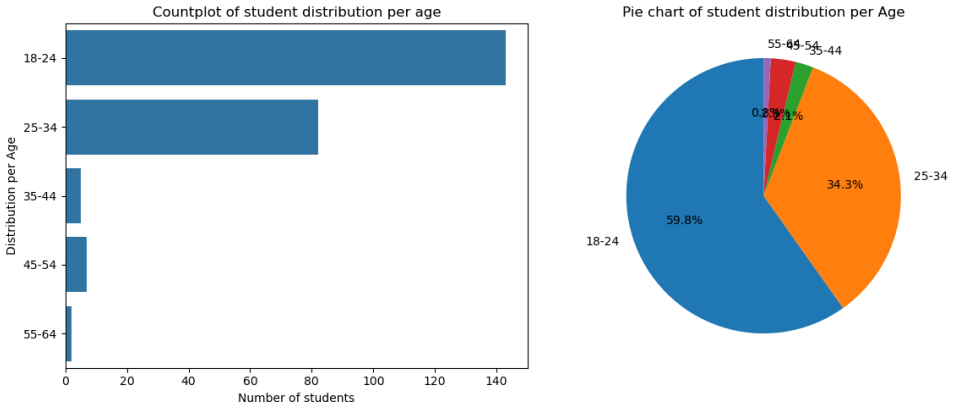


Figure 15: Age Groups Distribution Illustrated by Barplot and Pie Chart

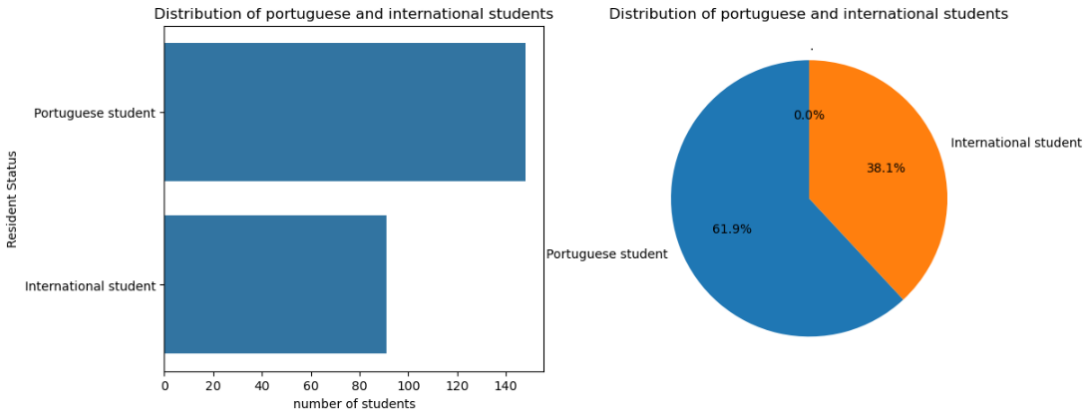


Figure 16: Portuguese and International Students Distribution Illustrated by Barplot and Pie Chart

Table 13: Comparison of Gender Variable Between Sample and Real Population

Gender	**Sample (%)**	**Population (%)**	**Difference (sample%-population %)**
Male	52.30	54.46	-2.16
Female	46.90	45.54	+1.36
Third gender/else	0.80	-	-

Table 14: Comparison of Academic Level Variable Between Sample and Real Population

Academic level	**Sample (%)**	**Population (%)**	**Difference (sample%-population %)**
Bachelor's	32.30	18.86	+13.44
Master's	55.64	59.59	-3.95
PhD	3.89	3.08	-0.81
Post-graduation	8.17	18.47	-10.30

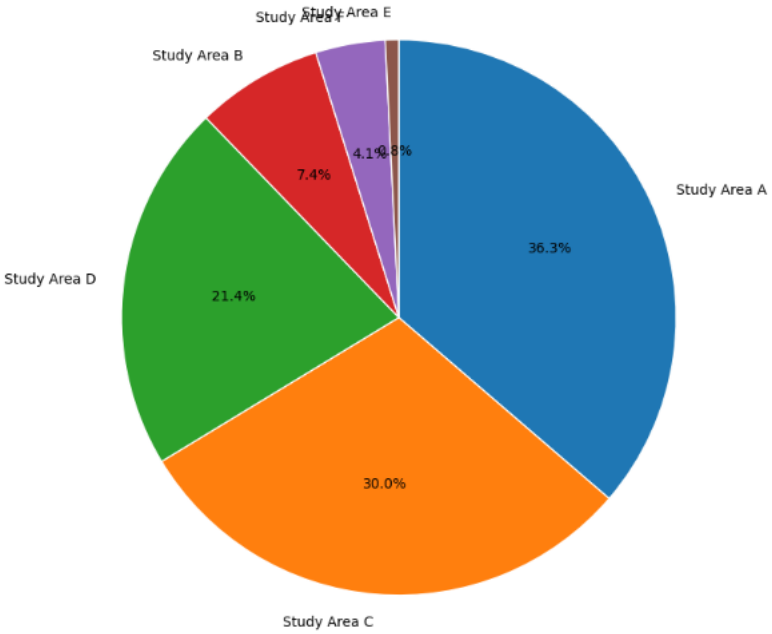


Figure 17: Distribution of Real Sample Population Across Programs for Academic Year 2024–2025 (Pie Chart)

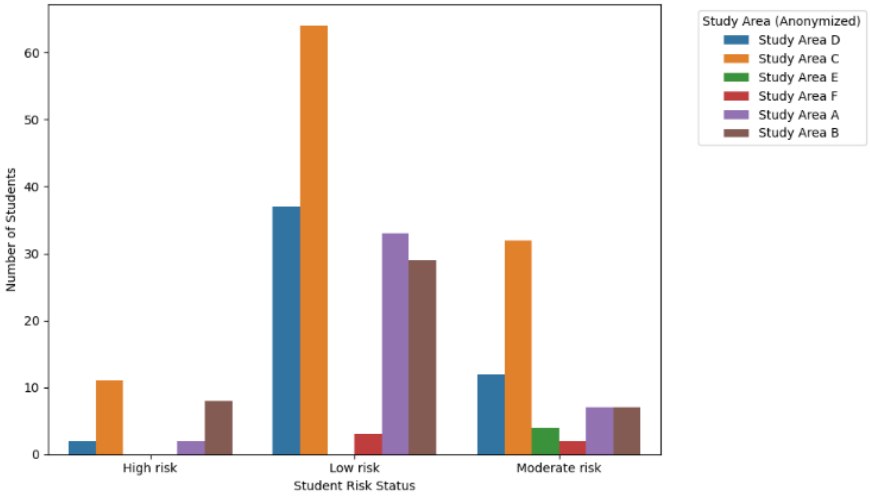


Figure 18: Distribution of Study Area across Risk Class in Barplot

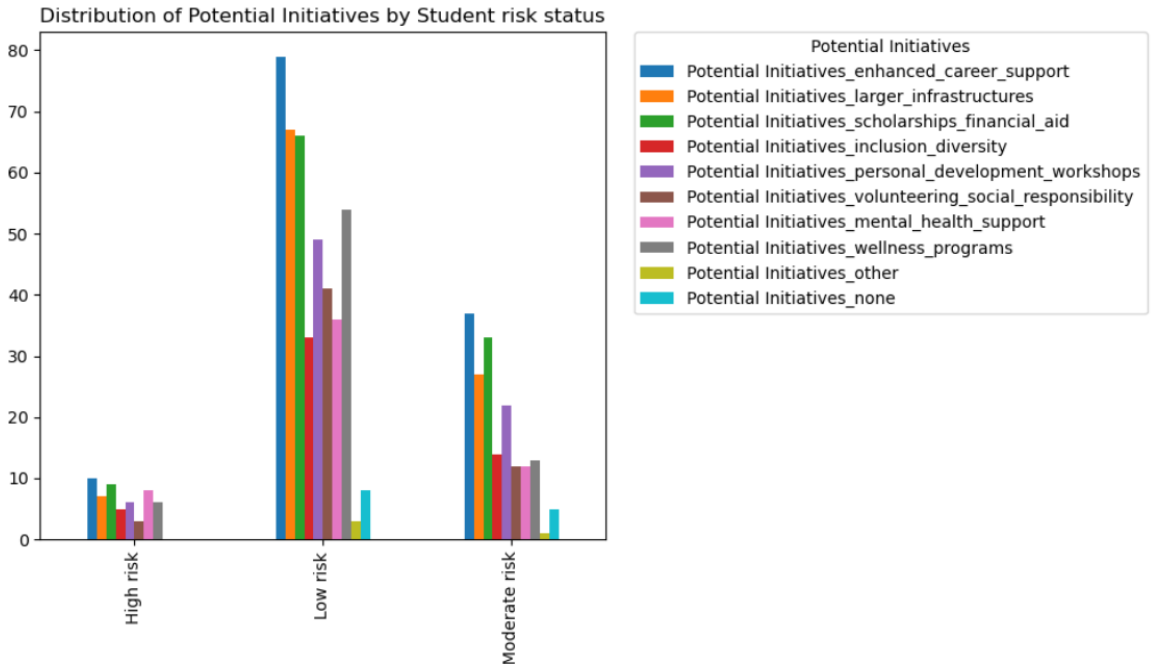


Figure 19: Distribution of Potential Intitiatives Across student Risk Classes

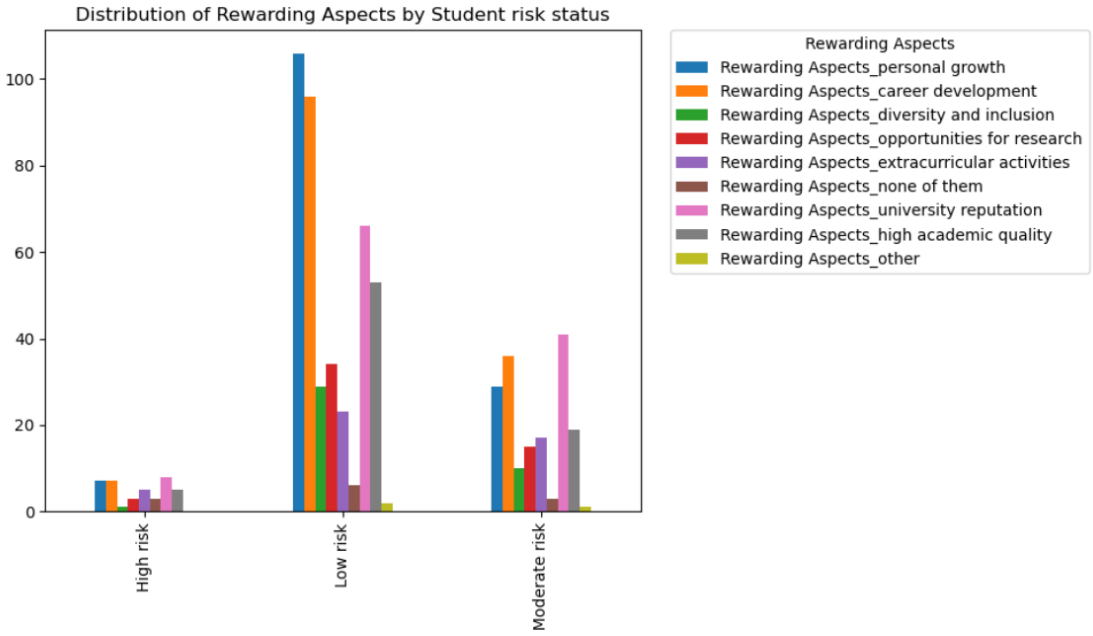


Figure 20: Distribution of Rewarding Aspects Across student Risk Classes

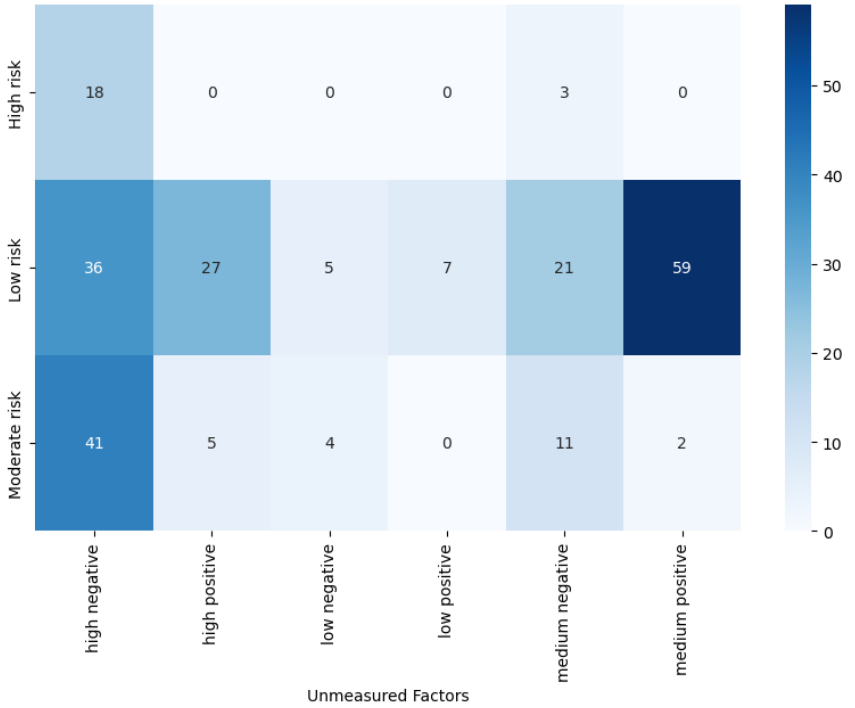


Figure 21: Correlation Between “Unmeasured Factors” and “Student Risk Status”.

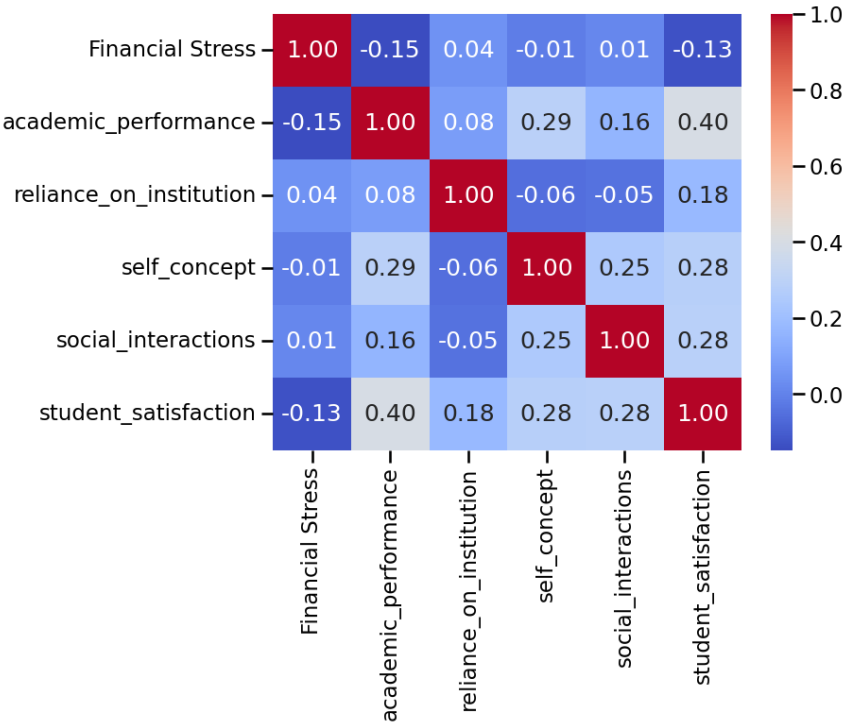


Figure 22: Correlation Among Selected Influencing Factors of Well-being

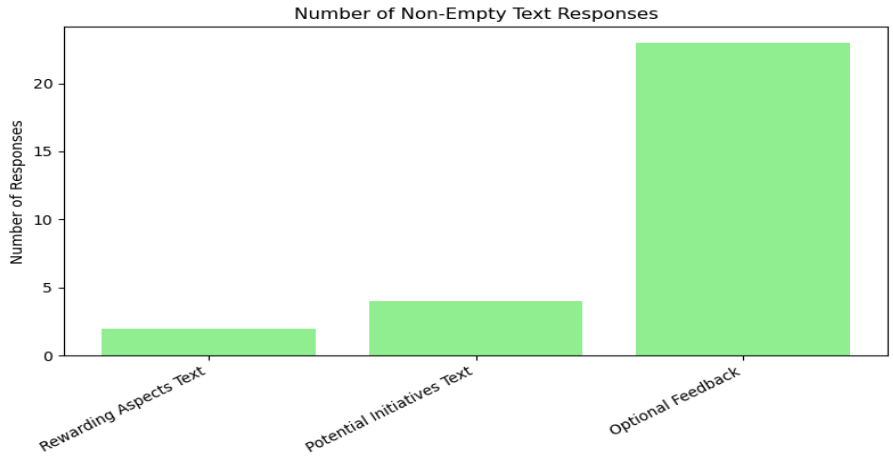


Figure 23: Distribution of Participant Answers for Open-Ended Questions

Table 15: Summary of Descriptive Statistics for Likert Scale Variables

Variable	mean	SD	Min	Max	Skewness
Well being	3.34	0.79	1.33	4.67	-0.46
Social interactions	3.72	0.90	1.00	5.00	-0.59
Student satisfaction	3.78	0.83	1.00	5.00	-0.99
Self-concept	3.64	0.86	1.00	5.00	-0.71
Academic performance	3.43	0.85	1.00	5.00	-0.40
Financial stress	3.28	1.24	1.00	5.00	-0.26
Reliance on institution	2.81	1.00	1.00	5.00	-0.16

APPENDIX E: MODELS AND FEATURES IMPORTANCE

```
# Random Forest Pipeline
rf_pipe = Pipeline(steps=[
    ('pre', preprocess_full),
    ('clf', RandomForestClassifier(
        n_estimators=600,
        random_state=42, # the seed is fixed at 42
        class_weight='balanced', # for the class imbalance
        bootstrap=True
    ))
])

# Hyperparameter Search
param_dist = {
    'clf__max_depth'        : randint(3, 8),
    'clf__min_samples_leaf' : randint(2, 10),
    'clf__max_features'     : uniform(0.3, 0.7)
}

cv = StratifiedKFold(n_splits=5, shuffle=True, random_state=42)
scoring = {'f1_macro': 'f1_macro', 'accuracy': 'accuracy'}

search = RandomizedSearchCV(
    estimator          = rf_pipe,
    param_distributions= param_dist,
    n_iter             = 25,
    cv                 = cv,
    scoring            = scoring,
    refit              = 'f1_macro',
    n_jobs             = -1,
    random_state       = 42,
    verbose            = 1
)
```

Figure 24: Random Forest Pipeline Code and Hyperparameter Tuning

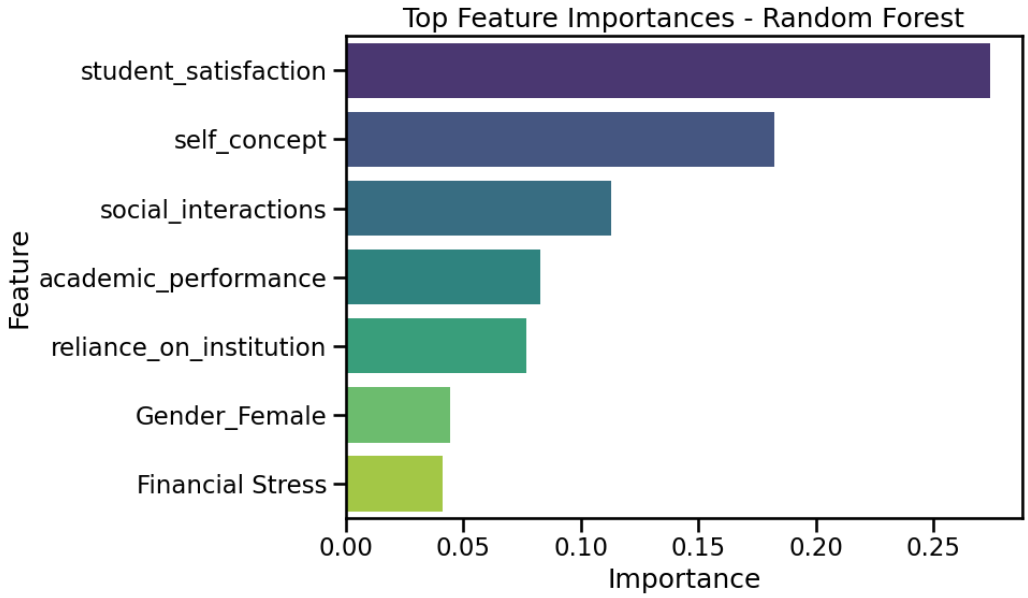


Figure 25: Random Forest Feature Importance Plot Based On Mean Decrease Impurity

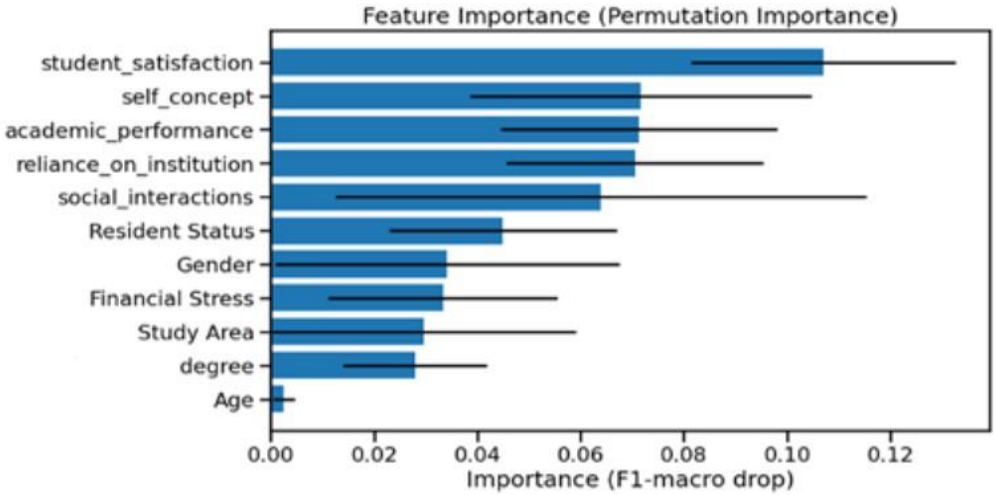


Figure 26: SVM Feature Importance Plot based on Permutation Importance

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa