Enumerators' Characteristics and Data Quality in Susenas 2025: A Logistic Regression Study

Iin Jumsinah* and Mahlia Muis

Hasanuddin University, Makassar, Indonesia

*jumsinahi24a@student.unhas.ac.id

Abstract. This study employs data from the 2025 National Socio-economic Survey (Susenas) conducted in South Sulawesi Province, Indonesia. This research examines the causal effect of enumerators' biographical characteristics on data quality. A quantitative approach was adopted, with binary logistic regression applied to evaluate the influence of five key enumerator characteristics age, gender, educational attainment, prior Susenas experience, and primary occupation on data quality outcomes. The results indicate that enumerator biographical attributes collectively exert a statistically significant effect on the likelihood of producing high-quality survey data. At the individual level, gender, prior Susenas experience, and occupation demonstrated statistically significant associations with improved data quality; notably, homemakers and self-employed individuals contributed most prominently to this effect. Female enumerators demonstrated a higher probability of generating accurate responses compared to their male counterparts. In line with these findings, prior Susenas' experience and occupational background were positively correlated with data reliability. These results underscore the relevance of personal attributes in shaping field performance and suggest that recruitment strategies should incorporate both experiential and occupational considerations. Enhancing enumerator training is further recommended to ensure the consistency and integrity of future Susenas data collection efforts.

Keywords: Biographic-characteristic, Data-Quality, Susenas-Enumerator, Logistic-Regression, Statistic-Indonesia.

1 Introduction

As the national authority responsible for core statistical operations, BPS Statistics Indonesia plays a pivotal role not only in conducting large-scale surveys and censuses but also in supplying data that aligns with the informational needs of both the public and government institutions. According to the 2024 Data Needs Survey, the National Socio-economic Survey (Susenas) microdata emerged as the most frequently accessed dataset, comprising 30.27% of total microdata requests [1]. This substantial demand is likely attributable to the breadth and depth of information captured through Susenas, which encompasses a wide range of socioeconomic indicators. These data have been

M. Nohong et al. (eds.), *Proceedings of the 10th International Conference on Accounting, Management, and Economics (10th ICAME 2025)*, Advances in Economics, Business and Management Research 388,
https://doi.org/10.2991/978-94-6239-709-5_89

extensively utilized as empirical evidence to support the formulation, monitoring, and evaluation of national development programs. Moreover, Susenas outputs are recognized as essential inputs for advancing sustainable development initiatives in Indonesia [2].

Susenas produces a range of strategic indicators, including the poverty rate and the Gini ratio. Prior research by [3] suggests that these two indicators are positively correlated, such that improvements in poverty outcomes are typically associated with declining income inequality, and vice versa. However, the 2025 Susenas results for South Sulawesi Province reveal a divergent pattern. Between 2024 and 2025, the poverty rate fell from 8.06% to 7.60%, in contrast to the Gini ratio, which increased from 0.360 to 0.363 [4]. This deviation from the expected trend raises concerns regarding the consistency and validity of the data produced. The observed anomaly contradicts established empirical patterns and prompts critical reflection on the reliability of the survey outcomes.

High-quality data is essential for both tactical and strategic decision-making. As emphasized by [5], managers and policymakers rely on accurate data to address performance-related challenges effectively. The integrity and reliability of data collection activities are predicated on the competence, workload, and operational capacity of enumerators. This finding is consistent with the evidence presented in [6], which suggests that staff capacity has a substantial influence on the quality of outputs. Furthermore, [7] emphasizes that technical factors and the enumerator's personal attributes determine data quality. In practice, data quality in the Susenas survey depends not only on the survey instrument and data processing system but also on the personal attributes of enumerators responsible for data collection [8].

As stated by [9,10], the accuracy of survey results is largely dependent on the enumerators, where errors can arise due to improper interpretation of instructions, variations in interview techniques, or inaccurate recording of answers. This finding is reinforced by the research of [11], which highlights the importance of enumerators' social skills, negotiation skills in the field, and adaptation to the local linguistic and cultural context in influencing who is willing to participate and how answers are constructed during the interview process. Moreover, biographical characteristics like age, gender, education level, and work experience have a big impact on the precision of data collection and the caliber of interactions between enumerators and respondents [12].

This study focuses on the biographical characteristics of enumerators involved in the 2025 Susenas data collection in South Sulawesi Province, specifically examining age, gender, education level, work experience, and primary occupation. Three core research questions are addressed: (1) What is the general profile of enumerators based on their biographical characteristics? (2) Do these characteristics collectively influence the quality of Susenas data? and (3) Do individual biographical variables exert a partial effect on data quality?

The study intends to answer these issues by thoroughly examining the ways in which the biographical characteristics of enumerators such as namely, age, gender, education, prior experience, and vocational background, affect the caliber of data gathered in the 2025 Susenas. The primary objective is to generate a comprehensive profile of

enumerators, assess both the simultaneous and individual effects of biographical factors on data quality, and identify the most influential variables. The findings are expected to contribute not only to the theoretical advancement of survey data quality research but also to provide practical implications for BPS Statistics Indonesia. To improve the precision, consistency, and timeliness of data collection, the findings may specifically guide the creation of recruitment plans, capacity-building programs, and enumerator management systems. Ultimately, improvements in the quality of Susenas data will enhance statistical accuracy, which serves as a critical foundation for effective regional and national planning, assessment, and policymaking.

2 Literature Review

2.1 Biographical Characteristics of Enumerators

Biographical characteristics refer to inherent or acquired attributes that individuals possess throughout their lives. According to [13], these traits include age, gender, marital status, tenure, and educational attainment. In the context of survey implementation, such features may influence an enumerator's ability to communicate effectively, adapt to respondents' environments, and complete assigned tasks. Enumerators with extensive experience in specific surveys tend to employ more comprehensive strategies to elicit information from respondents [14]. [15] said that enumerators who have other jobs and are prepared to work on the weekends generally stay in touch with respondents more often. [16] emphasized that these traits represent personal dimensions that shape individual work patterns.

International studies have further explored the implications of enumerator characteristics.[17] found that respondent gender may influence their willingness to disclose sensitive information, such as income or reproductive health status. In the field of selection psychology, [18] argued that dispositional traits and role clarity mediate task effectiveness in observational fieldwork. Educational level, meanwhile, plays a critical role in an enumerator's ability to comprehend survey concepts and procedures, as well as in ensuring accurate documentation of field data.

Additional evidence from a longitudinal study conducted in the Sleman HDSS, as documented in [19], demonstrated that respondents consistently preferred enumerators who exhibited key interpersonal qualities, including politeness, clarity in explaining research objectives, and cultural sensitivity.

A complementary study by [20] confirmed that enumerator attributes may account for up to 12% of the variation in data collection outcomes. Internal documentation from BPS further indicates that younger, well-educated enumerators with prior survey experience are more likely to produce data with lower error rates.

2.2 Survey Data Quality

According to [21], the quality of field survey data is determined by three key dimensions: quantity (completeness of responses), quality (accuracy of answers), and timeliness (promptness of data delivery). Evidence from [22] shows that enumerators'

work experience is positively correlated with the speed of data collection, which, in turn, contributes to improved data quality. In field survey contexts, data quality encompasses the completeness of entries, the precision of recorded responses, and the timeliness of submission. As posited by [23], work quality can be assessed through output, time efficiency, and procedural compliance dimensions that are often shaped by the enumerator's personal attributes. Empirical evidence from Sub-Saharan Africa, as presented in [24], demonstrates that enumerator biographical attributes significantly influence respondent trust and the reliability of household asset and reproductive event reporting. This study establishes a framework for examining whether similar patterns are observable within the sociocultural context of South Sulawesi. Given the potential variability in respondent–enumerator dynamics across regions, understanding the role of biographical traits in shaping data quality remains a critical area of inquiry.

3 Research Methodology

This study examines how enumerators' biographical characteristics impact the caliber of data produced during field survey operations using an explanatory, quantitative research approach [25]. The research was conducted across the operational areas of Statistics Indonesia (BPS) at the regency/municipality level throughout South Sulawesi Province, leveraging both primary and secondary datasets.

The research utilized primary data to measure the quality of enumerator outputs, as appraised by supervisors operating at the regency and municipality levels, including Susenas coordinators and enumerator training instructors. Secondary data were obtained from the existing "Sobat-BPS" data, which contains detailed personal information on enumerators. The study population comprised all enumerators involved in the 2025 Susenas data collection period, totaling 971 individuals. Applying Slovin's formula at a 95% confidence level, the calculated minimum sample size was 284 respondents.

Sampling was conducted using purposive multistage sampling, a non-probability technique that combines staged selection with predefined criteria [26]. In the first stage, seven regencies were purposively selected to represent geographic diversity, urban and rural areas, islands, coastal zones, and mountainous regions out of the total 24 regencies/municipalities in South Sulawesi. In the subsequent stage, Enumerators from the selected regencies or municipalities were systematically identified and incorporated into the sample one by one, continuing until the desired sample size of 284 was reached.

Data collection was administered online via Google Forms™ and Google Sheets™. The instrument included a roster of enumerators and a set of indicators to evaluate data quality, encompassing accuracy, completeness, and consistency. This method was chosen to facilitate coordination, expedite the data entry process, and minimize the risk of data loss. Moreover, it ensured that assessments were conducted directly by authorized personnel with comprehensive knowledge of field operations.

The designated statistical procedures were then used to export and analyze the collected data. The research timeline spanned two months, covering preparation, data collection, statistical analysis, and reporting of findings.

3.1 Conceptual Framework

The conceptual foundation of this work is informed by theoretical viewpoints and empirical findings that highlight the relationship between person traits and data quality, particularly in the context of survey data collection. Previous studies suggest that data quality depends not only on survey instruments and technical procedures but also on enumerators' biographical attributes, including age, gender, educational attainment, prior work experience, and primary occupation [9, 13, 20]. Such attributes can shape enumerator–respondent interactions, affect the accuracy of recorded responses, and influence compliance with assigned tasks.

Building on this theoretical foundation, the present study constructs a conceptual framework that positions data quality as the dependent variable and enumerators' biographical characteristics as independent variables. This framework serves as the basis for hypothesis testing, which is conducted using binary logistic regression analysis. To identify the most influential biographical factors, the hypotheses are formulated to assess both simultaneous and partial effects. Accordingly, the study hypothesizes that enumerators' biographical characteristics exert a statistically significant influence on Susenas data quality, both collectively and individually. Each dimension from demographic traits to professional experienceis expected to exhibit a meaningful association with the probability of generating high-quality survey data.

3.2 Operational Definition

This study employs five independent variables (X) and one dependent variable (Y), namely the quality of data collected by enumerators during the implementation of the National Socio-Economic Survey (Susenas). Data quality is evaluated across three dimensions: accuracy, completeness, and validity of questionnaire responses [27]. As detailed in Table 1, each indicator was rated on a 0–4 scale. The cumulative score was then classified into two categories: totals of 9 or above were designated as high-quality data, whereas totals below 9 were categorized as low-quality data.

Table 1. Measurement Indicators of the Dependent Variable

Indicator	Answer	Class/Category
Number of Revisions Made by Supervisors Following Data Collection (y11)	1. Extensive Revisions (Multiple corrections required post-submission) 2. Moderate Revisions (Several adjustments made by supervisors) 3. Minor Revisions (Few corrections with minimal impact) 4. No Revisions (Data accepted without any modifications)	0 = Low Quality/ Non-Compliant Data 1 = High Quality/ Compliant Data

Indicator	Answer	Class/Category
Timeliness of Task Execution Relative to Official Survey Schedule (y12)	1. Delayed Completion (Tasks completed beyond the scheduled timeframe) 2. Moderately Timely Completion (Tasks completed with minor delays relative to the schedule) 3. Timely Completion (Tasks completed within the designed timeframe) 4. Highly Timely Completion (Tasks completed ahead of schedule or with exceptional punctuality)	
Overall Quality of Collected Data (y13)	1. Low Quality (Data exhibits significant inaccuracies, incompleteness, or inconsistency) 2. Moderately Acceptable (Data meets minimum standards with some minor issues in accuracy or completeness) 3. Acceptable Quality (Data is generally accurate, complete, and consistent) 4. High Quality (Data is exceptionally accurate, complete, and consistent across all dimensions)	

Table 2. Independent Variables (X): Biographical Attributes of Susenas Enumerators

Variable	Category
Enumerator Age	Continuous variable (measured in years)
Sex	1 = Male 2 = Female
Primary Occupation	1. Village/Subdistrict Official 2. PKK Cadre / Other Community Cadres 3. Civil Servant / Contract Staff 4. Homemaker 5. Entrepreneur 6. Student 7. Others (e.g. Farmer)

Variable	Category
Highest Educational Attainment	1= Primary School or Equivalent 2= Junior High School or Equivalent 3= Senior High School or Equivalent 4= Diploma Level (D1/D2/D3) 5= Undergraduate Degree (D4/S1) 6= Master's Degree (S2) 7= Doctoral Degree (S3)
Enumerators' Prior Experience in Susenas	0= No Prior Experience 1= Has Prior Experience

3.3 Analytical Methods

This study employs two analytical approaches: descriptive and inferential analysis. The descriptive analysis provides an overview of the biographical characteristics of Susenas' enumerators, presented through graphical representations. The inferential analysis utilizes binary logistic regression to identify key variables that significantly influence the quality of Susenas data.

As noted in [28], when the dependent variable is binary, with a value of 1 signifying high-quality data and a value of 0 signifying low-quality data, binary logistic regression is appropriate.This method enables the estimation of the likelihood that enumerators' biographical characteristics affect the probability of producing high-quality data, assessed in terms of accuracy, completeness, and consistency.

The logistic regression model adopted in this study follows the formulation proposed by [27] as follows:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 \quad (1)$$

In this model, p represents the probability of producing high-quality data, β_0 de-notes the intercept, and β_1 refers to the regression coefficient that captures the effect of each independent variable on the log-odds of generating high-quality data. To facilitate interpretation, the estimated coefficients are transformed into odds ratios.

The Likelihood Ratio and Wald tests are used to determine statistical significance at the 5% level. Furthermore, to evaluate whether the logistic regression model adequately captures the observed data, the Hosmer-Lemeshow goodness of fit test is utilized.

3.4 Instrument Reliability and Validity Testing

Before performing binary logistic regression analysis, the research instrument must undergo reliability testing to confirm its validity in measuring the target variables. Cronbach's Alpha coefficient is frequently utilized to evaluate internal consistency. An instrument is deemed dependable if the alpha value is 0.60 or above [29]. Reliability

pertains to the stability of measurement outcomes; as noted by [30], repeated application of the instrument to the same subjects should yield relatively consistent results. Consequently, only variables verified for reliability should be incorporated into the logistic regression model to guarantee the validity of analytical results.

Items with extremely low item–total correlation values, specifically those below 0.30, should be excluded to enhance the internal consistency of the measurement instrument [31]. Beyond reliability, the instrument must also be evaluated for validity to ensure that each item accurately captures the intended construct. Validity assessment can be conducted via the Pearson Product Moment correlation between individual item scores and the aggregate score.An item is considered valid if it has a positive association and a significance level below 0.05. According to [32] instrument validity refers to the degree to which a measurement tool accurately reflects the concept it is intended to assess..

Binary logistic regression necessitates proper independent variables to accurately assess the association with a dichotomous dependent variable (coded as 0 or 1); the validity test is a critical prerequisite in this analytical process.

4 Results And Discussion

4.1 Descriptive Analysis

Based on the findings of this study regarding the biographical characteristics of Susenas enumerators in BPS South Sulawesi, out of a total of 284 enumerators, 190 individuals or approximately 67% were found to have produced high-quality Susenas data.

Table 3. Distribution of Susenas Enumerators by Gender and Data Quality Classification in South Sulawesi Province, 2025

Sex	Quality Category (%)		Total (%)
	Low	High	
Male	44%	56%	100%
Female	30%	70%	100%
Total	33%	67%	100%

Source: Processed Survey Data

The distribution of data quality levels by sex reveals a notable disparity between male and female enumerators. As shown in Table 3, 56% of male respondents produced high-quality data, while 44% fell into the low-quality category. In contrast, female enumerators demonstrated a higher proportion of high-quality outcomes, with 70% classified as producing high-quality data and only 30% categorized as producing low-quality data.

Overall, 67% of the total enumerator population analyzed were found to have generated high-quality data, whereas the remaining 33% did not meet the established criteria. These findings suggest that, within the context of this study, female

enumerators exhibited a greater tendency to achieve higher data quality levels compared to their male counterparts.

Table 4. Percentage of Susenas Enumerators by Occupation and Data Quality Level in South Sulawesi Province, 2025

Primary Occupation	Quality Category (%)		Total
	Low	High	
1. Village/Subdistrict Official	36%	64%	100%
2. PKK Cadre / Other Community Cadres	40%	60%	100%
3. Civil Servant / Contract Staff	35%	65%	100%
4. Homemaker	21%	79%	100%
5. Self-employed	23%	77%	100%
6. Student	62%	38%	100%
7. Others (e.g., Farmer)	45%	55%	100%
Total	33%	67%	100%

Source: Processed Survey Data

As shown in Table 4, 67% of the enumerators analyzed were classified as producing high-quality data, while the remaining 33% fell into the low-quality category. However, notable variation exists across occupational groups.

Enumerators whose primary occupation is homemaking exhibited the highest proportion of high-quality data at 79%, followed by self-employed individuals (77%) and civil servants or contract staff (65%). These findings suggest that individuals from non-formal and semi-formal occupational backgrounds tend to demonstrate a greater capacity for generating reliable survey data.

Conversely, enumerators from student backgrounds showed the highest proportion of low quality data, with 62% falling into this category and only 38% classified as high-quality. This may reflect limited field experience or insufficient technical understanding of survey procedures among student enumerators.

Village/subdistrict officials and PKK/community cadres reported high-quality proportions of 64% and 60%, respectively both slightly below the overall average. Meanwhile, the "Others" category (e.g., farmers and miscellaneous professions) yielded a high-quality proportion of 55%.

Table 5. Percentage of Susenas Enumerators by Prior Susenas Experience and Data Quality Level in South Sulawesi Province, 2025

Prior Experience in Susenas	Quality Category (%)		Total (%)
	Low	High	

No Experience	61%	39%	100%
Has Experience	22%	78%	100%
Total	33%	67%	100%

Source: Processed Survey Data

Enumerators without prior Susenas experience demonstrated relatively lower performance, with only 39% producing high-quality data and 61% falling into the low-quality category. In contrast, those with previous Susenas experience achieved significantly better outcomes, with 78% classified as high-quality and only 22% as low-quality. Overall, 67% of the enumerators analyzed were found to produce high-quality data, while 33% did not meet the expected quality standards.

Table 6. Percentage of Susenas Enumerators by Educational Attainment and Data Quality Level in South Sulawesi Province, 2025

Education Level	Quality Category (%)		Total (%)
	Low	High	
Senior High School	37%	63%	100%
Diploma (D1/D2/D3)	21%	79%	100%
Undergraduate (D4/S1)	31%	69%	100%
Master's Degree (S2)	50%	50%	100%
Total	33%	67%	100%

Source: Processed Survey Data

Table 6 presents the distribution of data quality by enumerators' highest level of education. While 67% of the total enumerators produced high-quality data, notable variation exists across educational groups. Enumerators with diploma-level education (D1/D2/D3) exhibited the highest performance, with 79% producing high-quality data. This was followed by those with undergraduate degrees (69%) and senior high school graduates (63%). Interestingly, enumerators with master's degrees (S2) showed a balanced outcome, with 50% in each category, despite having the highest formal education level. This may suggest that formal education alone does not guarantee superior field performance.

Table 7. Percentage of Susenas Enumerators by Age Group and Data Quality Level in South Sulawesi Province, 2025

Age Group	Quality Category (%)		Total (%)
	Low	High	
20–24	50%	50%	100%
25–29	32%	68%	100%

30–34	31%	69%	100%
35–39	32%	68%	100%
40–44	33%	67%	100%
45–50	27%	73%	100%
Total	33%	67%	100%

Source: Processed Survey Data

Table 7 illustrates the distribution of data quality across age groups. While 67% of enumerators overall produced high-quality data, performance varied notably by age. The 45–50 age group demonstrated the highest quality, with 73% producing high-quality data and only 27% classified as low-quality. This was followed by the 30–34 age group (69%), and both the 25–29 and 35–39 age groups (68%). The youngest group, aged 20–24, showed the lowest performance, with an equal split between high and low quality. These findings suggest a potential relationship between age and data quality, indicating that older enumerators tend to produce more reliable survey data.

4.2 Instrument Testing (Validity And Reliability)

Validity Assessment. y11 stand for Number of Revisions Made by Supervisors Following Data Collection, y12 is Timeliness of Task Execution Relative to Official Survey Schedule, and y13 is Overall Quality of Collected Data.

Table 8. Validity Test of Enumerator Data Quality Items

	Items	y11	y12	y13	Total
y11	Pearson Correlation	1	.315**	.487**	.824**
	Sig. (2-tailed)		.000	.000	.000
	N	284	284	284	284
y12	Pearson Correlation	.315**	1	.328**	.656**
	Sig. (2-tailed)	.000		.000	.000
	N	284	284	284	284
y13	Pearson Correlation	.487**	.328**	1	.805**
	Sig. (2-tailed)	.000	.000		.000
	N	284	284	284	284
Total	Pearson Correlation	.824**	.656**	.805**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	284	284	284	284

**. Correlation is significant at the 0.01 level (2-tailed).

Source: Processed Survey Data

Table 8 presents the results of the validity test conducted on three data quality items (y11, y12, y13) using Pearson correlation with the total score. Item validity was

evaluated based on the strength and statistical significance of the correlation between each item and the overall score. All correlation values exceeded the commonly accepted threshold of r > 0.30 and were statistically significant at the 0.01 level. These results indicate that all three items possess strong construct validity.

Reliability Assessment. To evaluate reliability, the Cronbach's Alpha coefficient was employed. Higher Cronbach's Alpha scores (which range from 0 to 1) indicate greater internal consistency. A Cronbach's Alpha rating of 0.60 or higher generally means that a questionnaire is considered credible.

Table 9. Reliability Test of Data Quality Items for Susenas Enumerators

Reliability Statistics	
Cronbach's Alpha	N of Items
.810	4

Source: Processed Survey Data

The reliability analysis yielded a Cronbach's Alpha of 0.810, surpassing the widely accepted minimum standard of 0.60. This suggests that the instrument possesses strong reliability, consistently generating stable outcomes when repeatedly applied to evaluate the quality of Susenas data. Such reliability supports the appropriateness of using this instrument in subsequent analyses, including binary logistic regression, to identify factors influencing data quality. It also suggests that respondents interpreted the questions consistently, with no substantial fluctuations in responses attributable to instrument inconsistency. Consequently, analytical results derived from this instrument can be regarded as both stable and trustworthy.

Inferential Analysis. The binary logistic regression equation derived from the results of this study is formulated as follows:

$$ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 \quad (2)$$

$$ln\left(\frac{p}{1-p}\right) = 1.416 - 0.000X_1 - 0.840X_2 + 0.477X_{31} + 0.216X_{32} + 0.000X_{33} + 0.951X_{34} + 1.348X_{35} + 0.133X_{36} - 0.766X_{41} - 0.463X_{42} - 0.051X_{43} - 1.887X_5 \quad (3)$$

Table 10. Variable Description

Variable Code	Description

X_1	Enumerator's Age (continuous variable measured in years)
X_2	Sex (1 = Male, 2 = Female)
X_{31}	Occupation: Village/Subdistrict Official
X_{32}	Occupation: PKK Cadre / Other Community Cadres
X_{33}	Occupation: Civil Servant / Contract Staff
X_{34}	Occupation: Homemaker
X_{35}	Occupation: Self-employed (Entrepreneur)
X_{36}	Occupation: Others (e.g., Farmer, Miscellaneous)
X_{41}	Educational Attainment: Senior High School or Equivalent
X_{42}	Educational Attainment: Diploma Level (D1/D2/D3)
X_{43}	Educational Attainment: Undergraduate Degree (D4/S1)
X_5	Prior Susenas Work Experience (0 = No, 1 = Yes)

Goodness of Fit Test. Before applying binary logistic regression analysis, the suitability of the regression model was evaluated using SPSS. The Hosmer and Lemeshow Goodness-of-Fit Test was employed to assess model adequacy.

Table 11. Hosmer and Lemeshow Goodness-of-Fit Test Results

Chi-square	Df	*Sig*
8.130	8	0.421

Source: Processed Data

As shown in Table 11, the chi-square significance value is 0.421, which exceeds the conventional threshold of 0.05. This result leads to the decision to fail to reject the null hypothesis (H_0), indicating that, at the 95% confidence level, the logistic regression model is adequately specified to explain the influence of enumerators' biographical attributes on Susenas data quality.

Additionally, the classification table and model estimation output reveal that the model achieves an overall predictive accuracy of 73.9%. This suggests that the logistic regression model possesses a satisfactory ability to classify and predict the quality of Susenas data based on the enumerators' characteristics.

Simultaneous Testing. Simultaneous testing is intended to evaluate the collective ability of a logistic regression model, which incorporates multiple independent variables, to explain variability in the data [28]. Within the Omnibus Test of Model

Coefficients, the calculated G-statistic is tested against the critical threshold at a significance level of 0.05, based on 12 degrees of freedo.

Table 12. Omnibus Test of Model Coefficients

	Chi-square	Df	Sig
Model	59,571	12	0,000

Source: Processed Data

As shown in Table 12, the significance value of the G-test is 0.000, which is less than the 0.05 threshold. This leads to the rejection of the null hypothesis (H_0), indicating that at least one independent variable significantly explains the influence of enumerators' biographical characteristics on Susenas data quality in South Sulawesi Province in 2025.

Partial Testing. Following the results of the simultaneous test, partial testing was conducted to identify which specific independent variables significantly affect the status of Susenas data quality. This analysis was performed using the Wald statistic.

Table 13. Wald Test Statistics for Individual Predictors

Independent Variable	Coefficient B	S.E.	Wald	df	Sig.
(1)	(2)	(3)	(4)	(5)	(6)
Age	.000	.021	.000	1	.996
Sex (Female)	-.840	.387	4.718	1	**.030**
Education (overall)	-	-	4.497	3	.213
Education: Senior High School	-.766	1.519	.255	1	.614
Education: Diploma (D1/D2/D3)	-.463	1.667	.077	1	.781
Education: Undergraduate (D4/S1)	-.051	1.504	.001	1	.973
Occupation (overall)	-	-	10.669	6	.099
Occupation: Village Official	.477	.512	.866	1	.352
Occupation: PKK Cadre/Others	.216	.520	.173	1	.677
Occupation: Civil Servant	.000	.554	.000	1	1.000
Occupation: Homemaker	.951	.465	4.192	1	**.041**
Occupation: Self-employed	1.348	.526	6.567	1	**.010**
Occupation: Student/Others	.133	.764	.030	1	.861
Susenas Experience (Yes)	-1.887	.325	33.642	1	**.000**
Constant	1.416	1.728	.672	1	.413

Source: Processed Data

Independent variables with Wald test significance values of less than 0.05 are considered to have a statistically significant effect on the data quality of Susenas. As shown in Table 13, variables such as age, education, and several occupation categories

(village official, PKK cadre, civil servant, and student/others) have significance values above 0.05, leading to the decision to accept H_0.

Conversely, sex (female), occupation as homemaker and self-employed, and prior Susenas experience all have significance values below 0.05, resulting in the rejection of H_0.These findings suggest that these variables have a significant influence on the likelihood of producing high-quality Susenas data in South Sulawesi Province in 2025.

Odds Ratio Interpretation.

Table 14 presents the odds ratios for significant variables affecting the status of Susenas data quality. These values indicate the likelihood that enumerators will produce high-quality data relative to reference categories.

Table 14. Odds Ratios for Significant Variables

Significant Variable	Odds Ratio	Interpretation
(1)	(2)	(3)
Sex (Female)	0.432	Female enumerators are 0.432 times as likely to produce high-quality data compared to male enumerators.
Occupation (Homemaker)	2.589	Enumerators whose primary occupation is homemaking are 2.589 times more likely to produce high-quality data than those in other occupations.
Occupation (Self-employed)	3.851	Self-employed enumerators are 3.851 times more likely to produce high-quality data compared to those in other occupations.
Prior Susenas Experience	0.151	Enumerators with prior Susenas experience are 0.151 times as likely to produce high-quality data compared to those without such experience.

5 Discussion

The findings indicate that enumerator characteristics are materially associated with the probability of producing high-quality Susenas data, particularly with respect to sex, prior Susenas experience, and occupational background. This pattern is consistent with the view that data quality in field surveys is not determined solely by instruments and procedures, but is also shaped by enumerators' personal attributes and their capacity to manage respondent interaction in situ [13, 23]. The significance of sex is plausibly linked to respondent comfort and disclosure behaviours, especially when interviews involve socio-economic information that may be perceived as sensitive [17]. Likewise, prior experience appears to strengthen procedural compliance, improve the interpretation of survey concepts, and reduce recording errors through accumulated task familiarity and role clarity, which aligns with evidence that experienced enumerators adopt more effective strategies for eliciting complete and accurate responses [14, 18]. The occupational effects, especially for homemakers and self-employed enumerators, may reflect greater schedule flexibility and sustained engagement with respondents,

which has been noted as beneficial for maintaining contact and completing interviews under field constraints [15]. More broadly, these results resonate with prior work suggesting that enumerator attributes can account for non-trivial variation in data collection outcomes and can influence respondent trust, thereby shaping the reliability of household reporting [20, 24]. By contrast, the non-significance of age and formal education suggests that demographic maturity and credential level do not automatically translate into superior field performance, reinforcing arguments that interpersonal competence, cultural sensitivity, and practical task execution may matter more than formal attainment alone [19, 23].

From a methodological and policy perspective, the evidence supports a shift from purely credential-based selection towards recruitment and deployment strategies that explicitly incorporate experiential and occupational considerations, while safeguarding representativeness and fairness in enumerator hiring [13, 20]. Training programmes should therefore prioritise applied interviewing skills, standardised probing techniques, and field simulations that develop negotiation and adaptive communication competencies, which have been shown to shape respondent cooperation and response construction [19]. In addition, experience should be leveraged through structured mentoring, whereby seasoned enumerators support less experienced staff in managing difficult interviews and maintaining consistency across villages and subdistricts, in line with insights from selection psychology on the importance of role clarity and task effectiveness in fieldwork [18]. Finally, given that survey data quality is multidimensional, encompassing accuracy, completeness, and timeliness [21], continuous monitoring and feedback mechanisms should be aligned to these dimensions and used to refine training content and supervision intensity. Future research could strengthen inference by adopting multilevel specifications to account for clustering by area and supervisory units, and by examining whether experience and occupational background interact with local respondent dynamics to influence reporting reliability across contexts [22, 24].

6 Conclusion

Based on the results of the binary logistic regression analysis, several biographical characteristics of enumerators were found to significantly influence the quality of Susenas 2025 data in South Sulawesi. The variables with statistically significant effects ($p < 0.05$) include sex, occupation (specifically homemaker and self-employed categories), and prior Susenas experience.

Specifically, female enumerators demonstrated a higher probability of producing high-quality data compared to their male counterparts (odds ratio = 0.432). Enumerators whose primary occupation was homemaking or self-employment exhibited substantially greater odds of generating high-quality data 2.589 and 3.851 times, respectively than those in other occupational categories. Another key finding is that prior experience with Susenas surveys is positively associated with data quality, although the odds ratio of 0.151 suggests a complex relationship that may warrant further investigation.

Age, educational level, and other occupational categories did not have statistically significant impacts on the quality of the data. These results highlight how enumerator performance is influenced by personal characteristics, specifically sex and employment, particularly in terms of respondent engagement and data recording accuracy. This finding is consistent with earlier survey methodological research that emphasizes the impact of enumerator characteristics on data dependability.

6.1 Recommendations and Suggestions

Based on the study's findings, improving the quality of Susenas data should be supported by a reformulation of the enumerator recruitment strategy to broaden opportunities for individuals from diverse professional backgrounds, treating prior survey experience not as a primary requirement but as an added advantage. BPS is also advised to strengthen practice-based training and field simulations, including refresher sessions for experienced enumerators, and to enhance the selection process by incorporating assessments of interpersonal skills and continuous feedback-based monitoring. Furthermore, developing a performance appraisal system that integrates quality scores with quality-based incentives is essential for maintaining data consistency. For future research, spatial and ordinal regression approaches could be employed to capture quality variations in greater detail. At the same time, qualitative studies may explore the motivations and behaviors of enumerators in producing high-quality data.

References

1. Badan Pusat Statistik: Hasil Survei Kebutuhan Data BPS Tahun 2024. Badan Pusat Statistik (2024).
2. Badan Pusat Statistik: Indikator SDGs Kesejahteraan Rakyat Tahun 2024. Badan Pusat Statistik (2024).
3. Fauziana, H., Wardhana, A., Rusgianto, S.: The Effect of Education, Income, Unemployment, and Poverty Toward the Gini Ratio in Member of OIC Countries. Daengku Journal of Humanities and Social Sciences Innovation, 2(2), 181-191 (2022). https://doi.org/10.35877/454RI.daengku874
4. Badan Pusat Statistik: Data Kemiskinan dan Ketimpangan Provinsi Sulawesi Selatan Tahun 2024–2025. Badan Pusat Statistik (2025).
5. Batini, C., Cappiello, C., Francalanci, C., Maurino, A.: Methodologies for Data Quality Assessment and Improvement, ACM Computing Surveys, 41(3), 1–52 (2009). https://doi.org/10.1145/1541880.1541883
6. Nasrul, R., Buulolo, R., Ratnasari, E.: Pengaruh Kapasitas Pegawai Terhadap Kualitas Pekerjaan, Jurnal Administrasi dan Manajemen 10(2), 45–56 (2021).
7. Sarifudding: Karakteristik Individu dan Kinerja Pegawai, Jurnal Manajemen dan Bisnis, 5(1), 23–34 (2016).
8. Blattman, C., Jamison, J., Koroknay-Palicz, T., Rodrigues, K., Sheridan, M.: Measuring the Measurement Error: A Method to Qualitatively Validate Survey Data, Journal of Development Economics, 120, 99–112 (2016). https://doi.org/10.1016/j.jdeveco.2016.01.005
9. Groves, R. M.: Survey Methodology: Wiley-Interscience (2004).

10. West, B. T., Blom, A. G.: Explaining Interviewer Effects: A Research Synthesis, Journal of Survey Statistics and Methodology, 5(2), 175–211 (2017). https://doi.org/10.1093/jssam/smw024
11. Randall, S., Coast, E., Compaore, N., Antoine, P.: The Power of the Interviewer, Demographic Research, 28(27), 763–792 (2013). https://www.demographic-research.org/articles/volume/28/27/
12. Flores-Macias, F., Lawson, C.: Effects of Interviewer Gender on Survey Responses: Findings from a Household Survey in Mexico, International Journal of Public Opinion Research, 20(1), 101–113 (2008). https://doi.org/10.1093/ijpor/edn007
13. Robbins, S. P., Judge, T. A.: Organizational Behavior (19th ed., Global Edition). Pearson Education (2023).
14. Purdon, S., Campanelli, P., Sturgis, P.: Interviewers' Calling Strategies on Face-to-Face Interview Surveys, Journal of Official Statistics, 15(2), 199–216 (1999).
15. Pietrelli, R., d'Errico, M., Dassesse, K.: Measuring Household Food Security Through Surveys: Do the Characteristics of the Enumerators Matter? Development Policy Review (2021). https://doi.org/10.1111/DPR.12534
16. Siagian, S. P.: Manajemen Sumber Daya Manusia. Bumi Aksara (2008).
17. Álvarez-Aragón, P., Champeaux, H.: Measuring Norms and Enumerator Effects: Survey Method Matters, DeFiPP Working Paper No. 2024-01, University of Namur & University of Cagliari (2024). https://ssrn.com/abstract=4693068
18. Barends, A. J., de Vries, R. E.: Construct Validity of a Personality Assessment Game in a Simulated Selection Situation and the Moderating Roles of the Ability to Identify Criteria and Dispositional Insight, International Journal of Selection and Assessment, 31(1), 120–134 (2022). https://doi.org/10.1111/ijsa.12404
19. Harahap, W., Dewi, F.S.T.: Karakter Enumerator yang diinginkan Responden: Studi HDSS Sleman, Jurnal Kesehatan Masyarakat, 14(2), 120–132 (2018). https://jurnal.ugm.ac.id/bkm/article/view/29338/20779
20. Schueler, R., Rodriguez-Segura, D.: Enumerator Effects and Impact Evaluations, 2023. https://doi.org/10.26300/dz3n-d153 [Accessed: 6 October 2025].
21. Yusuf, M., Prasetyo, E.: Manajemen Kualitas Data Survei Lapangan: Konsep, Dimensi, dan Implementasi. Penerbit Mitra Cendekia (2021).
22. Nasution, A. E., Lesmana, M. T.: Analysis of Factors Affecting the Performance of Lecturers In Conducting Research: A Case Study on Private Universities in Medan, Asia Pacific Journal of Management and Education (APJME), 3(2), 33–44 (2020).
23. Mangkunegara, A. P.: Manajemen Sumber Daya Manusia Perusahaan. Remaja Rosdakarya (2018).
24. Haber, N., Robyn, P. J., Hamadou, S., Yama, G., Hien, H., Louvouezo, D., Fink, G.: Surveyor Gender Modifies Average Survey Responses: Evidence From Household Surveys in Four Sub-Saharan African Countries, 2018. https://arxiv.org/abs/1810.01981 [Accessed: 6 October 2025].
25. Creswell, J. W., Creswell, J. D.: Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (5th ed.). SAGE Publications (2018).
26. Cochran, W. G.: Sampling Techniques (3rd ed.). John Wiley & Sons (1977).
27. Biemer, P. P.: Total Survey Error: Design, Implementation, and Evaluation, Public Opinion Quarterly, 74(5), 817–848 (2010). https://doi.org/10.1093/poq/nfq058
28. Hosmer, D. W., Lemeshow, S., Sturdivant, R. X.: Applied Logistic Regression (3rd ed.). Wiley (2013).
29. Cronbach, L. J.: Coefficient Alpha and the Internal Structure of Tests, Psychometrika, 16(3), 297–334 (1951). https://doi.org/10.1007/BF02310555

30. Malhotra, N. K.: Marketing Research: An Applied Orientation (7th ed.). Pearson Education (2016).
31. Ghozali, I.: Aplikasi Analisis Multivariate dengan Program IBM SPSS 25. Universitas Diponegoro Press (2018).
32. Arikunto, S.: Prosedur Penelitian: Suatu Pendekatan Praktik. Rineka Cipta (2019).