Asian Journal of Social Science and Management Technology
ISSN: 2313-7410
Volume 8, Issue 5, September-October, 2026
Available at www.ajssmt.com

Integration of SERVQUAL, Importance–Performance Analysis, and Google Review Sentiment Analysis for Academic Administrative Service Improvement in Indonesian Higher Education

Triyadi Joko Purwanto[1], Agus Dwi Priyanto[2], Sigit Prastowo[3], Retno Wulan Damayanti[4], Suryadi Budi Utomo[5], Hasna Wahidaturrasyidah Anshori[6]

[1]*Directorate of Academic Affairs, Sebelas Maret University, Indonesia*
[2]*English Study Program, Vocational School, Universitas Sebelas Maret, Indonesia*
[3]*Animal Science Study Program, Faculty of Animal Science, Universitas Sebelas Maret, Indonesia*
[4, 6]*Industrial Engineering Study Program, Faculty of Engineering, Universitas Sebelas Maret, Indonesia*
[5]*Chemistry Education Study Program, Faculty of Teacher Training and Education, Universitas Sebelas Maret, Indonesia*

ABSTRACT: Academic administrative services contribute to the quality of higher education, but are rarely evaluated using more than one data source. This study integrates a SERVQUAL survey, Importance–Performance Analysis, and sentiment analysis of Google reviews to prioritize service improvements at the Directorate of Academic Affairs, University of XYZ, Indonesia. A 36-item questionnaire measuring perception and expectation was completed by 300 purposively selected students in June–July 2026, while 59 textual reviews from 117 Google ratings (2018–2025) were manually labeled by three independent coders, mapped to the five SERVQUAL dimensions, and grouped into complaint themes. The overall gap was small (−0.033), but responsiveness (−0.465), tangibles (−0.282), and reliability (−0.185) fell below expectations, and 13 items entered the top-priority quadrant. The largest gaps concerned digital queuing, response speed of online channels, and system integration, although 66.7% of respondents used the service offline. Of the textual reviews, 81.4% were negative, dominated by complaints about operating hours, waiting time, and queuing. Matching complaint themes with survey items showed convergence on time-related and digital-channel problems but divergence on assurance, whose positive gap reflected low expectations. The study recommends six improvements that can be incorporated into updated standard operating procedures to support Sustainable Development Goal 4.

Keywords – SERVQUAL, importance–performance analysis, sentiment analysis, online reviews; academic administrative services, SDG 4

---

1. INTRODUCTION

The quality of higher education is determined not only by the classroom learning process, but also by the support services students receive during their studies. One such support service is academic administration, which includes diploma processing, graduation, data correction, and certificate issuance. Poor administrative service quality, such as slow processing, unclear information, or inefficient procedures, can hinder students

from completing various academic requirements. The impact can also continue after students graduate, for example, when documents are needed for job applications or other purposes. This situation demonstrates that the quality of administrative services is a crucial part of the student experience and contributes to their perception of the quality of higher education.

Improving the quality of academic administration services is also in line with the Sustainable Development Goals (SDGs) agenda, particularly SDG 4, which emphasizes quality education and equal access to higher education [1]. Previous studies have shown that non-academic services influence students' perceptions of higher education quality [2], [3]. Given this role, measuring the quality of academic administration services is necessary to assess the extent to which the services meet student needs and to identify aspects that require improvement.

One of the instruments that is widely used to measure service quality is SERVQUAL which was developed by Parasuraman et al. [4]. This model measures the gap between user expectations and perceptions across five dimensions: tangibles, reliability, responsiveness, assurance, and empathy. Its application to higher education demonstrates a gap between user expectations and perceptions across various service attributes [5], [6], [7], [8]. However, survey-based measurements have limitations because user experience is represented through pre-defined attributes, so specific experiences or complaints that emerge spontaneously are not necessarily fully reflected in the measurement results.

In contrast, online reviews contain spontaneous user experiences, including complaints and aspects of the service that are considered problematic, so they can be used as the voice of the customer that complements structured surveys [9]. This data also has limitations because not all users are motivated to provide reviews, and the characteristics of reviewers may differ from those of the overall population. Hu et al. [10] shows that ratings tend to be clustered at extremes because users with average experiences are rarely motivated to write reviews. Therefore, online reviews are better read as indicators of the most memorable experiences, rather than as a measure of overall user satisfaction, thus complementing surveys that are measured but limited to predetermined attributes.

A number of studies have utilized online review analysis to identify user perceptions of service quality in various sectors, such as e-commerce [11], airline [12], and hospitality [13]. Sentiment analysis has also been used in Google Maps reviews in Indonesia [14]. However, in the context of higher education academic administration services, research that systematically combines SERVQUAL measurement with digital review analysis is still rare. However, these two approaches are complementary because each captures different aspects of the user experience. SERVQUAL reveals gaps in structured attributes, while digital reviews reveal spontaneously emerging problem themes. Furthermore, prioritizing improvements is not sufficient based solely on the magnitude of the gap value. Therefore, Importance–Performance Analysis (IPA) is needed to determine which attributes need to be prioritized based on their level of importance and perceived user performance [15], [16]. The position of this research in relation to previous studies is summarized in Table 1.

This need is evident in the academic administration services of the Directorate of Academic Affairs of the University of XYZ, which serves all students across all faculties. As of October 2025, the Google Maps page of the Directorate of Academic Affairs at the University of XYZ received an average rating of 2.7 out of 5, with the majority of reviews being negative and containing complaints about the service process. Meanwhile, evaluations of service quality at universities generally rely on satisfaction surveys, while public assessments on digital channels have not been systematically utilized for improvement. Exploring both sources is necessary to ensure service improvements are directed at aspects that are truly perceived by users.

Based on the description above, this study aims to prioritize improvements in academic administration services by combining the results of the SERVQUAL survey analyzed through IPA and the results of sentiment analysis of Google reviews. More specifically, this study measures the gap between student expectations and perceptions of academic administration services, then determines service attributes that need to be prioritized through IPA. Furthermore, this study identifies sentiment and complaint themes in Google reviews, then combines these two sources of evidence to create service improvement priorities. Theoretically, this study is expected to enrich the study of service quality measurement by combining surveys and online reviews based

on problem themes, rather than simply matching dimension labels. Practically, the results can be used as a basis for prioritizing improvements and updating standard operating procedures (SOP) for academic services.

Table 1. Position of the research in relation to previous studies

Researchers	Context	Data and methods	Main findings
Soares et al. [8]	Administrative services of state-owned companies, Brazil	SERVQUAL Survey	Only the Tangibles dimension has a positive value
Rizos et al. [7]	PT administration services, Greece	SERVQUAL Survey	There is a negative gap in all dimensions
Magasi et al. [17]	PT in Dodoma, Tanzania	SERVQUAL survey, logistic regression	Service quality below student expectations
Raharjo et al. [18]	PT in Indonesia	HEISQUAL survey, ACSI, IPA, PGCV	IPA and PGCV produce a priority order of improvement
Palese and Usai [11]	E-commerce	Topic modeling online reviews	All five dimensions of SERVQUAL can be extracted; responsiveness complaints dominate negative reviews.
Chatterjee et al. [12]	Airline	Online reviews of 362 airlines; text mining and machine learning	SERVQUAL dimensions can be measured from online reviews
Ipmawati et al. [14]	Tourist destinations, Yogyakarta	Google Maps Review; SVM	Google Maps review sentiment classification
This research	XYZ university, Indonesia	SERVQUAL survey + IPA + sentiment of Google reviews	Priority improvements for theme-based integration results

2. LITERATURE REVIEW

2.1 SERVQUAL and its critics

SERVQUAL measures service quality as the difference between perceptions (P) of the service received and expectations (E) of that service in five dimensions, namely Tangibles (physical evidence), Reliability (reliability), Responsiveness (responsiveness), Assurance (guarantee), and Empathy (empathy) [4]. A negative gap indicates that the service has not met expectations. The main advantage of this model is its ability to identify the service areas that fall short of user expectations.

This model has also drawn criticism Cronin & Taylor [19] argue that the performance score alone (SERVPERF) is sufficient to explain service quality, whereas Babakus & Boller [20] questioning the use of difference scores which can give rise to measurement problems. In the context of higher education, Abdullah [2] developed HEdPERF as a sector-specific instrument, and other models such as HESQUAL have also been proposed [21]. However, SERVQUAL was chosen for this study for two reasons. First, the focus of the study is administrative services, whose characteristics are similar to those of service providers in general. Second, the expectation component is necessary to construct the IPA and to determine whether low scores are due to weak performance or high expectations.

2.2 Importance Performance Analysis (IPA)

IPA maps each service attribute on a Cartesian diagram with a performance axis and an importance axis [16]. The boundaries of the two axes divide the attributes into four quadrants: high priority, maintain, low priority, and redundant. O'Neill & Palmer [15] shows that science is useful for directing continuous quality

improvement in higher education, and Raharjo et al. [18] Combining it with student satisfaction measurements in Indonesia. The validity of IPA as a management tool has also been questioned by some researchers [22]. Therefore, in this study, IPA is used as a tool to determine priorities, read together with the gap size, not as the sole basis for decisions.

2.3 Online reviews and sentiment analysis

Online reviews are a form of electronic word-of-mouth written voluntarily. One motive for writing them is to vent negative feelings after a bad experience [9]. Reviews have also been shown to influence other people's decisions [23]. Sentiment analysis is an effort to recognize opinions and assessments in text, for example positive or negative [24].

The main limitation of online reviews is self-selection bias, Hu et al. [10] shows that the distribution of ratings tends to be J-shaped, with many extreme values and few middle values, because people with average experiences rarely write reviews. Therefore, online reviews are more appropriately treated as supplementary evidence indicating the types of problems that most affect users, rather than as a measure of population satisfaction.

2.4 Integration of Online Surveys and Reviews

Mixed methods research with a convergent design collects and analyzes two types of data separately, then compares the results to see similarities and differences [25]. In this study, the SERVQUAL survey provided a measurable picture of service quality gaps across the respondent population, while Google reviews provided insights into the most memorable experiences for users. Consistent findings reinforce confidence in the priority of improvement, while divergent findings provide a basis for examining the limitations of each source.

3. RESEARCH METHODS

3.1 Research design

This study used a convergent mixed-methods design. Survey and review data were analyzed separately and then combined in the integration stage to develop improvement priorities. The research flow is shown in Fig. 1.

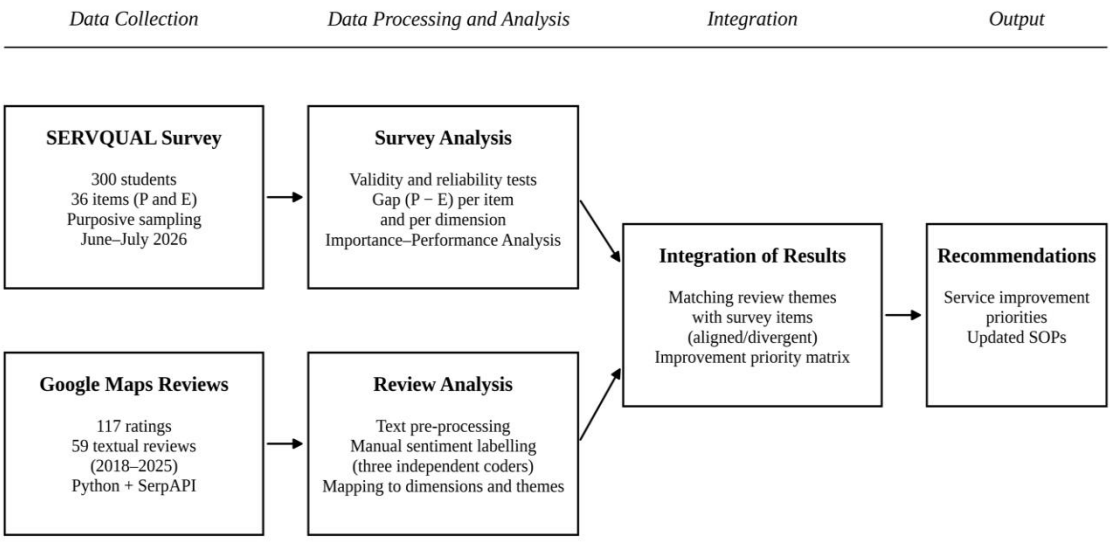


Figure 1. Research flow

3.2 Google review data

Review data were collected from the Google Maps profile page of the Directorate of Academic Affairs of the University of XYZ in November 2025 through web scraping using Python with the help of SerpAPI, covering usernames, ratings, and review texts. The dataset consisted of 117 ratings posted between February 2018 and

October 2025, of which 59 contained text and were analyzed further, while the remaining 58 consisted only of star ratings.

Review text was cleaned through five pre-processing stages: case folding, cleansing to remove punctuation, emojis, and numbers, stemming to return words with affixes to their base form, stopword removal, and tokenizing. Each review was then assigned a positive or negative sentiment label by three independent labelers based on the overall context of the text, and discrepancies were resolved through collaborative discussion until consensus was reached. A lexicon-based approach using a sentiment dictionary compiled from the vocabulary of academic service reviews was tested, but the results were less suitable for texts containing a lot of informal language, abbreviations, and combinations of praise and complaints within a single sentence. Therefore, manual labeling was used as the basis for the analysis. This labeling was based on the text content, not solely on the number of stars, so high-star reviews containing complaints were still labeled negative.

Each review was also mapped to the five SERVQUAL dimensions on a scale of 0–1 using the operational definitions in Table 2. A single review could fit more than one dimension, and reviews that did not address a specific service aspect were not mapped to any dimension. Additionally, reviews were grouped into problem themes based on emerging keywords, such as opening hours, queues, wait times, and staff attitude. Reviewer names were anonymized throughout the reporting.

Table 2. Operational definitions of mapping reviews to SERVQUAL dimensions

Dimensions	Mapped review criteria
Tangibles	Mentioning the physical facilities of the service, such as waiting room, cleanliness, parking, room comfort, and supporting facilities.
Reliability	Addressing timeliness of service, consistency of operating hours, clarity of procedures, and certainty of resolution.
Responsiveness	Related to the speed of service, the response of officers to requests or complaints, and the alertness of service
Assurance	Reflects professionalism, officer competence, clarity of information, and trust in the service.
Empathy	Contains assessments of staff attitudes, friendliness, caring, politeness, and personal attention.

3.3 SERVQUAL Survey

The questionnaire consisted of 36 items divided into five dimensions, namely Tangibles (10 items), Reliability (6), Responsiveness (7), Assurance (7), and Empathy (6). Each item was scored twice, for perception and expectation, using a five-point Likert scale. The items were structured according to the context of academic administrative services, including online channels such as the university's academic information system, email, WhatsApp, and Telegram. Sampling was conducted using a non-probability purposive technique, with respondents selected based on the criterion of having used the administrative services of the Directorate of Academic Affairs. The questionnaire was administered through Google Forms and distributed via the Directorate's official social media accounts and to students visiting the service office between June and July 2026. A total of 300 respondents were analyzed. All respondents participated voluntarily, and their identities were not recorded in the analysis.

3.4 Analysis techniques

Item validity was tested using Pearson correlation between item scores and total scores, calculated separately for perceptions and expectations. An item was declared valid if the calculated r exceeded the table r at a two-sided 5% significance level. With n = 300, the table r value used was 0.113. Reliability was tested using Cronbach's alpha per dimension and overall, with an acceptance limit of 0.60 [26], [27].

The gap for each item is calculated as the difference between the average perception score and the average expectation score (P – E), while the dimension gap is the average gap for items within that dimension. The

level of conformity is calculated as P/E × 100%, with values below 100% indicating that the service has not met user expectations. In IPA, the X-axis uses the average perception score as a measure of performance, while the Y-axis uses the average expectation score as a proxy for the level of importance, considering that the expectation score reflects the level of service that users perceive they should receive. The quadrant boundaries are set at the average of all items, which is 3.988 for perception and 4.021 for expectation.

3.5 Integration of Results

Integration was conducted based on problem themes. Each complaint theme derived from the review was paired with a questionnaire item that measured the same issue, and then the direction of the findings was compared. A theme was considered aligned if the complaint in the review was accompanied by a negative gap and the corresponding item was in Quadrant I. A theme was considered divergent if the complaint in the review was not accompanied by a negative gap in the survey. The comparison results were compiled into a priority improvement matrix that served as the basis for developing recommendations.

4. RESULTS

4.1 Respondent profile

Respondents were predominantly female (70.3%) and undergraduate students (97.0%), with the majority coming from the Faculty of Teacher Training and Education (27.7%), followed by the Faculty of Agriculture and the Faculty of Engineering (9.7 each). Most respondents accessed services offline by visiting the counter (66.7%), while only 6.3% used online channels. The most frequently used services were graduation and graduation services (50.0%), followed by academic financial services (25.7%). The majority of respondents used services once (43.0%) or two to three times (42.0%). A complete respondent profile is presented in Table 3.

Table 3. Respondent Profile

Characteristics	Category	Number	%
Gender	Female	211	70.3
	Male	89	29.7
Study level	Undergraduate (S1)	291	97.0
	Master (S2)	9	3.0
Faculty/school	Faculty of Teacher Training and Education	83	27.7
	Agriculture	29	9.7
	FSRD	29	9.7
	Technique	25	8.3
	Vocational School	23	7.7
	FEB	19	6.3
	Faculty of Mathematics and Natural Sciences; Faculty of Social and Political Sciences; Faculty of Psychology; Faculty of Cultural Sciences (each)	14	4.7
	Law	11	3.7
	Others (Faculty of Animal Husbandry, Faculty of Sport Sciences, Faculty of Information Technology and Data Science, Postgraduate School, Faculty of Medicine)	25	8.3
Service media	Offline (come to the counter)	200	66.7
	Online	19	6.3

Characteristics	Category	Number	%
	Offline and online	81	27.0
Frequency of visits to the service office	1 time	129	43.0
	2–3 times	126	42.0
	4–5 times	26	8.7
	More than 5 times	19	6.3
Service type	Graduation	150	50.0
	Academic finance	77	25.7
	Study status	50	16.7
	Special academic programs (MBKM, KKN, grants)	48	16.0
	Changes to student data	38	12.7
	Mutation/resignation	7	2.3

4.2 Validity and reliability of the instrument

All 36 items were found to be valid for both perception and expectation. The correlation coefficients for perception items ranged from 0.653 to 0.863, while for expectation items, the correlation coefficients ranged from 0.214 to 0.463, all above the r value of 0.113. All dimensions were also reliable (Table 4). Perception reliability was very high for all dimensions, while expectation reliability was high for three dimensions and moderate for Responsiveness and Empathy.

Table 4. Reliability test results (Cronbach's alpha)

Dimensions	Number of items	α Perception	Category	α Expectation	Category
Tangibles	10	0.934	Very high	0.813	High
Reliability	6	0.930	Very high	0.818	High
Responsiveness	7	0.926	Very high	0.789	Enough
Assurance	7	0.939	Very high	0.819	High
Empathy	6	0.950	Very high	0.778	Enough
Overall	36	0.983	Very high	0.814	High

4.3 Google review sentiment and themes

The distribution of 117 Google ratings shows an extreme pattern. Fifty-six ratings (47.9%) were 1-star and 36 ratings (30.8%) were 5-star, while only 5, 8, and 12 were 2-, 3-, and 4-star, respectively. The average rating was 2.72. Of the 59 text reviews, 48 (81.4%) were negative and 11 (18.6%) were positive.

Mapping the reviews to the SERVQUAL dimensions (Table 5) shows that the largest number of reviews fell into the Reliability category, namely 33 reviews, 28 of which were negative. This is because complaints about delayed opening hours and uncertainty of resolution fall under the operational definition of Reliability. Seven reviews were categorized as Assurance, all of which were negative, particularly regarding the professionalism of the staff. In contrast, reviews categorized as Tangibles tended to be positive, with 4 out of 5 praising the comfort of the waiting room, parking area, and supporting facilities.

Table 5. Distribution of review sentiment per SERVQUAL dimension

Dimensions	Number of reviews	Positive	Negative	Proportion negative
Reliability	33	5	28	84.8%
Assurance	7	0	7	100.0%

Dimensions	Number of reviews	Positive	Negative	Proportion negative
Tangibles	5	4	1	20.0%
Responsiveness	5	2	3	60.0%
Empathy	3	1	2	66.7%

Grouping by theme through keyword searches provides a more concrete picture of the issues complained about. The most frequent theme was inconvenient operating hours, mentioned in 18 reviews, all of which were negative. Reviewers described services not opening after lunch hours, while queues were already long. Next on the list were wait times and slow responses, including unanswered calls (16 reviews, 15 negative), unorganized queues (13 reviews, 11 negative), staff perceived as unfriendly or unprofessional (13 reviews, 12 negative), and unclear information and procedures (6 reviews, all negative). Physical facilities appeared in seven reviews, and most were positive. A single review could contain more than one theme, so the total number of themes exceeded the number of text-based reviews.

4.4 SERVQUAL gap per dimension

Overall, the average perception (3.988) was slightly lower than the average expectation (4.021), resulting in an overall gap of only −0.033 with a 99.2% agreement rate (Table 6). However, this overall figure masks large differences between dimensions. Three dimensions had negative gaps, with Responsiveness being the worst (−0.465). The other two dimensions, Assurance (+0.564) and Empathy (+0.340), exceeded expectations. The positive gaps in these two dimensions were more due to low expectations (3.467 and 3.688) than high perceptions, as perceptions across all five dimensions were relatively equal, at around 3.95–4.03.

Table 6. SERVQUAL gap per dimension

Dimensions	Perception (P)	Expectation I	Gap (P – E)	Conformity (P/E)	Ranking
Responsiveness	3,954	4,419	−0.465	89.5%	1
Tangibles	3,978	4,260	−0.282	93.4%	2
Reliability	3,954	4,139	−0.185	95.5%	3
Empathy	4,028	3,688	+0.340	109.2%	4
Assurance	4,030	3,467	+0.564	116.3%	5
Overall	3,988	4,021	−0.033	99.2%	–

4.5 SERVQUAL gap per item

Eighteen of the 36 items had a negative gap. The ten items with the worst gaps are presented in Table 7. The worst items were the availability of an orderly digital queuing system (-0.953), the speed of response to online channels such as WhatsApp, email, and Telegram (-0.900), and the integration of online service systems between platforms and between units (-0.800). Seven of the ten worst items related to service time and digital channels. Interestingly, the items regarding officers providing fast and accurate service (RS1, +0.137) and officers responding to requests for assistance (RS2, +0.193) actually had a positive gap. This means that responsiveness issues lie more in the system and service channels than in face-to-face interactions with officers.

Table 7. Ten items with the worst gaps

Code	Item (abbreviated)	P	E	Gap	IPA Quadrant
TG7	Orderly and clear queuing system (number/digital)	3,687	4,640	−0.953	I
RS4	Fast response via online channels (WhatsApp/e-mail/Telegram)	3,793	4,693	−0.900	I
TG9	Integrated online service system across	3,883	4,683	−0.800	I

Code	Item (abbreviated)	P	E	Gap	IPA Quadrant
	platforms/units				
RS7	Online help desk available	3,963	4,703	−0.740	I
RS5	The service is opened on time according to the specified hours.	3,980	4,663	−0.683	I
RS6	Information when there is a service delay	3,810	4,487	−0.677	I
RS3	Queue waiting times and processes are relatively short	3,977	4,560	−0.583	I
TG8	Service flow and SOP are clearly displayed	3,897	4,430	−0.533	I
TG4	Online services are easily accessible	4,000	4,473	−0.473	II
TG5	The officers are well-dressed and professional	4,180	4,643	−0.463	II

Source: processed primary data (2026).

4.6 Importance–Performance Analysis

The IPA results are shown in Fig. 2 and Table 8. Thirteen items are in Quadrant I (top priority), which is high expectations but low performance. All negative Responsiveness items (RS3–RS7) fall into this quadrant, along with four Tangibles items on disability-friendly facilities, queuing systems, service flow and SOP presentation, and online system integration. This quadrant also includes three Reliability items on information consistency, application status monitoring, and policy change explanations, and one Assurance item on clarity of operating hours. The disability-friendly facilities item (TG3) has a gap of -0.423 and does not appear in Google reviews, so this need was only identified through the survey. Quadrant IV (excessive) contains 14 items dominated by Assurance and Empathy. These items have good scores but are considered less important by respondents, so existing resources can be gradually shifted to Quadrant I items.

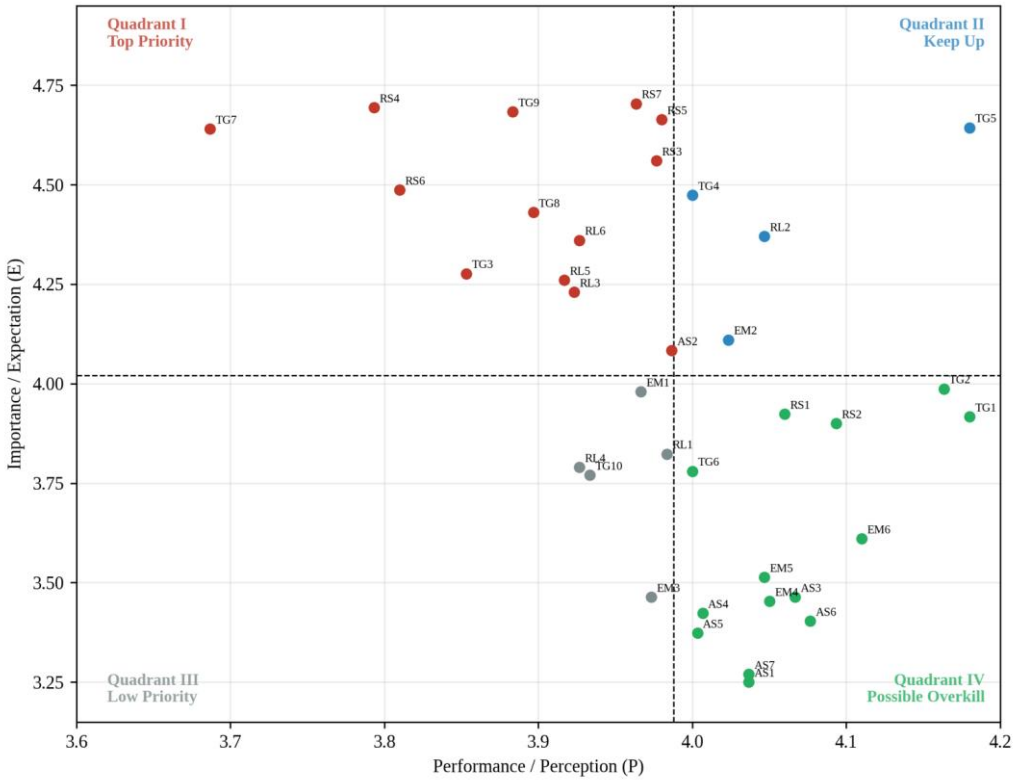


Figure 2. Cartesian diagram of IPA (dotted line = average P 3.988 and E 4.021)

Table 8. Distribution of items based on the IPA quadrant

Quadrant	Meaning	Number	Item code
I – Top priority	High expectations, low performance	13	TG3, TG7, TG8, TG9, RL3, RL5, RL6, RS3, RS4, RS5, RS6, RS7, AS2
II – Keep up the good work	High expectations, high performance	4	TG4, TG5, RL2, EM2
III – Low priority	Low expectations, low performance	5	TG10, RL1, RL4, EM1, EM3
IV – Possible overkill	Low expectations, high performance	14	TG1, TG2, TG6, RS1, RS2, AS1, AS3, AS4, AS5, AS6, AS7, EM4, EM5, EM6

4.7 Integration of surveys and reviews

At the dimension label level, the two sources appear to be inconsistent. The most frequently complained-about dimension in reviews is Reliability, while the worst survey gap is in Responsiveness. This discrepancy is largely due to differences in grouping. Operating hour complaints were included under Reliability in the review analysis, but were measured through the Responsiveness (RS5) and Assurance (AS2) items in the questionnaire. Therefore, integration was conducted based on themes (Table 8). The results show four themes aligned on the issues side, one theme aligned on the strengths side, and one theme that differed.

Table 9. Theme-based integration matrix

Theme	Review evidence	Survey evidence (gap; quadrant)	Status	Priority
Inaccurate opening hours	18 reviews (18 negative, 0 positive)	RS5 −0.683 (I); AS2 −0.097 (I)	Aligned	High
Disorderly queues	13 reviews (11 negative, 2 positive)	TG7 −0.953 (I); RS3 −0.583 (I)	Aligned	High
Slow response and online channels	16 reviews (15 negative, 1 positive)	RS4 −0.900 (I); RS7 −0.740 (I); RS6 −0.677 (I)	Aligned	High
Information, procedures, and process certainty	6 reviews (6 negative, 0 positive)	TG9 −0.800 (I); TG8 −0.533 (I); RL6 −0.433 (I); RL5 −0.343 (I); RL3 −0.307 (I)	Aligned	High
Staff attitude and professionalism	13 reviews (12 negative, 1 positive)	Assurance +0.564; Empathy +0.340; EM2 −0.087 (II)	Divergent	Medium
Physical facilities	7 reviews (2 negative, 5 positive)	TG1 +0.263 (IV); TG2 +0.177 (IV)	Aligned (strength)	Maintain

5. DISCUSSION

The most important finding of this study is the convergence between the two sources of evidence on problems related to service time and service channels. Complaints about delayed opening hours, disorderly queues, and slow responses recurred in the reviews, while the survey items measuring the same issues showed the largest negative gaps and fell into Quadrant I. This pattern is consistent with the finding of Palese & Usai [11] that responsiveness complaints dominate negative online reviews, as well as with studies of higher education administrative services that reported negative gaps in dimensions related to service processes [7], [8]. Because two sources with different collection methods point to the same problems, the priority for improvement on this theme rests on a strong basis.

The results for each item also indicate that responsiveness issues lie more with the service system than with the staff. Items regarding staff speed and accuracy in providing service showed positive gaps, while items

regarding digital queues, online channel responsiveness, online help desks, and system integration were showed the largest negative gaps. This suggests that students expect administrative services to be delivered through a fast and connected digital system, but the existing system has not met these expectations. This finding aligns with studies of higher education digital transformation, which show that institutional digitalization efforts are rarely comprehensively designed [28].

This pattern is also relevant to the respondent profile. A total of 66.7% of respondents only used offline services, while the item with the worst gap actually related to online services. This condition indicates that student expectations for digital services are high, even though utilization rates remain low. This low utilization is suspected to be related to the unresponsive online channels, leading students to prefer to go directly to the counter to handle their needs. This assumption is supported by several reviews complaining about the difficulty of contacting services, although further research is needed to confirm this through further research.

The two data sources show different tones. A total of 74.5% of perception responses in the survey were rated 4 or 5, whereas 81.4% of textual reviews were negative. Google ratings were also concentrated at the extremes, with 47.9% being 1-star and 30.8% being 5-star. This pattern corresponds to a J-shaped distribution resulting from self-selection bias, the tendency for highly disappointed or highly satisfied users to leave reviews [10]. Venting negative feelings is also one of the motives for writing online reviews [9]. The two data sources were also collected in different periods. The reviews cover 2018 to 2025, whereas the survey was conducted in 2026, so some reviews may reflect service conditions that have since changed. Therefore, Google reviews cannot be treated as a measure of overall student satisfaction. They are more likely to uncover the most memorable negative experiences and provide context behind the survey figures, such as gaps in the operating hours item related to services not being open after the break.

The most striking difference is found in the Assurance dimension. The survey ranked it as the dimension with the most positive gap (+0.564), while the seven reviews that touched on this dimension were all negative. This positive gap resulted more from low expectations, which ranged from 3.25 to 3.47 on six of the seven items, than from high perceptions. The difference may also be influenced by the scope of the definition, as Assurance in the review analysis encompasses clarity of information and professionalism in general and is therefore not identical to the questionnaire item. Furthermore, the number of reviews that touched on this dimension was relatively small, so the complaints that emerged may stem from specific experiences not shared by most respondents. These possibilities require further exploration through in-depth interviews.

The low expectations for the Assurance and Empathy dimensions have implications for how SERVQUAL results are interpreted. Within the SERVQUAL framework, a positive gap is generally interpreted as service exceeding expectations. However, perceptions across the five dimensions in this study were relatively equal, ranging from 3.954 to 4.030, suggesting that differences in gaps between dimensions are more determined by differences in expectations than differences in performance. This finding aligns with criticisms that the use of gap scores in SERVQUAL is problematic [20] and that performance scores have an important role in assessing service quality [19]. Therefore, gap readings should be accompanied by separate readings of perception and expectation scores. IPA fulfills this need by placing the two components on separate axes, allowing for a more comprehensive assessment of each attribute's position [16].

The overall gap of -0.033 indicates that the service, on average, nearly meets expectations. However, this figure is formed because the negative gap in 18 items is offset by positive gaps in the Assurance and Empathy dimensions. Behind this average figure, there are 13 items in Quadrant I, with the worst gaps in digital queues (-0.953) and online channel responsiveness (-0.900). Evaluations relying solely on the overall figure risk overlooking these shortcomings. Therefore, item-by-item analysis and IPA are necessary so that service evaluations can identify specific improvement targets [15], [18].

Based on the integration matrix, this study formulated six recommendations for improvement that can be incorporated into updated SOPs. Problems with unruly queues and long wait times can be addressed through a digital queue system that displays numbers and estimated wait times. Slow responses via WhatsApp, email, and Telegram, as well as the lack of information when services are delayed, can be addressed by establishing

an online help desk with written response time standards. Difficulty monitoring the progress of applications can be addressed through cross-platform service system integration with status tracking features.

Complaints about services not opening on time require certainty about operating hours, including rotating break schedules and announcements of changes to service hours. Unclear and frequently conflicting information across study programs, faculties, and university needs to be addressed by transparently displaying service flows, aligning information across units, and providing explanations for any policy changes. Furthermore, existing facilities for people with disabilities need to be improved to accommodate various user groups, such as priority queues, queue systems that visually display numbers and announce them audibly, and online services accessible to screen readers. Complaints about staff behavior also need to be addressed through service training, although this is assigned a medium priority because it is not supported by the survey results.

These recommendations are expected to contribute to the achievement of SDG 4, particularly target 4.3 on equal access to quality higher education and target 4.a on disability-friendly educational facilities [1]. Timely, easily accessible online administrative services with guaranteed processes can reduce non-academic barriers for students, especially those with busy schedules or those unable to visit the service office in person.

6. CONCLUSION

The overall quality of the administrative services at the Directorate of Academic Affairs of the University of XYZ nearly met student expectations, with an overall gap of −0.033. However, the Responsiveness, Tangibles, and Reliability dimensions remained below expectations, with 13 items in Quadrant I of the IPA matrix as the top priority for improvement. The largest gaps were found in digital services, namely the queuing system, online channel responsiveness, and system integration, although most respondents still used the services offline. Meanwhile, 81.4% of Google reviews were negative, dominated by complaints regarding inaccurate operating hours, slow responses, and disorderly queues. Theme-based integration showed that both data sources were aligned on issues of service time, queues, online channels, and process certainty, making these themes worthy of priority in updating the standard operating procedures. The difference in findings on the Assurance dimension indicates that a positive gap in the survey does not necessarily mean complaint-free service, as it may be influenced by respondents' low expectations.

This study has several limitations. The number of text-based reviews is relatively small and spread over a long period, so some reviews may reflect service conditions that have since changed. Review themes were grouped through keyword searches, and therefore still depend on the researchers' accuracy. Survey respondents were selected purposively rather than randomly, with the largest proportion coming from the Faculty of Teacher Training and Education, so the results cannot be generalized to all students of the University of XYZ. The study was also conducted at a single institution using a cross-sectional design, so it cannot yet demonstrate changes in service quality over time.

Further research is recommended to use a larger set of reviews and machine learning-based sentiment classification, to test differences in perceptions between respondent groups based on faculty and the service channel used, to add interviews that explain the divergent findings on the Assurance dimension, and to re-measure service quality after the standard operating procedures have been updated in order to assess the impact of the improvements.

7. Acknowledgements

This research was funded by Sebelas Maret University through the Policy Governance Research Grant (PTK-UNS) scheme for the 2026 Fiscal Year with contract number 460/UN27.22/PT.01.03/2026. The authors would like to thank the Academic Directorate of Sebelas Maret University for their support in conducting the research, as well as all students who participated as respondents.

8. REFERENCES

[1] United Nations, "Transforming our world: the 2030 Agenda for Sustainable Development (A/RES/70/1)," Oct. 2015. Accessed: Sep. 22, 2026. [Online]. Available: https://digitallibrary.un.org/record/3923923?v=pdf#files.

[2] F. Abdullah, "The development of HEdPERF: a new measuring instrument of service quality for the higher education sector," *Int. J. Consum. Stud.*, vol. 30, no. 6, pp. 569–581, Nov. 2006, doi: 10.1111/j.1470-6431.2005.00480.x.

[3] P. Sultan and H. Yin Wong, "Service quality in higher education – a review and research agenda," *International Journal of Quality and Service Sciences*, vol. 2, no. 2, pp. 259–272, Jul. 2010, doi: 10.1108/17566691011057393.

[4] A. ; Parasuraman, V. A. ; Zeithaml, and L. L. Berry, "Servqual: A Multiple-Item Scale For Measuring Consumer Perc," 1988.

[5] A. Brochado, "Comparing alternative instruments to measure service quality in higher education," *Quality Assurance in Education*, vol. 17, no. 2, pp. 174–190, Apr. 2009, doi: 10.1108/09684880910951381.

[6] R. B. Galeeva, "SERVQUAL application and adaptation for educational service quality assessments in Russian higher education," *Quality Assurance in Education*, vol. 24, no. 3, pp. 329–348, Jul. 2016, doi: 10.1108/QAE-06-2015-0024.

[7] S. Rizos, E. Sfakianaki, and A. Kakouris, "Quality of Administrative Services in Higher Education," *European Journal of Educational Management*, vol. volume-5-2022, no. 2, pp. 115–128, Dec. 2022, doi: 10.12973/eujem.5.2.115.

[8] M. C. Soares, O. Novaski, and R. Anholon, "SERVQUAL model applied to higher education public administrative services," *Brazilian Journal of Operations & Production Management*, vol. 14, no. 3, pp. 338–349, Sep. 2017, doi: 10.14488/BJOPM.2017.v14.n3.a7.

[9] T. Hennig-Thurau, K. P. Gwinner, G. Walsh, and D. D. Gremler, "Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet?," *Journal of Interactive Marketing*, vol. 18, no. 1, pp. 38–52, Feb. 2004, doi: 10.1002/dir.10073.

[10] N. Hu, J. Zhang, and P. A. Pavlou, "Overcoming the J-shaped distribution of product reviews," *Commun. ACM*, vol. 52, no. 10, pp. 144–147, Oct. 2009, doi: 10.1145/1562764.1562800.

[11] B. Palese and A. Usai, "The relative importance of service quality dimensions in E-commerce experiences," *Int. J. Inf. Manage.*, vol. 40, pp. 132–140, Jun. 2018, doi: 10.1016/j.ijinfomgt.2018.02.001.

[12] S. Chatterjee, A. Ghatak, R. Nikte, S. Gupta, and A. Kumar, "Measuring SERVQUAL dimensions and their importance for customer-satisfaction using online reviews: a text mining approach," *Journal of Enterprise Information Management*, vol. 36, no. 1, pp. 22–44, Jan. 2023, doi: 10.1108/JEIM-06-2021-0252.

[13] X. Xu and Y. Li, "The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: A text mining approach," *Int. J. Hosp. Manag.*, vol. 55, pp. 57–69, May 2016, doi: 10.1016/j.ijhm.2016.03.003.

[14] J. Ipmawati, S. Saifulloh, and K. Kusnawi, "Analisis Sentimen Tempat Wisata Berdasarkan Ulasan pada Google Maps Menggunakan Algoritma Support Vector Machine," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, pp. 247–256, Jan. 2024, doi: 10.57152/malcom.v4i1.1066.

[15] M. A. O'Neill and A. Palmer, "Importance-performance analysis: a useful tool for directing continuous quality improvement in higher education," *Quality Assurance in Education*, vol. 12, no. 1, pp. 39–52, Mar. 2004, doi: 10.1108/09684880410517423.

[16] J. A. Martilla and J. C. James, "Importance-Performance Analysis," *J. Mark.*, vol. 41, no. 1, pp. 77–79, Jan. 1977, doi: 10.1177/002224297704100112.

[17] C. Magasi, R. G. Mashenene, and D. M. Ndengenesa, "Service Quality and Students' Satisfaction in Tanzania's Higher Education: A Re-examination of SERVQUAL Model," *International Review of Management and Marketing*, vol. 12, no. 3, pp. 18–25, May 2022, doi: 10.32479/irmm.13040.

[18] T. H. Raharjo, K. B. Mulyono, I. Ismiyati, and A. Jaenudin, "HEISQUAL – ACSI – IPA – PGCV: Synthesis of higher education service satisfaction measurements," *Asian Management and Business Review*, pp. 121–137, Aug. 2023, doi: 10.20885/AMBR.vol3.iss2.art2.

[19] J. J. Cronin and S. A. Taylor, "Measuring Service Quality: A Reexamination and Extension," *J. Mark.*, vol. 56, no. 3, p. 55, Jul. 1992, doi: 10.2307/1252296.

[20] E. Babakus and G. W. Boller, "An empirical assessment of the SERVQUAL scale," *J. Bus. Res.*, vol. 24, no. 3, pp. 253–268, May 1992, doi: 10.1016/0148-2963(92)90022-4.

[21] V. Teeroovengadum, T. J. Kamalanabhan, and A. K. Seebaluck, "Measuring service quality in higher education," *Quality Assurance in Education*, vol. 24, no. 2, pp. 244–258, Apr. 2016, doi: 10.1108/QAE-06-2014-0028.

[22] I. Sever, "Importance-performance analysis: A valid management tool?," *Tour. Manag.*, vol. 48, pp. 43–53, Jun. 2015, doi: 10.1016/j.tourman.2014.10.022.

[23] J. A. Chevalier and D. Mayzlin, "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, vol. 43, no. 3, pp. 345–354, Aug. 2006, doi: 10.1509/jmkr.43.3.345.

[24] B. Liu, "Sentiment Analysis and Opinion Mining," *Synthesis Lectures on Human Language Technologies*, vol. 5, no. 1, pp. 1–167, May 2012, doi: 10.2200/S00416ED1V01Y201204HLT016.

[25] J. W. Creswell and V. L. P. Clark, *Designing and Conducting Mixed Methodes Reseach*. Inc: SAGE Publication, 2017.

[26] L. J. Cronbach, "Coefficient Alpha and the Internal Structure of Tests," *Psychometrika*, vol. 16, no. 3, pp. 297–334, Sep. 1951, doi: 10.1007/BF02310555.

[27] K. S. Taber, "The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education," *Res. Sci. Educ.*, vol. 48, no. 6, pp. 1273–1296, Dec. 2018, doi: 10.1007/s11165-016-9602-2.

[28] L. Benavides, J. Tamayo Arias, M. Arango Serna, J. Branch Bedoya, and D. Burgos, "Digital Transformation in Higher Education Institutions: A Systematic Literature Review," *Sensors*, vol. 20, no. 11, p. 3291, Jun. 2020, doi: 10.3390/s20113291.

INFO

Corresponding Author: Retno Wulan Damayanti, Industrial Engineering Study Program, Faculty of Engineering, Universitas Sebelas Maret, Indonesia.

How to cite/reference this article: Triyadi Joko Purwanto, Agus Dwi Priyanto, Sigit Prastowo, Retno Wulan Damayanti, Suryadi Budi Utomo, Hasna Wahidaturrasyidah Anshori, Integration of SERVQUAL, Importance–Performance Analysis, and Google Review Sentiment Analysis for Academic Administrative Service Improvement in Indonesian Higher Education, Asian. Jour. Social. Scie. Mgmt. Tech. 2026; 8(5): 227-240.