

Segmentation: As an Effort to Optimize the Number of New Student applicants at Private Universities

¹Totok Wibisono, ²DC. Kuswardani, ³Tri Endang Yani, ⁴Wawan Setiawan

^{1,2,3,4} Program Studi Manajemen, Fakultas Ekonomi, Universitas Semarang, Indonesia

E-mail: totokwibisono69@gmail.com

ABSTRACT

The purpose of this study is to cluster (segment) users of higher education services at FE-USM to develop effective communication so as to increase the number of users of higher education services. The research method used was descriptive research, using convenience sampling with a sample size of 373 respondents from a population of 5,589. Data collection was conducted using a survey via Google Forms, and data processing was performed using the K-Means clustering method with SPSS. This study resulted in three clusters with the following membership: cluster 1 with 11 members, cluster 2 with 259 members, and cluster 3 with 3 members, where the cluster-forming variables were influenced by six variables, initially using eight variables. The six (6) variables that formed the clusters were age, occupation, marital status, income, reasons for choosing a university, and reasons for choosing a study program. The communication media that need to be developed are word of mouth, websites and social media, as well as brochures and banners with content tailored to each cluster.

Keywords: Segmentation, Higher Education Institution, Marketing Communication Strategy, K-means

INTRODUCTION

Market segmentation is a crucial strategy for universities to implement in order to ensure their survival amid increasing competition among universities to find users of available university services (Vázquez et al., 2023). Market segmentation has been widely practiced by companies in the tourism industry (Hajibaba et al., 2020), the product industry, the transportation industry (Soto et al., 2021) and others, energy (Gazdecki Michał and Goryńska-Goldmann et al., 2021), Suppliers (Rahiminia et al., 2023).

Research on segmentation (market grouping) for universities has been conducted by researchers in an effort to improve marketing efforts so that the universities concerned can optimize their resources to be used effectively in recruiting and graduating their students. This research has been conducted both abroad and in Indonesia, for example, Student Choice Criteria for Selecting an Indonesian Understanding University: Conjoint Analysis Approach (Kusumawati 2011), Identifying Market Segment of University Student Using Cluster Analysis Based on Psychographics Variable (Khusnuliawati & Putri, 2021), and The Potential of Promotional Media as Marketing Instrument for Open and Distance Higher Education: Segmentation Study (Budiman et al., 2023).

From the research that has been

conducted, several research objectives have been achieved: first, grouping student segmentation to predict performance and failure in completing studies (more evaluative in nature to measure the performance of the university (Davari et al., 2019), second, conducting student segmentation merely as an application of segmentation analysis in higher education institutions (Miguéis et al., 2018) and third, segmenting students to develop higher education strategies for more effective student recruitment amid high competition (Wilkins, 2020).

Semarang University is one of the universities in Central Java within the scope of LLDIKTI VI, located in the city of Semarang. It was established in June 1987, starting with higher education at the POLYTECHNIC under the name Politeknik Semarang, and in 1994 it changed to higher education at the university level under the name Universitas Semarang. At the time of its establishment, it offered four vocational D3 programs, namely Clerical Studies, Business Management, Civil Engineering, and Agricultural Technology. Currently, Semarang University offers 16 programs, namely D3 Business Management, S1-Law, S1-Management, S1-Accounting, S1-Civil Engineering, S1-Electrical Engineering, S1-Regional and Urban Planning, S1-Agricultural Product Technology, S1-Psychology, S1-Computer Engineering, S1-Communication Science, S1-Tourism, S2-Management, S2-Law,

and S2-Psychology, with a total of over 16,511 students who have chosen to pursue their studies at Semarang University.

The Faculty of Economics at Semarang University, as part of Semarang University, offers a Diploma 3 program in Business Management (vocational), a Bachelor's degree program in Management, and a Bachelor's degree program in Accounting. It is one of the faculties with the highest number of students at Semarang University, with a total of 5,589 active students currently enrolled in these three study programs.

In the last decade, competition among higher education institutions in the city of Semarang has increased, as evidenced by the growth of higher education infrastructure in nearby locations such as UNIMUS, UPGRIS, STIE BPD, and UNISULA. This illustrates the high level of competition to attract students to use these higher education services, with strategic location being the main consideration in choosing a higher education institution.

Therefore, it is necessary to conduct a study (research) on the segmentation of students in the Faculty of Economics, University of Semarang, with the following considerations: first, the Faculty of Economics, USM, contributes the most to USM; second, there is high competition to attract students to use the available study programs; third, USM has never conducted this study as an effort to optimize the use of available resources, and fourth, previous studies have examined university segmentation for full-time students, while the current study examines university segmentation for part-time students.

METHOD

This research is descriptive quantitative research, the results of which are to produce a student segment profile and the appropriate communication mix for prospective students. In this study, the data used is primary data obtained through a survey with a population of part-time students (regular employee classes at the Faculty of Economics, University of Semarang). The population in this study is part-time students (employee classes) who are currently enrolled in the learning program at the Faculty of Economics, Semarang University. The current population is 5,589 active students (source: FE-USM database). The sampling technique used in this study is Convenience Sampling, which involves distributing questionnaires to students after they finish attending classes to respond to the questionnaires provided. The sample size to be taken in this study is calculated using the Slovin formula, which is:

$$n = \frac{N}{1 + Ne^2}$$

n is the sample size

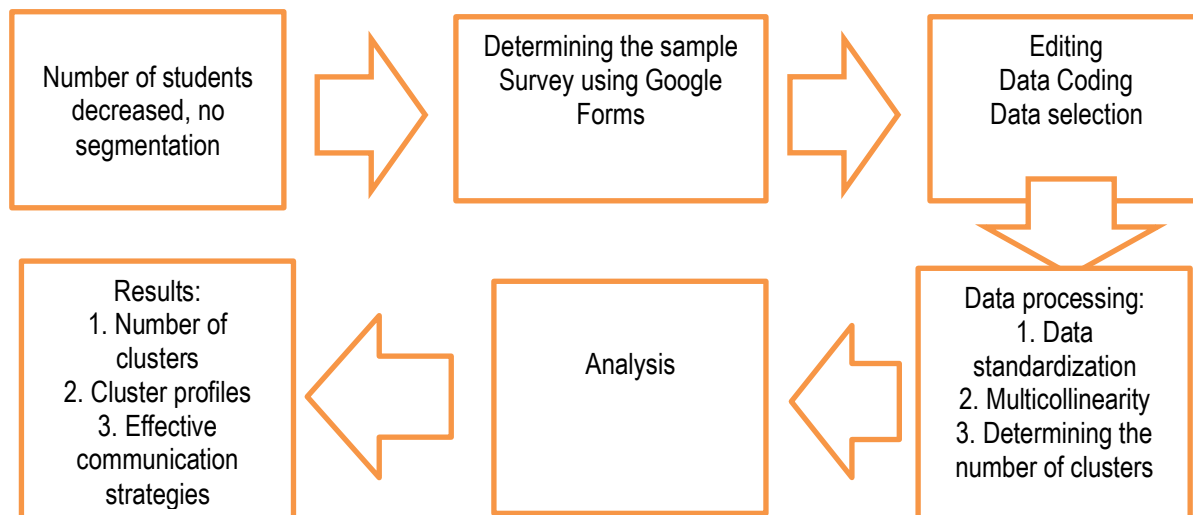
N is the population size (5,589)

E is the tolerance limit (5%)

Therefore, the sample size is $5,589 / 1 + (5,589 \times 0.05^2) = 373$

The attributes to be tested in the formation of clustering or segmentation include 1. Student profiles, including: gender, age, marital status, occupation, and income. 2. Motivational attributes considered in choosing a university as a place of study. 3. Rational attributes considered in choosing a university. The collected data was processed using SPSS with Cluster Analysis, namely using the K-Mean algorithm to produce Clustering.

Figure. Research Stages



RESULT and DICUSSION

The planned sample size for this study was 373 samples. The researcher distributed questionnaires to 497 respondents, which should be sufficient to distribute questionnaires equal to the sample size plus 10%, meaning 373 (1+10%) equals 373 + 37, so the total number of questionnaires distributed is sufficient at 410.

The researcher used convenience sampling to collect data from active students who took courses offered at each semester level in each employee study program, namely D3 Business Management, S1-Accounting, and S1-Management. The researcher then distributed the

questionnaires through the lecturers teaching the relevant courses (as enumerators) and the questionnaires were given to the respondents at the end of the first lecture (usually the lecture contract material) through a Google form uploaded to the WhatsApp group for the course, with the hope of not disrupting the lecture. Out of the 495 responses obtained, after confirmation with the faculty as the organizer by matching the names and student ID numbers listed, the 495 responses were reduced to 475 responses. After verification regarding the completeness of the responses, the sample size became 373 valid responses. For more details, see the table below (Table 1).

Table 1. Description of the Processed Sample

Questionnaire Distributed	Questionnaire Returned	Questionnaire verified	Questionnaire Valid
495	495	475	373

Respondent Description

The Regular Evening D3-Business Management program had 6.7% (25 respondents), the Regular Evening S1-Accounting program had 29% (108 respondents), and the Regular Evening S1-Management program had 48.3% (180

respondents), while for the Employee class, the D3-Business Management Week-End program had 0.8% (3 respondents), the S1-Accounting Week-End program had 3.8% (14 respondents), and the S1-Management Week-End program had 11.5% (43 respondents)

Table 2. Respondent Description

Program	Frequency	Percent	Valid percentage
Diploma 3 Afternoon	25	6.7	6.7
Bachelor's Degree in Accounting Evening	108	29.0	29.0
Bachelor of Management (Evening)	180	48.3	48.3
Diploma 3 Weekend	3	.8	.8
Bachelor's Degree in Accounting Weekend	14	3.8	3.8
Bachelor's Degree in Management Weekend	43	11.5	11.5
Total	373	100.0	100.0

In this study, the important variables to be processed include: Respondent Name, Respondent Student ID Number to determine the respondent's study program, gender, age, marital status, occupation, income/earnings, residential status, and tuition payer. These are the socio-demographic variables of the respondents to describe their socio-demographic segmentation, as well as behavioral and psychographic variables (Kotler, 2012) or rational variables and motivational variables (Angulo et al., 2010) which are variables considered in choosing a university and study program.

Data Standardization

As the data obtained from the survey,

after tabulation and coding of the data obtained from all variables, it turned out that the sizes of the variables differed from one variable to another, so the researcher standardized the data using descriptive statistics to calculate the Z-score for each variable, including: age variable, Gender, Marital Status, Tuition Costs, Income, Residence Status, Occupation, Main Considerations for Choosing a University, and Considerations for Choosing a Study Program. With this data standardization, the variables could be compared efficiently (Zhang, 2025).

Multicollinearity

After standardizing the data, the next step is to test for multicollinearity by correlating the

bivariate data to determine whether the variables used contain multicollinearity by measuring the level of correlation between variables. If the correlation is greater than 0.6 and/or close to 1, it means that the data contains multicollinearity and the variables that have multicollinearity must be eliminated according to the researcher's consideration. It turns out that the data does not have multicollinearity, so it can proceed to the next stage.

Determining the Number of Clusters

To determine the ideal number of clusters, the researcher determines the agglomeration schedule through hierarchical clustering to obtain the agglomeration schedule.

From the Agglomeration table, simply read the coefficient section to see how the coefficient changes from each stage until finding the coefficient with the largest change. In the table, at stages 368 to 370, there was a significant change in the coefficient, which initially changed only from 0.043 to 0.246 at stage 368 to stage 369, where the coefficient increased to 0.967, then at stage 370, it increased to (1.384). This increase indicates an extraordinary change in the coefficient from stage 367 to stage 370. Therefore, the number of clusters is determined based on the difference between stage 367 and stage 370, meaning there are 3 clusters.

Table 3. Agglomeration Schedule Table

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
361	21	43	5,12	345	320	362
362	3	21	5,299	360	361	366
363	2	206	5,377	359	0	368
364	8	295	5,573	353	0	366
365	26	334	6,488	0	0	370
366	3	8	6,525	362	364	367
367	3	238	6,682	366	342	369
368	2	241	6,844	363	0	369
369	2	3	7,809	368	367	370
370	2	26	9,193	369	365	0

Source : SPSS output

Cluster Member Determination

The basis for determining the K-Mean algorithm cluster is by determining the proximity distance using Euclidean Distance with the following formula $D(a,b) =$

$$\sqrt{\sum_{k=1}^n (a_k - b_k)^2}$$

In determining the clusters, the researcher used SPSS version 18, which produced the following output, as shown in Table 4:

Table 4. Number of Cluster Members

Cluster Name	Number of Members
Cluster 1	11
Cluster 2	358
Cluster 3	2
Valid	371

Source: SPSS output

Determining Cluster Components

In determining cluster components,

researchers used varimax factor analysis, which produced the output shown in the table 5.below:

Table 5. Rotated Component Matrix ^a

	Component		
	1	2	3
Zscore: student gender	..172	..366	.287
Zscore: age of respondents	-.798	-.322	.030

Zscore: respondent's marital status	..873	-.037	-.015
Zscore: student employment	..054	..749	-.052
Zscore: main consideration in choosing a company	-.214	..113	. .611
Zscore: tuition costs covered	..032	..812	-.064
Zscore: consideration of choosing a study program	..032	..026	-.606
Zscore: residential status	..268	-.151	.517

Source: SPSS output

Table 5 shows that component 1 is formed by the age of respondents (with an absolute value of 0.798) and marital status (absolute value of 0.873). Component 2 is formed by the occupation of students (with an absolute value of 0.749) and tuition fees borne by (with an absolute value of 0.812). For component 3, the formers are consideration in choosing a university

(absolute value 0.611) and consideration in choosing a study program (0.606).

Clusters Graphically

In the graphical explanation, the researcher must determine the magnitude of the influence of these components in determining clusters with factor analysis, the output of which is as follows:

Table 6. Total Variance Explained

Componen	Initial eigenvalues			Etraction Sum of square loading			Rotation Sums of Square Loading		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.814	22.673	22.673	1,814	22,673	22,673	1,551	19,383	19,383
2	1,233	15,410	38,083	1,233	15,410	38,083	1,495	18,682	38,065
3	1,097	13,718	51,801	1,097	13,718	51,801	1,099	13,736	51,801
4	.986	12,322	64,123						
5	.922	11,522	75,645						
6	.902	11,279	86,924						
7	.640	7,997	94,921						
8	.406	5,079	100,000						

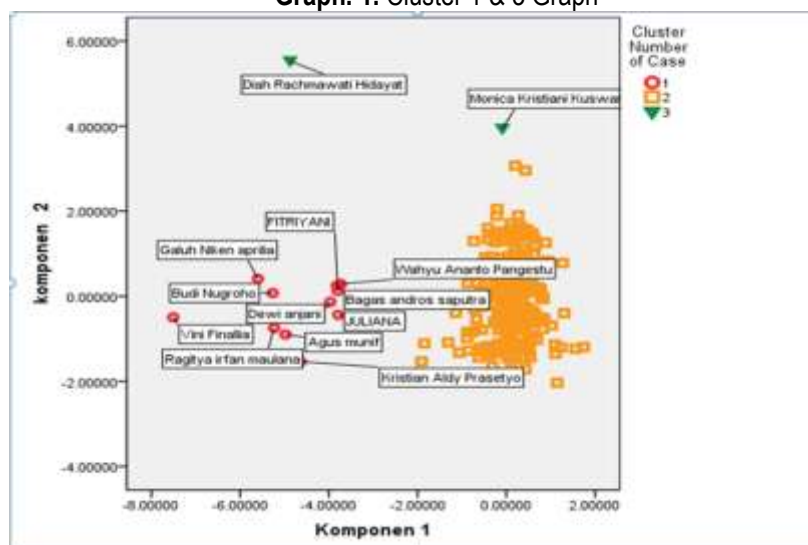
Source: SPSS Analysis

From Table 6, we observe that in the component column, three components are formed, but only two components are used, namely component 1 as the x-axis and component 2 as

the y-axis, both of which contribute to explaining 38.083% of the clusters.

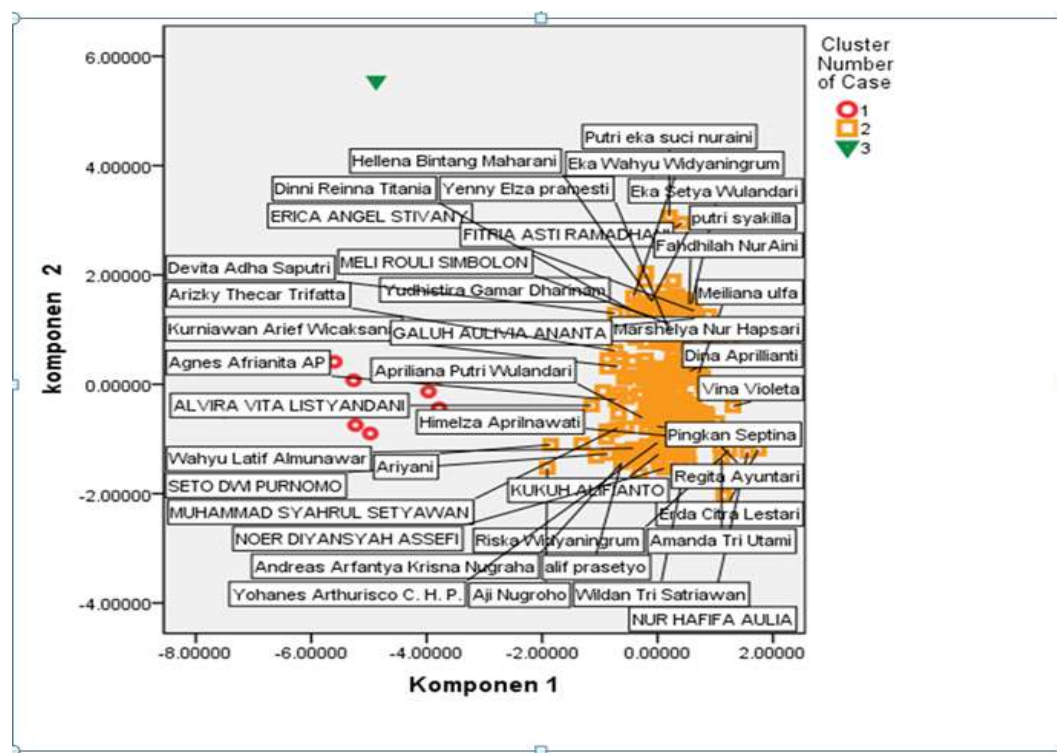
This can be visually observed in the graph below (Graph 1)

Graph. 1. Cluster 1 & 3 Graph



Source : SPSS output

Graph 2. Cluster 2 Graph



Source : SPSS output

Description of Each Cluster

Table 7. Description of Each Cluster

Variable Name	Cluster 1	Cluster 2	Cluster 3
Respondent Age	23 - 40	18 - 25	20 - 24
Marital Status	Married	Not married	Married / not married
Student	Private employee/	Private employee	Private employee
Employee	Entrepreneur		
Tuition	Self	Parents/Self/Combined	Sponsor
Choosing	Accreditation, Affordable	Affordable, Accreditation	Accreditation,
A university		Location	Affordable
Choosing a Major	Suitable for your field Of study, Suitable for your career	Suitable for your field of study	Suitable for your field of study

Source : SPSS Output

From graph 1, it can be seen that cluster 1 has 11 members who have the following characteristics (see Table 7): Respondents aged 23-40 years (adults), married, employed in the private sector and self-employed, self-funded their studies, and chose a university based on institutional accreditation and affordability of tuition fees, while choosing a major based on their field of interest and career goals. For Cluster 2, which has 357 members (see Figure 2), the characteristics

are (see Table 7): ages 18-25 (youngest) with marital status unmarried, employed as private sector employees, partially self-funded tuition, and most parents still involved in funding, choosing a university based on affordability, accreditation, and location, while choosing a major because it aligns with their field of interest. Meanwhile, Cluster 3 has 2 members (see graph 1) with the following characteristics (see table 7): aged 20-25 years (middle age) with marital status

married/unmarried, employed in the private sector, studying with sponsorship funding, with the reason for choosing the university based on accreditation

and affordability of tuition fees, while choosing the major because it aligns with their field of study.

Information Media Used in Each Cluster

Table 8. Communication Media for Cluster 1

Media Name	Frequency	Percent	Valid Percent	Cumulative Percent
Friends/relatives	9	81.8	81.8	81.8
Website	1	9.1	9.1	90.9
Brochure/banner	1	9.1	9.1	100.0
Total	11	100.0	100.0	

Source : SPSS output

In cluster 1, 81.8% of the media used was information from friends/relatives (word of mouth) and websites and banners, each accounting for 9.1%.

Table 9. Communication media in cluster 2

Media	Frequency	Percent	Valid Percent	Cumulative Percent
Friends	227	63.6	63.6	63.6
Website	77	21.6	21.6	85.2
3.00	1	.3	.3	85.4
Instagram	5	1.4	1.4	86.8
5.00	1	.3	.3	87.1
Brochures/banner	45	12.6	12.6	99.7
7.00	1	.3	.3	100.0
Total	357	100.0	100.0	

Source : SPSS output

In cluster 2, 63.6% of respondents used communication media from friends/relatives (Word of Mouth), 21.6% used websites, 12.6% used brochures and banners, and 1.4% used Instagram.

Table 10. Communication media in cluster 3

Media	Frequency	Percent	Valid Percent	Cumulative Percent
Friends	1	50.0	50.0	50.0
Website	1	50.0	50.0	100.0
Total	2	100.00.00	100.0	

In terms of communication media, cluster 3 uses communication media through friends and websites.

DISCUSSION

In selecting the variables that form the segments, the researcher used demographic variables with 6 demographic variables including age, gender, marital status, occupation, income, residential status, 1 behavioral variable (tuition fees covered) and 2 preference variables (the variable for reasons for choosing a university and the variable for reasons for choosing a major). The reasons the author chose 6 demographic variables are: the reason for the age variable is because this

class does not limit enrollment to those who have just graduated from high school; even those who graduated from high school more than 10 years ago are allowed to enter this level of education. the reason for the marital status variable is because those who are married are allowed to enter this level of education, the reason for the occupation variable is because this class was originally intended for those who work, the reason for the income variable is because those who work have an income that influences their choice of place of study, while the reason for the residence variable is because those who work may live in their own homes. The reason the researcher chose one

behavioral variable, namely the variable (tuition fees are covered by), is first, it is possible that parents are still involved in financing tuition even though they have an income, and second, considering the age requirement for admission to this level, it also allows those who have just graduated from high school, who are usually working in positions with relatively small salaries at this age, to enroll. The last two variables are one motivational variable, namely the main reason for choosing a university, and one rational variable, namely the reason for choosing a study program. The reason for this is that in determining their choice of study location, prospective students usually have to consider the criteria of the university and the criteria of the study program.

Based on these nine variables, after grouping them, three clusters were formed, which were formed from the components of age, marital status, student employment, tuition fee payer, reasons for choosing a university, and reasons for choosing a study program, where these six components only contributed 38.1% of the existing clusters. Cluster 1 has a profile of 23-40 years of age, married marital status, private employee/entrepreneur employment (eligible position), tuition fees paid independently, and reasons for choosing the university due to accreditation and affordable costs, and choosing the study program because it is in line with their knowledge and work. Cluster 2 has a profile of age 18-25, unmarried, private sector employment (non-eligible position), tuition fees mostly still covered by parents, reasons for choosing the university are affordable costs, accreditation, and location, and reasons for choosing the study program are because it is in line with their knowledge. Cluster 3 has the following profile: age 20-24, marital status can be married/unmarried, private sector employment, tuition fees covered by sponsors, reasons for choosing the university are accreditation and affordable costs.

Previous research focused more on students who study using all their time for learning activities (Prastyabudi et al., 2024) (Susilo, 2016) (Lozano et al., 2021) (Khusnuliawati & Putri, 2021) (Khusnuliawati & Putri, 2021), while the current research focuses on students who study using part-time (time after work) for learning activities. The results of previous research, the variables that form segments are more dominant socio-demographic variables, geo-demographics of the student's family while the current research student

demographic variables influence the formation of segments such as student age variables, marital status, student employment (all 3 variables are demographic variables), the purpose of choosing a college and the purpose of choosing a study program (are motivational variables and rational variables). The appropriate communication media for cluster 1 is word of mouth because the main choice of communication tool is 81.8% friends/relatives, plus the use of websites and brochures/banners. The appropriate communication media for cluster 2 is word of mouth because the main choice of communication tool for 63.6% of the median is friends/relatives, plus the use of websites (21.6%), followed by brochures/banners (12.6%) and Instagram (1.4%). The appropriate communication medium for cluster 3 is word of mouth, as communication comes from friends/relatives, supplemented by websites.

CONCLUSION

This study focuses on a private university, namely Semarang University, specifically the Faculty of Economics, for students who study part-time (after work). To obtain the selected segments, this study uses K-Mean Cluster as a tool to find the segments in question so that the institution can develop effective communication strategies. There are three segments that can be developed with forms of communication appropriate to each segment: Segment 1, referred to as the independent student segment, has a profile of students aged 23-40 years old, married, employed as private employees/entrepreneurs (with decent incomes), paying their own tuition fees, and choosing the university because of its accreditation and affordable costs and choosing a study program because it is relevant to their field of study and work.

Communication uses word of mouth through friends/relatives, websites, and brochures/banners. Segment 2, referred to as the less independent student segment, has a profile of being aged 18-25 years old, unmarried, employed in the private sector (with a lower income), with tuition fees mostly still covered by parents, choosing the university because of its affordability, accreditation, and location, and choosing the study program because it is in line with their field of study. Communication uses word of mouth through friends/relatives, websites, brochures/banners, and Instagram. Segment 3 (Three), referred to as the working student segment, has the following profile: age 20-24, marital status can be married or unmarried, private sector employment, tuition costs covered by sponsors,

reasons for choosing the university include accreditation and affordability. Communication uses word of mouth through friends/relatives, supplemented by websites. The Faculty of Economics at Semarang University, by focusing on these three segments and adapting the chosen communication

REFERENCES

- Angulo, F., Pergelova, A., & Rialp, J. (2010). A market segmentation approach for higher education based on rational and emotional factors. *Journal of Marketing for Higher Education*, 20(1), 1–17. <https://doi.org/10.1080/08841241003788029>
- Budiman, R., Wiradharma, G., Harsasi, M., Riana, K. E., & Riady, Y. (2023). The potential of promotional media as a marketing instrument for open and distance higher education: a segmentation study. *JMKSP (Jurnal Manajemen, Kepemimpinan, Dan Supervisi Pendidikan)*, 8(1), 515–528.
- Davari, M., Noursalehi, P., & Keramati, A. (2019). Data mining approach to professional education market segmentation: A case study. *Journal of Marketing for Higher Education*, 29(1), 45–66.
- Gazdecki Michał and Goryńska-Goldmann, E., Kiss, M., & Szakály, Z. (2021). Segmentation of food consumers based on their sustainable attitude. *Energies*, 14(11), 3179.
- Hajibaba, H., Grün, B., & Dolnicar, S. (2020). Improving the stability of market segmentation analysis. *International Journal of Contemporary Hospitality Management*, 32(4), 1393–1411.
- Khusnuliawati, H., & Putri, D. R. (2021). Hybrid clustering based on multi-criteria segmentation for higher education marketing. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 19(5), 1498–1506.
- Kotler, P. (2012). *Marketing Management/Philip Kotler, Kevin Lane Keller.-. Marketing, 1009, 4Ps.*
- Lozano, F. A. M., Cruz Pulido, J. M., & García Rodríguez, J. F. (2021). The market segmentation of higher education in Colombia reveals social inequalities. *Cogent Education*, 8(1). <https://doi.org/10.1080/2331186X.2021.1877885>
- Miguéis, V. L., Freitas, A., Garcia, P. J. V., & Silva, A. (2018). Early segmentation of students according to their academic performance: A predictive modelling approach. *Decision Support Systems*, 115, 36–51.
- Prastyabudi, W. A., Alifah, A. N., & Nurdin, A. (2024). Segmenting the Higher Education Market: An Analysis of Admissions Data Using K-Means Clustering. *Procedia Computer Science*, 234(2023), 96–105. <https://doi.org/10.1016/j.procs.2024.02.156>
- Rahiminia, M., Razmi, J., Shahrabi Farahani, S., & Sabbaghnia, A. (2023). Cluster-based supplier segmentation: a sustainable data-driven approach. *Modern Supply Chain Research and Applications*, 5(3), 209–228.
- Soto, J. J., Cantillo, V., & Arellana, J. (2021). Market segmentation for incentivising sustainable transport policies. *Transportation Research Part D: Transport and Environment*, 99, 103013.
- Susilo, W. H. (2016). An Impact of Behavioral Segmentation to Increase Consumer Loyalty: Empirical Study in Higher Education of Postgraduate Institutions at Jakarta. *Procedia - Social and Behavioral Sciences*, 229, 183–195. <https://doi.org/10.1016/j.sbspro.2016.07.128>
- Vázquez, J.-L., Lanero, A., García, J. A., & Morano, X. (2023). Segmentation of consumers based on awareness, attitudes and use of sustainability labels in the purchase of commonly used products. *Sustainable Production and Consumption*, 38, 115–129.
- Wilkins, S. (2020). The positioning and competitive strategies of higher education institutions in the United Arab Emirates. *International Journal of Educational Management*, 34(1), 139–153.
- Zhang, H. (2025). AI-driven innovation and entrepreneurship education: a K-means clustering approach for Chinese university students. *Discover Artificial Intelligence*, 5(1), 1–19.