JESICA: Journal of Enhance Studies in Informatics and Computer Applications
Journal homepage: jesica.itsk-soepraoen.ac.id
e-ISSN: 3046-6938 p-ISSN: 3046-6997
Vol. 2, No. 2, July 2025, Pp. 62-69

Analyzing the Distribution of Household Electricity Usage in Indonesia Using Two-Way ANOVA

Fonda Leviany*[1], Mayang Anglingsari Putri[2]
1 Data Science Study Program, Faculty of Science and Technology, Universitas Terbuka, Indonesia
2 Information Systems Study Program, Faculty of Science and Technology, Universitas Terbuka, Indonesia

***Corresponding author**
E-mail address:
fonda.leviany@ecampus.ut.ac.id

Keywords:
Household electricity usage, Indonesia, location, two-way ANOVA, years

Abstract
Indonesia continues to experience notable challenges in providing equal access to electricity across its vast territory. In recent years, the government has made substantial progress through its national electrification initiatives. By the end of 2024, the National Electrification Ratio had reached 99.83%. However, despite this remarkable development, many remote and inland areas—particularly in certain regions of Papua—remain disconnected from the national electricity grid. Recent data on the percentage of households using electricity as the main lighting source in 38 provinces during 2023 and 2024 presents an opportunity for comparative analysis to assess equity in electricity access across regions and over time. This study proposes an analytical approach using Two-Way ANOVA, with location (urban vs. rural) and year (2023 vs. 2024) as the independent variables. The two-way ANOVA further confirmed that the main effect of location (urban vs. rural) was statistically significant ($F = 10.049$, $p = 0.002$), indicating a persistent gap in access between the two locations. However, the overall change from 2023 to 2024 was not statistically significant ($p = 0.273$), and the urban-rural gap remained relatively stable. These findings emphasize the importance of targeted policies to improve rural electrification and reduce regional disparities in access to electricity.

1. Introduction

As the world's largest archipelagic state, Indonesia continues to experience notable challenges in providing equal access to electricity across its vast territory, such as locations far from existing networks, settlements in remote areas, and the low economic capacity of the community to install and pay for electricity [1]. In recent years, the government has made substantial progress through its national electrification initiatives. By the end of 2024, the National Electrification Ratio had reached 99.83% [2]. However, despite this remarkable development, many remote and inland areas—particularly in certain regions of Papua—remain disconnected from the national electricity grid [3].

Persistent disparities in electricity access across Indonesia underscore the necessity for comprehensive evaluations of electrification initiatives by policymakers and development stakeholders. Research aimed at assessing the distribution of electricity access is crucial, not only to measure development outcomes, but also to inform the formulation of more inclusive and equitable energy policies. Data-driven evaluations play a key role in supporting the achievement of global objectives, particularly Sustainable Development Goal 7, which emphasizes access to affordable, reliable, sustainable, and modern energy for everyone. As highlighted in the SDG 7 Roadmap for Indonesia, "the key objective is to assist the Government of Indonesia in developing policy measures that will enable the achievements SGD 7 targets" [4]. The government has undertaken several initiatives to improve electricity access in Indonesia, such as providing assistance for new electricity installations and constructing infrastructure to extend electricity supply to remote residential areas. However, despite progress in both policy and infrastructure development, there remains a lack of empirical statistical studies that examine whether disparities in electricity access differ significantly across various locations and time periods.

A study by [5] evaluated the equitable distribution of electricity in Indonesia using a qualitative approach involving the 5 Whys method; the Urgency, Seriousness, and Growth (USG) model; and Bardach's eightfold path framework. The findings indicated that unequal access to electricity has adverse effects on education, healthcare, and economic development, ultimately slowing progress in underdeveloped areas. The study recommended solutions such as promotions of renewable energy, decentralization of energy policies, and strengthening collaboration between government and the private sector [5]. Meanwhile, a qualitative study by [6] showed that to achieve fair access to electricity in Indonesia's rural electrification program, it is important to address all elements of energy injustice such as unequal income to pay for electricity, inadequate access, and violations of the non-recognition of socio-economic characteristics that are typical in various locations in Indonesia. [6].

Recent data on the percentage of households using electricity as the main source of lighting in 38 provinces during 2023 and 2024 presents an opportunity for comparative analysis to assess equity in electricity access across locations and over time. This study proposes the use of a Two-Way ANOVA to analyze the data. Two-Way ANOVA is a quantitative method that examines the impact of two categorical independent variables [7]. This method is proposed as an analysis tool that can answer whether there is a significant difference in the distribution of electricity use as the primary lighting source in urban vs rural areas and in 2023 vs 2024. It provides stronger inferential power to identify whether disparities in electrification are systemic and policy-sensitive, thus bridging a critical methodological gap in the current literature. Previous research by [8] applied Two-Way ANOVA to investigate the interaction between learning models and class schedules on student exam results [8]. Previous research by [9] used it to evaluate the effect of creativity and activeness on student understanding [9], while [10] applied the method to explore how teaching methods and initial mathematical abilities affect students' problem-solving skills [10]. A study by [11] determined the effect of year, treatment, replication, and interaction on rice and wheat yields using the same method.

Although previous studies have explored Indonesia's electricity distribution from various perspectives, there is still a notable research gap regarding statistical comparisons of household electrification rates across both location types (urban vs. rural) and over time (2023 vs. 2024). Prior research often lacks an integrated analysis of these two dimensions. The decision to use Two-Way ANOVA in this study is based on its strength in simultaneously evaluating the main effects of two independent variables. This approach is particularly relevant given the distinct policy treatments for urban and rural areas, such as differing subsidies and infrastructure programs. By analyzing both variables together, this method allows for a clearer understanding of whether disparities persist and how they may evolve over time—insights that are crucial for data-driven, regionally-targeted policy interventions.

Therefore, this study proposes an analytical approach using Two-Way ANOVA, with location (urban vs. rural) and year (2023 vs. 2024) as the independent variables. The objective is to assess whether significant differences exist in the percentage of electricity usage based on location and years. These research questions can be effectively addressed using the Two-Way ANOVA method.

2. Research Method

This study adopted a quantitative research design, structured through a systematic sequence of analytical steps. The primary data source comprised secondary data on Percentage of Household Electricity Usage in Indonesia, retrieved from Badan Pusat Statistik (Central Bureau of Statistics of Indonesia). The dataset included the percentage of households using electricity as their main source of lighting, disaggregated by province, location (urban and rural), and year (2023 and 2024). Prior to analysis, the data underwent preprocessing procedures, including the identification and treatment of missing values and potential outliers to ensure data quality. Missing values were handled through listwise deletion, due to significantly outperformed for missing data when the proportion of missing value is small [12]. Outliers were detected using the interquartile range (IQR) method and retained in the analysis, as they reflected valid differences between rural and urban regions rather than data errors. This approach was chosen to preserve the integrity of regional disparities in electrification, which are central to the study's objectives.

The core statistical analysis employed a Two-Way ANOVA to examine the effects of two independent variables—year and residential classification—on the dependent variable, namely the percentage of electrified households. This method also allowed for the assessment of interaction effects between the two factors. The significance of the results was evaluated using F-statistics and corresponding p-values. Where significant effects were found, post-hoc comparisons were conducted to explore specific group differences. The findings of this study provide empirical insights into the dynamics of household electricity usage across time and settlement types, contributing to a more nuanced understanding of energy equity and infrastructure development in Indonesia.

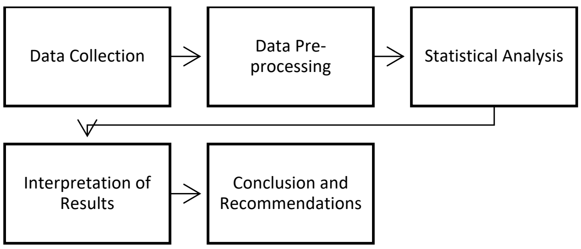

Figure 1. Research method

2.1 Pre-processing

Data pre-processing is an essential step in ensuring the quality and reliability of the dataset before conducting statistical analysis. This step transforms the data into a format that is more easily and effectively processed in data mining, machine learning, and other data science tasks [13]. In this study, pre-processing step involved identifying and addressing missing values and potential outliers within the secondary data on household electricity usage.

Handling missing data was crucial to maintain the accuracy and completeness of the analysis, as unaddressed gaps could undermine the reliability of the findings. Similarly, identifying and handling outliers was important to minimize the impact of extreme values that might otherwise skew the outcomes. Thorough data cleaning improves the reliability and precision of research results, supporting evidence-based decisions that contribute to better patient outcomes [14]. These pre-processing procedures ensured that the data used in the Two-Way ANOVA were clean, consistent, and suitable for accurate examination of the effects of year and location on household electrification usage.

2.2 Descriptive Statistic

Descriptive statistics are used to summarize and present data through numerical measures and graphical representations. Numerical measures include central tendency such as frequency, mean, and median, and measures of dispersion such as variance and standard deviation. Data visualization can be done using tools like boxplots, histograms, bar charts, and pie charts. In this research, a boxplot and histogram are used to display the data.

2.3 Two-Way ANOVA

Two-Way ANOVA (Analysis of Variance) is a statistical method used to examine the effect of two different independent variables on one dependent variable at the same time. In this study, two-way ANOVA is used to analyze how both years and location simultaneously influence the percentage of households with electricity. Here is the effect model [15]:

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} \quad (1)$$

Where, μ is the grand mean, τ_i is the additive main effect of level i from the first factor, β_j is the additive main effect of level j from the second factor, and ϵ_{ij} is the error. The Two-Way Analysis of Variance (ANOVA) table presented in this study illustrates the decomposition of total variability in the dataset into components attributable to two independent categorical variables, referred to as locations (residential classification: urban and rural) and year (2023 and 2024), along with the residual error terms.

Table 1 ANOVA table

Source of variation	Sum of squares	df	Mean square	F-ratio
Locations	$SSA = \sum_{i-1}^{a} \frac{y_{i.}^2}{b} - \frac{y_{..}^2}{ab}$	$a-1$	$MSA = \frac{SSA}{a-1}$	$\frac{MSA}{MSE}$
Years	$SSB = \sum_{j-1}^{b} \frac{y_{.j}^2}{a} - \frac{y_{..}^2}{ab}$	$b-1$	$MSB = \frac{SSB}{b-1}$	$\frac{MSB}{MSE}$
Error	$SSE = SST - SSA - SSB$	$(a-1)(b-1)$	$MSE = \frac{SSE}{(a-1)(b-1)}$	
Total	$SST = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{ab}$	$ab-1$		

Following the Two-Way ANOVA, which determines whether statistically significant differences exist among group means, a Tukey's Honestly Significant Difference (HSD) post hoc test was conducted to further examine the pairwise comparisons between group means [16]. This procedure is essential when the ANOVA yields a significant F-ratio, as it identifies specifically which groups differ while controlling for the Type I error rate across multiple comparisons.

$$HSD = T_\alpha = q_\alpha(p, f)\sqrt{\frac{MSE}{n}} \quad (2)$$

Where, $q = \frac{\bar{y}_{max} - \bar{y}_{min}}{\sqrt{\frac{MSE}{n}}}$, α is significance level, p is number of groups, f is number of degrees of freedom, and $q_\alpha(p, f)$ obtained from studentized range statistic table.

3. Results and Discussion

Rising electricity demand—still predominantly fueled by fossil sources—and uneven progress in household energy transition efforts highlight the urgency of analyzing electricity accessibility patterns more closely [17]. The following table presents the primary data of the percentage of households in Indonesia, categorized by province and location (urban and rural), based on their main source of lighting derived from electricity. The data is expressed in percentages and has been obtained from Badan Pusat Statistik. This information provides insight into the distribution and accessibility of electricity across different regions and types of residential areas in the country.

Table 2 Data on the percentage of households by province, village classification (urban and rural), and main source of lighting from electricity, expressed in percent

Province	2023		2024	
	Urban	Rural	Urban	Rural
Aceh	99.91	99.83	99.9	99.84
Sumatera Utara	99.96	98.94	99.99	99.4
Sumatera Barat	99.9	99.15	99.85	98.93
Riau	99.83	98.87	99.97	99.22
Jambi	99.95	99.18	99.78	99.48
Sumatera Selatan	99.99	99.54	99.99	99.4
Bengkulu	99.77	99.65	99.89	99.87
Lampung	100	99.65	100	99.91
Kep. Bangka Belitung	99.95	99.89	99.86	99.9
Kep. Riau	99.98	98.37	99.99	99.23
DKI Jakarta	99.98	-	100	-
Jawa Barat	99.99	99.94	99.99	100
Jawa Tengah	99.98	100	99.98	99.99
DI Yogyakarta	100	100	100	100
Jawa Timur	100	99.96	100	99.99
Banten	99.92	99.31	100	99.15
Bali	100	99.67	100	99.85
Nusa Tenggara Barat	99.89	99.87	100	99.97
Nusa Tenggara Timur	99.22	93.41	99.73	95.74
Kalimantan Barat	99.82	96.88	99.96	98.55
Kalimantan Tengah	99.99	98.08	99.93	98.74
Kalimantan Selatan	100	99.84	100	99.85
Kalimantan Timur	99.96	99.69	100	99.92
Kalimantan Utara	99.95	98.57	100	98.56
Sulawesi Utara	99.97	99.47	99.96	99.76
Sulawesi Tengah	99.88	98.55	99.67	99.07
Sulawesi Selatan	99.82	99.53	99.96	99.53
Sulawesi Tenggara	99.98	99.65	100	99.81
Gorontalo	99.76	98.84	100	99.44
Sulawesi Barat	100	99.27	100	99.52
Maluku	99.86	93.99	99.81	95.16
Maluku Utara	100	96.18	100	98.16
Papua Barat	99.95	92.39	99.56	94.9
Papua Barat Daya	-	-	100	95.87
Papua	99.15	68.51	100	90.09
Papua Selatan	-	-	97.16	71.35
Papua Tengah	-	-	100	43.46
Papua Pegunungan	-	-	99.68	71.6

3.1 Python Code

The data obtained from the Badan Pusat Statistik was further processed using Python programming. The data underwent missing value and outlier checks through boxplot visualization. Furthermore, Two-Way ANOVA was implemented using Python programming to analyze differences in the average percentage of household electricity usage across urban and rural areas in Indonesia during 2023 and 2024. Python enables efficient processing and statistical evaluation of datasets, allowing an objective assessment of variance among multiple groups through computational techniques. This method supports rigorous hypothesis testing by calculating the F-statistic and corresponding p-value, thereby facilitating evidence-based conclusions regarding differences in group means. The Python code used is as follows:

```
# 1. Cek missing values
print("Missing values after filtering:")
print(df_filtered.isnull().sum())

# 2. Boxplot
df_melt = df_filtered.melt(id_vars=['Province'],
                           value_vars=['Urban_2023', 'Rural_2023', 'Urban_2024', 'Rural_2
                           var_name='Category',
                           value_name='Coverage')

df_melt = df_melt.dropna()

sns.set(style="whitegrid")
plt.figure(figsize=(12,7))
ax = sns.boxplot(x='Category', y='Coverage', data=df_melt, palette='pastel', showfliers=T
ax.set_title('Boxplot of Household Electricity Coverage (%) by Category (Filtered Data)',
ax.set_xlabel('Category', fontsize=14)
ax.set_ylabel('Coverage (%)', fontsize=14)
ax.grid(True, linestyle='--', alpha=0.7)
ax.set_ylim(96, 101)
plt.xticks(rotation=15)
plt.show()

# 3. Histogram
plt.figure(figsize=(12,7))
sns.histplot(data=df_melt, x='Coverage', hue='Category', multiple='stack', palette='paste
plt.title('Histogram of Household Electricity Coverage (%) by Category (Filtered Data)',
plt.xlabel('Coverage (%)', fontsize=14)
plt.ylabel('Count', fontsize=14)
plt.show()

# 4. Statdesk
print("\nStatistical Summary by Category (Filtered Data):")
print(df_melt.groupby('Category')['Coverage'].describe())
```

Figure 2. Python code program

3.2 Interpretation of Results

The results of data pre-processing by checking for missing values and outliers can be seen in Table 3 as follows. Based on Table 3, it can be referred that there are some missing values on data, while Figure 3 shows that there are some outliers. To address this, some observations that contains missing value were removed which is the data from provinces of DKI Jakarta, Papua Barat Daya, Papua Selatan, Papua Tengah, and Papua Pegunungan. The removal of data with missing values and outliers aims to maintain balance in terms of location and years factors. Subsequently, the data that has undergone the pre-processing stage is presented in the following histogram.

Table 3. Missing value and outliers observation

Data checking	Observation
Missing Value	DKI Jakarta, Papua Barat Daya, Papua Selatan, Papua Tengah, Papua Pegunungan
Outliers	Nusa Tenggara Timur, Maluku, Maluku Utara, Papua Barat, Papua Barat Daya, Papua, Papua Selatan, Papua Tengah, Papua Pegunungan

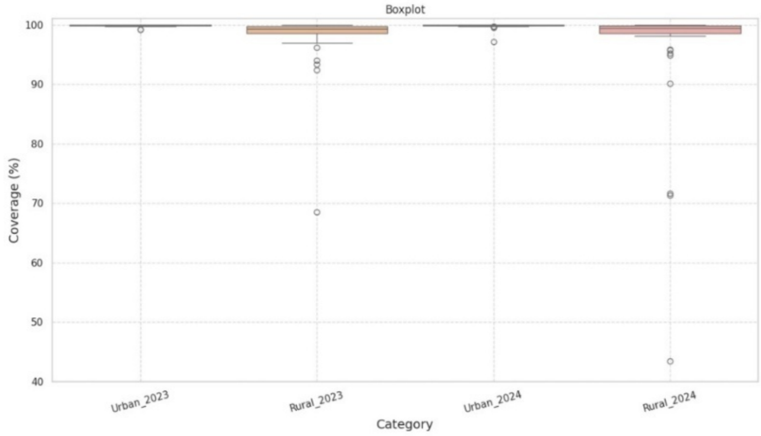

Figure 3. Boxplot

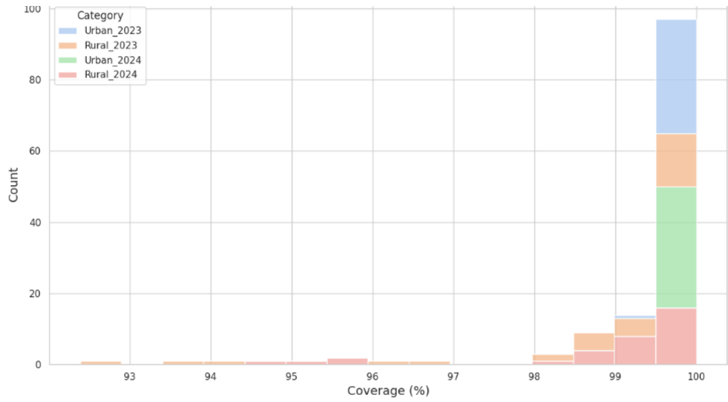

Figure 4. Histogram of household electricity usage by category

The histogram illustrates the distribution of household electrification percentages across Indonesian provinces in 2023 and 2024 and rural-urban locations. The x-axis represents the percentage of households using electricity as their main source of lighting, while the y-axis indicates the number of provinces falling within each percentage range. All the data distributions are heavily skewed to the right, with most provinces clustered near the 100% mark. This indicates that most regions in Indonesia have achieved near-universal access to electricity. The similarity in the shape and concentration of the distributions suggests that there was only a marginal improvement in electrification between 2023 and 2024. A few provinces remain at lower percentage levels, likely to correspond to remote or underdeveloped areas. Overall, the histogram supports the conclusion that while national electrification rates are high, targeted efforts are still needed to address the remaining disparities in specific regions.

Table 4. Descriptive statistics

Location	Mean		Standard deviation		Variance	
	2023	2024	2023	2024	2023	2024
Urban	99.9	99.9	0.195	0.110	0.038	0.012
Rural	97.7	98.8	5.59	2.05	31.2	4.21

The descriptive statistics reveal important insights into the distribution and progression of household electrification in Indonesia across urban and rural areas in 2023 and 2024. In urban areas, the mean electrification rate remained consistently high at 99.9% in both years, indicating near-universal access to electricity. Furthermore, the standard deviation and variance in urban areas decreased from 0.195 and 0.038 in 2023 to 0.110 and 0.012 in 2024, respectively. This suggests that not only is access nearly complete, but it has also become more uniform across provinces. In contrast, rural areas showed a notable improvement. The mean electrification rate increased from 97.7% in 2023 to 98.8% in 2024, reflecting progress in expanding access. More significantly, the standard deviation dropped from 5.59 to 2.05, and the variance from 31.2 to 4.21, indicating a substantial reduction in disparity among rural provinces. These findings suggest that while urban areas have reached a saturation point in electrification, rural areas are catching up, both in terms of average access and consistency. The narrowing gap between urban and rural electrification rates highlights the effectiveness of recent efforts to promote energy equity across regions.

Table 5. ANOVA table

Sources	df	Sum square	Mean square	F-value	Prob.
Location	1	89.2	89.18	10.049	0.002
Years	1	10.8	10.77	1.213	0.273
Residuals	129	1144.8	8.87		

The results of the two-way ANOVA in Table 5 provide important insights into the factors influencing household electrification usage across Indonesian provinces. The analysis revealed a statistically significant main effect of location, with an F-value of 10.049 and a p-value of 0.002. This indicates that there is a meaningful difference in electrification rates between urban and rural areas, with urban areas likely having higher and more consistent access to electricity. In contrast, the main effect of year was not statistically significant, suggesting that there was no substantial change in electrification rates between 2023 and 2024 across all locations. These findings highlight that while electrification levels are generally high, disparities between urban and rural areas persist, and the pace of change over time has not been sufficient to significantly alter this gap. Therefore, targeted efforts are still needed to accelerate rural electrification and ensure more equitable access across regions.

The ANOVA results, which show a statistically significant difference in electrification rates between urban and rural areas, underscore the urgent need for more targeted energy policies. This disparity highlights that while urban areas have nearly achieved universal access, rural regions still lag behind, indicating a persistent inequality. Policymakers should interpret these findings as a call to strengthen interventions in rural electrifications such as expanding off-grid renewable energy programs, increasing subsidies for rural infrastructure, and enhancing monitoring systems to ensure consistent service delivery. By addressing these gaps, the government can support a more equitable energy transition and accelerate progress toward national and global energy access targets, particularly those outlined in Sustainable Development Goal 7.

Table 6. Tukey's test

Sources	diff	lower	upper	p adj
Location	1.644	0.618	2.670	0.002
Years	0.571	-0.455	1.597	0.273

The post-hoc comparisons provide further insight into the specific group differences following the Two-Way ANOVA. The main effect of location (urban vs rural) was statistically significant, with a mean difference of 1.644, indicating that, on average, urban areas had significantly higher household electrification rates than rural areas. In contrast, the main effect of year (2023 vs 2024) was not significant (mean difference = 0.571, p = 0.273), reaffirming that there was no substantial change in electrification rates over time.

4. Conclusion

The findings of this study provide a detailed picture of household electrification usage across Indonesia's provinces, with a focus on differences between urban and rural areas over 2023 and 2024. Descriptive statistics revealed that urban areas consistently achieved near-universal access to electricity, with a mean electrification rate of 99.9% in both years and minimal variation across provinces. In contrast, rural areas showed a meaningful improvement, with the average rate increasing from 97.7% in 2023 to 98.8% in 2024. More notably, the standard deviation and variance in rural areas decreased significantly, indicating a reduction in disparities and a more equitable distribution of electricity access.

The histogram analysis supported these findings, showing that most provinces were clustered near 100% electrification, with only a few outliers at lower levels. This suggests that while national electrification is high, the remaining challenges are concentrated in specific regions. The two-way ANOVA further confirmed that the main effect of location (urban vs rural) was statistically significant, indicating a persistent gap in access between the two locations. However, the overall change from 2023 to 2024 was not substantial and that the urban-rural gap remained relatively stable.

Based on these findings, several policy recommendations can be made. First, while urban electrification is effectively complete, rural areas still require targeted interventions to close the remaining gap. Programs should prioritize provinces that consistently fall below the national average, particularly those in remote or underdeveloped regions. Second, the lack of significant year-over-year improvement suggests a need to evaluate the effectiveness and reach of current electrification initiatives. Finally, future policies should not only aim to expand access but also ensure reliability and quality of electricity supply, especially in rural settings where infrastructure may be less robust. Then, out suggestion for further research is to conduct further analysis by addressing missing values and outliers using other techniques such as imputation.

References

[1] Kementerian ESDM, "Upaya Peningkatan Rasio Elektrifikasi dan Pemerataan Listrik Masyarakat Melalui BPBL," 2023. [Online]. Available: https://www.esdm.go.id/id/berita-unit/direktorat-jenderal-ketenagalistrikan/upaya-peningkatan-rasio-elektrifikasi-dan-pemerataan-listrik-masyarakat-melalui-bpbl.

[2] R. Primadya, "PT PLN (Persero)," 15 01 2025. [Online]. Available: https://web.pln.co.id/cms/media/siaran-pers/2025/01/wujud-negara-hadir-pemerintah-dan-pln-berhasil-listriki-9992-persen-desa-di-seluruh-indonesia/.

[3] Badan Pusat Statistik, "Persentase Rumah Tangga menurut Provinsi, Klasifikasi Desa, dan Sumber Penerangan Utama dari Listrik (Persen), 2024," Badan Pusat Statistik, 2024.

[4] UN.ESCAP, "Energy transition pathways for the 2030 agenda : SDG7 roadmap for Indonesia," 2020. [Online]. Available: https://hdl.handle.net/20.500.12870/260..

[5] Sansuadi, "Evaluasi Ketimpangan Pembangunan Infrastruktur Ketenagalistrikan di Indonesia: Dampak dan Upaya Pemerataan Akses Energi," *Journal of Innovation Research and Knowledge,* vol. 4, no. 11, pp. 8209-8224, 2025.

[6] A. P. Muyasyaroh, "'Just' access to electricity: Energy justice in Indonesia's rural electrification (LISDES) program," dalam *IOP Conf. Ser.: Earth Environ. Sci.*, 2023.

[7] H. J. Seltman, Experimental Design and Analysis, Pennsylvania: Carnegie Mellon University, 2018.

[8] S. F. Sihotang, "Analisis Hasil Belajar Mahasiswa Universitas Potensi Utama Menggunakan Two Way Anova," *MES: Journal of Mathematics Education and Science,* vol. 8, no. 1, pp. 173-179, 2022.

[9] Indriyani, U. Rizqi dan U. Mahmudah, "Bagaimana Kreativitas dan Keaktifan Mahasiswa Mempengaruhi Pemahaman Materi Abstrak Matematika Melalui E-Learning," *Al Khawarizmi: Jurnal Pendidikan dan Pembelajaran Matematika,* vol. 4, no. 2, pp. 112-131, 2020.

[10] T. Noryanti, Z. A. MZ dan H. Nufus, "Pengaruh Penerapan Metode Tutor Sebaya terhadap Kemampuan Pemecahan Masalah Matematis Berdasarkan Kemampuan Awal Matematis," *Jurnal Pijar MIPA,* vol. 14, no. 3, pp. 102-107, 2019.

[11] S. Sharma, P. Singh dan S. Kumar, "Responses of Soil Carbon Pools, Enzymatic Activity, and Crop Yields to Nitrogen and Straw Incorporation in a Rice-Wheat Cropping System in North-Western India," *Frontiers in Sustainable Food System,* vol. 4, pp. 1-19, 2020.

[12] W. Junthopas dan C. Wongoutong, "Comparison of Listwise Deletion and Imputation Methods for Handling a Single Missing Response Value in a Central Composite Design," *Thailand Statistician,* vol. 20, no. 3, pp. 545-561, 2022.

[13] D. Varma, A. Nehansh dan P. Swathy, "Data Preprocessing Toolkit : An Approach to automate data preprocessing," *International Journal of Scientific Research in Engineering and Management (IJSREM),* vol. 07, no. 03, pp. 1-5, 2023.

[14] A. M. Sharifnia, D. E. Kpormegbey, D. K. Thapa dan M. Cleary, "A Primer of Data Cleaning in Quantitative Research: Handling Missing Values and Outliers," *Journal of Advanced Nursing,* pp. 1-6, 2025.

[15] D. C. Montgomery, Design and Analysis of Experiments 9th Edition, 8 penyunt., Arizona: Wiley, 2019.

[16] C. E. Agbangba, E. S. Aide, H. Honfo dan R. G. Kakai, "On the use of post-hoc tests in environmental and biological sciences: A critical review," *Heliyon,* vol. 10, no. 3, pp. 1-12, 2024.

[17] Institute for Essential Services Reform (IESR), "Indonesia Energy Transition Outlook 2025: Navigating Indonesia's Energy Transition at the Crossroads – A Pivotal Moment for Resdefining the Future (Vol. 5)," IESR, Jakarta, 2025.