“Weather” to Rent a Bike

Jonathan Pufall
WISCONSIN SCHOOL OF BUSINESS
December 2024

1. Executive Summary

This report examines bike rental patterns for both casual and registered users, focusing on seasonal trends, weekly usage patterns, and factors influencing rental demand. The analysis utilizes statistical measures such as average rentals, standard deviation, quartiles, and regression analysis to uncover key drivers of rental behavior. Registered users make up the majority of bike rentals, averaging 155 rentals compared to 37 for casual users. Their usage is more predictable, with higher variability linked to commuting and specific time-based needs. In contrast, casual users display lesser variability and lower average rentals, suggesting more leisure-driven usage. A pie chart confirms that registered users account for 81% of all rentals, emphasizing their central role in demand.

Seasonal patterns reveal that summer consistently drives the highest rental demand, with a sharp decline in winter. While both user types are greatly affected by seasonal changes, registered users maintain relatively steady usage year-round, while casual users are more influenced. Key months, like June 2012 and September 2011, show peak demand likely driven by favorable weather and events. Weekly and hourly patterns indicate that registered users primarily rent bikes on weekdays during commute hours (8 AM and 5 PM), while casual users rent on weekends for recreational purposes. Despite these differences, both user groups share similar peak times on weekends, with registered users still dominating weekday rentals.

Regression analysis with an R-squared value of 0.33 shows that temperature, season, and year are significant drivers of bike rentals. Each degree rise in temperature results in approximately four additional bike rentals, and rentals increase during fall and spring. Interestingly, the summer season itself showed no significant effect when temperature was included as a variable. A two-tailed t-test confirms significant differences in rental behavior between weekdays and weekends for both user groups, with casual users showing higher rentals on weekends, and registered users predominantly using bikes for commuting during weekdays.

These findings reveal the distinct behaviors of registered and casual users, with registered users exhibiting predictable, commute-driven patterns and casual users displaying more variable, leisure-based behavior. The analysis offers insights into the factors that drive bike rental demand, including weather conditions, seasons, and user types.

Based on these findings, several recommendations are made: improving bike availability during peak seasons, offering seasonal pricing and promotions to incentivize rentals during off-peak months, optimizing fleet distribution for weekdays and weekends, responding to weather patterns with real-time data to adjust availability, and enhancing the mobile app for casual users with features like bike reservations and weather notifications. These strategies can help optimize bike availability, meet demand, and improve overall service for both user segments.

2. Introduction

This analysis focuses on understanding the impact of weather patterns on bike rental behavior for both casual and registered users. As urban bike-sharing systems continue to grow in popularity, understanding the factors that drive rental demand is crucial for optimizing operations, ensuring adequate bike availability, and providing tailored services to different user segments. One of the most significant factors influencing bike rental behavior is weather, as variations in temperature, humidity, and seasonal changes can directly affect user demand. Additionally, understanding the patterns of casual and registered users helps inform strategies for balancing bike availability and improving service delivery across different times of the year.

The primary objective of this report is to identify the relationship between weather patterns, time variables (such as weekdays versus weekends), and user demand for bike rentals. By examining how temperature, seasonality, and time of day affect rental trends, the report aims to provide actionable insights for optimizing bike rental operations, improving customer satisfaction, and predicting future rental behavior. Furthermore, the report intends to uncover any key differences between casual and registered users in terms of their rental patterns and how weather influences these behaviors. This understanding will enable operators to make informed decisions on fleet distribution, pricing strategies, and promotional campaigns tailored to different user groups.

The dataset used in this analysis spans several years and includes metrics related to bike rental behavior, such as the number of rentals per day, the type of users (casual or registered), and the weather conditions on those days. Key weather factors such as temperature, humidity, and seasonality are included, as well as user behavior metrics that capture rental frequencies, usage patterns, and peak times. This rich dataset allows for a comprehensive analysis of how different weather conditions impact the demand for bike rentals, enabling a more nuanced understanding of how to optimize bike-sharing services throughout the year.

3. Analysis

The analysis section of this report delves into the patterns and trends observed in bike rentals, with a particular focus on the impact of user type (casual vs. registered), differing time frames, and temperature. By examining rental data across various time periods and correlating it with key weather variables, this analysis aims to identify the key drivers influencing demand and how different factors, such as seasonality, day of the week, and weather, contribute to rental behavior. Understanding these trends is essential for optimizing bike availability, improving operational efficiency, and forecasting future demand.

This section will cover the findings of statistical analyses, including measures of central tendency (average rentals), variability (standard deviation), and the distribution of rentals (quartiles), as well as regression modeling to explore the relationship between weather conditions and rental patterns. Additionally, the section will present insights into how these factors vary across user groups and time periods, providing a comprehensive understanding of the dynamics driving bike rental usage.

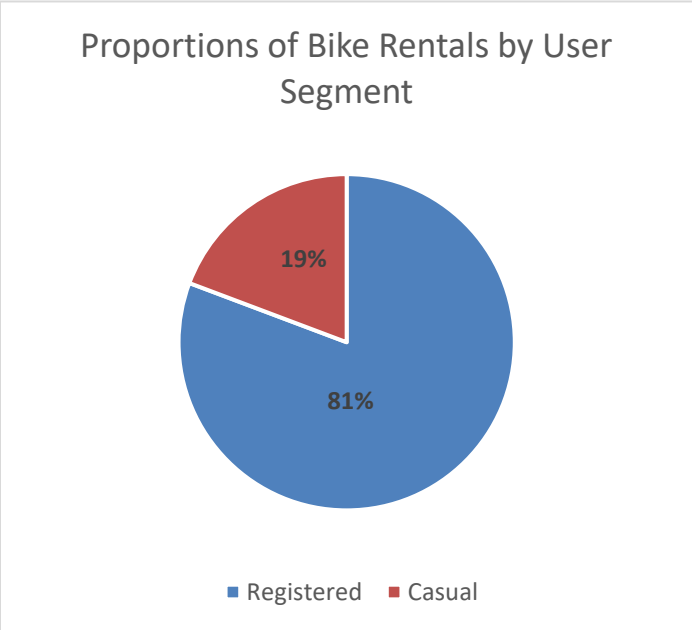
3.1 Bike Rental Patterns

Summary Statistics

Statistic Type	Registered User Data	Casual User Data
Average	155	37
1st and 3rd Quartiles	34 / 223	4 / 50
Standard Deviation	149.96	51.20

The analysis of bike rental summary data provides valuable insights into user behavior using various statistical measures. The average number of rentals for casual users is 37, while for registered users it is much higher at 155. This indicates that registered users contribute significantly more to the overall demand for bike rentals. The standard deviation for casual users is 51.20, compared to 149.96 for registered users. The higher standard deviation for registered users reflects greater variability in their rental patterns, likely influenced by commuting needs and time-specific usage. The first and third quartiles for casual users are 4 and 50, respectively, while for registered users, they are 34 and 223. These quartiles offer insight into the spread of rental data, helping to identify typical rental behaviors while minimizing the impact of outliers. They show that while casual users exhibit more variability, registered users maintain a more consistent base level of rentals, even on days with lower demand. The quartiles also suggest that casual usage tends to be more sporadic and dependent on specific factors like weekends or weather conditions.

Overall, these measures reveal that registered users consistently contribute the majority of rentals, with more predictable patterns, while casual users show less consistent demand.



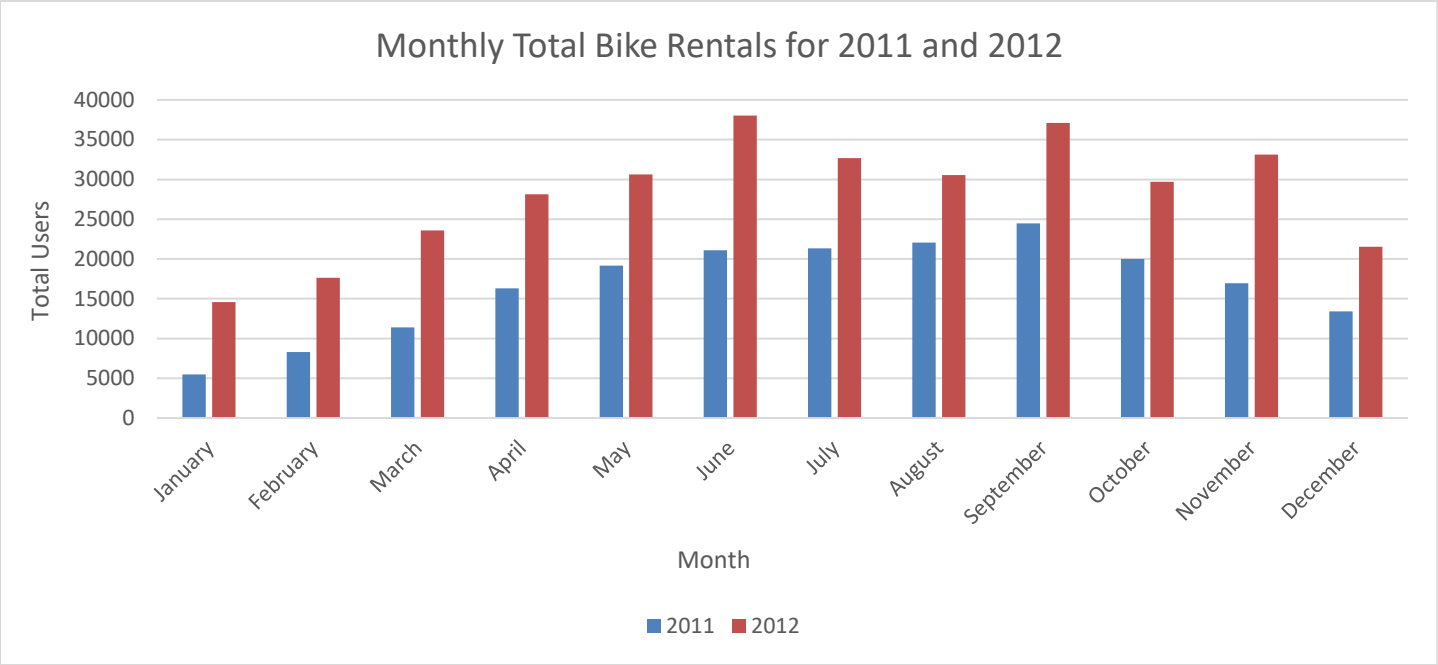
User Distribution and Rental Demand Breakdown

The pie chart illustrates a significant disparity in bike rental demand between registered and casual users, with registered users accounting for 81% of all rentals and casual users comprising only 19%. This data suggests that registered users are the primary drivers of overall demand. The larger share of rentals by registered users can be attributed to their regular, predictable usage, which is often driven by commuting needs, such as daily work or school travel. This group's consistent rental behavior likely reflects a need for reliable, convenient transportation.

The dominance of registered users suggests that bike rental services should focus on optimizing their offerings for this group, ensuring availability and reliability, especially during peak commuting hours. Casual users, while contributing less overall demand, still represent an important segment, and targeted marketing or promotions could encourage more frequent use. By understanding these two distinct user groups, bike rental operations can better tailor their services to maximize usage and improve user satisfaction across both segments.

Analysis of Seasonal Demand Patterns

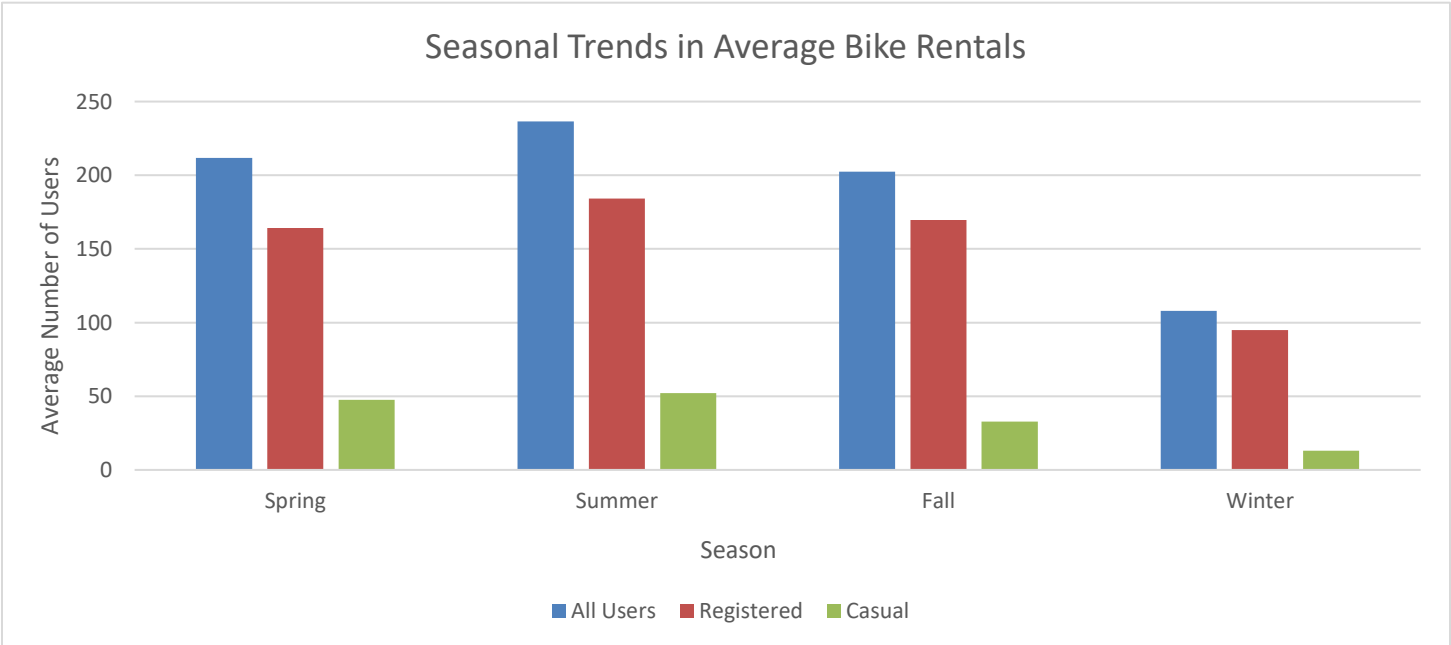
The following column chart reveals clear seasonal trends in bike rentals, with higher totals during the summer months in both years. June 2012 stands out as the busiest month overall, reflecting peak demand during warmer weather. Similarly, September 2011 sees a significant peak, suggesting that seasonal conditions and potentially other factors like events influenced demand. In contrast, winter months like January and February exhibit the lowest rental totals, reflecting reduced usage during colder weather.



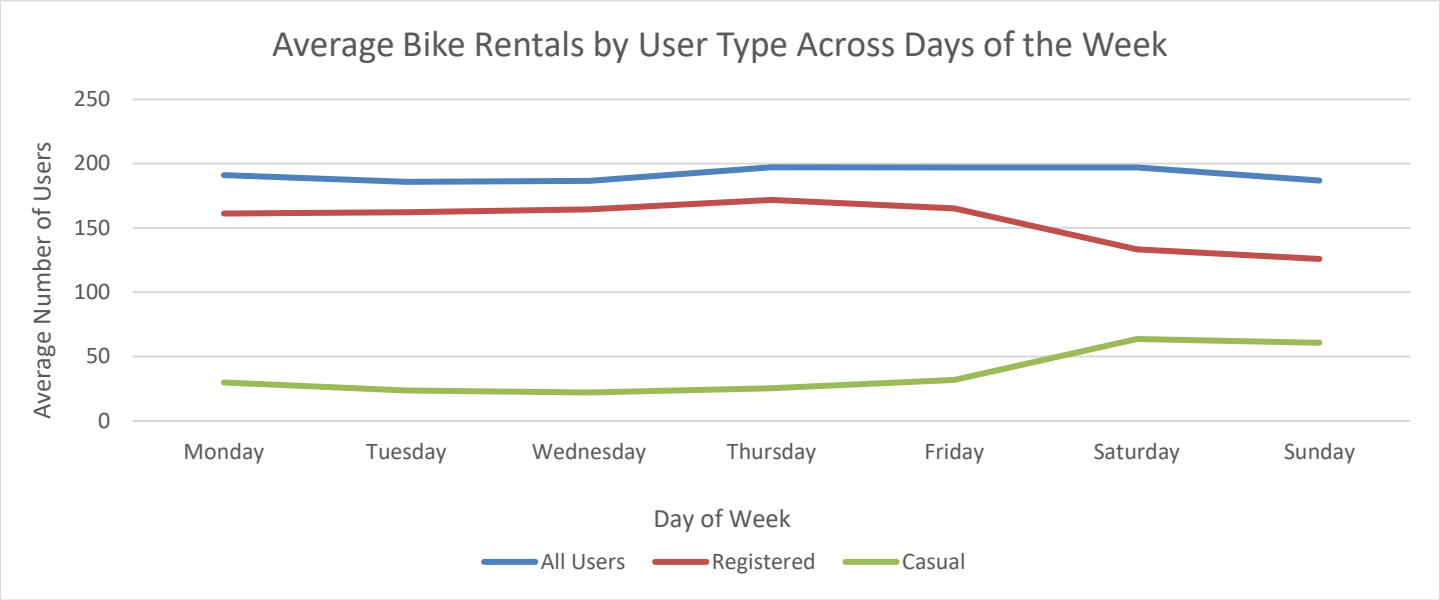
The column chart below illustrates a clear seasonal fluctuation in bike rental demand, with summer showing the highest rentals, averaging well over 200. This reflects the natural increase in outdoor activities during warmer weather, appealing to both casual and registered users. Fall and spring show moderate declines, with rentals dropping slightly due to cooler and less predictable weather, although still higher than winter.

Winter sees a sharp drop in rentals, especially among casual users, who are more affected by cold weather. Registered users also experience a decrease, but their usage is more driven by commuting needs, making them less sensitive to seasonal changes.

The chart also highlights that registered users consistently dominate the rental market across all seasons, accounting for the majority of rentals. This underscores the more predictable and year-round usage patterns of registered users compared to casual users, whose rentals peak in warmer months and fluctuate with the weather.

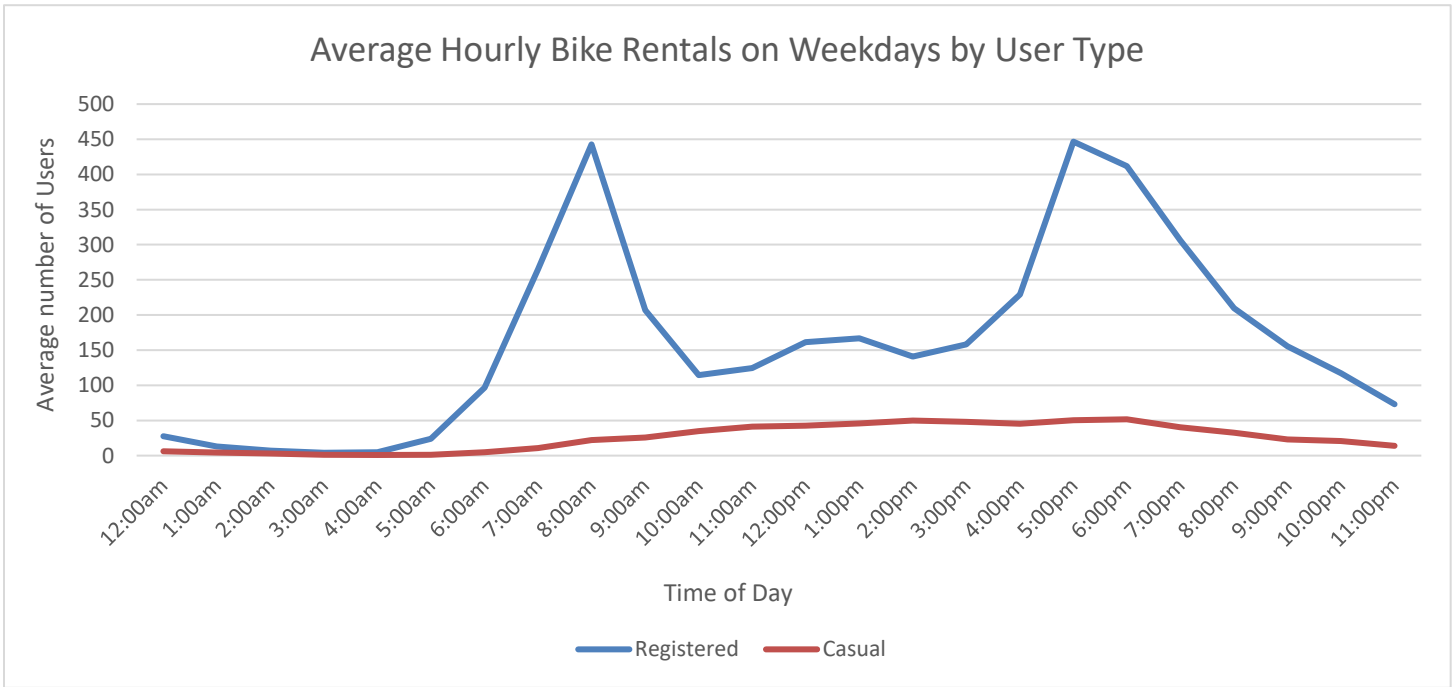


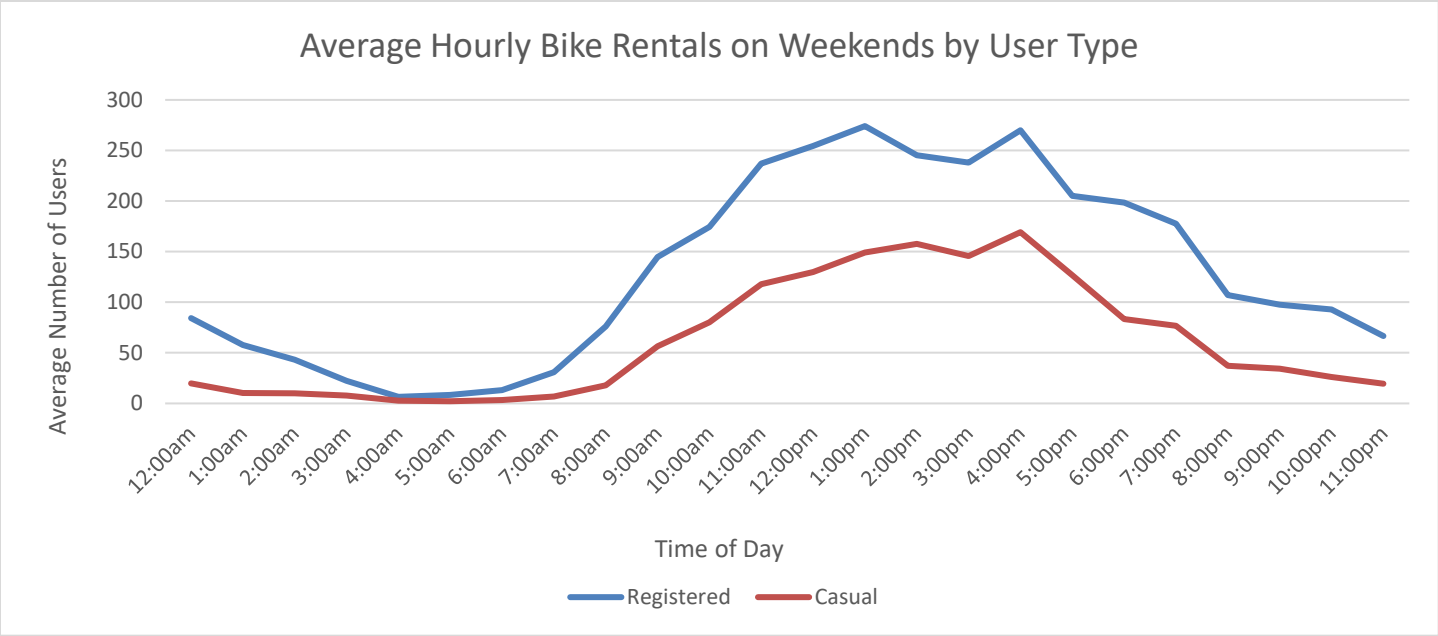
Weekday and Hourly Rental Patterns



The line chart above reveals clear distinctions in bike rental behavior across the days of the week for the three user segments. The registered users consistently have higher rental volumes from Monday to Friday, likely due to work-related commutes. Their rentals drop significantly on the weekend, which could indicate that registered users primarily rely on bikes for weekday activities. In contrast, casual users show a noticeable pattern of low rentals during the weekdays, but their rentals sharply rise on the weekend. This suggests that casual users may rent bikes more for leisure activities, especially on Saturdays and Sundays.

The “all users” line, which represents the combined total of both user types, remains fairly consistent throughout the week, with no rise on the weekends. This indicates that, although casual users experience a surge in rentals over the weekend, the consistent volume of registered users throughout the weekdays keeps the overall rental numbers stable. The overall pattern reinforces the idea that bike rentals are predominantly work-related for registered users, while casual users are more likely to rent bikes for recreational purposes during the weekends.





The two line charts above illustrate distinct patterns in hourly bike rental behavior for casual and registered users on weekdays versus weekends:

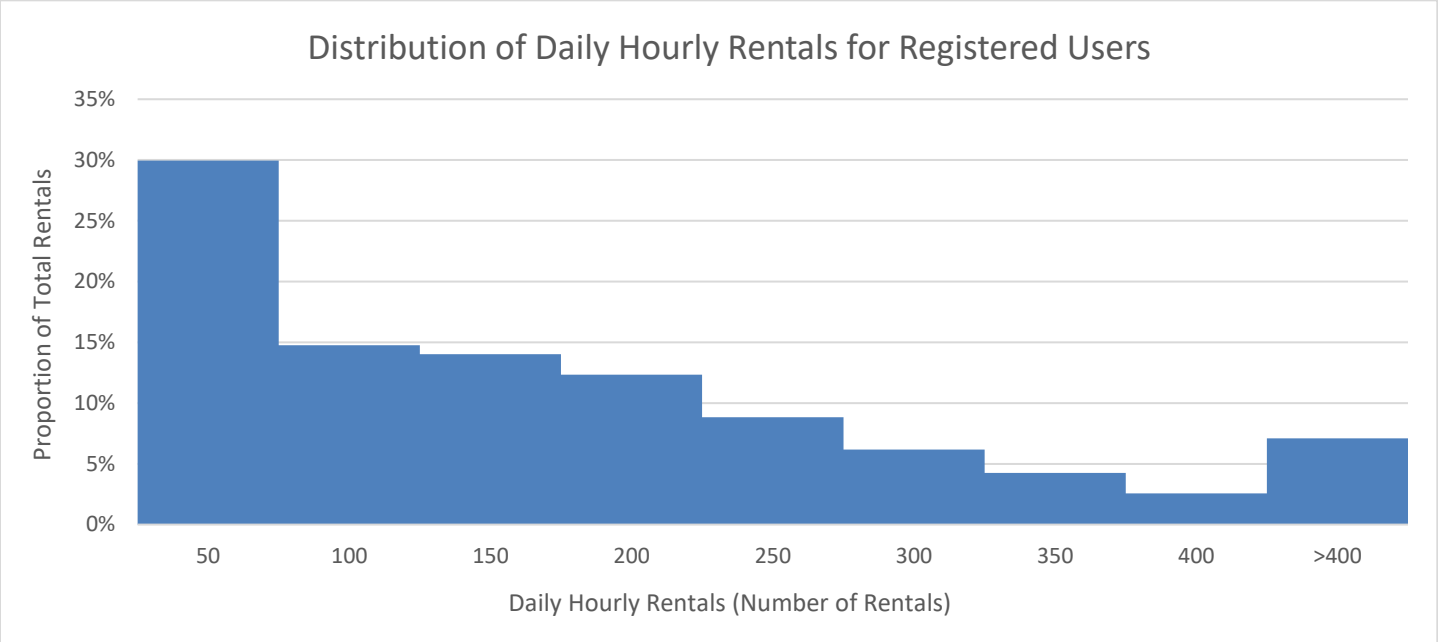
On weekdays, registered users exhibit pronounced peaks in rental activity at 8:00 AM and 5:00 PM, aligning with typical commuting hours. This reflects that registered users are likely using bikes as a primary mode of transportation to and from work or school. In contrast, casual users show a relatively steady rental volume throughout the day, with a shallow peak between 3:00 PM and 6:00 PM, likely driven by leisure or non-commute activities. On weekends, the patterns for casual and registered users become more similar. Both lines peak between 1:00 PM and 4:00 PM, suggesting increased recreational bike use during the afternoons when weather conditions and free time are favorable. However, the casual users' line remains lower than the registered users' line, consistent with their smaller share of the total user base.

The variability in the registered users' weekday pattern, driven by commuting behavior, contrasts with the more stable and leisure-oriented usage by casual users. This difference diminishes on weekends, as registered users shift their focus away from commuting, resembling the casual users' behavior.

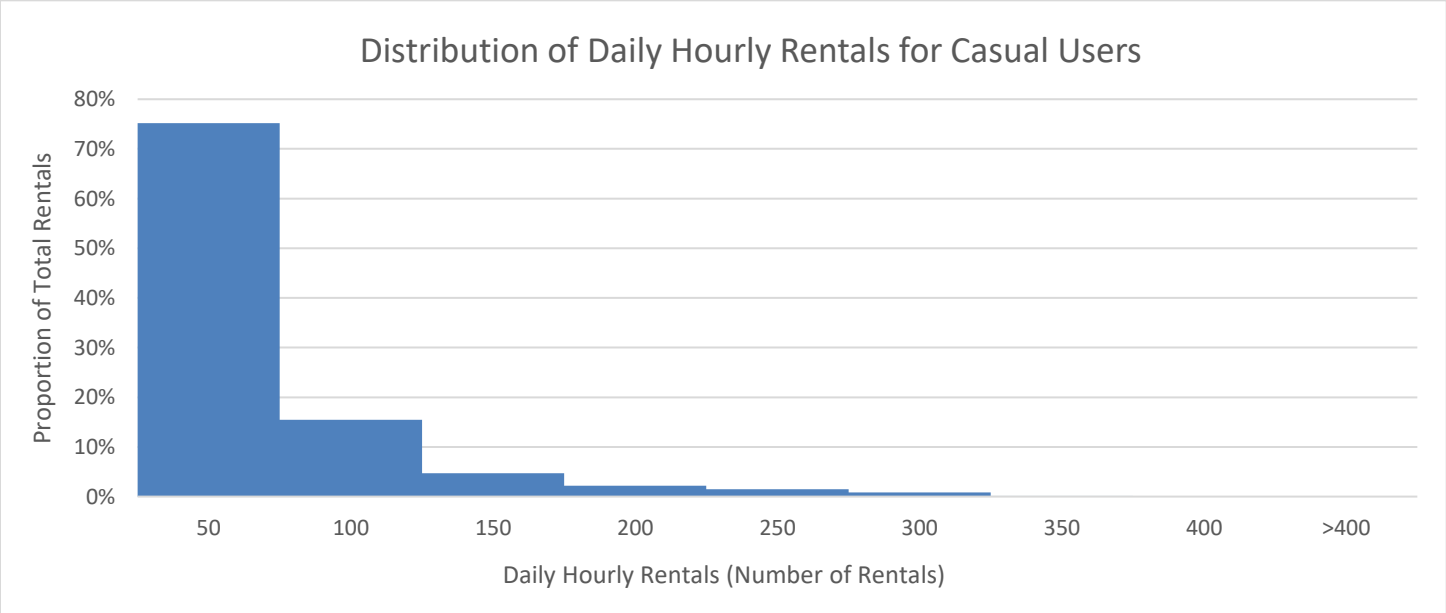
Analysis of Hourly Rental Distribution by User Type

The histograms provide a detailed view of the distribution of daily hourly rentals for both casual and registered users, with the y-axis representing proportions (percentages) of the total number of rentals. This analysis helps to highlight the distinct rental behaviors between the two groups by displaying how rentals are spread across different levels of daily usage. By focusing on the proportion of rentals within each daily rental bin, the histograms reveal the variation in demand and rental frequency for each user type, providing a clear comparison of the overall rental patterns.

For registered users, the distribution is relatively spread out, with a mode in the smallest bin (50 rentals) at approximately 30% of the observations. The proportions steadily decline as the bin size increases, with notable activity across bins up to 400 rentals and a smaller proportion (7%) exceeding 400 rentals. This indicates that registered users have a wide range of daily rental patterns, possibly driven by consistent use during commuting hours or varying levels of demand across the week.

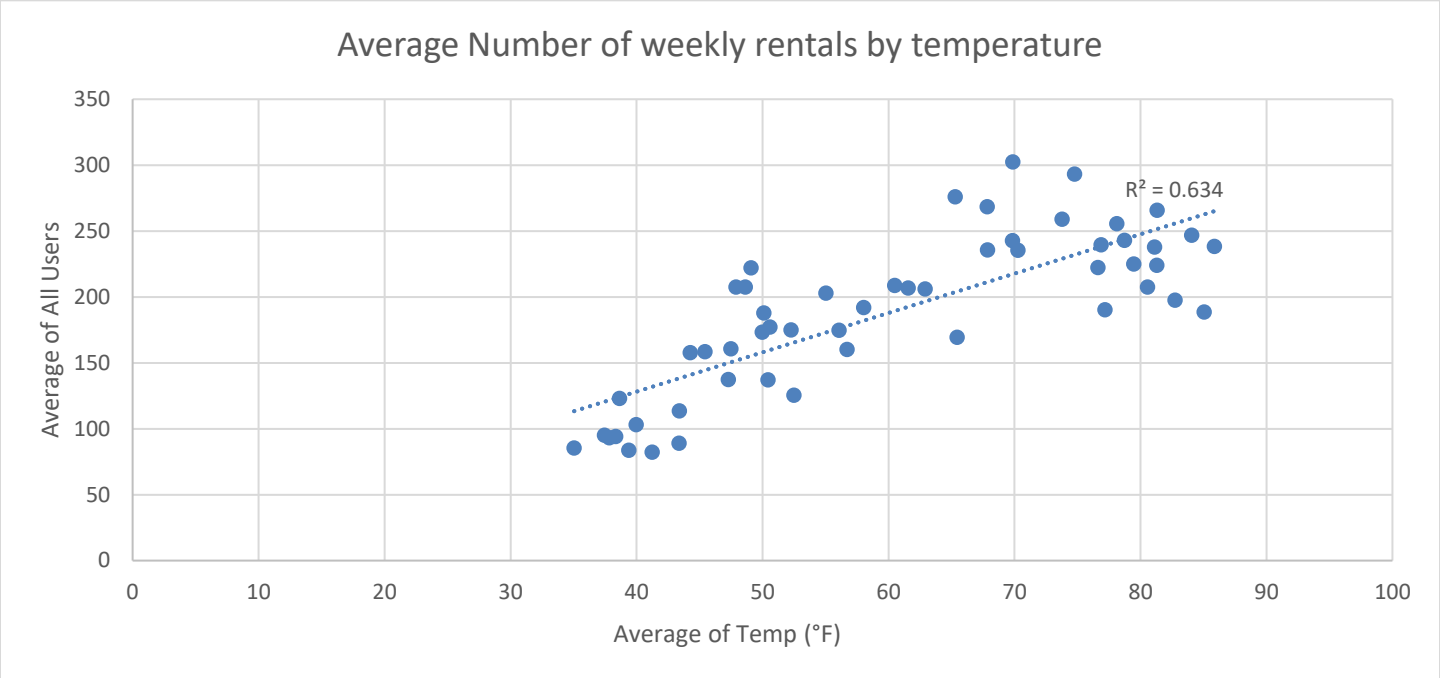


In contrast, for casual users, the distribution is much more concentrated. Over 75% of daily rentals fall within the smallest bin (50 rentals), and the proportions sharply decline as the bins increase. There is virtually no activity beyond the 350 rentals bin. This highly skewed distribution highlights that casual users typically rent bikes in small volumes, likely for short, infrequent leisure activities.



The stark contrast between the two histograms highlights the fundamentally different usage patterns of casual and registered users. Registered users show greater variability in their rental behavior, with a wider spread of daily rental volumes and higher maximum rentals. This variability likely reflects their regular, intentional use of bikes, often driven by commuting needs or routine activities. In contrast, casual users exhibit a highly skewed distribution, with the vast majority of their rentals concentrated in the lower rental bins, signifying a much more sporadic, low-volume usage pattern. Their rentals are typically infrequent and driven by leisure or seasonal factors, rather than the consistent, demand-driven usage seen in registered users.

Relationship Between Temperature and Bike Rentals



The scatterplot above displays the relationship between temperature (in Fahrenheit) and the average number of bike rentals. The regression equation derived from the data is: $y = 2.9854x + 8.7362$

In the equation, y represents the average number of bike rentals, and x represents the temperature. This equation indicates that for each 1°F increase in temperature, the average number of bike rentals increases by approximately 2.9854. The y-intercept of 8.7362 suggests that even at very low temperatures, there are minimal rentals, though this value may not be meaningful outside the observed range.

The R-squared value of 0.634 signifies that approximately 63.4% of the variability in bike rentals can be explained by changes in temperature. This indicates a moderately strong positive correlation between temperature and bike rentals. While other factors might also influence bike rentals, temperature appears to be a significant driver of demand.

3.2 Regression Analysis

The analysis investigates whether there is a significant difference in average bike rentals between weekdays and weekends for three user groups: all users, casual users, and registered users. A two-tailed t-test at a 5% significance level is applied to evaluate these differences, with the following conclusions:

User Group	Sample Difference	Confidence Interval	Conclusion
All Users	0.38	(-14.86, 14.86)	No significant difference
Registered Users	35.75	(-5.37, 5.37)	Significant difference
Casual Users	-35.38	(-10.52, 10.52)	Significant difference

All Users: Since the sample difference (0.38) lies within the confidence interval, we fail to reject the null hypothesis. This means there is no significant difference in the number of bike rentals for all users between weekdays and weekends. For all users, rentals appear to be consistent between weekdays and weekends, indicating steady demand across the week.

Registered Users: Since the sample difference (35.75) lies outside the confidence interval, we reject the null hypothesis. This indicates a significant difference in rentals for casual users between weekdays and weekends. Casual users rent significantly more bikes on weekends than on weekdays. This pattern could reflect the tendency of casual users to rent bikes for leisure activities, which are more common during weekends.

Casual Users: Since the sample difference (-35.38) lies outside the confidence interval, we reject the null hypothesis. This indicates a significant difference in rentals for registered users between weekdays and weekends. Registered users rent significantly more bikes on weekdays than on weekends. This trend likely reflects the commuting habits of registered users, who may primarily use bikes for work-related travel during the week.

Description of Regression Analysis

The regression model predicts the number of bike rentals based on several independent variables: weather conditions, seasons, weekend vs. weekday, and year. Key statistics from the regression output include:

- R-squared = 0.33: This indicates that 33% of the variation in bike rentals is explained by the model.
- Significance F = 0.00: The model is statistically significant overall.
- Standard Error = 148.63: This represents the average distance that the observed values fall from the regression line.

Regression Equation

The regression equation, based on the provided coefficients, is:

All Rentals = 36.98 + (4.11 * Weekend) + (4.32 * Temperature) + (11.20 * Mist) + (-4.18 * Precipitation) + (25.89 * Spring) + (-8.49 * Summer) + (78.15 * Fall) + (76.88 * Year 2012) + (-2.92 * Humidity) + (0.51 * Windspeed)

Regression Table

Variable	Coefficient	Std. Error	t-Stat	P-value	Lower 95%	Upper 95%
Intercept	36.98	18.91	1.96	0.05	-0.10	74.06
Weekend	4.11	6.19	0.66	0.51	-8.03	16.26
Temp (F)	4.32	0.27	15.87	0.00	3.79	4.85
Mist	11.20	6.74	1.66	0.10	-2.02	24.42
Precipitation	-4.18	11.68	-0.36	0.72	-27.07	18.72
Spring	25.89	10.28	2.52	0.01	5.74	46.05
Summer	-8.49	13.15	-0.65	0.52	-34.28	17.30
Fall	78.15	8.83	8.85	0.00	60.84	95.46
Year 2012	76.88	5.67	13.56	0.00	65.77	88.00
Humidity	-2.92	0.17	-16.99	0.00	-3.26	-2.58
Windspeed	0.51	0.38	1.35	0.18	-0.23	1.25

Interpretation of Regression Results

Significant Factors

- Significant at the 5% level (P-value < 0.05):
    - Temperature (Temp F): Higher temperatures significantly increase bike rentals. For each additional degree Fahrenheit, rentals increase by 4.32 bikes.
    - Spring and Fall seasons: Rentals increase significantly in Spring (+25.89 rentals) and even more in Fall (+78.15 rentals), compared to the baseline (Winter).
    - Year 2012: Rentals in 2012 were significantly higher (+76.88 rentals) than in the baseline year.
    - Humidity: Rentals decrease significantly as humidity increases; for each percentage point increase in humidity, rentals drop by 2.92 bikes.

R-Squared

- The R-squared value (0.33) indicates that the model explains 33% of the variance in bike rentals. While this shows that the model captures some meaningful patterns, other unexplained factors likely influence bike rentals.

Non-Significant Coefficients

- Weekend: The coefficient (+4.11) is not statistically significant (P-value = 0.51), indicating no meaningful difference between weekend and weekday rentals.
- Mist and Precipitation: Mist has a positive but non-significant effect on rentals, while precipitation has a negative but also non-significant effect.
- Summer: Surprisingly, summer shows a non-significant coefficient (-8.49, P-value = 0.52). This suggests that rentals in Summer are not significantly different from Winter, contrary to expectations.

Summer Coefficient

- The non-significant Summer coefficient is unexpected because one might assume higher bike rentals during warmer months. However, the inclusion of Temperature as a variable likely captures much of the seasonal effect. Since Summer often coincides with high temperatures, the separate influence of Summer as a season is diminished.

4. Conclusion

The analysis of bike rental patterns reveals distinct behaviors between registered and casual users, as well as seasonal and daily variations in rental trends. The following key insights were identified:

1. User Segment Behavior:
    - Registered Users: Registered users dominate bike rentals, accounting for 81% of all usage. Their patterns are more predictable and significantly influenced by work-related commutes, as indicated by peak activity during weekday mornings and evenings. Additionally, registered users demonstrate greater variability in usage, likely driven by differing weekday and weekend needs.
    - Casual Users: Casual users contribute 19% of rentals and exhibit more sporadic patterns. Their usage is concentrated on weekends and warmer seasons, reflecting recreational purposes rather than consistent commuting needs.

2. Seasonal and Temporal Trends:
    - Seasonal analysis shows that demand is highest during summer and lowest in winter, aligning with expectations tied to weather conditions. Registered users maintain a consistent presence year-round, while casual users are more sensitive to seasonal changes.
    - Weekday vs. weekend patterns highlight distinct differences: registered users rent more bikes during weekdays for commuting, whereas casual users peak on weekends for leisure.
3. Hourly Usage Patterns:
    - On weekdays, registered users' rentals spike during commuting hours, while casual users maintain steady, low volumes. On weekends, both groups exhibit a more similar pattern, with afternoon peaks driven by recreational activity.
4. Regression and Correlation Analysis:
    - A positive correlation between temperature and bike rentals (R-squared = 0.634) indicates that warmer weather significantly boosts demand. Each 1°F increase in temperature leads to an average increase of nearly 3 rentals.
    - The multiple regression model explains 33% of the variation in bike rentals, identifying factors like temperature, season, and weather conditions as significant predictors.
5. Statistical Insights:
    - T-test results confirm significant differences in weekday vs. weekend rentals for both user groups, with registered users renting more on weekdays and casual users more on weekends.
    - Histogram distributions highlight the variability in registered users' rentals and the concentrated, low-volume nature of casual users' activity.

Recommendations

Based on these findings, the following recommendations are proposed:

1. Enhance Services for Registered Users: Focus on optimizing weekday operations, particularly during commuting hours, to cater to the consistent demand from registered users. Additional incentives for registered users during weekends could also help balance usage.
2. Target Casual Users During Peak Seasons: Develop targeted marketing campaigns for casual users in summer and fall, highlighting recreational opportunities and discounts for weekend rentals.
3. Seasonal Resource Allocation: Adjust resource allocation to account for seasonal fluctuations, ensuring sufficient capacity during high-demand summer months and cost-effective operations during the winter.
4. Data-Driven Strategies: Utilize the regression insights to predict rental demand based on weather forecasts, allowing dynamic pricing or fleet adjustments to maximize efficiency and revenue.

These conclusions underscore the importance of tailoring services to the distinct needs of registered and casual users while leveraging seasonal and environmental trends to optimize bike rental operations.

5. Appendix

The appendix includes notes on the data preparation process, detailing the steps taken to clean, transform, and organize the data for analysis. It also contains the elevator charts that highlight the most impactful trends and findings discussed in the case study.

5.1 Notes on Data Preparation

In preparing the data for analysis, I focused on ensuring its accuracy and completeness. This involved identifying and removing any duplicate or erroneous records that could distort the results. Thankfully, the dataset had only a few duplicates and errors. I addressed records with missing or incorrectly formatted values, especially in the datetime, weather, and temp columns, to ensure consistency and accuracy. Additionally, any outliers in the temperature and humidity readings were handled appropriately to avoid skewing the analysis.

I reviewed key fields such as season, holiday, windspeed, and user data (casual, registered, and all users) to ensure they were correctly recorded. Some records had missing values in columns like Mist and Precipitation, which I either excluded from the analysis or imputed based on surrounding data to preserve data integrity.

Once cleaned, I proceeded with the analysis, examining seasonal trends, weather impacts on bike rentals, and correlations between temperature and rental patterns. Overall, the dataset was of good quality, with only minor issues that didn't significantly affect the overall findings. However, attention to these small inconsistencies is important when interpreting the final results.

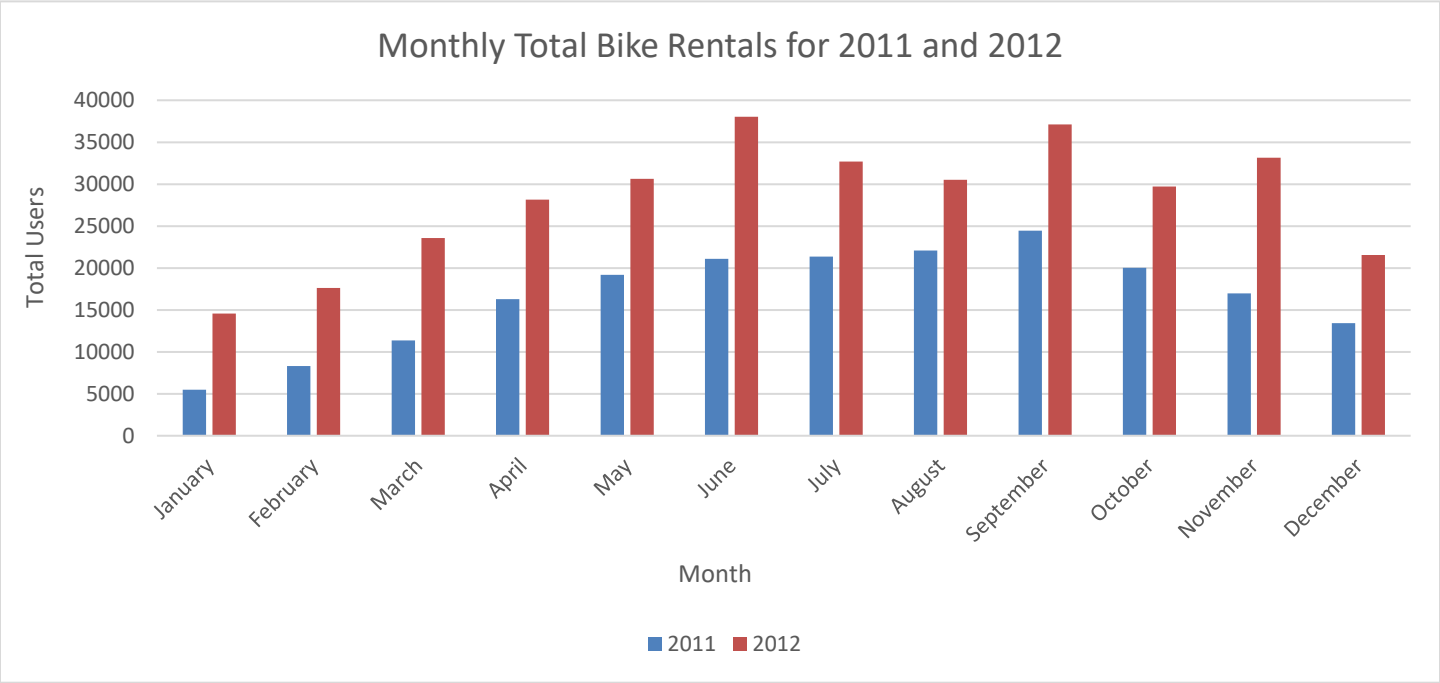
5.2 Elevator Charts

These three charts offer a comprehensive blend of high-level trends and actionable insights, effectively distilling complex data into clear, visual representations. By highlighting key patterns and correlations, such as seasonal demand fluctuations and the influence of temperature on rentals, they empower executives with the essential information needed to drive strategic decision-making. The insights gleaned from these charts enable informed decisions that can optimize operations, enhance customer experience, and drive business growth, ensuring that the company is well-positioned to respond to demand shifts and maximize efficiency across all user segments.

1. Regression Table: This chart provides the most direct insights into the key drivers of bike rental activity. It highlights the relationships between variables such as temperature, weather, and user type with rental counts. By showcasing the regression coefficients and their significance, this chart can quickly demonstrate how factors like temperature and weather affect bike usage. This is essential for executives making strategic decisions about resource allocation and pricing based on these variables.

Variable	Coefficient	Std. Error	t-Stat	P-value	Lower 95%	Upper 95%
Intercept	36.98	18.91	1.96	0.05	-0.10	74.06
Weekend	4.11	6.19	0.66	0.51	-8.03	16.26
Temp (F)	4.32	0.27	15.87	0.00	3.79	4.85
Mist	11.20	6.74	1.66	0.10	-2.02	24.42
Precipitation	-4.18	11.68	-0.36	0.72	-27.07	18.72
Spring	25.89	10.28	2.52	0.01	5.74	46.05
Summer	-8.49	13.15	-0.65	0.52	-34.28	17.30
Fall	78.15	8.83	8.85	0.00	60.84	95.46
Year 2012	76.88	5.67	13.56	0.00	65.77	88.00
Humidity	-2.92	0.17	-16.99	0.00	-3.26	-2.58
Windspeed	0.51	0.38	1.35	0.18	-0.23	1.25

2. Monthly Total Bike Rentals for 2011 and 2012: This chart provides a comprehensive view of bike rental trends over time. By comparing rentals across two years, it reveals seasonal patterns, growth or decline, and potential areas for improvement. Monthly data is particularly useful for planning marketing and operational strategies, such as adjusting bike availability for peak months or understanding how external factors (like weather or holidays) impact bike usage.



3. Average Hourly Bike Rentals on Weekdays by User Type: This chart offers granular insights into user behavior throughout the week. It helps executives understand peak times for casual versus registered users and tailor services accordingly, such as adjusting rental locations or improving fleet management during busy hours.