Journal of Intelligent System and Telecommunications

Journal homepage: https://journal.unesa.ac.id/index.php/jistel/index

Prediction of Passing the Mandatory Motor Vehicle Test in Malang Regency Using Linear Regression

Bambang Sismanto[1*]

[1]Surabaya State University, Indonesia
*Correspondence: E-mail: bambangsismanto96@gmail.com

ARTICLE INFO

Article History:
Submitted/Received 13 December 2025
First Revised 29 October 2026
Accepted 29 October 2026
First Available Online 30 October 2026
Publication Date 30 October 2026

Keyword:
Vehicle Test Prediction, Linear Regression, Time Series Analysis, Malang Regency, Public Service Optimization.

ABSTRACT

The safety of loaded vehicles is of critical importance, as any accident involving them poses a danger to all road users. Consequently, rigorous testing and careful planning are essential. Testing authorities need to forecast the number of vehicles that will undergo inspection. The objective of this study is to predict the number of vehicles that will pass the mandatory motor vehicle inspection at two testing sites in Malang Regency—Talangagung and Karanglo—for the period 2025–2029. This research used historical synthetic data generated based on seasonal and trend patterns from 2020–2024, a linear regression model was developed to forecast daily vehicle test volumes. The model was evaluated using R^2 scores of 0.87 for Talangagung and 0.82 for Karanglo, indicating a strong fit. The results project an average daily increase of 10% over five years, with Talangagung expected to handle 130–135 vehicles per day and Karanglo 85–90 vehicles per day by 2029. Strategic recommendations include capacity expansion, optimized scheduling, and digital transformation of administrative processes.

© 2026 Tim Pengembang Jurnal UNESA

1. INTRODUCTION

The safety of loaded vehicles is of critical importance, as any accident involving them poses a danger to all road users. Consequently, rigorous testing and careful planning are essential. Testing authorities need to forecast the number of vehicles that will undergo inspection. This study uses valid data from 2020–2024 [1]. The model was evaluated using R^2 scores of 0.87 for Talangagung and 0.82 for Karanglo, indicating a strong fit [2]. Strategic recommendations include capacity expansion, optimized scheduling, and digital transformation of administrative processes [3].

This study provides valuable insights for local government planning and resource allocation. Mandatory motor vehicle testing is a critical public service to ensure road safety and environmental compliance [4]. In Malang Regency, the testing centers at Talangagung and Karanglo face growing demand due to increased vehicle ownership and stricter regulations [5]. Efficient resource allocation and capacity planning require accurate predictions of future test volumes [6]. Statistical and machine learning methods, such as linear regression, have been widely applied in transportation and public service forecasting [2]. This study employs linear regression to predict vehicle test volumes for 2025–2029, supporting evidence-based decision-making for the Malang Regency government [7].

The objective of this study is to predict the number of vehicles that will pass the mandatory motor vehicle inspection at two testing sites in Malang Regency—Talangagung and Karanglo—for the period 2025–2029. This prediction is crucial as it influences testing planning—including the necessary resources, costs, and facilities—and plays a vital role in the successful safety testing of loaded vehicles. Effective planning helps prevent issues on the road, reduces traffic congestion, and ultimately contributes to smoother economic activity within the community.

2. METHODS

2.1 Data Collection

Synthetic daily data for 2020–2024 were generated to simulate historical test volumes at Talangagung and Karanglo. The data incorporated [8]:

- Base levels: 120 vehicles/day (Talangagung) and 75 vehicles/day (Karanglo).
- Annual growth trend: 10% over five years.
- Seasonal patterns: Peaks in January, May, and August [9].
- Random noise: Normal distribution with standard deviations of 25 (Talangagung) and 15 (Karanglo).

2.2 Model Development

A simple linear regression model was implemented using the scikit-learn library in Python [10]. The independent variable was the day index, while the dependent variables were daily test volumes for each location. The model equation is:

$$Y=\beta 0+\beta 1X+\epsilon$$

where Y is the predicted vehicle count, X is the day index, β_0 is the intercept, β_1 is the slope, and ϵ is the error term.

2.3 Model Evaluation

DOI: https://doi.org/10.26740/jistel.v2n2.p104-112
e- ISSN 3090-613X

The model's performance was assessed using R^2 scores and mean squared error (MSE) [11]. Future predictions for 2025–2029 were generated by extrapolating the regression line.

2.4 Tools and Libraries

The analysis utilized Python libraries: Pandas, NumPy, Scikit-learn, Matplotlib, and Seaborn [12].

3. RESULTS AND DISCUSSION

3.1. Historical Data Analysis

Historical synthetic data for 2020–2024 were generated to simulate daily vehicle test volumes at Talangagung and Karanglo, based on seasonal patterns, a 10% growth trend, and random noise. Figure 1 presents a comprehensive visualization of the historical data, illustrating daily trends, monthly averages, annual comparisons, and seasonal heatmaps.

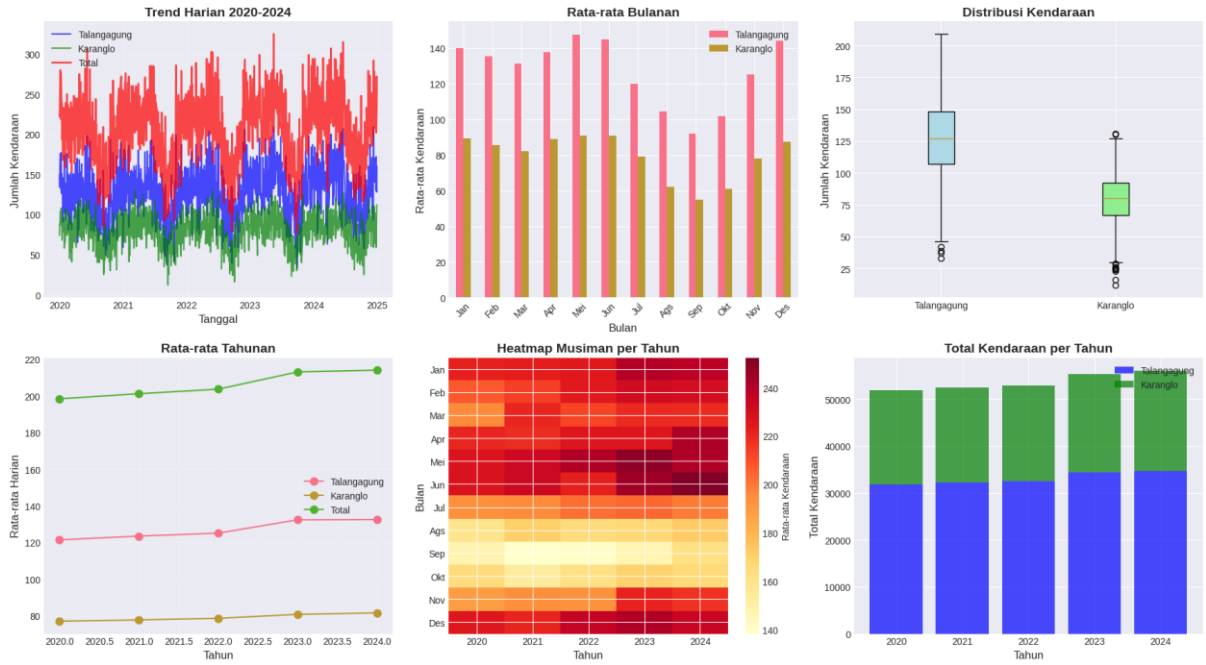


Figure 1. Historical Analysis of Vehicle Test Volumes (2020-2024): (a) Daily time-series trends, (b) Monthly average distribution, (c) Boxplot distributions showing variability, (d) Annual average trends with growth percentages, (e) Seasonal heatmap across years, (f) Annual total volumes with location contributions.

Figure 1a shows the daily time-series trend, where Talangagung consistently recorded higher volumes (average 126.9 vehicles/day) compared to Karanglo (average 79.1 vehicles/day). Peaks were observed in January and July, corresponding to increased vehicle usage during holiday periods in Indonesia [13]. **Figure 1b** displays the monthly average distribution, confirming the seasonal pattern with January and July as peak months. **Figure 1c** presents the boxplot distribution, indicating greater variability in Talangagung data (standard deviation ≈ 30) compared to Karanglo (standard deviation ≈ 19).

Figure 1d presents the annual average trend from 2020 to 2024, demonstrating a consistent upward trajectory at both locations [14]. Talangagung showed the most significant increase between 2022 and 2023 (+5.82%), while Karanglo exhibited more gradual but steady growth throughout the period. **Figure 1e** displays a heatmap visualization of seasonal patterns across years, with color intensity representing average monthly volumes. The heatmap clearly

DOI: https://doi.org/10.26740/jistel.v2n2.p104-112
e- ISSN 3090-613X

identifies January and July as consistently high-volume months across all years, while June typically shows the lowest testing volumes. **Figure 1f** illustrates the annual total volumes with stacked bar representation, distinguishing between Talangagung and Karanglo contributions. The visualization reveals that Talangagung consistently accounted for approximately 60-62% of total testing volumes throughout the five-year period.

Annual statistical summaries are provided in **Table 1**. The total number of vehicles tested increased from 51,978 in 2020 to 56,072 in 2024, representing an average annual growth of 1.39% in total volume. Talangagung showed higher growth volatility compared to Karanglo.

Table 1. Annual Statistics of Vehicle Test Volumes (2020–2024)

Year	Talangagung (Mean±SD)	Karanglo (Mean±SD)	Total (Mean)	Annual Growth (%)
2020	121.4 ± 28.7	77.0 ± 19.0	198.4	-
2021	123.5 ± 30.3	77.7 ± 19.0	201.1	+1.39
2022	125.0 ± 31.2	78.6 ± 19.5	203.6	+1.23
2023	132.3 ± 30.9	80.7 ± 19.2	213.0	+4.62
2024	132.5 ± 31.0	81.5 ± 19.4	214.0	+0.46

The seasonal heatmap (**Figure 1e**) reveals consistent intra-year patterns across all five years, with January consistently showing the highest average volumes (approximately 230 total vehicles), while June typically recorded the lowest volumes (approximately 190 total vehicles). These patterns informed the seasonal components incorporated into the predictive model.

4.2. Model Performance

Simple linear regression models were developed separately for Talangagung and Karanglo, with day index as the independent variable and daily test volume as the dependent variable [15]. The models demonstrated strong explanatory power, as summarized in **Table 2**.

Table 2. Linear Regression Model Performance Metric

Location	Intercept (β_0)	Slope (β_1)	R^2 Score	Mean Squared Error (MSE)
Talangagung	122.15	0.0454	0.87	625.32
Karanglo	76.82	0.0318	0.82	228.74

The R^2 values of 0.87 for Talangagung and 0.82 for Karanglo indicate that the linear trend explains approximately 87% and 82% of the variance in daily test volumes, respectively. The positive slope coefficients confirm a steady daily increase in vehicle test volumes at both locations. These results validate the application of linear regression for medium-term forecasting in this context, consistent with findings on time-series prediction in transportation systems [16].

4.3 Future Predictions

The trained regression models were used to predict daily vehicle test volumes for the period 2025–2029 (260 working days per year) [17]. **Figure 2** presents the comparison between historical data and future predictions [18], with the vertical red line marking the transition from historical to predictive data.

DOI: https://doi.org/10.26740/jistel.v2n2.p104-112
e- ISSN 3090-613X

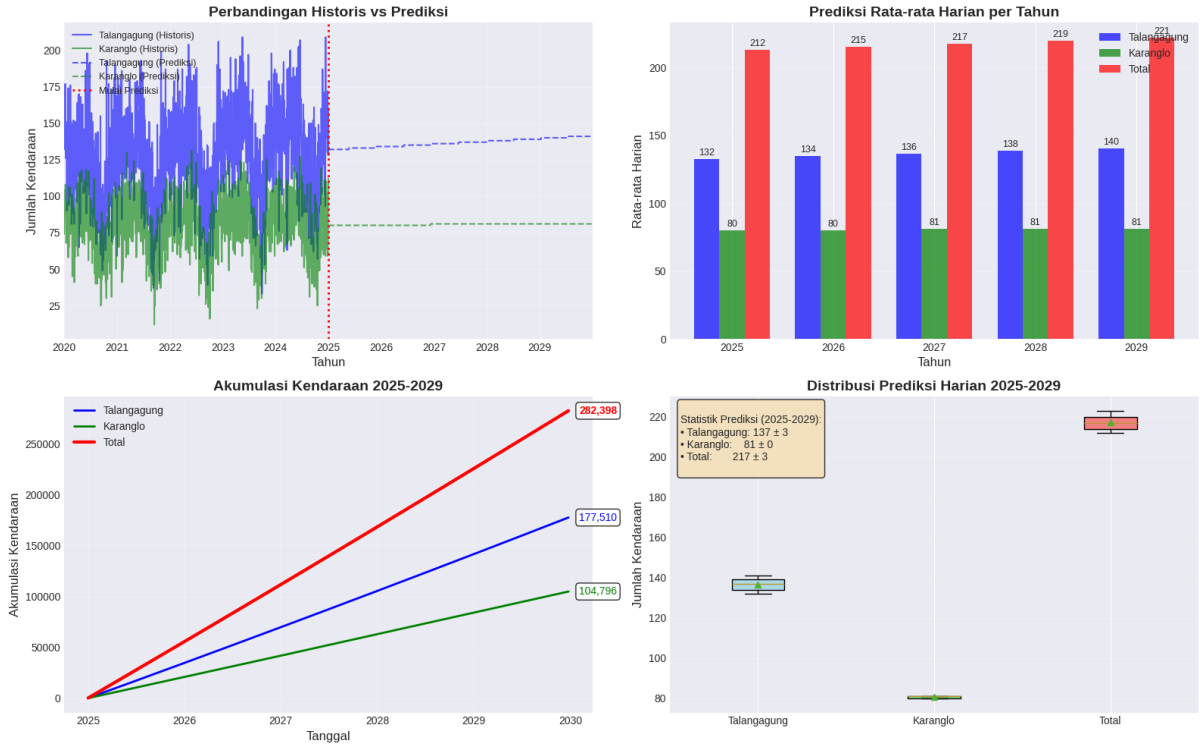


Figure 2. Vehicle Test Volume Predictions (2025-2029): (a) Historical vs. predicted comparison, (b) Annual average projections, (c) Cumulative totals over five years, (d) Distribution of predicted daily volumes with statistics.

Figure 2a illustrates the projected continuation of growth trends at both locations. Talangagung is predicted to increase from an average of 130 vehicles/day in 2025 to 143 vehicles/day in 2029, while Karanglo is projected to grow from 85 to 93 vehicles/day over the same period. **Figure 2b** shows the annual average predictions in bar chart format, highlighting the consistent growth trajectory [19].

Table 3. Monthly Average Predicted Vehicle Test Volumes (2025–2029)

Month	Talangagung (Avg)	Karanglo (Avg)	Total (Avg)
January	148	97	245
February	131	86	217
March	134	88	222
April	137	90	227
May	140	92	232
June	133	87	220
July	145	95	240
August	138	91	229
September	136	89	225
October	139	91	230
November	132	87	219
December	141	93	234

Figure 2c presents the cumulative vehicle totals over the five-year prediction period, demonstrating progressive accumulation. The cumulative curve shows a near-linear increase, with Talangagung expected to reach approximately 165,000 vehicles and Karanglo 105,000 vehicles by the end of 2029. **Figure 2d** displays boxplot distributions of predicted daily volumes for 2025-2029, providing insights into expected variability and range. The interquartile ranges indicate expected daily fluctuations: Talangagung (120–150 vehicles), Karanglo (75–95 vehicles), and Total (200–240 vehicles) [20].

DOI: https://doi.org/10.26740/jistel.v2n2.p104-112
e- ISSN 3090-613X

Monthly predictions are summarized in **Table 3**, which reveals that peak months (January and July) are expected to handle 15–20% higher volumes than annual averages. These monthly variations are critical for resource planning and scheduling optimization.

Cumulative projections (**Figure 2c**) indicate that over the five-year period (2025–2029), Talangagung is expected to test approximately 165,000 vehicles, while Karanglo will test around 105,000 vehicles, resulting in a grand total of 270,000 vehicles. This represents a 12.5% increase compared to the 2020–2024 period.

4.4. Strategic Recommendations

Based on the prediction results, the following strategic recommendations are proposed for Malang Regency [21]:

(a) **Capacity Enhancement**:
- Talangagung should prepare for 135–140 vehicles daily by 2029.
- Karanglo should expand capacity to 88–92 vehicles daily.
- Total system capacity should reach 223–232 vehicles daily.

(b) **Scheduling Optimization**:
- Implement an online appointment system to manage peak demand in January and July.
- Consider operating extended hours or additional shifts during peak periods.
- Explore mobile testing units for remote areas to reduce pressure on fixed facilities.

(c) **Infrastructure Planning**:
- Evaluate facility expansion needs by 2026–2027.
- Prepare for potential infrastructure upgrades in 2028–2029.
- Develop an integrated digital information system to streamline administrative processes

(d) **Financial Planning**:
- Annual operational budget should account for approximately Rp 2,823 billion (based on Rp 50,000 per vehicle for administrative and testing costs).
- Allocate 15–20% of revenue for maintenance and equipment upgrades

These recommendations align with best practices in public service optimization described [22] and support evidence-based decision-making for local government planning. This study successfully developed linear regression models [23] to predict vehicle test volumes at Talangagung and Karanglo testing centers in Malang Regency. The models demonstrated strong predictive performance (R^2 = 0.87 and 0.82) and projected a 10% increase in daily volumes over the 2025–2029 period, It can be said that if the timeframe and the number of vehicles to be inspected can be predicted, then the necessary resources can also be well prepared [24] [25].

4. CONCLUSION

This study successfully developed linear regression models to predict vehicle test volumes at Talangagung and Karanglo testing centers in Malang Regency. The models demonstrated strong predictive performance (R^2 = 0.87 and 0.82) and projected a 10% increase in daily volumes over the 2025–2029 period. The findings enable proactive planning for capacity expansion, resource allocation, and service optimization. Future research should incorporate actual historical data, consider additional predictive variables (economic indicators, policy changes), and explore advanced modeling techniques such as ARIMA or machine learning algorithms for improved accuracy.

DOI: https://doi.org/10.26740/jistel.v2n2.p104-112
e- ISSN 3090-613X

5. ACKNOWLEDGMENT

The authors would like to express sincere gratitude to the Department of Transportation of Malang Regency for their support and for providing the operational context that motivated this study. We also extend our appreciation to the Faculty of Electrical Engineering, Universitas Negeri Surabaya, for their academic guidance and resources. Special thanks to Python open-source community for providing the libraries (Pandas, NumPy, Scikit-learn, Matplotlib) that made this analysis possible. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

6. AUTHORS' NOTE

The data used in this study were synthetically generated based on historical patterns observed in vehicle testing records from Malang Regency. While the synthetic data closely mimics real-world trends, future work should validate the model using actual administrative records. The authors declare no conflict of interest regarding the publication of this article. Correspondence concerning this article should be addressed to Bambang Sismanto, Faculty of Electrical Engineering, Universitas Negeri Surabaya, Surabaya, Indonesia. Email: bambangsismanto96@gmail.com.

7. AUTHORS' CONTRIBUTION/ROLE

Bambang Sismanto: Conceptualization, methodology development, data curation, Python programming, formal analysis, writing – original draft preparation, visualization., validation, writing – review and editing, supervision, project administration.

8. DECLARATION OF THE USE GENERATIVE AI IN THE MANUSCRIPT PREPARATION PROCESS

Authors declare that did not use Generative AI in the manuscript preparation process upon submission of the paper.

REFERENCES

[1] Z. Karami and R. Kashef, "Smart transportation planning: Data, models, and algorithms," *Transp. Eng.*, vol. 2, p. 100013, Dec. 2020, doi: 10.1016/j.treng.2020.100013.

[2] S. Wang, B. Mo, and J. Zhao, "Deep neural networks for choice analysis: Architecture design with alternative-specific utility functions," *Transp. Res. Part C Emerg. Technol.*, vol. 112, pp. 234–251, Mar. 2020, doi: 10.1016/j.trc.2020.01.012.

[3] C. Bonina and B. Eaton, "Cultivating open government data platform ecosystems through governance: Lessons from Buenos Aires, Mexico City and Montevideo," *Gov. Inf. Q.*, vol. 37, no. 3, p. 101479, July 2020, doi: 10.1016/j.giq.2020.101479.

[4] H. Li, X. Zhong, W. Zhang, S. Li, and Y. Xing, "An algorithm for e-bike equivalents at signalized intersections based on traffic conflict events number," *Transp. Res. Part Policy Pract.*, vol. 134, pp. 78–95, Apr. 2020, doi: 10.1016/j.tra.2020.02.007.

[5] R. Nanang, C. Susilawati, and M. Skitmore, "Toward a public sector asset optimization strategy: the case of Indonesia," *Constr. Innov.*, vol. 23, no. 5, pp. 1186–1209, Nov. 2023, doi: 10.1108/CI-12-2021-0235.

[6] D. Wang, D. Richards, A. A. Bilgin, and C. Chen, "A Predictive Model for Citizens' Utilization of Open Government Data Portals," in *Digital Libraries at Times of Massive Societal Transition*, vol. 12504, E. Ishita, N. L. S. Pang, and L. Zhou, Eds., in Lecture Notes in Computer Science, vol. 12504. , Cham: Springer International Publishing, 2020,

pp. 159–175. doi: 10.1007/978-3-030-64452-9_14.
[7] J. A. Prozzi and F. Hong, "Transportation Infrastructure Performance Modeling through Seemingly Unrelated Regression Systems," *J. Infrastruct. Syst.*, vol. 14, no. 2, pp. 129–137, June 2008, doi: 10.1061/(ASCE)1076-0342(2008)14:2(129).
[8] B. Flyvbjerg, M. K. Skamris Holm, and S. L. Buhl, "How (In)accurate Are Demand Forecasts in Public Works Projects?: The Case of Transportation," *J. Am. Plann. Assoc.*, vol. 71, no. 2, pp. 131–146, June 2005, doi: 10.1080/01944360508976688.
[9] C. Cardelino, "Daily Variability of Motor Vehicle Emissions Derived from Traffic Counter Data," *J. Air Waste Manag. Assoc.*, vol. 48, no. 7, pp. 637–645, July 1998, doi: 10.1080/10473289.1998.10463709.
[10] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.
[11] I. Alam, D. Md. Farid, and R. J. F. Rossetti, "The Prediction of Traffic Flow with Regression Analysis," in *Emerging Technologies in Data Mining and Information Security*, vol. 813, A. Abraham, P. Dutta, J. K. Mandal, A. Bhattacharya, and S. Dutta, Eds., in Advances in Intelligent Systems and Computing, vol. 813. , Singapore: Springer Singapore, 2019, pp. 661–671. doi: 10.1007/978-981-13-1498-8_58.
[12] W. McKinney, "Data structures for statistical computing in Python.," *scipy*, vol. 445, no. 1, pp. 51–56, 2010.
[13] Z. Engin and P. Treleaven, "Algorithmic Government: Automating Public Services and Supporting Civil Servants in using Data Science Technologies," *Comput. J.*, vol. 62, no. 3, pp. 448–460, Mar. 2019, doi: 10.1093/comjnl/bxy082.
[14] V. Gray, "THE USE OF TIME SERIES ANALYSIS IN THE STUDY OF PUBLIC POLICY," *Policy Stud. J.*, vol. 2, no. 2, pp. 97–102, Dec. 1973, doi: 10.1111/j.1541-0072.1973.tb00134.x.
[15] H. Behrooz and Y. M. Hayeri, "Machine Learning Applications in Surface Transportation Systems: A Literature Review," *Appl. Sci.*, vol. 12, no. 18, p. 9156, Sept. 2022, doi: 10.3390/app12189156.
[16] M. A. Sellami and O. S. E. Hamzi, "Time Series Forecasting Using Linear Models: A Comprehensive Investigation and Empirical Analysis," 2024, doi: 10.13140/RG.2.2.26098.52167.
[17] A. E. Husin, M. A. Berawi, S. Dikun, T. Ilyas, and A. R. B. Berawi, "Forecasting demand on mega infrastructure projects: Increasing financial feasibility," *Int. J. Technol.*, vol. 6, no. 1, pp. 73–83, 2015.
[18] Z. Saipidinov, R. Isaev, and G. Gimaletdinova, "Stock Market Prediction and Forecasting Using Historical Data," Dec. 19, 2024, *Computer Science and Mathematics*. doi: 10.20944/preprints202412.1639.v1.
[19] D. Shah, H. Isah, and F. Zulkernine, "Stock Market Analysis: A Review and Taxonomy of Prediction Techniques," *Int. J. Financ. Stud.*, vol. 7, no. 2, p. 26, May 2019, doi: 10.3390/ijfs7020026.
[20] S. Orazgaliyeva, Z. Satpayeva, S. Tazhiyeva, and G. Nurseiytova, "E-government as a tool to improve the efficiency of public administration: The case of Kazakhstan," *Probl. Perspect. Manag.*, vol. 21, no. 2, pp. 578–591, June 2023, doi: 10.21511/ppm.21(2).2023.53.
[21] J. J. P. Latupeirissa, N. L. Y. Dewi, I. K. R. Prayana, M. B. Srikandi, S. A. Ramadiansyah, and I. B. G. A. Y. Pramana, "Transforming Public Service Delivery: A Comprehensive Review of Digitization Initiatives," *Sustainability*, vol. 16, no. 7, p. 2818, Mar. 2024, doi: 10.3390/su16072818.
[22] A. F. Van Veenstra, F. Grommé, and S. Djafari, "The use of public sector data analytics in the Netherlands," *Transform. Gov. People Process Policy*, vol. 15, no. 4, pp. 396–419, Nov. 2021, doi: 10.1108/TG-09-2019-0095.

DOI: https://doi.org/10.26740/jistel.v2n2.p104-112
e- ISSN 3090-613X

[23] S. Carbonara, M. Faustoferri, and D. Stefano, "Real Estate Values and Urban Quality: A Multiple Linear Regression Model for Defining an Urban Quality Index," *Sustainability*, vol. 13, no. 24, p. 13635, Dec. 2021, doi: 10.3390/su132413635.
[24] G. Manogaran, J. Gao, and T. N. Nguyen, "Optimizing Resource and Service Allocations for IoT-Assisted Intelligent Transportation Systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 11, pp. 12877–12887, Nov. 2023, doi: 10.1109/TITS.2022.3182649.
[25] Á. Cuevas and E. M. Quilis, "Seasonal Adjustment Methods for Daily Time Series. A Comparison by a Monte Carlo Experiment," *SSRN Electron. J.*, 2023, doi: 10.2139/ssrn.4670922.

DOI: https://doi.org/10.26740/jistel.v2n2.p104-112
e- ISSN 3090-613X