

A Principal Component Analysis of Key Factors Influencing Enrollment Decisions in a Bachelor of Business Administration Program

Umaporn Chaisoong^{1,*}, Thitimaporn Waenphet², Nonthakan Nganjaturus³,
Jandara Suksam⁴, Thanet Yothasiri⁵, and Waraluck Maprasom⁶

^{1,2,3,4,5,6}Program in Digital Business Technology, Faculty of Management Technology, Rajamangala University of Technology Isan Surin Campus, Surin 32000, Thailand;

Email: ¹umaporn.ch@rmuti.ac.th, ²thitimaporn.wa@rmuti.ac.th, ³nonthakan.ng@rmuti.ac.th,
⁴jandara.su@rmuti.ac.th, ⁵prawit.yo@rmuti.ac.th, ⁶waraluck.ku@rmuti.ac.th

**Corresponding author:* Umaporn Chaisoong (umaporn.ch@rmuti.ac.th)

Received: 27 January 2026; **Revised:** 14 June 2026; **Accepted:** 1 July 2026; **Published:** 12 July 2026

Abstract

This study aims to identify and extract the key factors influencing students' enrollment decisions in higher education by adopting a data-driven analytical approach. Principal Component Analysis (PCA) was employed within the CRISP-DM framework to systematically analyze quantitative data collected from 571 prospective students across three provinces in Thailand. The survey instrument comprised 45 observed variables. Data suitability tests indicated strong adequacy for factor analysis, with a Kaiser–Meyer–Olkin (KMO) value of 0.963 and a statistically significant Bartlett's Test of Sphericity. The PCA results, using Varimax rotation, revealed five latent components shaping enrollment decisions: financial readiness, institutional support and safety, student identity formation, technological and curriculum relevance, and peer guidance. The findings demonstrate that enrollment decision-making is multidimensional and hierarchical, extending beyond academic reputation to include economic, psychosocial, and experiential considerations. Academically, this study contributes by integrating PCA with educational data mining to provide a structured interpretation of student choice behavior. Practically, the results offer evidence-based insights to support curriculum development, targeted communication strategies, and enrollment policy design in higher education institutions.

Keywords: Enrollment Decision, Principal Component Analysis, Data Mining, Higher Education, Surin



Copyright © 2026 Journal of Computer and Creative Technology. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC-BY-NC-ND).

Citation:

Chaisoong, U., Waenphet, T., Nganjaturus, N., Suksam, J., Yothasiri, T., & Maprasom, W. (2026). A Principal Component Analysis of Key Factors Influencing Enrollment Decisions in a Bachelor of Business Administration Program. *Journal of Computer and Creative Technology*, 4(2), e3446. <https://doi.org/10.65205/jcct.2026.e3446>

1. Introduction

Over the past decade, higher education institutions worldwide have faced increasingly intense competition for student enrollment, driven by demographic decline, reduced birth rates, and the rapid expansion of alternative learning pathways such as online education and lifelong learning models. These structural changes have reshaped how prospective students evaluate educational opportunities, making enrollment decisions more complex and multifaceted than in previous generations (Marginson, 2016; OECD, 2021). Rather than relying solely on institutional reputation or anticipated employment outcomes, students now consider a wider range of factors, including financial circumstances, perceived educational value, personal aspirations, and alignment between institutional offerings and individual goals (Hemsley-Brown & Oplatka, 2016). Moreover, Azzone & Soncin (2020) found that factors driving students' university choice in Italian higher education institutions included the university's geographical location, regional employment opportunities, university reputation, and ease of access. This research suggests that decision-making regarding higher education enrollment has become increasingly complex, involving multiple interconnected considerations rather than a single determining factor.

In Thailand, the continuing decline in the number of secondary school graduates has intensified competition among higher education institutions and increased pressure to develop more effective recruitment and curriculum management strategies. Many universities now use a public relations strategy that leverages new media and famous people to reach prospective students while developing international programs aligned with current trends and future labour market needs. (Somyoonsab, 2017). However, students' enrollment decisions ultimately depend on how well a program and an institution meet their needs and expectations (Kanjanasamranwong et al., 2017). As a result, understanding the factors that shape students' enrollment decisions has become a strategic priority rather than merely an academic concern (Daniel, 2015). Although previous studies have identified numerous determinants of educational choice, many have examined these variables independently or using descriptive approaches, resulting in fragmented findings and a limited understanding of the underlying decision-making structure. Moreover, educational decision-making data are often characterized by high dimensionality and substantial intercorrelations among variables, making it difficult to distinguish the most influential determinants using conventional analytical techniques (Romero & Ventura, 2020). Similar concerns have been highlighted in recent factor-analytic studies, which emphasize that students' educational choices emerge from the interaction of institutional, economic, social, family, and personal considerations rather than from isolated factors alone (Lee et al., 2026). Consequently, there is a growing need for analytical approaches capable of identifying latent structures that can explain how these factors collectively influence enrollment behavior.

To address this challenge, the present study employs Principal Component Analysis (PCA) within a structured data-analytic framework to identify and consolidate the principal factors influencing students' decisions to enroll in a Bachelor of Business Administration program. PCA is particularly suitable for reducing large sets of interrelated variables to a smaller number of interpretable components while preserving substantial information from the original dataset (Jolliffe & Cadima, 2016). By uncovering the latent structure underlying enrollment decisions, this study contributes to a deeper understanding of student choice behavior and provides evidence that can support institutional planning and decision-making. The findings are expected to inform curriculum development, student recruitment strategies, and targeted communication initiatives, thereby supporting more effective enrollment management and

enhancing institutional competitiveness in an increasingly challenging higher education environment (Daniel, 2015; Romero & Ventura, 2020).

2. Research Objectives

The objectives of this study are threefold:

1. To identify the principal components that influence students' decisions to enroll in a Bachelor of Business Administration program.
2. To classify and reduce the dimensionality of enrollment decision factors with statistical significance.
3. To apply data-driven analytical results to the development of targeted communication strategies aligned with students' needs.

3. Literature Review

Recent scholarship on higher education enrollment decisions increasingly conceptualizes student choice as a multidimensional and context-dependent process shaped by economic, psychosocial, and institutional factors. While earlier studies predominantly emphasized institutional reputation and employability outcomes, contemporary research demonstrates a shift toward more complex evaluative frameworks incorporating perceived value, student-institution fit, and experiential considerations (Hemsley-Brown & Oplatka, 2016). However, despite this conceptual advancement, empirical studies remain fragmented, often lacking integrative analytical approaches that capture underlying decision structures.

A recurring concern in the literature is that enrollment-related variables are often examined in isolation, resulting in a limited understanding of how economic, social, institutional, and personal influences interact within a broader decision-making framework. Although factor-based approaches have been increasingly adopted to classify related variables, many studies focus primarily on identifying groups of factors without adequately explaining the latent structure that connects them (Lee et al., 2026). Furthermore, the high dimensionality and intercorrelation of educational decision data pose methodological challenges that cannot be adequately addressed by conventional descriptive analyses alone (Romero & Ventura, 2020). These limitations suggest a need for analytical techniques that reduce complexity while preserving meaningful relationships among variables. Consequently, approaches such as Principal Component Analysis (PCA) offer a valuable means of uncovering interpretable latent structures that can support both theoretical understanding and evidence-based decision making in higher education contexts (Jolliffe & Cadima, 2016).

3.1 Changing Landscape of Higher Education Enrollment Decisions

The environment surrounding higher education enrollment has changed considerably over the last decade. Demographic decline, lower birth rates, and the expansion of alternative educational pathways have increased competition among institutions and diversified students' expectations (Marginson, 2016; OECD, 2021). Earlier studies frequently portrayed enrollment decisions as rational choices based primarily on academic reputation or future employment opportunities. More recent evidence, however, suggests that students evaluate educational options based on a broader set of considerations, including affordability, institutional support, personal development opportunities, and perceived compatibility with their future aspirations (Hemsley-Brown & Oplatka, 2016). These findings indicate that enrollment

decisions are increasingly shaped by multiple interconnected influences rather than by a single dominant factor, highlighting the need for analytical approaches capable of capturing complex decision structures.

3.2 Student Choice Theory and Hierarchical Decision-Making

Student Choice Theory, particularly the three-stage model proposed by Hossler & Gallagher (1987), remains a foundational framework for understanding enrollment behavior. Empirical evidence consistently supports the role of financial feasibility and accessibility as primary filters in early decision stages (Hossler et al., 2009). However, more recent studies challenge the linearity of this model, arguing that decision-making processes are increasingly iterative and influenced by dynamic interactions among institutional, social, and psychological factors (Paulsen, 2001). This suggests that while Student Choice Theory provides a useful structural lens, it may oversimplify the complexity of contemporary enrollment behavior, particularly in data-rich environments where multiple factors interact simultaneously.

3.3 Maslow's Hierarchy of Needs in Educational Decision Contexts

Maslow's Hierarchy of Needs offers a complementary explanatory framework by situating enrollment decisions within a layered motivational structure (Maslow, 1943). Empirical studies in higher education consistently demonstrate that financial stability and safety function as foundational conditions, while higher-order motivations such as belonging, esteem, and self-actualization influence later-stage decision preferences (Kuh et al., 2010). However, the application of Maslow's framework in educational research remains largely interpretive rather than empirically validated. While some studies have observed variables with hierarchical needs, few systematically test these relationships using structured multivariate techniques. This highlights a gap between theoretical interpretation and empirical validation.

3.4 Data Mining and Factor-Based Approaches in Enrollment Research

To address the complexity of enrollment decision-making, recent research has increasingly employed data mining and multivariate statistical techniques to uncover latent patterns in high-dimensional datasets (Romero & Ventura, 2020). Principal Component Analysis (PCA) has been widely used for dimensionality reduction and to enhance interpretability (Jolliffe & Cadima, 2016). While factor-based studies in the Thai context consistently identify economic, institutional, and social clusters (Santisirikul et al., 2023; Lertsahapan & Serttasangsri, 2020), most applications remain descriptive, focusing on variable grouping without linking findings to theoretical frameworks or strategic implications. Furthermore, few studies integrate these techniques within structured analytical processes such as CRISP-DM, limiting methodological transparency and reproducibility.

3.5 Mapping Prior Literature to the Five Extracted Components

Synthesizing prior research reveals both convergence and divergence across key dimensions. Financial readiness is consistently identified as a foundational determinant, aligning with both economic models and Maslow's lower-order needs. Institutional support and safety emerge as increasingly salient factors, reflecting a shift toward student well-being and trust-based decision-making. In contrast, identity formation and experiential engagement represent evolving dimensions that extend beyond traditional academic considerations, indicating a transition toward self-development-oriented motivations. Technological relevance and curriculum modernity further reinforce the importance of future-oriented competencies, while peer influence highlights the role of social networks in reducing uncertainty. Despite these consistencies, prior studies often treat these factors independently, lacking a unified structural representation.

3.6 Methodological Contribution: PCA within the CRISP-DM Framework

A critical limitation in existing research lies in the absence of integrated analytical frameworks that combine statistical rigor with practical applicability. While PCA and EFA are widely used, they are often applied in isolation rather than embedded in systematic data-mining processes. By incorporating PCA within the CRISP-DM framework, this study enhances methodological robustness, ensuring a structured progression from data understanding to model evaluation (Provost & Fawcett, 2013; Chaisoong & Tirakoat, 2020). This integration enables not only dimensionality reduction but also the translation of analytical findings into actionable insights.

3.7 Synthesis and Research Gap

In summary, prior literature confirms that enrollment decisions are multidimensional and influenced by interconnected economic, psychosocial, and institutional factors. However, three key gaps remain evident: (1) limited integration of theoretical frameworks such as Student Choice Theory and Maslow's Hierarchy with empirical analysis, (2) insufficient use of structured data-analytic methods to reduce dimensionality and reveal latent decision structures, and (3) a lack of translation from statistical findings to strategic and policy-oriented applications. Addressing these gaps, the present study aims to systematically identify and consolidate key factors in enrollment decisions using PCA within a structured analytical framework, thereby bridging theoretical interpretation and practical decision support in higher education.

4. Conceptual Framework

This research framework integrates theoretical perspectives with data-driven analytics to systematically explain students' enrollment decision-making processes. (see Figure 1).

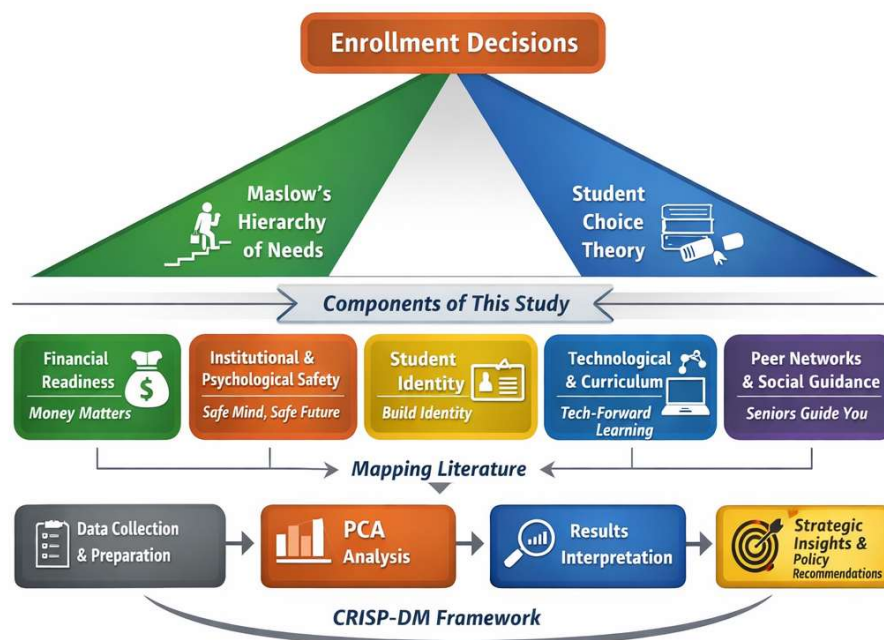


Figure 1. Research Framework

Note: Developed by the Authors and Graphic designed by ChatGPT

From Figure 1., it draws on Maslow's Hierarchy of Needs and Student Choice Theory to conceptualize the hierarchical structure of motivations and decision behaviors. Empirical data are processed within the CRISP-DM framework, encompassing data preparation, Principal Component Analysis (PCA) for analysis, and strategic interpretation of results. The factor extraction yields five core components: financial readiness; institutional safety and support; student identity formation; curriculum and technological modernity; and the role of student networks. Collectively, this framework establishes a coherent link between statistical evidence and higher education policy formulation, enabling institutions to translate analytical insights into strategic management actions in a rigorous, academically verifiable manner.

5. Research Methodology

5.1 Research Design

This study employs a quantitative, data-driven research design to identify and consolidate key factors influencing students' enrollment decisions in a Bachelor of Business Administration program. While the study follows the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework to ensure a systematic and replicable workflow—covering business understanding, data preparation, modeling, and evaluation—the core analytical technique is Principal Component Analysis (PCA), a multivariate statistical method appropriate for high-dimensional, intercorrelated educational data. In this context, CRISP-DM serves as a procedural framework, whereas PCA is used for dimensionality reduction and structural summarization rather than predictive data mining.

5.2 Population and Sample

The population comprised prospective students and individuals with the potential to apply for a Bachelor of Business Administration program within the study area. Data were collected via an online questionnaire (Google Forms) using purposive sampling to ensure alignment with the research objectives. A total of 571 valid responses were obtained from three provinces—Buriram, Surin, and Sisaket—representing the primary recruitment region of Rajamangala University of Technology Isan, Surin Campus.

The sample size meets established guidelines for factor/component analysis, which recommend at least 5–10 respondents per observed variable. With 45 observed variables, the achieved sample exceeds the minimum requirement, supporting stable component extraction and reliable estimation.

5.3 Research Instrument and Variables

Data were collected using a structured questionnaire developed from a comprehensive literature review on higher education enrollment behavior, educational marketing, Student Choice Theory, and Maslow's Hierarchy of Needs. The instrument employed a five-point Likert scale and included 45 observed variables (see Table 1.). To support international readability, the English version of all questionnaire items is provided alongside the original Thai items in Appendix A (and/or Bilingual Labels within Table 1.).

Content validity was assessed using the Index of Item–Objective Congruence (IOC) by three subject-matter experts, yielding values between 0.70 and 1.00, indicating strong alignment between items and constructs. The instrument was pilot-tested for clarity and consistency prior to full deployment.

Table 1. Question Items and Factor

No.	Question Items	Factor
1	Does your family have sufficient financial stability?	phy1
2	Does your family have enough savings for higher education expenses?	phy2
3	Does your family income adequately support your education?	phy3
4	Are educational expenses manageable for you?	phy4
5	Do you receive sufficient financial support?	phy5
6	Are tuition fees reasonable and affordable?	phy6
7	Are transportation costs manageable?	phy7
8	Is the institution located in a safe area?	safety8
9	Is the campus environment safe for students?	safety9
10	Are there adequate security systems in the institution?	safety10
11	Are academic support services sufficient?	safety11
12	Do you receive recommendations from teachers?	social12
13	Do your parents support your educational choices?	social13
14	Do your peers influence your decision?	social14
15	Are there activities that promote student interaction?	social15
16	Do alumni influence your decision?	social16
17	Is the program curriculum attractive?	esteem17
18	Is the institution well recognized?	esteem18
19	Do program rankings influence your decision?	esteem19
20	Does the institution offer diverse learning experiences?	esteem20
21	Does the institution have a good reputation?	esteem21
22	Does the curriculum align with current technology trends?	self22
23	Does the program enhance problem-solving skills?	self23
24	Does the institution support career opportunities?	self24
25	Does the curriculum align with labor market needs?	self25
26	Are modern technologies integrated into learning?	self26
27	Does the program promote innovation and creativity?	self27
28	Are learning resources sufficient?	self28
29	Are facilities modern and adequate?	self29
30	Are learning materials useful and up-to-date?	self30
31	Does the institution provide career guidance?	self31
32	Are extracurricular activities well organized?	self32
33	Is communication between students and staff effective?	self33
34	Are students supported in academic development?	self34
35	Are learning opportunities diverse?	self35
36	Are academic advising services available?	self36
37	Are internship opportunities provided?	self37
38	Does the institution support new student adaptation?	self38
39	Does the curriculum respond to societal needs?	self39

Table 1. Question Items and Factor (Cont.)

No.	Question Items	Factor
40	Are instructors knowledgeable and competent?	self40
41	Are collaborative learning activities encouraged?	self41
42	Are academic consultations accessible?	self42
43	Are learning resources regularly updated?	self43
44	Are student–faculty interactions encouraged?	self44
45	Does the institution provide career pathways after graduation?	self45

5.4 Data Collection Procedure

Data were collected online between February and March 2025 to ensure broad and efficient access to the target population. Prior to analysis, data quality checks were conducted. Responses with substantial incompleteness or patterned answering (e.g., Straight-Lining) were removed. The remaining dataset was screened to ensure adequacy for multivariate analysis.

5.5 Data Analysis Techniques

Data analysis followed a structured multi-stage procedure.

5.5.1 Data Preprocessing

1. Missing values: Cases with excessive missingness were excluded; remaining item-level missing values were minimal and handled using mean imputation due to the ordinal Likert structure and low missing rate.

2. Outlier detection: Multivariate outliers were examined using Mahalanobis distance ($p < .001$ Threshold), and influential cases were reviewed and removed where justified.

3. Assumptions and transformation: Variables were inspected for distributional anomalies; although PCA does not require strict normality, extreme skewness was mitigated where necessary (e.g., Winsorization). Variables were standardized (Z-scores) to ensure comparability and to prevent scale dominance.

5.5.2 Suitability for Component Analysis

The Kaiser–Meyer–Olkin (KMO) Measure of Sampling Adequacy and Bartlett’s Test of Sphericity were applied to evaluate factorability. The KMO statistic assesses the proportion of common variance among variables relative to partial correlations; values above 0.60 indicate adequacy, with values above 0.90 considered “excellent” (Kaiser, 1974; Hair et al., 2019). Bartlett’s Test assesses whether the correlation matrix differs significantly from the identity matrix, indicating sufficient intercorrelations for extraction.

5.5.3 Principal Component Analysis (PCA)

PCA was employed to reduce dimensionality and summarize the variance structure of the dataset (Jolliffe & Cadima, 2016). Components were retained based on eigenvalues greater than 1. To enhance interpretability for managerial applications, an orthogonal Varimax rotation was applied, maximizing variance of loadings within components and minimizing cross-loadings. A loading threshold of ≥ 0.50 was adopted to indicate meaningful associations between variables and components. The total variance explained and rotated component matrix were examined to ensure a stable and interpretable solution.

5.6 Strategic Application of Analytical Results

The extracted components provide empirically grounded dimensions for decision support in higher education management. Specifically, the results inform (i) targeted communication strategies aligned with student segments, (ii) curriculum development responsive to technological relevance and identity formation, and (iii) enrollment policy design addressing financial and psychosocial needs. By embedding PCA within a structured CRISP-DM workflow, the study ensures that statistical findings are translated into actionable insights, enhancing institutional competitiveness and evidence-based planning.

6. Results

This section presents the empirical findings in alignment with the three research objectives. The results are reported in a structured, step-by-step manner, with a focus on statistical procedures and numerical outcomes. Interpretations and theoretical implications are intentionally reserved for the Discussion section.

6.1 Identification of the Principal Components Influencing Students' Decisions to Enroll in the Bachelor of Business Administration Program

To address the first research objective, the dataset's suitability for component extraction was evaluated prior to analysis. The Kaiser–Meyer–Olkin (KMO) Measure of Sampling Adequacy yielded a value of 0.963, exceeding the recommended threshold of 0.60 and indicating excellent sampling adequacy.

Bartlett's Test of Sphericity was statistically significant ($\chi^2 = 12,840.517$, $df = 990$, $p < .001$), confirming that the correlation matrix significantly differed from an identity matrix.

Additionally, individual Measures of Sampling Adequacy (MSA) for all variables exceeded 0.90, indicating that all observed variables were suitable for inclusion and no variables required removal.

These results confirm that the dataset satisfies the statistical assumptions required for Principal Component Analysis and is appropriate for dimensionality reduction (see Table 2.).

Table 2. Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) & Bartlett's Test of Sphericity

KMO and Bartlett's Test		
Kaiser–Meyer–Olkin Measure of Sampling Adequacy.		0.963
Bartlett's Test of Sphericity	Approx. Chi-Square	12840.517
	df	990
	Sig.	.000

6.2 Classification and Dimensionality Reduction of Enrollment Decision Factors into Statistically Significant Components

Principal Component Analysis (PCA) was applied to 45 observed variables to reduce dimensionality and identify significant factor groupings. Components were retained based on the eigenvalue > 1 criterion, and Varimax rotation was used to enhance interpretability. A factor loading threshold of ≥ 0.50 was applied, and variables below this threshold were excluded from the final structure.

The analysis resulted in five principal components comprising 23 retained variables. All retained variables exhibited loadings ranging from 0.500 to 0.717, indicating strong associations with their respective components.

The five components collectively explained 50.549% of the total variance, demonstrating that the reduced structure preserved a substantial proportion of the dataset's variability. The first two components accounted for the largest shares of variance (13.60% and 13.40%), while the remaining components contributed between 5.80% and 8.90%.

The rotated component matrix revealed a clear structure: Component 1 (7 Variables), Component 2 (8 Variables), Component 3 (4 Variables), Component 4 (2 Variables), and Component 5 (2 Variables). These findings indicate that high-dimensional enrollment data can be effectively reduced into a smaller set of statistically meaningful components.

Table 3. Component Loadings Matrix Obtained from the Varimax-Rotated Solution

Component 1	Component 2	Component 3	Component 4	Component 5
self34 (0.651)	self42 (0.598)	phy3 (0.712)	esteem17 (0.663)	social12 (0.717)
self26 (0.591)	esteem21 (0.557)	phy1 (0.674)	self22 (0.658)	safety8 (0.564)
self41 (0.561)	self28 (0.550)	phy2 (0.663)		
self33 (0.532)	safety10 (0.546)	phy4 (0.662)		
safety11 (0.513)	self43 (0.525)			
self37 (0.506)	social14 (0.515)			
esteem19 (0.500)	esteem20 (0.512)			

Note. Varimax rotation was applied. Only factor loadings ≥ 0.50 are reported.

Table 3. presents the Varimax-rotated component loadings derived from Principal Component Analysis (PCA) applied to 45 observed variables. Components were retained based on the eigenvalue-greater-than-one criterion, and variables with factor loadings below 0.50 were excluded. The final solution consists of five principal components comprising 23 retained variables, indicating a clear and interpretable factor structure.

Table 4. Component Statistics and Explained Variance

Component	SS Loadings	% of Variance	Cumulative %
1	6.130	13.6	13.60
2	6.050	13.4	27.10
3	4.000	8.9	36.00
4	3.820	8.5	44.40
5	2.590	5.8	50.20

Table 4. presents the five principal components were retained based on their eigenvalues, collectively explaining 50.20% of the total variance. Component 1 accounts for the highest variance at 13.6%, followed closely by Component 2 at 13.4%, indicating their dominant contribution to the model. Components 3 to 5 contribute moderate proportions of variance, ranging from 5.8% to 8.9%, reflecting their complementary roles. The cumulative variance increases progressively across components, demonstrating an effective reduction of dimensionality while preserving a substantial portion of the dataset's information.

Table 5. Kaiser–Meyer–Olkin (KMO) Measure of Sampling Adequacy

Question Item	MSA	Question Item	MSA
phy1	0.943	Esteem21	0.934
phy3	0.941	self22	0.961
phy4	0.938	self23	0.949
Phy5	0.940	self24	0.931
Phy7	0.931	self37	0.928
safety8	0.933	self38	0.946
Safety9	0.946	self40	0.959
social12	0.939	self41	0.947
social14	0.951	self42	0.949
esteem16	0.928	self43	0.949
esteem17	0.967	self44	0.974
esteem19	0.935	self45	0.973
Overall = 0.963			

From Table 5., the findings further confirm the robustness of the dataset for PCA. The overall KMO value remained at 0.963, which is classified as “excellent,” while the individual Measures of Sampling Adequacy (MSA) for all variables exceeded 0.90. Collectively, these results indicate that all questionnaire items were of high quality and that no variables required removal from the analysis.

Table 6. Total Variance Explained

Component	Initial Eigenvalues		Extraction Sums of Squared Loadings		Rotation Sums of Squared Loadings	
	Total	% of Variance	Total	% of Variance	Total	% of Variance (Cumulative)
1	17.567	39.038	17.567	39.038	6.153	13.674
2	1.564	3.476	1.564	3.476	6.056	27.133
3	1.444	3.210	1.444	3.210	4.051	36.135
4	1.121	2.491	1.121	2.491	3.867	44.728
5	1.050	2.334	1.050	2.334	2.622	50.549

From Table 6., shows that five components with eigenvalues greater than one were retained, collectively accounting for 50.549% of the total variance. The Extraction Sums of Squared Loadings confirm that these retained components preserved the same level of explanatory power. After Varimax rotation, the variance was more evenly distributed across components, thereby enhancing interpretability while maintaining the overall cumulative explained variance.

Interpretation of Factor Extraction Results found that the extracted components provide a clear, system-level representation of students’ enrollment decision-making processes were as follows:

Component 1: Build Identity reflects institutional learning experiences and student identity formation through engagement, interaction, skill development, and institutional attachment.

Component 2: Safe Mind, Safe Future highlights the importance of psychological support, safety, recognition, and a caring institutional environment.

Component 3: Money Matters captures financial readiness and affordability as prerequisite conditions for enrollment decisions.

Component 4: Tech-Forward Learning emphasizes curriculum reputation and technological relevance as signals of future opportunity and social value.

Component 5: Seniors Guide You underscores the role of peer networks and perceived safety in reducing uncertainty during the decision process.

Evaluation of PCA and CRISP-DM: The evaluation of the analytical process within the CRISP-DM framework demonstrates that PCA was appropriately and effectively applied to behavioral educational data. During the Data Understanding and Data Preparation phases, questionnaire data were systematically transformed and validated, ensuring readiness for factor extraction. In the Modeling phase, PCA with Varimax rotation successfully reduced dimensionality while producing a stable and interpretable factor structure with minimal cross-loadings.

During the Evaluation phase, the extracted components collectively explained students' enrollment decisions across economic, social, psychological, and institutional dimensions, aligning well with the study's objectives and broader higher education policy contexts. These findings confirm that PCA, when embedded within the CRISP-DM framework, offers not only statistical rigor but also strong interpretive and strategic value.

6.3 The Data-Driven Analytical Findings Enabled the Development of Targeted Communication Strategies Tailored to Students' Needs

Regarding the third research objective, the extracted components provide a solid empirical basis for translating analytical findings into targeted communication strategies aligned with students' needs. Each principal component corresponds to a distinct decision dimension that can be directly mapped onto institutional messaging and policy design. Financial readiness underscores the importance of communicating affordability, scholarships, and financial support mechanisms, while institutional support and safety highlight the strategic value of emphasizing student well-being, counseling services, and secure learning environments. Components related to identity formation and technological relevance suggest that students increasingly respond to narratives centered on personal development, meaningful engagement, and future-oriented competencies. Additionally, the role of peer guidance indicates that student networks and senior mentorship function as effective mechanisms for reducing uncertainty during the enrollment process. Methodologically, embedding PCA within the CRISP-DM framework ensures that these insights emerge from a structured, transparent, and reproducible analytical workflow. Consequently, the findings not only satisfy the third objective but also demonstrate how data-driven analysis can inform evidence-based communication and enrollment strategies in a competitive higher education context.

7. Conclusion

This study employed a quantitative, data-driven approach to investigate the key factors influencing students' enrollment decisions in a Bachelor of Business Administration program. Principal Component Analysis (PCA) was applied to 45 observed variables to perform dimensionality reduction and identify underlying factor structures. Components were retained based on the eigenvalue-greater-than-one criterion, and variables with factor loadings below 0.50 were excluded from the analysis. As a result, five

principal components were extracted, comprising a total of 23 retained variables. The dataset was confirmed to be highly suitable for analysis ($KMO = 0.963$, $p < .001$), and the extracted components collectively explained 50.549% of the total variance.

The findings demonstrate that enrollment decisions can be represented through a structured set of interrelated factors derived from high-dimensional data. The components are presented according to their relative contribution to explained variance, with the first two components accounting for the largest proportions, followed by the remaining components. This ordering reflects their statistical importance within the model and enhances the transparency of the analytical results.

Methodologically, integrating PCA within the CRISP-DM framework strengthens analytical rigor by ensuring a systematic and reproducible process from data preparation to model evaluation. The resulting component structure provides a coherent and interpretable representation of enrollment decision factors, offering a robust empirical foundation for further interpretation and application in higher education research and strategic planning contexts.

8. Discussion

The results of this study provide empirical support for all three research objectives and offer a structured interpretation of students' enrollment decision-making processes. First, the identification of five principal components confirms that enrollment decisions are not determined by isolated variables but by multidimensional clusters of interrelated factors. This finding is consistent with prior research suggesting that higher education choice is influenced by a combination of economic, psychosocial, and institutional dimensions rather than purely academic considerations (Hemsley-Brown & Oplatka, 2016; Marginson, 2016; Azzone & Soncin, 2020).

Second, the dimensionality reduction achieved through PCA demonstrates that complex decision-related variables can be systematically grouped into a smaller number of statistically meaningful components. The retained five-component structure, explaining over 50% of the total variance, indicates that the extracted factors capture a substantial portion of the variability in students' decision patterns. Importantly, the extracted components are presented according to their relative contribution to the explained variance, with Component 1 and Component 2 accounting for the largest proportions, followed by Components 3, 4, and 5. This ordering is derived from the explained variance obtained through PCA rather than subjective interpretation, thereby enhancing the transparency and interpretability of the findings. These findings are consistent with Jolliffe and Cadima (2016), who emphasized that PCA is particularly effective for reducing data dimensionality while maintaining meaningful structural relationships among variables. Similarly, Romero and Ventura (2020) argued that dimensionality reduction techniques are increasingly valuable in Educational Data Mining because they simplify complex educational datasets and facilitate knowledge discovery without compromising analytical quality. Furthermore, the identified components are partially consistent with the findings of Kanjanasamranwong et al. (2017), who reported that institutional quality and reputation, institutional environment, social factors, and economic and financial factors were among the key determinants of students' decisions to enroll in higher education institutions.

From a theoretical perspective, the findings align with both Student Choice Theory and Maslow's Hierarchy of Needs (Maslow, 1943). The components associated with financial readiness and institutional safety reflect foundational decision conditions, consistent with cost-related filtering mechanisms and

lower-level needs. In contrast, components related to identity formation, peer influence, and technological relevance correspond to higher-order motivations, including belonging, esteem, and self-development. However, rather than implying a strictly linear hierarchy, the results suggest that these factors operate in an interconnected and context-dependent manner.

Third, the application of the extracted components demonstrates how data-driven analytical results can support decision-making processes. The transformation of high-dimensional survey data into interpretable structures enables clearer categorization of enrollment factors, facilitating their use in strategic planning and communication design. Unlike previous studies that relied primarily on descriptive analysis (Santisirikul et al., 2023; Lertsahapan & Serttasangsri, 2020), most applications remain descriptive, focusing on variable grouping without linking findings to theoretical frameworks or strategic implications. However, the present study employs a data-driven analytical approach, providing a more systematic and reproducible basis for understanding student behavior. These findings are consistent with previous research emphasizing the value of data-driven analytical techniques for identifying latent decision structures and supporting evidence-based decision-making in higher education (Jolliffe & Cadima, 2016; Romero & Ventura, 2020).

Overall, the findings highlight the importance of integrating statistical analysis with theoretical frameworks to capture the complexity of enrollment decisions. The results contribute to the literature by offering both methodological clarity and a structured representation of decision factors, while also providing a foundation for evidence-based applications in higher education contexts.

9. Recommendations

The findings offer several implications for higher education institutions seeking to improve enrollment management. The results indicate that students' decisions are influenced by a combination of financial, institutional, social, and personal factors rather than academic considerations alone. Consequently, universities should consider adopting a more student-centered approach by strengthening financial assistance, support services, opportunities for personal development, technology-enhanced learning environments, and peer-support initiatives. Such efforts may help institutions better respond to students' expectations and enhance their overall attractiveness.

The extracted components also provide a useful evidence base for strategic planning and communication. By understanding the factors that shape enrollment decisions, institutions can develop more targeted communication strategies, improve curriculum relevance, and design support systems that encourage student engagement and institutional commitment.

Several limitations should be acknowledged. The study was conducted within a specific geographical context using purposive sampling, which may restrict the generalizability of the findings. In addition, PCA is intended to identify latent structures and does not establish causal relationships among variables.

Future research should examine broader educational contexts and employ complementary analytical approaches. Techniques such as Association Rule Mining may reveal hidden patterns among enrollment factors, while Structural Equation Modeling can be used to investigate causal relationships and validate the component structure identified in this study.

10. Declaration of Generative AI

The authors declare that generative artificial intelligence was used exclusively for linguistic assistance, including grammar checking, wording improvement, and literature review. No AI tools were involved in data analysis, interpretation, or the generation of research findings. The authors retain full responsibility for the accuracy, integrity, and originality of the work.

11. CRediT Author Statement

Umaporn Chaisoong: Conceptualization, Methodology, Formal Analysis, Writing-Original Draft, Writing-Review and Editing, Supervision, and Project Administration.

Thitimaporn Waenphet: Data Curation, Investigation, Visualization.

Nonthakan Nganjaturas: Data Curation, Investigation, Visualization.

Jandara Suksam: Conceptualization, Methodology, Formal Analysis, Writing-Original Draft.

Thanet Yothasiri: Data Curation, Investigation, Visualization.

Waraluck Maprasom: Conceptualization, Methodology, Formal Analysis, Writing-Original Draft.

The author has read and approved the final manuscript.

12. References

- Azzone, G., & Soncin, M. (2020). Factors Driving University Choice: A Principal Component Analysis on Italian Institutions. *Studies in Higher Education*, 45(12), 2426–2438. <https://doi.org/10.1080/03075079.2019.1612354>
- Chaisoong, U., & Tirakoat, S. (2020). The Clustering of Questions Affect to Tourist's Decision Making for Chatbot Design. *2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, 784–787. Institute of Electrical and Electronics Engineers. <https://doi.org/10.1109/ecti-con49241.2020.9158069>
- Daniel, B. (2015). Big Data and Analytics in Higher Education: Opportunities and Challenges. *British Journal of Educational Technology*, 46(5), 904–920. <https://doi.org/10.1111/bjet.12230>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning.
- Hemsley-Brown, J., & Oplatka, I. (2016). *Higher Education Consumer Choice*. Palgrave Macmillan. <https://doi.org/10.1007/978-1-137-49720-8>
- Hossler, D., & Gallagher, K. S. (1987). Studying Student College Choice: A Three-Phase Model and the Implications for Policymakers. *College and University*, 62(3), 207–221.
- Hossler, D., Ziskin, M., Gross, J. P. K., Kim, S., & Cekic, O. (2009). Student Aid and Its Role in Encouraging Persistence. In Smart, J. C. (Ed.), *Higher Education: Handbook of Theory and Research* (24, 389–425). Springer. https://doi.org/10.1007/978-1-4020-9628-0_10
- Jolliffe, I. T., & Cadima, J. (2016). Principal Component Analysis: A Review and Recent Developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202. <https://doi.org/10.1098/rsta.2015.0202>
- Kaiser, H. F. (1974). An Index of Factorial Simplicity. *Psychometrika*, 39(1), 31–36. <https://doi.org/10.1007/bf02291575>

- Kanjanasamranwong, P., Urawong, K., Dam, R. M. A., & Dam, S. M. A. (2017). The Opportunities Forecast Model in Choosing to Attend the Prince of Songkhla University of Grade 12 Students, Songkhla Province. *Journal of Education Naresuan University*, 19(2), 12-24. (In Thai)
- Kuh, G. D., Kinzie, J., Schuh, J. H., & Whitt, E. J. (2010). *Student Success in College: Creating Conditions that Matter*. Jossey-Bass.
- Lee, T. A., Galway, G. J., Bdair, O. M., Sundly, A., Coombs, S. D., & Kennie-Kaulbach, N. R. (2026). Exploratory Factor Analysis of Decision-Making Domains Influencing Students' Pursuit of the PharmD Degree at 2 Canadian Universities. *American Journal of Pharmaceutical Education*, 90(7), 102005. <https://doi.org/10.1016/j.ajpe.2026.102005>
- Lertsahapan, S., & Serttasangri, C. (2020). A Study of Decision Making Factors Education in Private Vocational Institution Udon Thani Province. *Santapol College Academic Journal*, 6(2), 95-102. (In Thai)
- Marginson, S. (2016). *Higher Education and the Common Good*. MUP Academic.
- Maslow, A. H. (1943). A Theory of Human Motivation. *Psychological Review*, 50(4), 370-396. <https://doi.org/10.1037/h0054346>
- OECD. (2021). *Education at a Glance 2021: OECD indicators*. OECD Publishing. <https://doi.org/10.1787/b35a14e5-en>
- Paulsen, M. B. (2001). The Economics of Human Capital and Investment in Higher Education. In Paulsen, M. B., & Smart, J. C. (Eds.), *The Finance of Higher Education: Theory, Research, Policy, and Practice* (55-94). Agathon Press.
- Provost, F., & Fawcett, T. (2013). *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. O'Reilly Media.
- Romero, C., & Ventura, S. (2020). Educational Data Mining and Learning Analytics: An Updated Survey. *WIREs Data Mining and Knowledge Discovery*, 10(3), e1355. <https://doi.org/10.1002/widm.1355>
- Santisirikul, B., Arreerard, T., & Arreerard, W. (2023). Affecting Factors on the Choice of Pursuing a Bachelor's Degree Using Exploratory Factor Analysis. *Journal of Research and Development Institute, Rajabhat Maha Sarakham University*, 10(3), 549-566. (In Thai)
- Somyoonsab, N. (2017). The Influencing Promotional Strategies towards the Decision to Study in Private University, International Programs. *Journal of Humanities and Social Sciences, Rajapruk University*, 2(3), 73-83. (In Thai)