MJP Journal of Education and Teaching Learning
*Vol.4, No.1 (2026): 1244-1252*
*https://jurnalilmiah.co.id/index.php/MJPJETL*

# Vehicle Adoption Trend Analysis Based on Engine Displacement Segmentation Using Data Mining on Samsat Registration Data

**[1)]Muhammad Rafi Herdian, [2)]Fathoni**

[1,2)] Sriwijaya University, Palembang, Indonesia
Email: Rafiherdian06@gmail.com

**Corespondence Outhor**: Rafiherdian06@gmail.com

**Informasi Artikel**

**Key Word:**
Vehicle Registration; Mann-Kendall; STL Decomposition; Trend Analysis; Data Mining.

## ABSTRACT

Indonesia's South Sumatra province generates substantial four-wheeled vehicle registration activity, yet the administrative records that document this activity have rarely been subjected to systematic analytical scrutiny. This study draws on Samsat vehicle registration data from 2025 and the first quarter of 2026 to examine how new vehicle adoption patterns differ across engine displacement (CC) segments. The dataset spanned 15 Excel files: Dataset A covered the full year 2025 (17,867 units), while Dataset B captured Q1 2026 (6,190 units). Because column naming conventions varied across files, a keyword-matching-based preprocessing pipeline was developed to handle integration. The analysis combined the Mann-Kendall Trend Test, Sen's Slope Estimator, STL Decomposition, Kruskal-Wallis Test, and Mean Absolute Percentage Error (MAPE) as an out-of-sample projection benchmark.

The Mann-Kendall test returned no statistically significant monotonic trend ($p = 0.7105$), which indicates that the proportional split between engine displacement categories stayed broadly stable over the observation window. Low-displacement vehicles (⩽1,500 cc) held a consistent majority, averaging 66.10% of monthly registrations, throughout the period. STL Decomposition revealed a seasonal component responsible for roughly 50.02% of total data variance; the low-displacement segment peaked in November, while the high-displacement segment peaked in March. Projection accuracy, measured by MAPE, came in at 14.46% for the low-displacement segment (Accurate) and 37.92% for the high-displacement segment (Moderately Accurate). Based on these results, Samsat's administrative data proves to be highly useful for practical applications. Regional governments can utilize this data to improve public service planning and estimate regional own-source revenue (PAD) more accurately.

**Informasi Artikel**

**Kata Kunci:**

Registrasi Kendaraan; Mann-Kendall; Dekomposisi STL; Analisis Tren; Penambangan Data.

## ABSTRAK

Provinsi Sumatra Selatan menghasilkan aktivitas registrasi kendaraan roda empat yang cukup besar, namun catatan administratif yang mendokumentasikan aktivitas tersebut masih jarang diteliti secara analitis dan sistematis. Penelitian ini memanfaatkan data registrasi kendaraan Samsat dari tahun 2025 dan kuartal pertama tahun 2026 untuk menguji perbedaan pola adopsi kendaraan baru berdasarkan segmen kapasitas mesin (CC). Dataset yang digunakan mencakup 15 berkas Excel: Dataset A mencakup keseluruhan tahun 2025 (17.867 unit), sedangkan Dataset B mencakup kuartal pertama (Q1) tahun 2026 (6.190 unit). Karena konvensi penamaan kolom yang bervariasi antarberkas, sebuah alur prapemrosesan (*preprocessing pipeline*) berbasis pencocokan kata kunci (*keyword-matching*) dikembangkan untuk menangani integrasi data. Analisis yang digunakan mengombinasikan Uji Tren Mann-Kendall (*Mann-Kendall Trend Test*), Estimator Kemiringan Sen (*Sen's Slope Estimator*), Dekomposisi STL (*STL Decomposition*), Uji Kruskal-Wallis (*Kruskal-Wallis Test*), dan *Mean Absolute Percentage Error* (MAPE) sebagai tolok ukur proyeksi luar sampel (*out-of-sample projection*). Hasil Uji Mann-Kendall menunjukkan tidak adanya tren monoton yang signifikan secara statistik ($p$ = 0,7105), yang mengindikasikan bahwa pembagian proporsional antar-kategori kapasitas mesin secara umum tetap stabil selama periode pengamatan. Kendaraan berkapasitas mesin rendah (≤1.500 cc) secara konsisten mendominasi sebagai mayoritas dengan rata-rata **66,10%** dari total registrasi bulanan sepanjang periode tersebut. Dekomposisi STL menunjukkan adanya komponen musiman yang berkontribusi sebesar kurang lebih **50,02%** terhadap total varians data; segmen kapasitas mesin rendah mencapai puncaknya pada bulan November, sementara segmen kapasitas mesin tinggi memuncak pada bulan Maret. Akurasi proyeksi yang diukur dengan MAPE berada pada angka **14,46%** untuk segmen kapasitas mesin rendah (Akurat) dan **37,92%** untuk segmen kapasitas mesin tinggi (Cukup Akurat). Berdasarkan hasil penelitian ini, data administratif Samsat terbukti sangat bermanfaat untuk aplikasi praktis. Pemerintah daerah dapat memanfaatkan data tersebut untuk mengoptimalkan perencanaan pelayanan publik dan mengestimasi Pendapatan Asli Daerah (PAD) secara lebih akurat

***Article History***
Received: 15/05/2026
Revised: 28/06/2026
Accepted: 21/07/2026

✉ ***Corresponding Author:*** (1) Muhammad Rafi Herdian, (2) Department, (3) Institution of Sriwijaya University, (4) Palembang, Indonesia, (5) Email: Rafiherdian06@gmail.com

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.
Copyright © 2026 – [1)] Muhammad Rafi Herdian, [2)] Fathoni

## INTRODUCTION

The automotive industry in Indonesia is a massive sector that involves various stages, starting from component manufacturing to distribution and after-sales services. This extensive network employs millions of people across the country. According to Gaikindo data, Indonesia consistently holds the position of the second-largest automotive market in Southeast Asia [1]. That growth, however, is far from evenly distributed. Regions outside Java, South Sumatra among them, display vehicle adoption characteristics that have attracted relatively little scholarly attention, despite being economically significant. South Sumatra's economic base rests heavily on mining, plantation agriculture, and trade; these sectors collectively generate strong demand for four-wheeled vehicles, whether for personal mobility or as everyday instruments of commercial activity.

Furthermore, engine displacement (CC) segmentation is not only relevant for vehicle sales but also affects regional fiscal policy. It has direct consequences for Motor Vehicle Transfer Tax (BBN) and Annual Vehicle Tax (PKB) revenues flowing into regional budgets [2]. Low-displacement vehicles (≤1,500 cc), particularly those in the Low Cost Green Car (LCGC) and mid-range MPV segments, continue to dominate the national market, driven by their affordability and lower operating costs [3]. High-displacement vehicles (>1,500 cc) have been expanding their presence, though this growth tends to concentrate in areas with higher purchasing power and more intensive economic activity [4].

Prior research has approached the Indonesian automotive market dynamics from several angles. Nugroho and Setiawan [5] analyzed national vehicle sales trends using multiple linear regression on Gaikindo data, though their work could not be disaggregated to the provincial level. Halim and Permana [6] applied ARIMA models for ASEAN-region volume projections and confirmed that seasonal components matter considerably in such data. Rahmawati et al. [7] segmented dealer data in Java via K-Means clustering, but without drawing on administrative registration records. In adjacent fields, the Mann-Kendall test has established itself in hydrological and environmental time series analysis [8], while STL Decomposition has appeared in studies of regional economic indicators [9]. Urban transportation datasets have also begun to receive data mining attention [10][11]. As far as the authors can determine, no study has yet applied this particular combination of methods specifically to provincial-level Samsat registration data in Indonesia.

That gap motivates the present study. Three questions guide the analysis: whether any discernible shift in consumer preference occurred between the low- and high-displacement segments over 2025; what seasonal patterns emerge from the registration data; and how reliably those patterns project forward when tested against Q1 2026 actuals. Answers to these questions are expected to offer practical input for Samsat, regional governments, and automotive businesses operating in South Sumatra.

The data source was new four-wheeled vehicle registration records from authorized dealers, as documented by Samsat South Sumatra and stored in Microsoft Excel (.xlsx) format. Table 1 summarizes how the dataset was partitioned according to the analytical function.

Each file contained 15 columns: Payment Date (Tgl Bayar), Vehicle Type (Jenis), Brand (Merk), Model, Manufacturing Year (Thn Buat), Assembly Year (Thn Rakit), Engine Displacement (CC), Number of Wheels (Jml Roda), and the tax-related fields BBN Pokok, BBN Denda, Jml BBN, PKB Pokok, PKB Denda, and JML PKN. Of these, Tgl Bayar served as the temporal dimension and CC as the segmentation variable.

## METHOD

The study proceeded through four sequential stages: data collection and integration, preprocessing, trend and seasonal analysis, and model validation. All computations were carried out in Python 3.10+ using the pandas, pymannkendall, statsmodels, scipy, and matplotlib libraries [14].

Preprocessing followed the same five-step procedure for both datasets. The first step was multi-file integration via keyword-matching-based column detection, designed to handle naming variations across files, a common challenge in heterogeneous data integration [15]. Next, the Tgl Bayar column was parsed into a datetime format using the dayfirst=True parameter, consistent with the DD/MM/YYYY format of the source data.

CC values were then segmented according to those criteria. Finally, the data were aggregated by month, and the proportion of each segment relative to total monthly registrations was computed:

$$Proportion\_Low(t) = Volume\_Low(t) / (Volume\_Low(t) + Volume\_High(t))$$

Working with proportions rather than absolute volumes eliminates the confounding effect of overall market fluctuations, ensuring the analysis captures genuine shifts in consumer preference rather than swings in aggregate market size [16].

The Mann-Kendall Trend Test is a non-parametric method for detecting monotonic trends in time series, valued primarily because it does not require the data to follow a normal distribution [17]. The test statistic S is computed by comparing all pairwise observations; the null hypothesis ($H_0$) of no trend is rejected when $p < 0.05$. Sen's Slope Estimator [18] complements the test by quantifying the rate of change as the median slope across all observation pairs, a procedure considerably more robust to outliers than conventional regression-based estimators. The pymannkendall library in Python was used for both.

STL (Seasonal and Trend decomposition using Loess) [19] separates a time series into three components: trend T(t), seasonal S(t), and residual R(t), expressed as $Y(t) = T(t) + S(t) + R(t)$. The robust=True parameter was activated to reduce sensitivity to extreme values, and the period parameter was calibrated to the number of available observations. The statsmodels.tsa.seasonal.STL function handled implementation.

The Kruskal-Wallis Test is the non-parametric counterpart to one-way ANOVA [20]. Here, it was used to assess whether inter-month group differences identified through STL reached statistical significance, with $H_0$ rejected at $p < 0.05$.

Q1 2026 proportion projections were produced by extrapolating Sen's Slope from the December 2025 baseline: $\hat{Y}(t) = Y(\text{Dec } 2025) + \beta \times k$, where $k = 1, 2, 3$ correspond to January, February, and March 2026

## RESULTS AND DISCUSSION

Of the 15 Excel files processed, 11 loaded successfully in full. Four files, corresponding to February, April, June, and August 2025, lacked the mandatory CC and Tgl Bayar columns and were excluded from the integration pipeline. Dataset A, therefore, covers only eight months: January, March, May, July, September, October, November, and December 2025. Validation removed 650 rows with CC values outside the valid range (< 500 or > 8,000); no missing dates or empty CC entries were found. Table 4 presents the final segmentation distribution.

| CC Segment | Criterion | No. of Units | Proportion | Status |
|---|---|---|---|---|
| Low (≤1,500 cc) | ≤1,500 cc | 16,036 | 66.58% | Valid |
| High (>1,500 cc) | >1,500 cc | 8,021 | 33.30% | Valid |
| Invalid | < 500 or > 8,000 | 650 | - | Excluded |
| Total Valid (A+B) | - | 24,057 | 100% | - |

The dominance of low-displacement vehicles at 66.58% of total registrations aligns with broader national trends, where

E-ISSN: 2986-3015
P-ISSN: 2986-3791



LCGC and entry-level MPV segments continue to account for the majority of automotive sales [3][4]. This figure reflects the enduring role that affordability and fuel efficiency play in vehicle choice within the region.

Table 5 summarizes the monthly time series for Dataset A. Across eight valid months, the combined total reached 17,867 units and averaged 2,233 units per month. December 2025 recorded the highest single-month volume at 3,995 units, a peak plausibly linked to year-end dealer promotions and the acceleration of government fleet procurement before regional fiscal budgets closed.

However, the trend observed in Dataset B shows a slightly different pattern (Table 6). The low-displacement proportion in January 2026 reached 75.83%, well above the 2025 average of 66.10%. By February and March 2026, the figures had settled back into a more typical range, 66.90% and 69.82%, respectively, suggesting the January spike was transitory.

| Period | Low CC | High CC | Total | Prop. Low | Prop. High |
|---|---|---|---|---|---|
| Jan 2026 | ~276 | ~88 | ~364 | 75.83% | 24.17% |
| Feb 2026 | ~712 | ~352 | ~1,064 | 66.90% | 33.10% |
| Mar 2026 | ~703 | ~304 | ~1,007 | 69.82% | 30.18% |

| Segment | Tau (τ) | p-value | Trend Direction | Sen's Slope | Significant | $H_0$ |
|---|---|---|---|---|---|---|
| Low CC (≤1,500 cc) | −0.1429 | 0.7105 | No trend | −0.004868 | No | Not Rejected |
| High CC (>1,500 cc) | +0.1429 | 0.7105 | No trend | +0.004868 | No | Not Rejected |

With p-values of 0.7105, far above the $\alpha = 0.05$ threshold, $H_0$ was not rejected for either segment. No statistically significant monotonic trend in adoption proportions appeared across the eight months of observation. Sen's Slope of −0.004868 per month for the low-displacement segment translates to a decline of roughly 0.49 percentage points per month, a change too small to carry practical meaning [22]. The symmetric τ values, −0.1429 for low CC and +0.1429 for high CC, reflect opposing tendencies of equal magnitude, neither of which approaches significance.

Even though the statistical trend is not significant, the stability of the market composition provides valuable insights. Market composition in South Sumatra was broadly stable throughout 2025, with no pronounced shift between segments. The low-displacement proportion consistently ranging between 61% and 73% points to relatively durable consumer preferences within this observation window [3]

STL was applied with period=7, reduced from the standard period=12 because n=8, making the maximum allowable period n−1=7, and robust=True. Table 8 shows each component's contribution to total data variance.

| Segment | Trend (%) | Seasonal (%) | Residual (%) | Peak Month | Trough Month |
|---|---|---|---|---|---|
| Low CC (≤1,500 cc) | 49.98% | 50.02% | 0.00% | November (+0.0536) | March |
| High CC (>1,500 cc) | 49.98% | 50.02% | 0.00% | March (+0.0511) | November |

Trend and seasonal components each account for approximately 50% of the total variance in both segments. A residual of 0.00% indicates the model captured the entirety of data variation, a strong fit [19]. The seasonal pattern is precisely inverse: when the low-displacement segment peaks in November, the high-displacement segment troughs, and vice versa in March. This is a mathematical consequence of the sum-to-one complementary relationship between the two proportions.

The November peak in low-displacement registrations likely reflects purchasing behavior among middle-class consumers, who tend to commit to vehicle investments after major expenditure periods in the first half of the year have passed. The March peak in high-displacement registrations may reflect government budget disbursement at the start of the fiscal year and the release of performance-based allowances, which

**MJP Journal of Education and Teaching Learning**
Vol.4, No.1 July (2026): 1244-1252

prompt a segment of buyers toward higher-value vehicles [23].

| Segment | H-Statistic | p-value | Conclusion | $H_0$ |
|---|---|---|---|---|
| Low CC (≤1,500 cc) | 7.0000 | 0.4289 | Seasonal pattern not significant | Not Rejected |
| High CC (>1,500 cc) | 7.0000 | 0.4289 | Seasonal pattern not significant | Not Rejected |

The Kruskal-Wallis test returned p = 0.4289, leaving $H_0$ unrejected. This result requires careful interpretation. The non-significant outcome is primarily attributable to the limited statistical power that comes with only eight monthly observations, not to a genuine absence of seasonal structure [20]. STL's finding that seasonal variation accounts for 50.02% of total variance provides strong corroborating evidence for this reading. Had the dataset covered at least 24 months, the Kruskal-Wallis result would very likely have reached significance.

| Month | Proj. L | Actual L | APE L | Proj. H | Actual H | APE H | Note |
|---|---|---|---|---|---|---|---|
| Jan 2026 | 60.94% | 75.83% | 19.64% | 39.06% | 24.17% | 61.63% | Anomaly |
| Feb 2026 | 60.45% | 66.90% | 9.64% | 39.55% | 33.10% | 19.49% | Normal |
| Mar 2026 | 59.97% | 69.82% | 14.11% | 40.03% | 30.18% | 32.64% | Normal |
| MAPE | - | - | 14.46% | - | - | 37.92% | - |

The low-displacement segment achieved a MAPE of 14.46% (Accurate), confirming that the 2025 trend model provides a reasonably reliable basis for projecting this segment's proportions into Q1 2026. The high-displacement segment recorded 37.92% (Moderately Accurate), largely driven by the January 2026 anomaly: the actual high-displacement proportion (24.17%) deviated sharply from the projection (39.06%), producing an APE of 61.63%. February and March moderated substantially to 19.49% and 32.64%, respectively, suggesting January was a temporary disruption rather than a structural reversal.

At least three explanations can be offered for the January anomaly. The very low registration volume that month (~364 units) likely reflects a post-year-end holiday transition during which market activity had not yet recovered, and monthly proportions were statistically unstable. Consumers who take on installment financing at the start of the year also tend to favor the LCGC segment, given its lighter monthly payment obligations [3]. A third possibility is that some high-displacement units ordered in late 2025 were not registered until February 2026, temporarily suppressing that segment's January share.

By combining the results from these three analytical methods, a consistent trend can be observed. The Mann-Kendall result (p = 0.7105) establishes that market composition was stable across the observation period, an outcome directly useful for Samsat in projecting BBN revenues from historical patterns. The consistent dominance of low-displacement vehicles, ranging between 61% and 73%, reflects the purchasing power profile of South Sumatra's population, which remains predominantly lower-to-middle income; this is consistent with automotive market characteristics documented across other non-Java provinces [4][5].

The seasonal patterns from STL carry concrete operational implications. The November peak in low-displacement registrations can serve as an early signal for LCGC and MPV dealers to build adequate inventory from October onwards. For premium segment dealers, demand tends to strengthen around February and March, aligning with the high-displacement registration peak. One observation worth noting: December 2025 recorded the highest overall volume at 3,995 units, yet the low-displacement proportion fell to its lowest point at 61.43%. This likely reflects a disproportionate absolute increase in high-displacement registrations that month, potentially driven by year-end promotional campaigns among premium dealers [23].

From a methodological standpoint, the combination of Mann-Kendall, Sen's Slope, and STL Decomposition can yield interpretively meaningful findings even when data spans only eight months. The statistical power constraints imposed by a small sample must be acknowledged, but they do not invalidate the substantive conclusions. The MAPE of 14.46% for the low-displacement segment provides empirical support for the position that linear

**MJP Journal of Education and Teaching Learning**
Vol.4, No.1 July (2026): 1244-1252

extrapolation via Sen's Slope is sufficiently robust for short-term projections, particularly in segments exhibiting structurally stable behavior [21].

## CONCLUSION

This study analyzed vehicle adoption trends by engine displacement segmentation using Samsat registration data from South Sumatra across 2025. The Mann-Kendall test found no statistically significant monotonic trend ($p = 0.7105$), indicating that the relative market composition between the low- and high-displacement segments stayed broadly stable throughout the observation period. Low-displacement vehicles consistently dominated, averaging 66.10% of registrations, a figure in line with the automotive market profile of provinces outside Java.

STL Decomposition identified a seasonal component accounting for roughly 50% of total variance, with the low-displacement segment peaking in November and the high-displacement segment in March. The Kruskal-Wallis test did not return formal statistical significance; this outcome is more plausibly attributed to the limited sample of $n=8$ than to a genuine absence of seasonality. The seasonal patterns nonetheless retain practical value for dealer inventory planning and Samsat service capacity management.

Projection accuracy came in at a MAPE of 14.46% for the low-displacement segment (Accurate) and 37.92% for the high-displacement segment (Moderately Accurate). The January 2026 anomaly was confirmed as transitory, with proportions returning to normal levels in February and March.

The results show that there is considerable analytical potential in Samsat registration data that has not been fully utilized. There are also several limitations that need to be noted in this study: four of the twelve 2025 monthly files were unavailable; the eight-month observation window constrained statistical power; and linear extrapolation cannot capture more complex underlying dynamics. Future research should work from a complete 12-month dataset, extend the observation window to at least 24 months, explore advanced forecasting methods such as LSTM or Prophet, and conduct comparative analyses across other provinces in Sumatra [24].

## REFERENCES

[1] GAIKINDO, "Indonesian Automobile Industry Data." Accessed: May 11, 2026. [Online]. Available: https://www.gaikindo.or.id/indonesian-automobile-industry-data/

[2] Directorate General of Fiscal Balance, "Guidelines for the Collection of Motor Vehicle Tax and Motor Vehicle Transfer Tax Institution: Ministry of Finance of the Republic of Indonesia." Accessed: May 11, 2026. [Online]. Available: https://djpk.kemenkeu.go.id/?

[3] A. Meydian dan T. M. Z. S. Machmud, "Kebijakan Low Cost Green Car dalam Industri Otomotif Nasional," J. Ekon. Indones., vol. 10, no. 3, pp. 257–279, 2022, doi: 10.52813/jei.v10i3.149.

[4] F. Teknik and U. A. Jambi, "Imti Tsalil Amri, Santoso," Penerapan Data Min. Untuk Menganalisis Tingkat Kelarisan Penjualan Aksesori Mobil. Menggunakan Algoritma. K-Means, vol. 1, no. 2, pp. 25–30, 2018.

[5] F. G. Walelang, C. E. J. C. Montolalu, and D. Hatidja, "Penerapan Metode Autoregressive Integrated Moving Average ( ARIMA ) dalam Model Intervensi Fungsi Step untuk Memprediksi Penjualan Mobil di PT . Hasjrat Abadi Cabang Tendean Manado," Indones. J. Intell. Data Sci., vol. 01, no. 01, pp. 35–42, 2023.

[7] W. Afifi, D. R. Nastiti, and Q. Aini, "Clustering K-Means pada Data Ekspor (Studi Kasus: PT Gaikindo),"

Simetris J. Tek. Mesin, Elektro dan Ilmu Komput., vol. 11, no. 1, pp. 45–50, 2020, doi: 10.24176/simet.v11i1.3568.

[8] H. B. Mann, "Nonparametric Tests Against Trend," Econometrica, vol. 13, no. 3, pp. 245–259, Jul. 1945, doi: https://www.jstor.org/stable/1907187.

[9] Robert B Cleveland, William S. Cleveland, Jean E. McRae, and Irma Terpenning, "STL: A Seasonal-Trend decomposition Procedure Based on Loess," 1990.

[10] A. Hendrawan and N. P. Wijaya, "Data-Driven Traffic for Infrastructure Planning : An LSTM Approach Using Indonesian Road-Vehicle Trends," vol. 8, no. 2, pp. 1416–1432, 2026, doi: 10.63158/journalisi.v8i2.1516.

[11] E. Wijaya and Y. Tri Widayati, "Optimalisasi Performa Penjualan Di Pt Mustika Jaya Lestari Menggunakan Algoritma K-Nn Dan Arima (2021-2022) Optimizing Sales Performance At Pt Mustika Jaya Lestari Using K-Nn and Arima Algorithms (2021-2022)," J. Inf. Technol. Comput. Sci., vol. 7, no. 5, pp. 1628–1639, 2024.

[12] P. Candra Susanto, D. Ulfah Arini, D. Marlita, L. Yuntina, and E. Saribanon, "Mixed Methods Research Design Concepts: Quantitative, Qualitative, Exploratory Sequential, Embedded and Parallel Convergent," Int. J. Adv. Multidiscip., vol. 3, no. 3, pp. 471–485, 2024.

[13] A. R. Fadhilah, A. H. Nasution, and W. Monika, "Comparative Analysis of LSTM, FB Prophet, and Moving Average Methods for Fuel Sales Prediction: A Time Series Forecasting Approach," Build. Informatics, Technol. Sci., vol. 6, no. 3, pp. 1321–1331, 2024, doi: 10.47065/bits.v6i3.5877.

[14] Y. Kurnia, R. Arijanto, W. Layanda, and D. Surya Dwi Putra, "Web-Based Car Sales Prediction System Using the ARIMA (Autoregressive Integrated Moving Average) Model for Optimizing Automotive Marketing Strategies," Rubinstein Multidiscip. J., vol. 4, no. 1, 2025, doi: 10.31253/rubin.v4i1.4039.

[15] Y. Haruki, S. Ishikura, K. Demachi, and T. Hayashi, "Contextual Graph Embeddings: Accounting for Data Characteristics in Heterogeneous Data Integration," Proc. IEEE Int. Conf. Big Data, BigData, no. 2025, pp. 2843–2852, 2025, doi: 10.1109/BigData66926.2025.11400991.

[16] F. Sidqi, I. D. Sumitra, J. Dipati, U. No, and J. Barat, "5537-Article Text-17390-1-10-20210907".

[17] F. Aditya, E. Gusmayanti, and J. Sudrajat, "Rainfall trend analysis using Mann-Kendall and Sen's slope estimator test in West Kalimantan," IOP Conf. Ser. Earth Environ. Sci., vol. 893, no. 1, 2021, doi: 10.1088/1755-1315/893/1/012006.

[18] A. K. Gupta, "Estimation of the Mean and Standard Deviation of a Normal Population from a Censored Sample," Biometrika, vol. 39, no. 3/4, pp. 260–273, May 1952, doi: 10.2307/2334023.

[19] Rosmelina Deliani Satrisna, Aniq A. Rohmawati, and Siti Sa'adah, "Forecasting the COVID-19 Increment Rate in DKI Jakarta Using Non-Robust STL Decomposition and SARIMA Model," Int. J. Inf. Commun. Technol., vol. 7, no. 1, pp. 21–30, 2021, doi: 10.21108/ijoict.v7i1.554.

[20] R. A. K. Sherwani, H. Shakeel, W. B. Awan, M. Faheem, and M. Aslam, "Analysis of COVID-19 data using neutrosophic Kruskal Wallis H test," BMC Med. Res. Methodol., vol. 21, no. 1, pp. 1–7, 2021, doi: 10.1186/s12874-021-01410-x.

[21] S. A. Sinaga, "Implementasi Metode Arima (Autoregressive Integrated Moving Average) Untuk Prediksi Penjualan Mobil," J. Glob. Technol. Comput., vol. 2, no. 3, pp. 102–109, 2023, doi: 10.47065/jogtc.v2i3.4013.

[22] H. Khofaria, I. Iskandar, and Suhadi, "Analisis Curah Hujan Ekstrem di Kepulauan Maluku dan Hubungannya

dengan Fenomena ENSO Analysis of Extreme Rainfall in The Maluku Islands And its Relationship with The ENSO Phenomenon," J. Ilmu Fis. dan Pembelajarannya, vol. 8, no. 2, pp. 80–105, 2024.

[23] Y. A. Saputra and D. Mahdiana, "Analisis Penerapan Machine Learning , Deep Learning , dan Data Mining dalam Prediksi Penjualan di Industri Otomotif," Magister Ilmu Komputer , Fakultas Teknologi Informasi , Universitas Budi Luhur , Jakarta , Indonesia; Sistem Informasi , Fakultas Teknologi. Pendidik. dan Teknol. Indones., vol. 5, no. 6, pp. 1743–1755, 2025.

[24] I. G. T. Suryawan, I. K. N. Putra, P. M. Meliana, and I. G. I. Sudipa, "Performance Comparison of ARIMA, LSTM, and Prophet Methods in Sales Forecasting," Sinkron, vol. 8, no. 4, pp. 2410–2421, 2024, doi: 10.33395/sinkron.v8i4.14057.