Volume 02, Issue 01, June 2026 : 13-22 Received: April 21, 2026. Revised: May 07, 2026

International Journal of Electrical and Intelligent Engineering

Bicycle Rental Pattern Analysis Using Python and a Streamlit-Based Interactive Dashboard

Ratna Sari[1], Dewi Purnamasari[2*], Henny Prasetyani[3]

[1,2]Department of System and Information Technology, Universitas Ivet, Semarang, Indonesia
[3]Department of Informatics Education, Universitas Ivet, Semarang, Indonesia
* Corresponding author's Email: dewipurnamasari@unisvet.ac.id

Abstract: This study analyzes bike-sharing demand patterns and translates the analytical results into an interactive Streamlit dashboard for interpretable decision support. Using the Bike Sharing Dataset for 2011–2012 (17,379 hourly records), the workflow combines data-quality checking, temporal and environmental exploratory data analysis, aggregate usage indicators, variable-level Pearson correlation analysis, and dashboard implementation. The contribution of the study is not a new forecasting algorithm; rather, it is a reproducible and lightweight analytical pipeline that connects interpretable statistical evidence directly to an interactive interface for nontechnical exploration. Results show that rental activity peaks in the late afternoon and under clear weather conditions. Temperature (r = 0.404772) and apparent temperature (r = 0.400929) have moderate positive associations with total rentals, whereas humidity (r = -0.322910) has a weak negative association and weather severity (r = -0.142430) has a very weak negative association. The dashboard passed all defined black-box functional test scenarios. These findings demonstrate how descriptive analytics and interactive visualization can be combined to support transparent exploration of bike-sharing demand patterns while avoiding causal claims.

Keywords: Bike-Sharing, Data Analysis, Interactive Dashboard, Pearson Correlation, Python, Streamlit.

1. Introduction

Current technological developments increasingly encourage data-based decision-making and the systematic transformation of operational data into actionable information [1]–[4]. In urban mobility, bike-sharing systems generate large volumes of temporal and environmental data that can reveal demand patterns, peak periods, and operating conditions. Python is widely used for this purpose because its scientific ecosystem supports reproducible data manipulation, statistical analysis, and visualization [5].

Previous bike-sharing studies have predominantly focused on prediction and forecasting. Charles et al. [6] analyzed Seoul bike-rental data using several classification approaches, while Python-based data processing has also been shown to support efficient large-scale analysis [7]. Data visualization is important for translating complex numerical results into forms that are easier to interpret [8]. Exploratory Data Analysis (EDA) can expose distributions, temporal patterns, and anomalies before more advanced modeling is performed. Although RFM analysis is commonly used for customer segmentation [9], conventional RFM requires customer-level transaction histories and monetary values; therefore, this study does not treat the Bike Sharing Dataset as a customer-level RFM dataset. Pearson correlation is instead used to quantify variable-level linear associations with rental demand [10], and Streamlit is used to connect these analytical results to an interactive interface [11].

Shi et al. identified temperature, humidity, and peak-hour periods as influential factors in short-term bike-sharing demand and applied a Long Short-Term Memory network to capture temporal dependencies [12]. These findings indicate that bike-sharing demand should be examined across hourly, daily, seasonal, calendar-related, and environmental dimensions rather than through a single variable.

Ma et al. proposed a spatial-temporal graph attentional LSTM model that integrated historical trips, weather data, user information, and land-use characteristics to predict bike-sharing rental and return demand [13]. Subramanian et al. compared Random Forest, ARIMA, SARIMA, LSTM, and GRU models and emphasized the importance of exploratory analysis of patterns, seasonality, correlations, and outliers before predictive modeling [14]. Peláez-Rodríguez et al. compared machine-learning and deep-learning multivariate time-series methods using historical demand, calendar variables, and environmental data [15]. More recently, Betkier

Volume 02, Issue 01, June 2026 : 13-22 Received: April 21, 2026. Revised: May 07, 2026

IJEIE International Journal of Electrical and Intelligent Engineering

et al. employed graph neural networks and spatial aggregation to examine temporal, spatial, infrastructural, environmental, and weather-related factors in bike-sharing systems [16].

Besides demand forecasting, bike-sharing data analysis can support operational interpretation and communication. Grau-Escolano et al. examined electric- and mechanical-bike mobility patterns and showed that different bike types may exhibit distinct usage and maintenance characteristics [17]. Boufidis et al. developed a station-level demand prediction and visualization tool for bike-sharing operators, showing that analytical outputs become more operationally useful when users can explore demand patterns through visual interfaces [18].

The literature therefore reveals two complementary needs: accurate predictive modeling and interpretable analytical tools. Most cited studies prioritize forecasting performance or sophisticated spatial-temporal modeling, whereas comparatively less attention is given to lightweight applications that explicitly connect transparent descriptive/statistical analysis with an interactive dashboard for nontechnical exploration. This study addresses that practical-methodological gap.

The main contribution of this work is a reproducible, interpretable analytical pipeline that links (1) data-quality verification and temporal/environmental EDA, (2) aggregate usage indicators that are appropriate to the available data granularity, (3) variable-by-variable Pearson correlation analysis without averaging coefficients, and (4) a Streamlit dashboard that exposes the same analytical outputs through filters, charts, and summary tables. Unlike forecasting-oriented studies [12]–[16], the proposed approach prioritizes transparency and direct traceability from numerical evidence to dashboard visualization. The dashboard is functionally validated using black-box testing; usability is not claimed because no valid end-user SUS dataset was collected in the present study.

Accordingly, this research aims to (1) identify temporal and environmental patterns associated with bike-sharing demand, (2) quantify the strength and direction of relevant variable-level linear relationships with total rentals, and (3) implement the analytical results in a Streamlit-based interactive dashboard. The dashboard is therefore not a separate software output from the analysis; it is the presentation layer of the same analytical pipeline, enabling users to inspect the patterns and statistical evidence interactively.

2. Method

2.1 Dataset and Research Context

This research was developed from Practical Work activities in the Bangkit Academy 2024 Certified Independent Study and Internship Program. The research uses the Bike Sharing Dataset from the learning module "Learn Data Analysis with Python." The dataset records hourly bike-sharing activity during 2011–2012 and contains 17,379 observations [19]. The analytical unit is an hourly system-level observation rather than an individual customer transaction. This distinction is important because the dataset does not provide Customer ID, rental price, payment amount, or a complete user-level transaction history.

a. Time Dimension: dteday (date), yr (year), mnth (month), hr (hour), season (season), holiday (holiday), weekday (day of the week), and workingday (working day).
b. Environmental Conditions: weathersit (weather condition), temp (temperature), hum (humidity), and windspeed (wind speed).
c. Main Indicator: cnt (total number of bike-sharing rentals).

During preprocessing, the dataset was checked for missing values and duplicate records. No missing values or duplicated rows were identified, so no observations were removed. The dteday field was converted to a datetime data type. Categorical/time codes were retained for descriptive grouping, while numerical variables were used for correlation analysis. Derived aggregate indicators were created only for period-level exploration. After transformation, the analytical dataset contained 17,379 records and 19 variables. These steps were implemented programmatically so that the same transformations are applied consistently before visualization and dashboard deployment.

The analysis was performed using Python version 3.10.11 in Google Colaboratory (v1.0.0). The interactive dashboard was developed using Streamlit version 1.51.0. Data manipulation and numerical computation were carried out using Pandas (v2.2.2) and NumPy (v2.0.2), while Matplotlib (v3.10.0) and Seaborn (v0.13.2) were used for data visualization.

Volume 02, Issue 01, June 2026 : 13-22 Received: April 21, 2026. Revised: May 07, 2026

IJEIE International Journal of Electrical and Intelligent Engineering

Source code development was conducted using Visual Studio Code version 1.105.1.

2.2 Research Stages

The research was conducted through a sequential analytical-to-implementation workflow. Fig. 1 summarizes the relationship among data collection, preprocessing, EDA, aggregate usage analysis, Pearson correlation, dashboard development, deployment, and functional validation. This sequence ensures that the dashboard displays outputs generated from the same processed dataset and analytical definitions used in the statistical analysis.

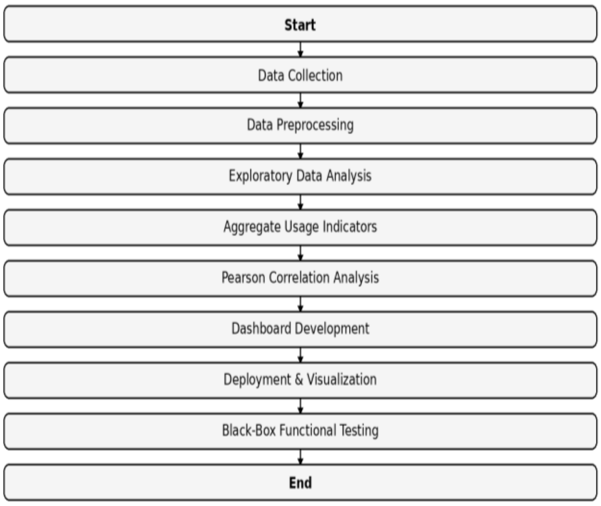

Figure 1. Research workflow

Based on Fig. 1, the research phases consist of the following interrelated steps:

a. Data Collection:
   The Bike Sharing Dataset was obtained from the data-analysis learning resource and used as a system-level case study of hourly bike-sharing demand.
b. Data Preprocessing
   Missing values and duplicates were checked; dteday was converted to datetime; variable types were verified; and analytical attributes needed for temporal grouping were generated. No records were deleted because the quality checks identified no missing or duplicate rows.
c. Exploratory Data Analysis (EDA)
   Descriptive statistics and graphical visualizations were used to inspect hourly, seasonal, weather-related, and calendar-based demand patterns before correlation interpretation.
d. Aggregate Usage Indicators
   Because the dataset does not contain customer identifiers or monetary transaction values, conventional customer-level RFM segmentation is not applicable. The earlier RFM-inspired outputs are therefore interpreted only as aggregate recency, frequency, and rental-volume indicators at period/season level, not as customer behavior or monetary value.
e. Pearson Correlation Analysis
   Pearson's r was calculated separately between each numerical variable and cnt (total rentals). The analysis emphasizes meaningful associations such as temperature, apparent temperature, hour, humidity, and weather condition. Coefficients were not averaged because averaging can conceal differences in magnitude and direction.
f. Dashboard Development and Deployment
   EDA charts, aggregate indicators, correlation results, and temporal filters were integrated into a Streamlit application and deployed online. The dashboard therefore functions as an interactive presentation layer for the analytical results.
g. System Testing
   Black-box testing was used to verify functional correctness of dashboard access, filters, analysis selection, and visualization outputs. This test validates software behavior but does not measure perceived usability; usability evaluation using instruments such as the System Usability Scale (SUS) is reserved for future user-based validation.

2.3 Performance Benchmarks and Validation

The analytical and software outputs were evaluated using two complementary criteria:

a. Statistical Validation: The strength and direction of linear association between each numerical explanatory variable and total bike rentals were evaluated using the Pearson correlation coefficient (r) [20]. The coefficient was calculated using Equation (1). Pearson correlation was interpreted as an

Volume 02, Issue 01, June 2026 : 13-22 Received: April 21, 2026. Revised: May 07, 2026

IJEIE International Journal of Electrical and Intelligent Engineering

association measure only and not as evidence of causality.

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (1)$$

In Equation (1), r is the Pearson correlation coefficient, n is the number of paired observations, x_i is the value of the explanatory variable x for observation i, and y_i is the corresponding value of the response variable y. In this study, y is cnt (total hourly bike rentals), while x is evaluated separately for each numerical variable, including temp, atemp, hr, hum, weathersit, windspeed, yr, and other encoded calendar variables. The classification in Table 1 follows the correlation-strength intervals adapted from Sugiyono [22].

Table 1. Correlation Coefficient Criteria

Correlation Coefficient (r)	**Interpretation**
0.00–0.199	Very Weak
0.20–0.399	Weak
0.40–0.599	Moderate
0.60–0.799	Strong
0.80–1.000	Very Strong

b. Functional Validation: Black-box testing was conducted to verify whether interface controls and navigation features produced the expected outputs without examining the internal program structure [23], [24]. The percentage of successful scenarios was used only as a functional-correctness indicator. It should not be interpreted as evidence of usability, user satisfaction, or dashboard effectiveness from an end-user perspective.

3. Results and Discussion

3.1 Research Findings

The results are presented in the same order as the analytical workflow so that each dashboard component can be traced to its underlying data-processing and analytical step.

3.1.1 Preprocessing Stage

The Bike Sharing Dataset contains 17,379 hourly records. Pandas and NumPy were used to inspect data types, missing values, duplicates, and numerical consistency, while Matplotlib and Seaborn supported exploratory visualization. The preprocessing procedure confirmed that no missing values or duplicate rows were present. The dteday variable was converted to datetime, and the resulting cleaned dataset was then used consistently for EDA, correlation analysis, aggregate indicators, and dashboard presentation.

Fig. 2 shows an excerpt of the cleaned dataset used in the subsequent analyses. Because no duplicate rows were identified, the .drop_duplicates() operation did not reduce the number of observations. The key transformation was the conversion and validation of date/time fields so that hourly, daily, monthly, seasonal, and yearly groupings could be generated reproducibly.

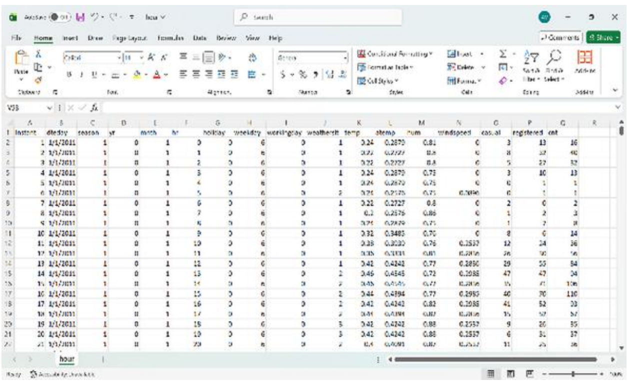

Figure 2. Excerpt of the cleaned hourly bike-sharing dataset

3.1.2 Visualization of Exploratory Data Analysis (EDA) Results

EDA was used to identify interpretable demand patterns before statistical testing. Fig. 3 – Fig. 5 show the principal weather, hourly, and seasonal patterns. The visual results are discussed together with previous bike-sharing studies to distinguish descriptive agreement from causal explanation.

a. Weather Condition Patterns

Fig. 3 shows that average rentals are highest under clear weather and decline as weather conditions become less favorable. This direction is consistent with Shi et al. [12], who identified environmental variables such as temperature and humidity as relevant to short-term bike-sharing demand. The present result is descriptive and does not establish that weather alone causes the

Volume 02, Issue 01, June 2026 : 13-22 Received: April 21, 2026. Revised: May 07, 2026

IJEIE International Journal of Electrical and Intelligent Engineering

observed change because temporal and behavioral factors may co-vary with weather conditions.

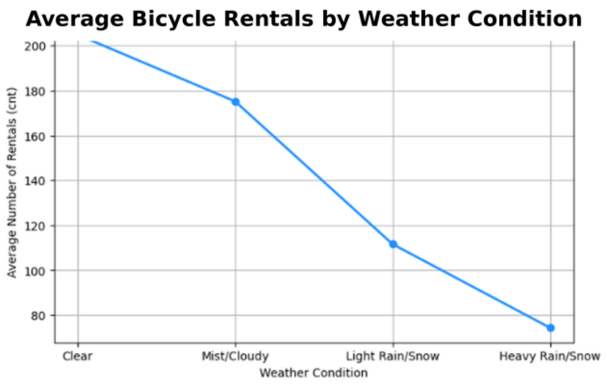

Figure 3. Average rentals by weather condition

b. Time Pattern

Fig. 4 shows a pronounced hourly structure, with the highest rental intensity occurring in the late afternoon, approximately 4:00–6:00 PM in the 2011 visualization. This supports the emphasis on peak-hour temporal dependencies reported by Shi et al. [12] and the importance of seasonality/pattern exploration highlighted by Subramanian et al. [14]. The weak positive correlation for hr (r = 0.394071) should therefore be interpreted carefully: a linear coefficient compresses a clearly nonlinear daily pattern into a single value.

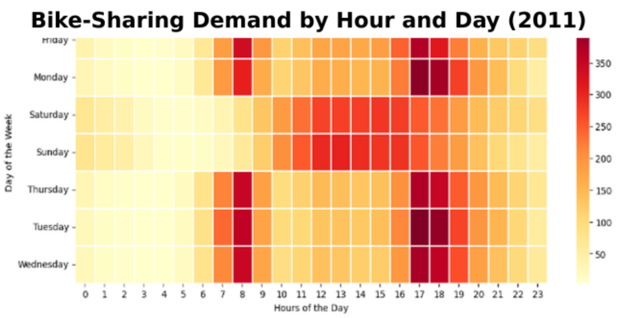

Figure 4. Bike-sharing demand by hour and day

c. Seasonal Trend Patterns

Fig. 5 indicates higher average rental activity during fall than in the other seasons. This reinforces the need to analyze bike-sharing demand across multiple time scales, as also emphasized in forecasting studies that combine calendar and environmental information [13]–[16]. Seasonal differences should be viewed as contextual patterns rather than isolated effects because season, temperature, daylight, and user activity are interrelated.

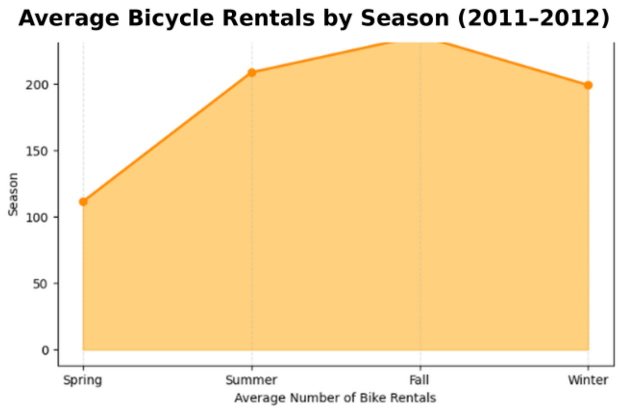

Figure 5. Average bike-sharing rentals by season

3.1.3 Aggregate Recency, Frequency, and Rental-Volume Indicators

The dataset is not suitable for conventional customer-level RFM segmentation because it contains no Customer ID and no rental price or payment amount. Therefore, the earlier RFM-inspired analysis is reframed as aggregate operational indicators. “Recency” represents elapsed time since the latest observed rental activity within an aggregate grouping, “frequency” represents the number of rental observations/activities within the selected period, and the former “monetary” component is renamed “rental volume” because it represents counts rather than currency. These indicators describe system-level usage intensity and must not be interpreted as individual customer behavior.

a. Average Aggregate Recency by Season

Fig. 6 presents the aggregate recency indicator by season. Lower values indicate that observed rental activity occurred more recently within the chosen grouping. The result is used only as a period-level operational descriptor; it does not identify repeat customers because customer-level identifiers are unavailable.

Volume 02, Issue 01, June 2026 : 13-22 Received: April 21, 2026. Revised: May 07, 2026

International Journal of Electrical and Intelligent Engineering

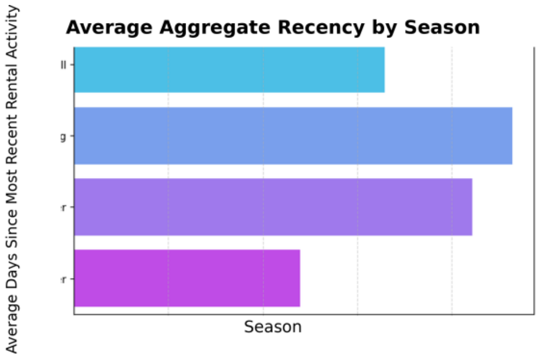

Figure 6. Aggregate recency by season

b. Relationship Between Aggregate Frequency and Total Rentals

Fig. 7 presents the relationship between the number of aggregated rental observations and total rental volume by month. The positive tendency is expected because periods with more observed rental activity can accumulate higher total counts. Accordingly, this plot is treated as a descriptive operational summary rather than evidence that individual users made repeated transactions.

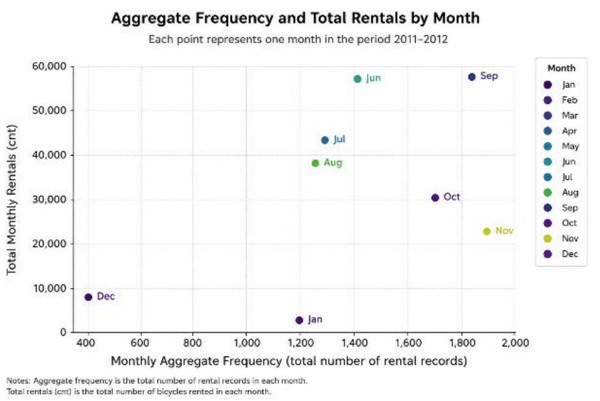

Figure 7. Aggregate frequency and total rentals

c. Distribution of Monthly Rental Volume

Fig. 8 shows the distribution of monthly total rental counts. The figure identifies periods with higher system-level demand, especially around the middle of the year. Because the variable is a count of rentals rather than revenue or expenditure, the term “rental volume” is used instead of “Monetary.”

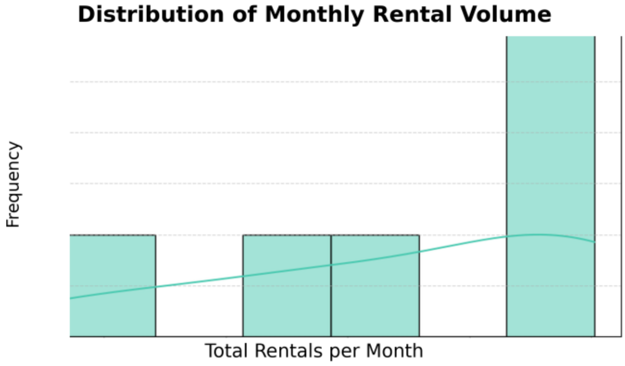

Figure 8. Monthly rental-volume distribution

3.1.4 Dashboard Development

The cleaned dataset and analytical outputs were integrated into a Streamlit application in Visual Studio Code. The dashboard was designed as the presentation layer of the analytical workflow: filters modify the displayed subset, while EDA charts, aggregate usage indicators, and correlation summaries are generated from the same processed data definitions described in Section 2.

Fig. 9 shows local execution and functional checking before deployment. This stage verified that the application launched correctly and that filters and visual outputs could be rendered without runtime errors. It is distinct from usability evaluation because no end-user satisfaction questionnaire was administered.

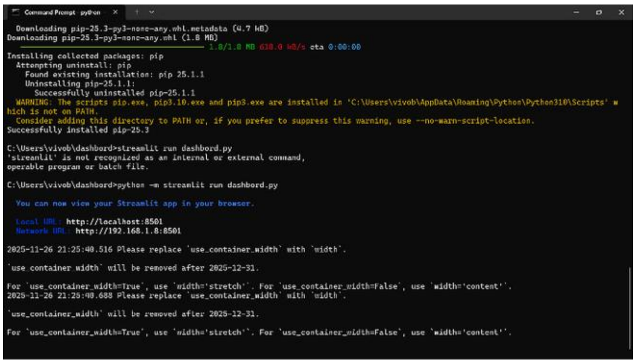

Figure 9. Local dashboard execution and functional checking

3.1.5 Deployment and Dashboard Display Results

The dashboard was deployed to Streamlit Cloud so that the analytical results could be accessed interactively. Deployment enables users to explore

Volume 02, Issue 01, June 2026 : 13-22 Received: April 21, 2026. Revised: May 07, 2026

International Journal of Electrical and Intelligent Engineering

the same temporal, environmental, and statistical patterns described in this article through an online interface.

Fig. 10 shows the English-language dashboard interface. The dashboard presents EDA visualizations, correlation information, aggregate usage indicators, and time-based filters in a single interface. Its value lies in making the analytical outputs directly explorable rather than replacing the underlying statistical analysis. Functional correctness was tested separately using black-box scenarios.

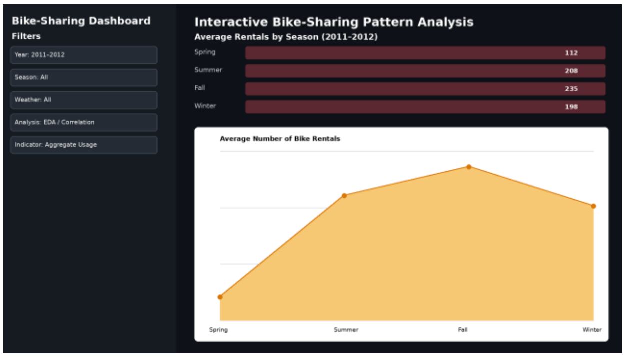

Figure 10. English-language Streamlit dashboard interface

3.1.6 Statistical and Functional Testing Results

a. Pearson Correlation Test Results

Pearson correlation was interpreted separately for each variable rather than through a single average coefficient. Temperature (temp, r = 0.404772) and apparent temperature (atemp, r = 0.400929) show moderate positive associations with total rentals. Hour (hr, r = 0.394071) shows a weak positive linear association, not a moderate one. Humidity (hum, r = -0.322910) shows a weak negative association, while weather condition severity (weathersit, r = -0.142430) shows a very weak negative association. The year variable (yr, r = 0.250495) has a weak positive association, indicating that rental levels were generally higher in the later year of the dataset. Calendar variables such as workingday, weekday, and holiday have very weak linear correlations, although they may still contribute to nonlinear or interaction-based patterns.

The variables registered and casual have very high/strong correlations with cnt because cnt is constructed from these component counts; they are therefore descriptive components of the outcome rather than independent explanatory predictors and should not be used to infer drivers of demand. Similarly, instant is an observation index and is not substantively interpreted as a demand factor. These distinctions are important because correlation strength alone does not establish explanatory importance. The temperature and humidity results are directionally consistent with prior work emphasizing environmental conditions [12], while the hourly pattern supports the temporal-dependency findings of forecasting studies [12]–[16]. The analysis does not average correlation coefficients, because doing so would obscure differences in sign and magnitude among variables.

Table 2. Pearson Correlation Test Results

Variable	r	Category	Direction
registered	0.972151	Very Strong	Positive
casual	0.694564	Strong	Positive
temp	0.404772	Moderate	Positive
atemp	0.400929	Moderate	Positive
hr	0.394071	Weak	Positive
instant	0.278379	Weak	Positive
yr	0.250495	Weak	Positive
season	0.178056	Very Weak	Positive
mnth	0.120638	Very Weak	Positive
windspeed	0.093234	Very Weak	Positive
workingday	0.030284	Very Weak	Positive
weekday	0.0269	Very Weak	Positive
holiday	-0.03093	Very Weak	Negative
weathersit	-0.14243	Very Weak	Negative
recency	-0.27775	Weak	Negative
hum	-0.32291	Weak	Negative

2. Black Box Testing Results

Black-box testing verified whether each dashboard control produced the expected output according to the defined scenarios. The scenarios covered dashboard access, date-range filtering, analysis selection, weather/time/seasonal options, aggregate usage indicators, and visualization output. The test focused on observable input-output behavior rather than internal implementation.

All defined black-box scenarios passed, yielding a 100% functional pass rate for the tested cases (Fig. 11). This result supports functional correctness only. It does not demonstrate perceived ease of use, learnability, satisfaction, or overall usability. A future study should recruit representative dashboard users and report a validated usability

Volume 02, Issue 01, June 2026 : 13-22 Received: April 21, 2026. Revised: May 07, 2026

IJEIE International Journal of Electrical and Intelligent Engineering

instrument such as the System Usability Scale (SUS), including participant characteristics, score distribution, mean, and variability.

Function	Action	Expected Result	Trial Iteration					Final Result
			1	2	3	4	5	
Dashboard Access	Open the dashboard URL	Displays the title "Bicycle Rental Analysis Dashboard" and the sidebar	✓	✓	✓	✓	✓	Passed
Date Range Filter	Select a date range	The data range is updated and the visualizations adjust accordingly	✓	✓	✓	✓	✓	Passed
Analysis Selection Dropdown	Select an analysis menu from the dropdown	The displayed content/visualization changes according to the selected menu	✓	✓	✓	✓	✓	Passed
Analysis Selection	Select "Weather ➔ ..."	A summary table and a weather line chart are displayed	✓	✓	✓	✓	✓	Passed
	Select "2011 Time Pattern ➔ ..."	An hour-versus-day heatmap is displayed	✓	✓	✓	✓	✓	Passed
	Select "2011 Monthly Pattern ➔ ..."	A bar chart showing the average monthly rentals in 2011 is displayed	✓	✓	✓	✓	✓	Passed
	Select "2011–2012 Seasonal Trend ➔ ..."	An area chart showing the average rentals by season is displayed	✓	✓	✓	✓	✓	Passed
	Select "RFM ➔ ..."	Recency bar chart, Frequency–Monetary scatter plot, and Monetary histogram are displayed	✓	✓	✓	✓	✓	Passed

Figure 11. Black-box testing results

4. Conclusion

This study contributes a lightweight and reproducible analytical pipeline that connects bike-sharing EDA and variable-level Pearson correlation directly to an interactive Streamlit dashboard. The scientific contribution lies in the transparent integration and interpretation of complementary analytical outputs rather than in proposing a new forecasting algorithm. Empirically, rentals peak in the late afternoon and are highest under clear weather; temperature ($r = 0.404772$) and apparent temperature ($r = 0.400929$) show moderate positive associations with total rentals, while humidity ($r = -0.322910$) shows a weak negative association. The hr coefficient ($r = 0.394071$) is weak and should be interpreted alongside the nonlinear hourly pattern visible in EDA. The dashboard passed all defined black-box functional scenarios, demonstrating that the analytical outputs can be delivered reliably through interactive filters and visualizations. Practically, this approach can support transparent exploration of demand patterns for teaching, preliminary operational analysis, and communication with nontechnical users. The study remains limited by its historical 2011–2012 dataset, the absence of customer-level identifiers and monetary transaction fields, and the lack of end-user usability data. Future work should add recent operational data, predictive models, explicit year/working-day subgroup comparisons, and SUS-based usability evaluation with representative users.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, R.S. and D.P.; methodology, R.S. and D.P.; software, R.S.; validation, R.S., D.P., and H.P.; formal analysis, R.S.; investigation, R.S.; data curation, R.S.; visualization, R.S.; writing—original draft preparation, R.S.; writing—review and editing, D.P. and H.P.; supervision, D.P. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

The authors would like to express their sincere gratitude to Bangkit Academy and Dicoding for the opportunity, learning experience, and valuable knowledge provided during the Practical Work program. The authors also extend their appreciation to Universitas Ivet Semarang for its academic support throughout this research.

References

[1] R. G. Wardhana, G. Wang, and F. Sibuea, "Penerapan Machine Learning dalam Prediksi Tingkat Kasus Penyakit di Indonesia," Journal of Information System Management (JOISM), vol. 5, no. 1, pp. 40–45, 2023.

[2] L. K. Bayu, "Studi Kasus Analisis Prediktif Emisi Carbon di Rwanda Menggunakan Machine Learning Model Regression," Kerja Praktik, Program Studi S1 Sistem Informasi, Fakultas Teknologi dan Informatika, Universitas Dinamika, Surabaya, Indonesia, 2024.

[3] H. Satria, "Implementasi Streamlit untuk Mengklasifikasikan Tingkat Pencemaran Udara di DKI Jakarta dengan Penggunaan Algoritma Regression," Undergraduate Thesis, Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana, Jakarta, Indonesia, 2022. [Online]. Available: https://repository.mercubuana.ac.id/68892/.

[4] D. A. Syafiq and Y. Sembiring, "Pengembangan Model Visualisasi Data Kependudukan Berbasis Teknologi Informasi untuk Mendukung

Pengambilan Keputusan di Kecamatan Medan Amplas," Jurnal Penelitian Inovatif (JUPIN), vol. 5, no. 2, pp. 1329–1338, 2025, doi:10.54082/jupin.1500.

[5] M. Suryati, T. A. D. Saputra, and I. H. Ikasari, "Penggunaan Python dalam Analisis Data dengan Machine Learning," JRIIN: Jurnal Riset Informatika dan Inovasi, vol. 2, no. 10, pp. 1959–1968, 2025.

[6] N. Charles, A. B. Wijaya, Marvin, and D. Jollyta, "Analisis Data Sewa Sepeda di Seoul dengan Variatif Metode Klasifikasi," Jurnal Algoritme, vol. 4, no. 2, pp. 53–62, 2024, doi:10.35957/algoritme.v4i2.6636.

[7] A. K. Putri and D. I. Nur, "Penggunaan Bahasa Python untuk Analisis dan Visualisasi Data Penduduk di Desa Sumberjo, Nganjuk," Karya: Jurnal Pengabdian Kepada Masyarakat, vol. 3, no. 3, pp. 206–217, 2023.

[8] R. G. Guntara, "Visualisasi Data Laporan Penjualan Toko Online Melalui Pendekatan Data Science Menggunakan Google Colab," ULIL ALBAB: Jurnal Ilmiah Multidisiplin, vol. 2, no. 6, pp. 2091–2100, 2023, doi:10.56799/jim.v2i6.1578.

[9] F. A. Maresti, G. M. Anugraheni, R. A. Hargiyanto, and K. Mustaqim, "Penerapan Exploratory Data Analysis (EDA) dan Analisis Recency, Frequency, and Monetary (RFM) untuk Segmentasi Pelanggan E-Commerce," Competitive, vol. 19, no. 1, pp. 14–25, 2024, doi:10.36618/competitive.v19i1.4059.

[10] P. P. N. Ardhaneswari and I. W. C. Suwitra, "Pearson Correlation Analysis in Determining the Relationship Between Price and Sales Volume of Wardah Matte Lip Cream on the Shopee E-Commerce Platform," Bisnis, Manajemen, dan Akuntansi, vol. 1, no. 2, pp. 85–92, 2024.

[11] K. N. Oktaviarini, E. D. Wahyuni, and R. P. Sari, "Transformasi Data Statistik Menjadi Visual Interaktif Menggunakan Streamlit: Studi Kasus BPS Kota Mojokerto," Jurnal Kridatama Sains dan Teknologi, vol. 6, no. 2, pp. 804–822, 2024, doi:10.53863/kst.v6i02.1426.

[12] Y. Shi, L. Zhang, S. Lu, and Q. Liu, "Short-term demand prediction of shared bikes based on LSTM network," *Electronics*, vol. 12, no. 6, Art. no. 1381, 2023, doi: 10.3390/electronics12061381.

[13] X. Ma, Y. Yin, Y. Jin, M. He, and M. Zhu, "Short-term prediction of bike-sharing demand using multi-source data: A spatial-temporal graph attentional LSTM approach," *Applied Sciences*, vol. 12, no. 3, Art. no. 1161, 2022, doi: 10.3390/app12031161.

[14] M. Subramanian, J. Cho, S. Veerappampalayam Easwaramoorthy, A. Murugesan, and R. Chinnasamy, "Enhancing sustainable transportation: AI-driven bike demand forecasting in smart cities," *Sustainability*, vol. 15, no. 18, Art. no. 13840, 2023, doi: 10.3390/su151813840.

[15] C. Peláez-Rodríguez, J. Pérez-Aracil, D. Fister, R. Torres-López, and S. Salcedo-Sanz, "Bike sharing and cable car demand forecasting using machine learning and deep learning multivariate time series approaches," *Expert Systems with Applications*, vol. 238, Art. no. 122264, 2024, doi: 10.1016/j.eswa.2023.122264.

[16] I. Betkier, M. Oszczypała, and W. Dawid, "Bike-sharing systems for sustainable cities: Demand prediction using graph neural networks and spatial aggregation," *Sustainable Cities and Society*, vol. 128, Art. no. 106479, 2025, doi: 10.1016/j.scs.2025.106479.

[17] J. Grau-Escolano, A. Bassolas, and J. Vicens, "Cycling into the workshop: E-bike and m-bike mobility patterns for predictive maintenance in Barcelona's bike-sharing system," *EPJ Data Science*, vol. 13, Art. no. 48, 2024, doi: 10.1140/epjds/s13688-024-00486-x.

[18] N. Boufidis, A. Nikiforiadis, K. Chrysostomou, and G. Aifadopoulou, "Development of a station-level demand prediction and visualization tool to support bike-sharing systems' operators," *Transportation Research Procedia*, vol. 47, pp. 51–58, 2020, doi: 10.1016/j.trpro.2020.03.072.

[19] N. Lakshmipathi, "Bike Sharing Dataset," Kaggle, 2013. [Online]. Available: https://www.kaggle.com/datasets/lakshmi25npathi/bike-sharing-dataset?select=hour.csv. [Accessed: Jan. 9, 2026].

[20] F. Jabnabillah and N. Margina, "Analisis Korelasi Pearson dalam Menentukan Hubungan antara Motivasi Belajar dengan Kemandirian Belajar pada Pembelajaran Daring," Jurnal Sintak, vol. 1, no. 1, pp. 14–18, 2022.

[21] R. Akbar, U. S. Sukmawati, and K. Katsirin, "Analisis Data Penelitian Kuantitatif (Pengujian

International Journal of Electrical and Intelligent Engineering

Hipotesis Asosiatif Korelasi)," Jurnal Pelita Nusantara, vol. 1, no. 3, pp. 430–448, 2023.

[22] Sugiyono, Quantitative, Qualitative, and R&D Research Methods (Metode Penelitian Kuantitatif, Kualitatif, dan R&D), 19th ed. Bandung, Indonesia: Alfabeta, 2013.

[23] S. A. H. Putri and H. P. Prasetiya, "Pengujian Software Testing Sistem ERP PT XYZ dengan Metode Black Box Testing," KURAWAL: Jurnal Teknologi, Informasi dan Industri, vol. 6, no. 1, pp. 15–29, 2023, doi:10.33479/kurawal.v6i1.604.

[24] A. F. Wijaya and B. A. Wardijono, "Pengukuran Kualitas Aplikasi Custody Berdasarkan ISO 25010 Menggunakan Otomatisasi Pengujian Black Box," KLIK: Kajian Ilmiah Informatika dan Komputer, vol. 4, no. 2, pp. 937–946, 2023, doi:10.30865/klik.v4i2.1215.