Comparative Analysis of Public Bicycle Usage Pattern

Wooyong SIM [a], Byongchan SHIN [b], Hyunju SHIN [c], Gunwoo LEE [d] and Hiroaki INOKUCHI [e]

[a,e] *Graduate School of Science and Engineering, Kansai University, Osaka, 564-8680, Japan*
[a] *E-mail: suy990135@hanyang.ac.kr*
[e] *E-mail: hiroaki@inokuchi.jp*
[b,c,d] *Department of Smart City Engineering, Hanyang University, Ansan, 15588, South Korea*
[b] *E-mail: bcshin@hanyang.ac.kr*
[c] *E-mail: sinzoo@hanyang.ac.kr*
[d] *E-mail: gunwoo@hanyang.ac.kr*

Abstract: This comparative study analyzes factors influencing the usage of public bicycles in Seoul and New York City, two cities with distinct cultural, environmental, and regional contexts. Multiple regression and Long Short-Term Memory analyses were conducted using rental history data from each bicycle-sharing system. Environmental factors, including weather and seasonal conditions, were assessed as independent variables. Results revealed significant impacts of seasonal variations and weekend/holiday effects on bicycle usage. Differences in coefficient values and variable significance between the cities suggest that regional climatic conditions distinctly influence usage patterns, despite commonalities in environmental factors. The study recommends incorporating personal and economic factors in future research to further refine strategies promoting public bicycle utilization.

Keywords: Public bicycle usage, multiple regression analysis, LSTM analysis, Ttareungi, Citi Bike

1. INTRODUCTION

The rising use of automobiles has intensified urban challenges such as traffic congestion and environmental degradation, underscoring the importance of public bicycles as sustainable urban transportation solutions (Kyung-Eun, 2019). Public bicycle systems offer practical alternatives for short-distance travel, reducing dependency on private vehicles and mitigating urban environmental impacts (PBSC Urban Solutions, 2021). As of August 2021, more than 3,000 bicycle-sharing programs operated across over 400 cities worldwide, reflecting their significance in urban sustainability efforts (PBSC Urban Solutions, 2021). Successful implementation, however, depends on several interconnected factors. Consequently, cities and urban planners frequently undertake feasibility studies to evaluate bicycle-sharing systems' effectiveness (F.-Imani et al., 2017). Seoul's "Ttareungi" and New York City's "Citi Bike" exemplify highly successful public bicycle initiatives. "Ttareungi" started in 2015 with 150 stations and 2,000 bicycles, rapidly expanding to 1,290 stations by January 2018, driven by increasing public demand and interest. By 2020, annual rentals reached approximately 23 million (Choi et al., 2021). Likewise, "Citi Bike" launched in May 2013 with 332 stations and 6,000 bicycles, becoming the largest bicycle-sharing system in the U.S., achieving over 100 million cumulative uses by July 2020 (Citi Bike, 2020). The notable growth of these programs highlights their significance as sustainable transport options. These two systems were selected for comparison due to their representation of distinct geographic and cultural contexts: Seoul as a major Asian metropolitan area and New York City as a significant North American urban environment. Both offer extensive and accessible datasets, facilitating a detailed comparative analysis of how varying urban and environmental conditions influence bicycle usage patterns. This study primarily investigates environmental influences specifically weather and seasonal factors on public bicycle usage in both cities. Through comparative analysis, this research aims to identify how public bicycle systems adapt effectively to diverse urban environments. The findings are expected to inform urban planners and policymakers, aiding them in developing targeted strategies to enhance sustainability and efficiency. Additionally, the study highlights the potential benefits of incorporating personal and economic characteristics in future analyses to further promote public bicycle use and system improvements.

2. PRIOR RESEARCH REVIEW

2.1 Literature Review

Since its introduction in 2015, there has been ongoing research on Seoul's public bicycle system, "Ttareungi." Jang et al. (2018) identified the impact of seasons, membership types, and public transportation accessibility on "Ttareungi" usage behavior. Joo et al. (2018) proposed a methodology to establish the area of influence of each station based on the usage frequency by incorporating distance data, time, and spatial characteristics. Choi et al. (2021) compared various network clustering approaches to analyze the optimal positioning of Seoul's public bike stations, emphasizing the importance of efficient infrastructure planning in improving system usage. Since its introduction in 2013, there has also been ongoing research on New York City's public bicycle system, "Citi Bike." An et al. (2019) investigated the correlation between weather conditions and bicycle usage. Wang (2015) developed a bicycle rental demand-forecasting model using Citi Bike data. Feng and Wang (2017) used multiple linear regression and random forest analyses to predict bike rental demand. They found that random forest improved predictive accuracy to 82%, highlighting the importance of variables such as temperature, humidity, and wind speed in demand forecasting. Singhvi et al. (2015) also predicted bike usage for New York City's bike-sharing system, integrating spatial and temporal variables, and demonstrated how neighborhood-level predictions could enhance operational management. Research on public bicycles has also addressed socio-economic factors influencing bicycle usage patterns. Park et al. (2023) explored changes in Seoul's public bike usage following the COVID-19 outbreak through user surveys, identifying increased weekday usage among female users and those with higher income levels. Kumar Dey et al. (2021) proposed a comprehensive framework utilizing a multiple discrete-continuous model to estimate origin-destination flows for bikeshare systems, contributing valuable insights into user demand patterns and operational strategies. Safety factors have also been extensively studied. Hwang and Lee (2018) analyzed neighborhood environmental factors affecting bicycle accidents and their severity in Seoul, emphasizing the significance of infrastructure improvements for enhanced safety. Similarly, Branion-Calles et al. (2019) conducted a cross-sectional survey across Canadian and U.S. cities, associating bicycle infrastructure availability and individual characteristics with city-wide bicycling safety perceptions, reinforcing the need for safety-oriented infrastructure development.

2.2 Limitations of Prior Research

Research on public bicycles has been actively conducted to explore factors affecting usage patterns, accident rates, and utilization rates in specific cities. However, previous studies have some limitations. Most studies have focused on a single city, providing detailed analyses of unique environments and user behaviors. However, there is a lack of discussion on the application of these findings to other cities. As each city has unique user characteristics and climatic conditions that influence the operation and usage patterns of public bicycle systems, additional analyses that consider these regional characteristics are necessary to apply the research results from Seoul and New York City to other cities. Given these limitations, a comprehensive and multifaceted approach to public bicycle systems in multiple cities is required. This will provide a deeper understanding of their efficient design and operation and will offer essential guidelines for successful implementation in various urban environments.

2.3 Research Distinction

This study distinguishes itself by conducting a comparative analysis of Seoul's "Ttareungi" and New York City's "Citi Bike." Unlike prior studies that focused on a single city, this study compared the public bicycle systems of these two major cities to provide new insights into how their environments influence usage patterns and system operations.

This study analyzed daily bicycle usage counts and regional weather data and contributed significantly to the understanding of how public transportation systems can adapt to changing urban environments. By comparing the systems in Seoul and New York City, this study examined the impact of regional characteristics on the design and operation of public bicycle systems, offering valuable insights into their efficient design and operation in various urban environments.

Furthermore, this study plays a crucial role in evaluating the impact of public bicycle systems on urban transportation. By understanding their usage patterns, urban planners and policymakers can develop strategies to enhance sustainability and efficiency. In addition, this study explored practical alternatives to promote public bicycle usage, thereby, contributing to their overall improvement.

3. RESEARCH METHODOLOGY AND DATA PREPROCESSING

3.1 Multiple Regression Analysis

The multiple regression analysis used in this study is a statistical method that determines the linear relationship between several independent variables and a single dependent variable to identify the influencing factors (Jae-Min *et al.*, 2018). This method quantifies the relative influence of independent variables on the dependent variable, enabling a clearer understanding of specific phenomena.

The multiple regression model is expressed as follows:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon \tag{1}$$

Where,

Y	: dependent variable,
$x_1, x_2, \cdots, x_n$	: independent variables,
β_0	: intercept,
$\beta_1, \beta_2, \cdots, \beta_n$	: regression coefficients and
ε	: error term.

To address potential multicollinearity among the independent variables, we used correlation coefficient analysis and the variance inflation factor (VIF). The model fit was evaluated using the coefficient of determination (R^2), adjusted R^2, and F-statistics.

The independent variables in this study included weather and seasonal characteristics, which were selected to analyze their impact on public bicycle usage behavior. We focused on the magnitude, direction, and statistical significance of each regression coefficient.

3.2 Long Short-Term Memory (LSTM) Analysis

This study employed an LSTM neural network to model the time-dependent and nonlinear patterns in public bicycle usage. Unlike traditional RNNs, LSTM units incorporate gating mechanisms such as input, forget, and output gates to preserve relevant information across longer periods. This design helps mitigate issues such as vanishing or exploding gradients, enabling the effective learning of long-range temporal dependencies.

The modeling approach involved converting historical bicycle usage records into sequences of past observations. Each sequence was fed into the LSTM, which processed these inputs through its recurrent architecture. The final hidden state was then passed to a fully connected layer to predict its future usage. Hyperparameters, including the number of LSTM units and the time window length, were selected empirically.

Training minimizes the suitable loss function (MSE), and early stopping prevents overfitting. Using this method, the model can capture complex temporal dynamics, offering more accurate forecasts than conventional linear methods (Shi et al., 2023).

3.3 Data Collection

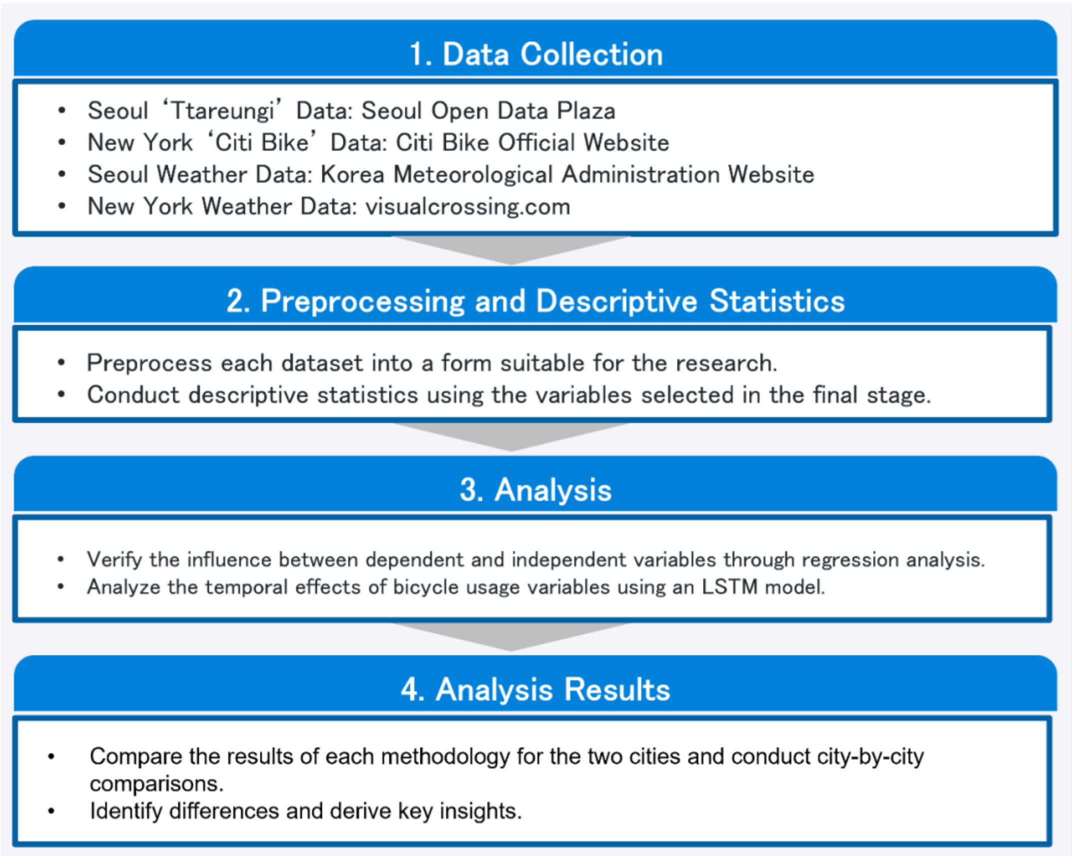

Figure 1. Research flowchart

In this study, to analyze the usage patterns of public bicycles in Seoul and New York City, data were collected from representative bicycle-sharing services in these cities: "Ttareungi" in Seoul and "Citi Bike" in New York City. The data included records of bicycle rentals and returns from 2018. The data for Seoul were collected from the Open Data Plaza, while those for New York City were obtained from the official "Citi Bike" website.

The original usage records used in this study comprise approximately 3.8 million records for "Ttareungi" and over 17 million records for "Citi Bike," amounting to a substantial volume of

Figure . Analysis Flowchart

data. Data were collected from 1,235 and 919 rental stations in Seoul and New York City, respectively. These rental stations covered key areas in each city and were visualized based on their location information. After excluding missing values, 1,229 "Ttareungi" and 919 "Citi Bike" rental stations in Seoul and New York City, respectively, were mapped. This comprehensive dataset allowed for a detailed analysis of public bicycle usage patterns, contributing to a deeper understanding of how these systems operate in different urban environments. The dependent variable in this study was the frequency of public bicycle usage based on records of rentals and returns. These data are essential for understanding the usage patterns and frequency of bicycle-sharing services and provide crucial information regarding urban transportation and leisure activities.

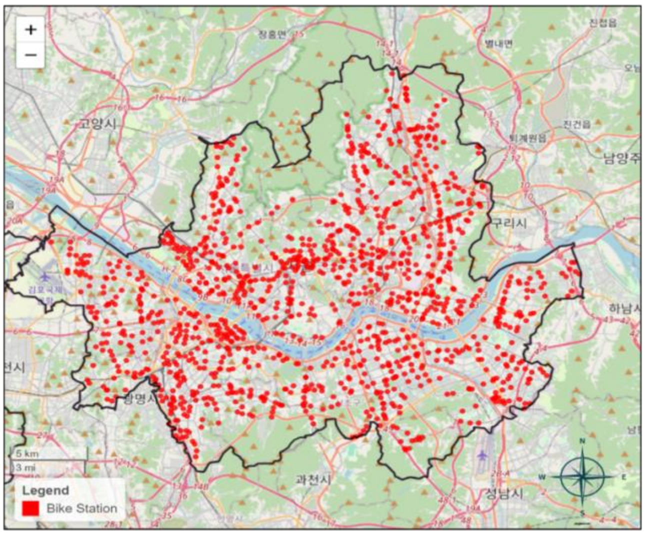

Figure 2. "Ttareungi" rental station locations

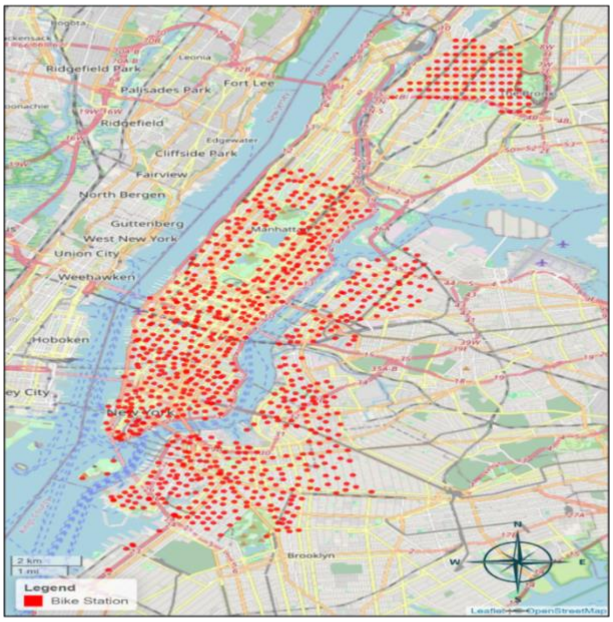

Figure 3. "Citi Bike" rental station locations

The primary independent variables were weather characteristics. Weather conditions significantly impact public bicycle usage, and it is, thus, important to understand how various weather factors, such as temperature, precipitation, and wind speed, influence the frequency of bicycle usage.

For the data related to Seoul's public bicycle "Ttareungi," daily usage information was obtained from the Seoul Open Data Plaza. Weather data for Seoul were obtained from the official website of the Korea Meteorological Administration.

For New York City's public bicycle "Citi Bike," data were collected using the "Citi Bike" system data available on the official Citi Bike website. Weather data for New York City were obtained from visualcrossing.com.

Table 1. Variable Explanation and Data Sources for Seoul

Variable			Description	Source
DV	Daily Usage Count		Daily Usage Count by Rental Station	Seoul Public Bicycle Usage Information (Daily)
IV	WC	Average Temperature	Daily Average Temperature(℃)	Korea Meteorological Administration
		Minimum Temperature	Daily Minimum Temperature(℃)	
		Maximum Temperature	Daily Maximum Temperature(℃)	
		Maximum Wind Speed	Daily Maximum Wind Speed(m/s)	
		Average Wind Speed	Daily Average Wind Speed(m/s)	
		Average Relative Humidity	Daily Average Relative Humidity(%)	
		Minimum Relative Humidity	Daily Minimum Relative Humidity(%)	
		Precipitation	Precipitation	

Table 2. Variable Explanation and Data Sources for New York

Variable			Description	Source
DV	rental start		trip record data	'CitiBike' System Data
	return			
IV	WC	Average Temperature	Daily Average Temperature(℃)	visualcrossing.com
		Minimum Temperature	Daily Minimum Temperature(℃)	
		Maximum Temperature	Daily Maximum Temperature(℃)	
		Maximum Wind Speed	Daily Maximum Wind Speed(m/s)	
		Average Wind Speed	Daily Average Wind Speed(m/s)	
		Average Relative Humidity	Daily Average Relative Humidity(%)	
		Minimum Relative Humidity	Daily Minimum Relative Humidity(%)	
		Precipitation	Precipitation	

The corresponding descriptions of data used for "Ttareungi" and "Citi Bike" are presented in Table 1 and Table 2, respectively. Data analysis was carried out using R programming language version 4.3.0 and Python 3.12.6.

3.4 Data Preprocessing and Descriptive Statistics

The corresponding preprocessing steps and descriptive statistics for Ttareungi and Citi Bike are presented in Table 3 and Table 4, respectively. Prior to conducting the main analysis, preprocessing and descriptive statistics were performed on the 2018 usage datasets of Seoul's "Ttareungi" and New York City's "Citi Bike." For "Ttareungi," among the total 3,776,972 records, daily rental counts were aggregated. Records from April 11 and May 10 were excluded due to missing data, resulting in 363 daily entries used in the analysis. The dependent variable, daily usage count, was normalized before being input into both regression and LSTM models. Similarly, for "Citi Bike," among the total 17,528,318 rental records, daily rental counts were calculated. The same dates (April 11 and May 10) were excluded to match the Seoul dataset, resulting in 363 daily entries. The dependent variable was also normalized prior to modeling. Independent variables included the normalized daily rental count (DV) and the following weather and seasonal characteristics: maximum temperature (°C), maximum wind speed (m/s), average wind speed (m/s), average relative humidity (%), precipitation (clear: 0, precipitation: 1), and weekend/holiday (weekday: 0, weekend/holiday: 1). In addition, a seasonal dummy variable was constructed based on the seasonality observed in northern hemisphere cities, using autumn (September, October, November) as the reference category. Other seasons spring (March, April, May), summer (June, July, August), and winter (December, January, February) were each assigned a value of 1 accordingly, following the climatological categorization suggested by Kevin E. Trenberth (1983).

Table 3. Descriptive Statistics for Seoul's 'Ttareungi' Variables

Variable		Cnt	Mean	SD	MinV	MaxV
DV	Number of Uses (bikes/day)	363	27,718	16,490	1,037	64,645
WC	Max Temp(℃)	363	17.95	11.67	-10.70	39.60
	Max Wind(m/s)	363	3.96	1.12	1.60	8.80
	Avg Wind(m/s)	363	1.73	0.61	0.70	4.10
	Avg Hum(%)	363	57.47	15.10	22.90	97.00
	Precipi(Precip:1,Clear:0)	363	0.35	0.48	0.00	1.00
Weekend, Holiday(W/H:1, WD:0)		363	0.32	0.47	0.00	1.00
SV	Spring(Spring : 1, other : 0)	363	0.25	0.43	0.00	1.00
	Summer(Summer : 1, other : 0)	363	0.25	0.44	0.00	1.00
	Winter(Winter : 1, other : 0)	363	0.25	0.43	0.00	1.00

Table 4. Descriptive Statistics for New York's 'Citi Bike' Variables

Variable		Cnt	Mean	SD	MinV	MaxV
DV	Number of Uses (bikes/day)	363	48,022	19,480	1,922	80,651
WC	Max Temp(℃)	363	17.59	10.03	-4.60	35.60
	Max Wind(m/s)	363	25.18	7.96	11.6	48.8
	Avg Wind(m/s)	363	9.54	3.27	3.9	20.7
	Avg Hum(%)	363	67.66	13.94	29.2	96.5
	Precipi(Precip:1,Clear:0)	363	0.69	0.46	0.00	1.00
Weekend, Holiday(W/H:1, WD:0)		363	363	0.31	0.46	0.00
SV	Spring(Spring : 1, other : 0)	363	0.25	0.43	0.00	1.00
	Summer(Summer : 1, other : 0)	363	0.25	0.44	0.00	1.00
	Winter(Winter : 1, other : 0)	363	0.25	0.43	0.00	1.00

4. RESULTS ANALYSIS

4.1 Results of "Ttareungi" Data Analysis by Multiple Regression Analysis

The analysis of Seoul's public bicycle "Ttareungi" data shows an R-squared value of 0.794, indicating that the model explains approximately 79.4% of the variance in the dependent variable. The adjusted R-squared value was 0.789, accounting for the number of variables and the sample size (Table 1). The regression analysis results were interpreted at a 95% confidence level.

Table 5. "Ttareungi" data regression analysis results

Variable	Estimate	Std. Error	t-value	p-value	VIF
Intercept (const)	0.408	0.031	13.052	<0.001	-
Maximum Temperature (°C)	0.010	0.001	13.757	<0.001	4.968
Maximum Wind Speed (m/s)	-0.014	0.006	-2.526	0.012	2.434
Average Wind Speed (m/s)	0.024	0.010	2.326	0.021	2.604
Average Relative Humidity (%)	-0.003	0.000	-7.235	<0.001	2.084
Precipitation	-0.019	0.011	-1.702	0.090	1.734
Weekend, Holiday	-0.029	0.009	-3.340	<0.001	1.017
Spring	-0.181	0.012	-15.716	<0.001	1.579
Summer	-0.165	0.015	-11.282	<0.001	2.580
Winter	-0.190	0.017	-11.279	<0.001	3.394
R-squared	0.794				
Adjusted R-squared	0.789				

The coefficients derived from the multiple regression analysis show the effect of each variable on the usage of "Ttareungi." The analysis excluded variables that were statistically insignificant or suspected of multicollinearity, based on their p-values and VIF values.

The statistical significance of each coefficient was evaluated using the p-value, which determines whether the effect of the variable on the dependent variable is statistically significant (i.e., not due to chance). Statistical significance was set at $p < 0.05$, whereas no statistical significance was set at $p > 0.05$. In addition, VIF values close to 1 and above 5, respectively, indicate a low and a high likelihood of multicollinearity among the independent variables.

According to the analysis, the maximum temperature (°C) positively influenced bicycle usage, with a standardized coefficient value of 0.010. The p-value was less than 0.001, indicating statistical significance and that higher temperatures led to increased bicycle usage. The maximum wind speed (m/s) negatively affects bicycle usage, with a standardized coefficient of -0.014 and a p-value of 0.012, indicating statistical significance and that strong winds hindered bicycle usage.

The average wind speed (m/s) positively affected bicycle usage, with a standardized coefficient of 0.024. The p-value was 0.021, indicating that moderate wind speed positively influenced bicycle usage.

The average relative humidity (%) negatively affected bicycle usage, with a standardized coefficient of -0.003. The p-value was less than 0.001, indicating statistical significance and that high humidity decreased bicycle usage.

Precipitation negatively affects bicycle usage, with a coefficient of -0.019 and a non-significant p-value of 0.09, suggesting that further research is needed to confirm if rainy weather significantly reduces bicycle usage.

Weekends and holidays negatively impact bicycle usage, with a coefficient of -0.029. The p-

value was less than 0.001, indicating high statistical significance and a reduction in bicycle usage on weekends and holidays.
Seasonal variables significantly affected bicycle usage, with spring, summer, and winter negatively affecting usage compared with the base variable, autumn. The coefficients for spring, summer, and winter were -0.181, -0.165, and -0.190, respectively, with p-values less than 0.001, indicating statistical significance. The negative coefficients indicate that bicycle usage decreased in these seasons compared with autumn.

4.2 Results of "Citi Bike" Data Analysis by Multiple Regression Analysis

The analysis of New York City's public bicycle "Citi Bike" data shows an R-squared value of 0.811, indicating that the model explains approximately 81.1% of the variance in the dependent variable. The adjusted R-squared value was 0.806, accounting for the number of variables and the sample size (Table 2). The regression analysis results were interpreted at a 95% confidence level.

Table 6. "Citi Bike" data regression analysis results

Variable	Estimate	Std. Error	t-value	p-value	VIF
Intercept (const)	0.304	0.023	13.213	<0.001	-
Maximum Temperature (°C)	0.009	<0.001	20.346	<0.001	2.708
Maximum Wind Speed (m/s)	0.001	<0.001	3.818	<0.001	1.428
Average Wind Speed (m/s)	-0.002	<0.001	-2.777	0.006	1.132
Average Relative Humidity (%)	-0.002	<0.001	-7.900	<0.001	1.418
Precipitation	-0.024	0.006	-3.759	<0.001	1.337
Weekend, Holiday	-0.068	0.006	-11.869	<0.001	1.049
Spring	-0.051	0.008	-6.611	<0.001	1.654
Summer	-0.043	0.009	-4.710	<0.001	2.368
Winter	-0.064	0.009	-7.294	<0.001	2.128
R-squared	0.811				
Adjusted R-squared	0.806				

The coefficients derived from the multiple regression analysis explain the effect of each variable on the usage of "Citi Bike" in New York City.
The regression analysis results for the Citi Bike data revealed the specific impacts of various independent variables on bicycle usage. The maximum temperature (°C) positively influenced bicycle usage, with a standardized coefficient of 0.009. The p-value was less than 0.001, indicating statistical significance and that higher temperatures promoted bicycle usage.
The maximum wind speed (m/s) also positively influenced bicycle usage, with a coefficient of 0.001. The p-value was less than 0.001, indicating high statistical significance and that strong winds could positively influence bicycle usage.
Conversely, the average wind speed (m/s) negatively affects bicycle usage, with a coefficient of -0.002. The p-value was 0.006, indicating statistical significance and that increased average wind speeds reduced bicycle usage.
The coefficient for precipitation is -0.024, indicating that rainy days decrease bicycle usage. The p-value was <0.001, indicating statistical significance.
The coefficient for weekends and holidays is -0.068, indicating that bicycle usage decreases on weekends and holidays. The p-value was <0.001, indicating statistical significance.
The coefficients for seasonal variables are -0.051, -0.043, and 0.064 for spring, summer, and winter, respectively, indicating a decrease in bicycle usage in these seasons compared with the base variable, autumn. The p-values for spring, summer, and winter were all less than 0.001, indicating statistical significance.

4.3 Results of "Ttareungi" Data Analysis by LSTM

Prior to training, the input features and target values were normalized to a 0–1 range using a Min–Max scaler. Consequently, the reported Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) reflect differences in this scaled space. Interpreting these metrics in terms of the actual bicycle counts requires an inverse transformation of the predictions. Under these conditions, the LSTM model, when applied to the "Ttareungi" dataset, produced an MAE of 0.0648 and an RMSE of 0.0874. The model employed a 7-day input window (time_steps=7), two LSTM layers (64 and 32 units, both with "tanh" activation), Dropout layers (0.2), Adam optimization, and MSE loss. Training continued for up to 100 epochs with early stopping (patience=10) to mitigate overfitting.

The training and validation loss curves (Figure 4) generally decreased over time, indicating that the model learned identifiable patterns without severe overfitting. Although a comparison between the actual and predicted usages (Figure 5) showed that the LSTM model did not perfectly replicate all day-to-day fluctuations, it did capture broad temporal trends and responded to seasonal and environmental variations. These results suggest that the LSTM approach can provide reasonable estimates of bicycle usage under the given scaled conditions, potentially informing operational and planning decisions.

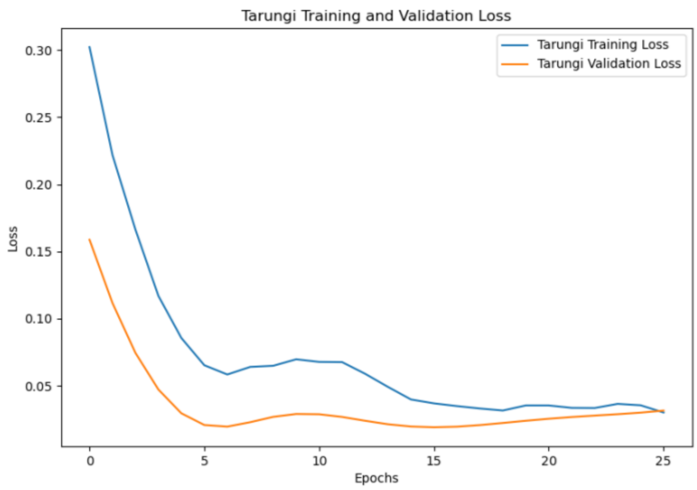

Figure 4. "Ttareungi" training and validation loss

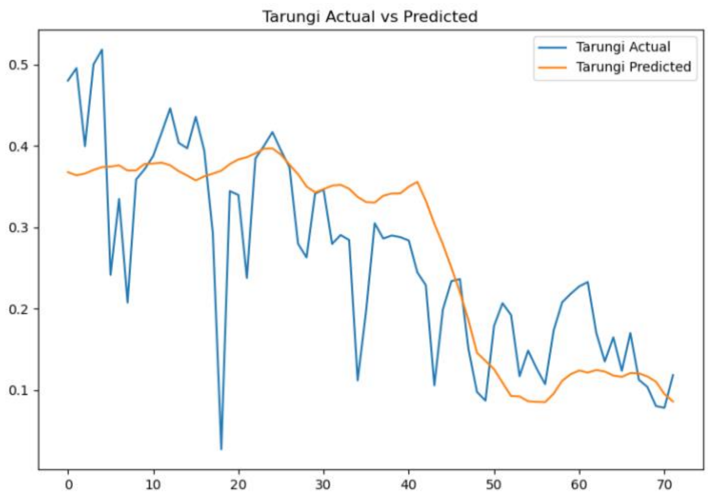

Figure 5. "Ttareungi" actual vs predicted

4.4 Results of “Citi Bike” Data Analysis by LSTM

The “Citi Bike” dataset underwent the same normalization process, scaling all inputs and the target to a 0–1 range. Under these conditions, an LSTM model with a 7-day input window, two LSTM layers (64 and 32 units, “tanh” activation), Dropout layers (0.2), Adam optimization, and MSE loss, trained for up to 100 epochs with early stopping, yielded an MAE of 0.0653 and an RMSE of 0.0790 on the scaled data. Although these metrics are not directly interpretable in the original units without inverse transformation, they indicate that the model′s predictions are reasonably close to the scaled ground truth values.

The training and validation loss curves (Figure 6) showed no pronounced overfitting, and the predicted values (Figure 7) aligned with the general usage patterns over time. Although the LSTM model did not capture every short-term fluctuation, it effectively responded to seasonal and environmental changes. Although this does not imply consistent superiority over simpler methods, the LSTM model provides valuable insights into the temporal dynamics of public bicycle usage, especially when interpreted within the scaled data context.

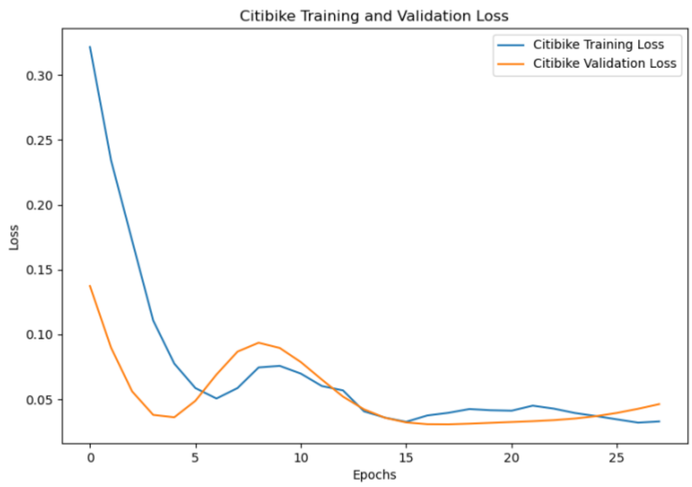

Figure 6. “Citi Bike” training and validation loss

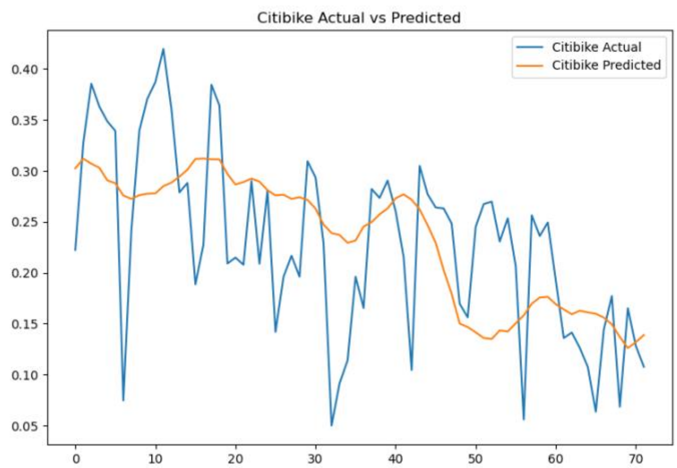

Figure 7. “Citi Bike” actual vs predicted

4.5 Comparative Analysis

This section compares factors influencing bicycle usage patterns of "Ttareungi" and "Citi Bike" using both regression and LSTM analyses. The regression results identified several key variables shaping usage patterns in both Seoul and New York City, including maximum temperature, wind speed, humidity, precipitation, weekends/holidays, and seasons.

In the regression analysis, the maximum temperature positively influenced bicycle usage in both cities, with standardized coefficients of 0.010 ($p < 0.001$) in Seoul and 0.009 ($p < 0.001$) in New York City. These findings align with those of the LSTM analysis, which also indicates that warmer conditions generally correspond to increased usage. Although the LSTM model does not directly provide standardized coefficients, the observed temporal patterns reinforce the importance of temperature as a consistent driver of bicycle demand across different urban contexts.

Wind speed presented contrasting effects. While maximum wind speed negatively affected bicycle usage in Seoul (coefficient -0.014, $p = 0.012$), it showed a minimal positive impact in New York City (coefficient 0.001, $p < 0.001$). The LSTM results similarly captured differing responses between cities over time, suggesting that variations in the local climate and infrastructure may influence how users react to wind conditions.

Humidity and precipitation both exhibited negative relationships with bicycle usage, with humidity reducing usage in Seoul and New York City at comparable levels (coefficients of -0.003 and -0.002, respectively, both $p < 0.001$) and precipitation having a more pronounced negative impact in New York City (-0.024, $p < 0.001$) than in Seoul (-0.019, $p = 0.09$). While the LSTM model did not directly quantify variable importance, its temporal forecasts responded to periods of increased humidity and rainfall by predicting lower usage, which was consistent with the regression findings.

Weekends and holidays were also linked to decreased bicycle usage in both cities. Although the regression analysis indicated negative coefficients, the LSTM forecasts reflected similar temporal drops in usage, highlighting a shared pattern despite potential cultural differences in leisure activities.

Seasonal variations were apparent as well. Both the regression and the LSTM analyses revealed that usage typically declined in seasons other than autumn, indicating that certain periods of the year — whether due to temperature, daylight hours, or weather patterns — were less conducive to cycling.

In summary, the regression approach provides a clear linear framework to quantify the influence of specific factors, whereas the LSTM model incorporates temporal and nonlinear aspects, offering a more dynamic understanding of how these factors interact over time. Together, these two methods underscore the significant role of environmental and temporal conditions in shaping bicycle usage patterns, informing strategies for more efficient and adaptive public bicycle system management.

5. CONCLUSION

5.1 Summary and Implications of the Present Study

The present study aims to clarify the characteristics of bicycle-sharing systems in Seoul and New York City by examining how various environmental factors influence bicycle usage patterns. The analysis identified temperature, wind speed, humidity, precipitation, weekends/holidays, and seasonal variation as significant factors in both cities.

Multiple regression analysis revealed that increasing temperatures generally led to higher bicycle usage, although this effect was more pronounced in Seoul. Mild weather conditions were consistently beneficial, aligning with existing findings that usage tends to rise until around 28 °C (Heaney *et al.*, 2019). Wind speed showed contrasting results: while higher maximum wind speeds reduced usage in Seoul, they had a minimal effect in New York City. Similarly, humidity and precipitation negatively affected usage in both cities, with precipitation exerting a stronger influence in New York City. Weekends and holidays were associated with decreased usage, more notably in New York City, possibly reflecting cultural and social differences. Seasonal variations also influenced demand, with spring, summer, and winter registering lower usage than autumn, particularly in Seoul.

The LSTM model complemented the regression analysis by accounting for the temporal dependencies and potential nonlinearities in the data. Although the regression model highlighted the linear relationships and provided clear quantitative estimates, the LSTM approach captured more dynamic patterns, reflecting how usage evolved over time and responded to changing environmental conditions. Although the LSTM model did not perfectly predict all day-to-day fluctuations, it provided valuable insights into the temporal nature of bicycle usage patterns.

Based on these findings, bicycle-sharing program operators and urban planners should consider seasonal and weather-related demand fluctuations when adjusting for bicycle availability, maintenance schedules, and infrastructure support. For instance, reducing the number of available bicycles in winter and improving waterproofing features during rainy seasons could enhance the overall user experience and operational efficiency. Furthermore, as long-term climate changes are likely to alter local weather patterns, proactive planning can help maintain sustainable and adaptive bicycle-sharing systems.

5.2 Study Limitations and Future Research Directions

This study primarily focused on environmental factors and their relationship with public bicycle usage. However, modern urban settings are influenced by a broader set of variables, including user demographics, digital infrastructure, integration with public transportation, and urban design. Future research should incorporate these additional factors, alongside more advanced modeling techniques, to develop a comprehensive understanding of bicycle usage patterns.

Expanding the range of data sources and variables could reveal deeper insights into the behavioral and infrastructural elements underlying bicycle demand. Such research may ultimately guide the design of more resilient, equitable, and efficient bicycle-sharing systems that adapt not only to environmental conditions but also to the evolving needs and preferences of urban populations.

REFERENCES

Kyung-Eun, S. (2019) A study on the characteristics of physical environments affecting the usage of public bike in Seoul, Korea: focusing on the different influences of factors by distance to bike station. *M.S. thesis, Graduate School, Hanyang University, Seoul, Korea*.

PBSC Urban Solutions (2021) The Meddin Bike-Sharing World Map, Oct. 27, 2021.

Faghih-Imani, A., Hampshire, R., Marla, L., Eluru, N. (2017) An empirical analysis of bike sharing usage and rebalancing: Evidence from Barcelona and Seville. *Transportation Research Part A: Policy and Practice*, 97, 177–191.

Choi, Y., Lee, J., Park, G. (2021) Comparison of attributed network clustering approaches

to analyzing Seoul public bike stations. *Stat*, 11, e415.
Citi Bike (2020) NYCBS December 2020 Monthly Report.
Jang, J., Lee, S. B., Lee, Y. I., Lee, M. Y. (2018) Effects of seasonal and membership characteristics on public bicycle traffic: Focusing on the Seoul Bike. *Journal of the Korean Society of Road Engineers*, 20(4), 47–58.
Joo, H.-M., Lee, Y.-I., Lee, J.-S. (2018) Analysis of users' influencing area and characteristics of public bicycle in Seoul. *Proceedings of the 2018 Autumn Conference Korean Society of Intelligent Transport Systems*, 565–573.
An, R., Zahnow, R., Pojani, D., Corcoran, J. (2019) Weather and cycling in New York: The case of Citibike. *Journal of Transport Geography*, 77, 97–112.
Wang, W. (2015) Forecasting bike rental demand using New York Citi Bike data. *M.S. thesis, Technological University Dublin, Dublin, Ireland*.
Trenberth, Y., Wang, S. (2017) A forecast for bicycle rental demand based on random forests and multiple linear regression. *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*, Wuhan, China, 101–105.
Heaney, A. K., Carrion, D., Burkart, K., Lesk, C., Jack, D. (2019) Climate change and physical activity: Estimated impacts of ambient temperature on bikeshare usage in New York City. *Environmental Health Perspectives*, 127(3), 037002.
Choi, Y., Lee, J., Park, G. (2021) Comparison of attributed network clustering approaches to analyzing Seoul public bike stations. Stat, 11, e415.
Singhvi, D., Singhvi, S., Frazier, P., Henderson, S. G., O'Mahony, E., Shmoys, D. B., Woodard, D. B. (2015) Predicting bike usage for New York City's bike sharing system. AAAI Workshop: Computational Sustainability.
Park, J., Namkung, O. S., Ko, J. (2023) Changes in public bike usage after the COVID-19 outbreak: A survey of Seoul public bike sharing users. Sustainable Cities and Society, 96, 104716.
Kumar Dey, B., Anowar, S., Eluru, N. (2021) A framework for estimating bikeshare origin destination flows using a multiple discrete continuous system. Transportation Research Part A: Policy and Practice, 144, 119–133.
Hwang, S.-G., Lee, S. (2018) Analysis of neighborhood environmental factors affecting bicycle accidents and accidental severity in Seoul, Korea. Journal of Korea Planning Association, 53(7), 49–66.
Branion-Calles, M., Nelson, T., Winters, M. (2019) Associations between individual characteristics, availability of bicycle infrastructure, and city-wide safety perceptions of bicycling: A cross-sectional survey of bicyclists in 6 Canadian and U.S. cities. Transportation Research Part A: Policy and Practice, 123, 229–239.
Trenberth, K. E. (1983) What are the seasons? *Bulletin of the American Meteorological Society*, 64(11), 1276–1282.
Shi, Y.; Zhang, L.; Lu, S.; Liu, Q. (2023). Short-Term Demand Prediction of Shared Bikes Based on LSTM Network. *Electronics*, 12(6), 1381. https://doi.org/10.3390/electronics12061381