UNIVERSITÀ DEGLI STUDI DI MILANO

Functional Data Analysis of Bike Rental Patterns

Course: Functional and Topological Data Analysis

Student: Emil Soltanov

Matricola: 11229A

Professor: Alessandra Micheletti

September 2024

Table of Contents

1 Summary

This report uses **Functional Data Analysis (FDA)** to examine hourly rental patterns in an urban bike-sharing system, focusing on factors such as temperature, season, and time of day. The analysis highlights significant demand peaks during commuter hours, a positive correlation between temperature and bike rentals, and clear seasonal usage variations. Consistent demand patterns between weekdays and weekends support streamlined operational strategies. Key recommendations include adjusting bike availability during peak hours, scaling fleet capacity seasonally, and offering promotions to encourage winter usage. These data-driven strategies aim to enhance bike availability, improve service reliability, and boost operational efficiency year-round.

2 Introduction

Shared mobility is rapidly becoming a norm in cities worldwide, with bike-sharing services offering convenient, efficient, and eco-friendly transport options. These systems generate substantial data on daily rental patterns under varying environmental conditions. Analyzing this data can provide valuable insights for bike-sharing operators, enabling improvements in bike availability, reduced wait times, and enhanced user satisfaction.

2.1 Objective

This project investigates the temporal behavior of bike rentals using Functional Data Analysis (FDA), a statistical approach well-suited for continuous, high-frequency data like hourly rental counts. By treating rental data as continuous functions rather than discrete points, FDA enables a deeper understanding of usage trends and reveals hidden patterns that traditional time-series analysis might overlook.

2.2 Relevance of FDA in Time-Series Analysis

Standard time series methods capture basic trends and seasonality, but they often struggle with high-dimensional, complex datasets that contain intricate time dependencies. Functional Data Analysis (FDA) addresses this challenge by representing time-varying data as smooth, continuous functions, allowing for the detection of latent structures and subtle variations within long-term trends. This approach enables a richer exploration of bike rental patterns, uncovering deeper insights into the factors that drive demand.

2.3 Scope of Analysis

This project applies various Functional Data Analysis (FDA) techniques—including data smoothing, basis function representation, **Functional Principal Component Analysis (FPCA)**, **Functional Regression**, and **Functional Depth Measures**—to interpret hourly bike rental behavior. These techniques enable a detailed analysis of rental patterns across different time periods, such as weekdays versus weekends, and examine how environmental factors, like temperature, influence rental demand.

2.4 Expected Insights

This FDA-based analysis aims to uncover meaningful patterns in bike rental data, including peak rental times and relationships with external variables. Insights from this study are expected to support operational strategies, optimize resource allocation, and enhance the efficiency of bike-sharing services. Additionally, these findings may inform policy-making and urban planning initiatives that promote sustainable mobility solutions.

3 Methodology

This project uses a publicly available bike-sharing dataset that includes both hourly and daily records of bike rentals. Key variables include date, season, weather conditions, temperature, humidity, wind speed, and rental counts. The hourly dataset, with over 17,000 observations, provides detailed insights into daily rental patterns, while the daily dataset, with 731 entries, offers an overview of longer-term trends.

All analyses were conducted in R, utilizing libraries such as:

- `fda` and `fdapace` for functional data analysis and smoothing,
- `refund` for functional regression modeling,
- `depthTools` and `TDA` for depth measures and outlier detection,
- `ggplot2` and `tidyverse` for data manipulation and visualization.

To explore temporal rental trends and relationships with environmental factors, the following Functional Data Analysis (FDA) techniques were applied:

- **Data Smoothing and Functional Representation:**

Hourly rental data was smoothed using B-spline and Fourier basis functions, transforming discrete rental counts into continuous curves. This allowed daily rental patterns to be represented as smooth functions over a 24-hour period, enabling more detailed temporal trend analysis.

- **Basis Function Representation:**

Both *B-spline* and *Fourier basis functions* were tested to represent the hourly rental curves. *B-splines* offer flexibility and are well-suited for irregular shapes, while *Fourier basis functions* are optimal for capturing periodic patterns, such as the cyclic nature of daily rental trends.

- **Functional Principal Component Analysis (FPCA):**

FPCA was applied to the smoothed data to identify primary patterns and sources of variability. By analyzing the main components, key rental behaviors—such as peak times and shifts in demand due to weather or time of day—were interpreted. A scree plot was used to assess variance explained by each component.

- **Functional Regression Analysis:**

Functional regression was used to examine the relationship between temperature (independent variable) and hourly bike rentals (dependent variable). Using Fourier basis functions, the model estimated the impact of temperature changes on rentals at different times of day.

- **Functional ANOVA (FANOVA):**

FANOVA was applied to assess the effect of seasons (Winter, Spring, Summer, Fall) on bike rentals. By treating season as a factor, this method highlighted cyclical and seasonal demand shifts.

- **Depth Measures and Outlier Detection:**

Modified Band Depth (MBD) was calculated to evaluate the "centrality" of each rental curve, identifying typical rental patterns and potential outliers. Functional boxplots with MBD measures visualized central tendencies and outliers, while separate analyses for weekdays and weekends revealed differences in user behavior.

- **Statistical Testing (Wilcoxon Test):**

A Wilcoxon rank-sum test was conducted to compare depth measures between weekdays and weekends, assessing any significant differences in rental distributions across these two groups.

4 Data Loading and Preparation

4.1 Loading the Dataset

The hourly and daily bike-sharing data were loaded from CSV files for this analysis. Each dataset includes records of bike rentals along with contextual information such as date, time, and weather conditions. The hourly dataset captures detailed rental activity throughout the day, while the daily dataset provides a broader overview with aggregated daily rental counts and other features. Together, these datasets form the foundation for analyzing temporal rental patterns.

4.2 Initial Data Exploration

To gain an understanding of the data structure, summary statistics and structural information were examined for both datasets:

- **Summary statistics** provided insights into the range, central tendency, and variability of each variable, including counts of casual and registered riders.

- **Structure checks** revealed the data types of each variable, which is essential for subsequent preprocessing steps.

This preliminary exploration helped ensure data consistency, identify any missing values, and assess the potential role of each feature in the analysis. No missing values were found in either dataset.

4.3 Data Preprocessing

1. **Categorical Variable Transformation**

Variables such as season, year, month, hour, weekday, and weather conditions were converted to categorical factors with meaningful labels. These transformations make categorical variables easier to interpret in visualizations and statistical models.

2. **Feature Scaling for Numeric Variables**

Numeric variables such as **temperature** (`temp`), **humidity** (`hum`), and **windspeed** (`windspeed`) are scaled to standardize them. This normalization is crucial when performing Functional Data Analysis (FDA) to ensure that differences in magnitude among features do not disproportionately influence the results.

3. **Handling Date Variables**

Date fields were converted to Date format for easier manipulation. Additional features, such as day of the year and week of the year, were extracted from the daily data to support seasonal analysis.

This systematic data preparation ensures that datasets are ready for Functional Data Analysis, with labeled categorical variables, standardized numeric features, and organized temporal information for robust analysis.

5 Functional Data Analysis (FDA)

This section analyzes the hourly rental data from the bike-sharing system using Functional Data Analysis (FDA). FDA treats rental data as continuous functions over time, allowing for detailed insights into temporal patterns, principal components, and the impact of external factors like temperature on bike rentals.

5.1 Smoothing of Hourly Bike Rentals

To start, a functional data object representing bike rentals across 24-hour periods was created. Smoothing was applied using B-splines to reduce noise in the hourly data and reveal general rental patterns throughout the day. The smoothed curves exhibit a clear diurnal pattern, with noticeable peaks during morning and evening hours, aligning with commuter behavior as people typically rent bikes for travel to and from work.

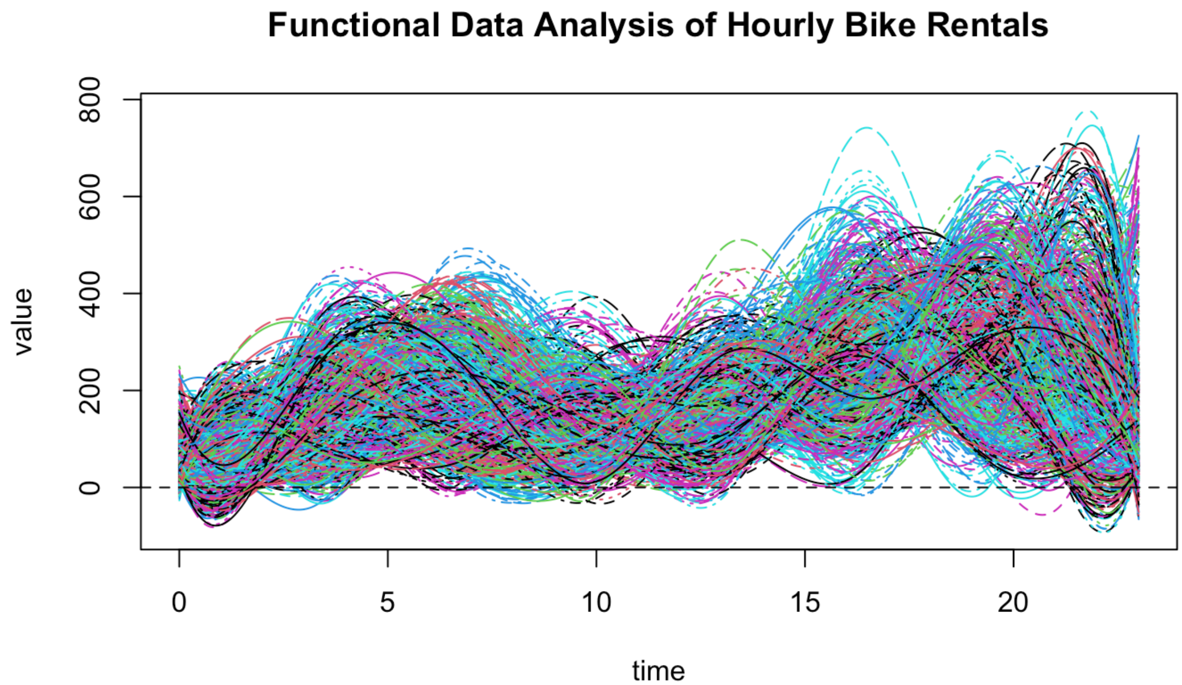

Figure 1: Functional Data Analysis of Hourly Bike Rentals

5.2 Smoothing of Hourly Temperature Data

Temperature was similarly treated as a functional variable, as it varies continuously over a 24-hour period. The smoothed temperature curves demonstrate a cyclical pattern, with higher values in the afternoon and lower values in the early morning and late evening. This temperature pattern is essential for understanding its relationship with bike rentals.

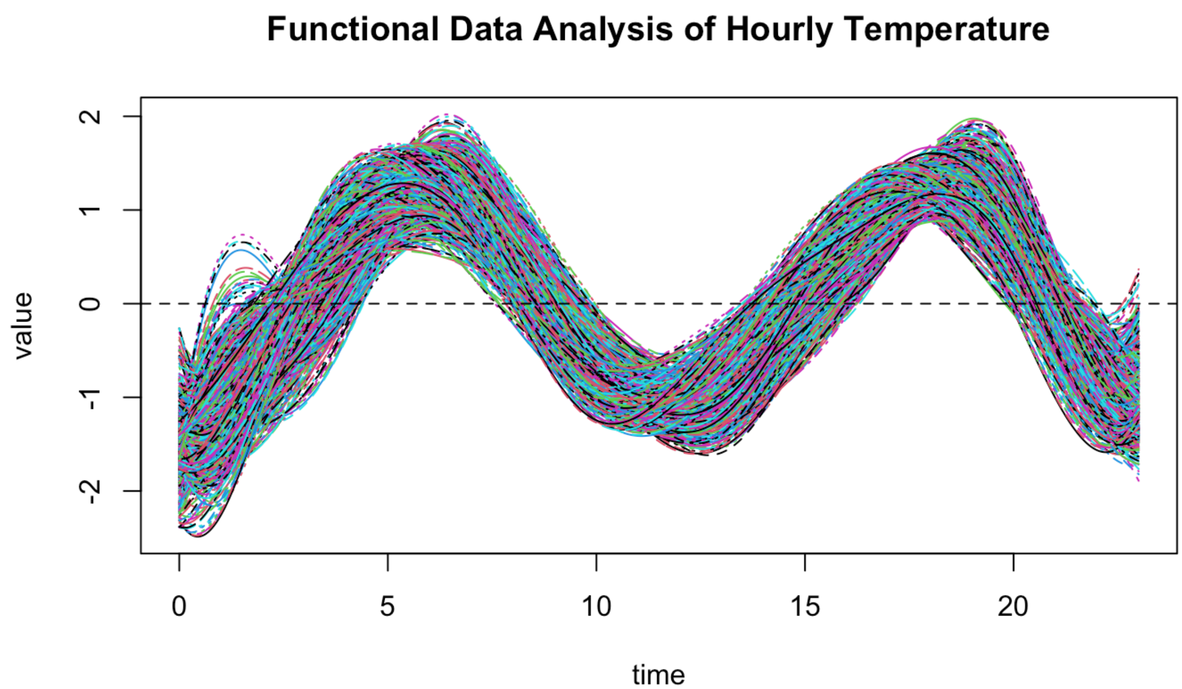

Figure 2: Functional Data Analysis of Hourly Temperature

5.3 Basis Function Representation

To further refine the functional data, a Fourier basis representation was employed for bike rentals over the 24-hour cycle. Fourier basis functions effectively capture periodic behavior, allowing a succinct representation of repeating daily rental patterns. This approach smooths out the hourly peaks and valleys in rental activity, providing a more continuous representation compared to observed curves.

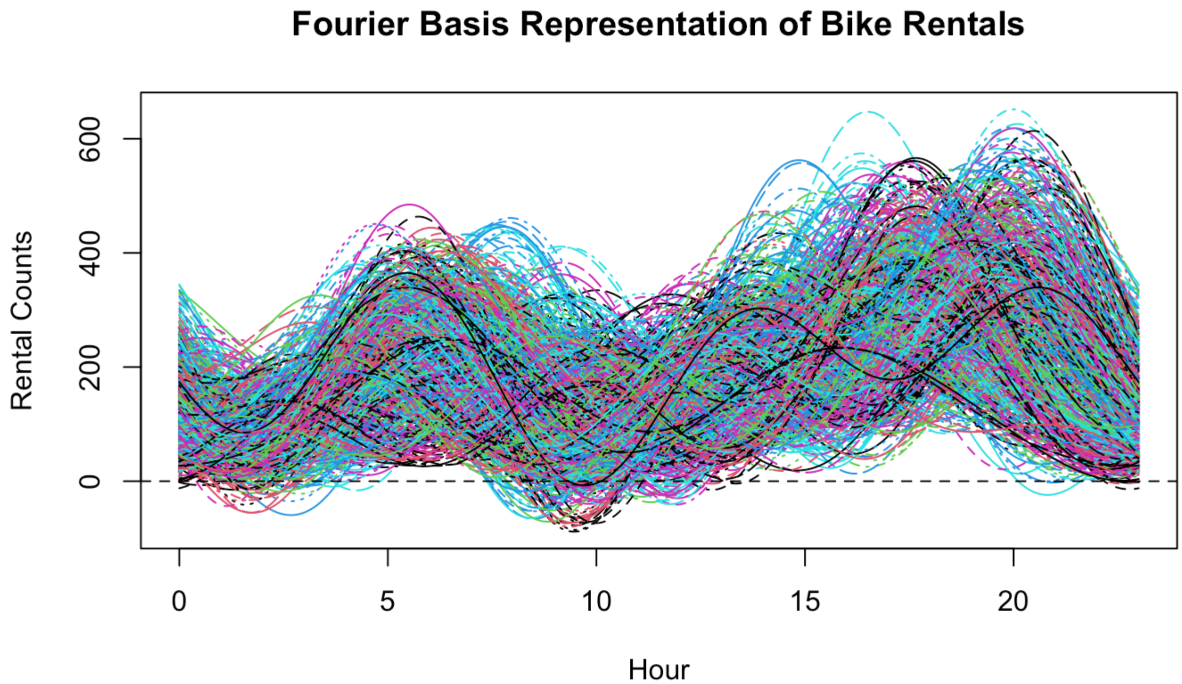

Figure 3: Fourier Basis Representation of Bike Rentals

5.4 Functional Principal Component Analysis (FPCA)

Functional Principal Component Analysis (FPCA) was applied to identify the primary components driving variation in bike rental patterns. The first few principal components (PCs) capture most of the variance, with the first component accounting for approximately 40%. This main component reflects the overall diurnal pattern, showing the rise and fall in rentals during morning and evening peaks. Subsequent components capture finer details, including fluctuations specific to certain times of day.

$$X_i(t) = \mu(t) + \sum_{k=1}^{K} \xi_{i,k} \varphi_k(t)$$

Explanation: This FPCA formula expresses the functional data $X_i(t)$ (e.g., bike rentals) for day i over time t as a sum of:

$\mu(t)$: the mean function across all days,

$\varphi_k(t)$: the principal component functions, representing main sources of variation in the data, and

$\xi_{i,k}$: the component scores for each day i, indicating how much each component k contributes to that day's pattern.

By decomposing $X_i(t)$ into these components, FPCA captures the main patterns of variability in the data.

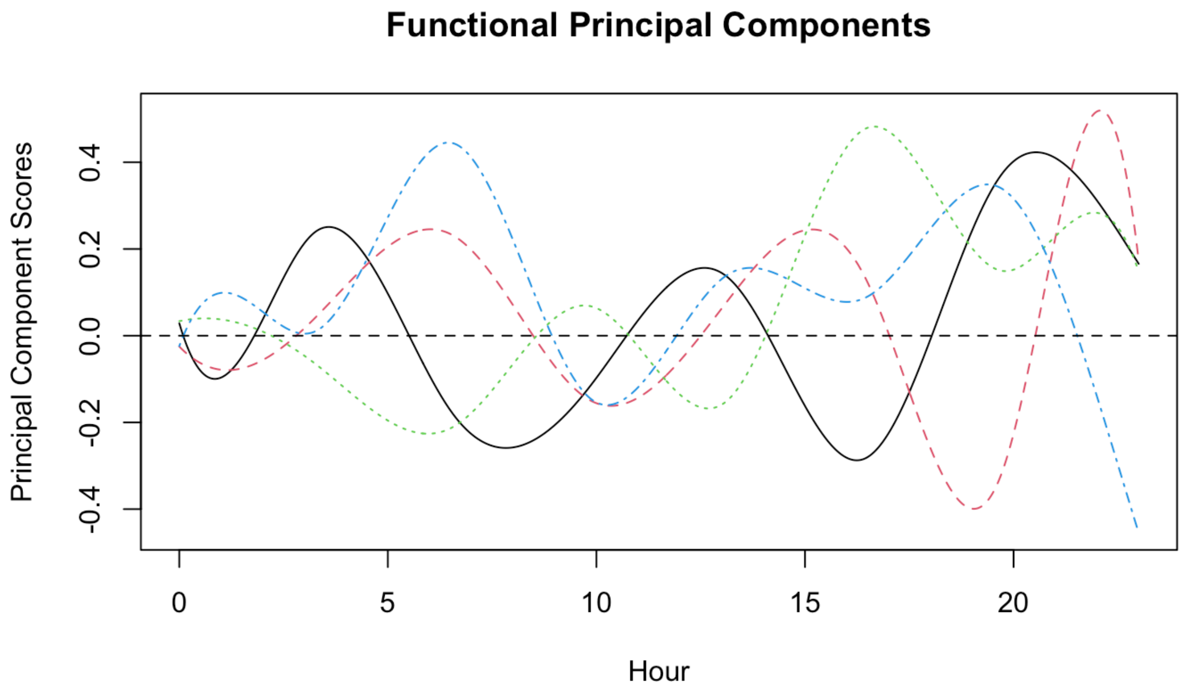

Figure 4: Functional Principal Components

To evaluate the cumulative variance explained by each component, a scree plot was generated. This plot shows the diminishing contributions of successive components, confirming that the first few components capture the majority of the rental data's variance. With the first three components accounting for over 80% of the variance, they were selected for further analysis. Additionally, the cumulative variance explained plot demonstrates a rapid increase within the first few components, justifying the choice of a limited number of components for effective data representation.

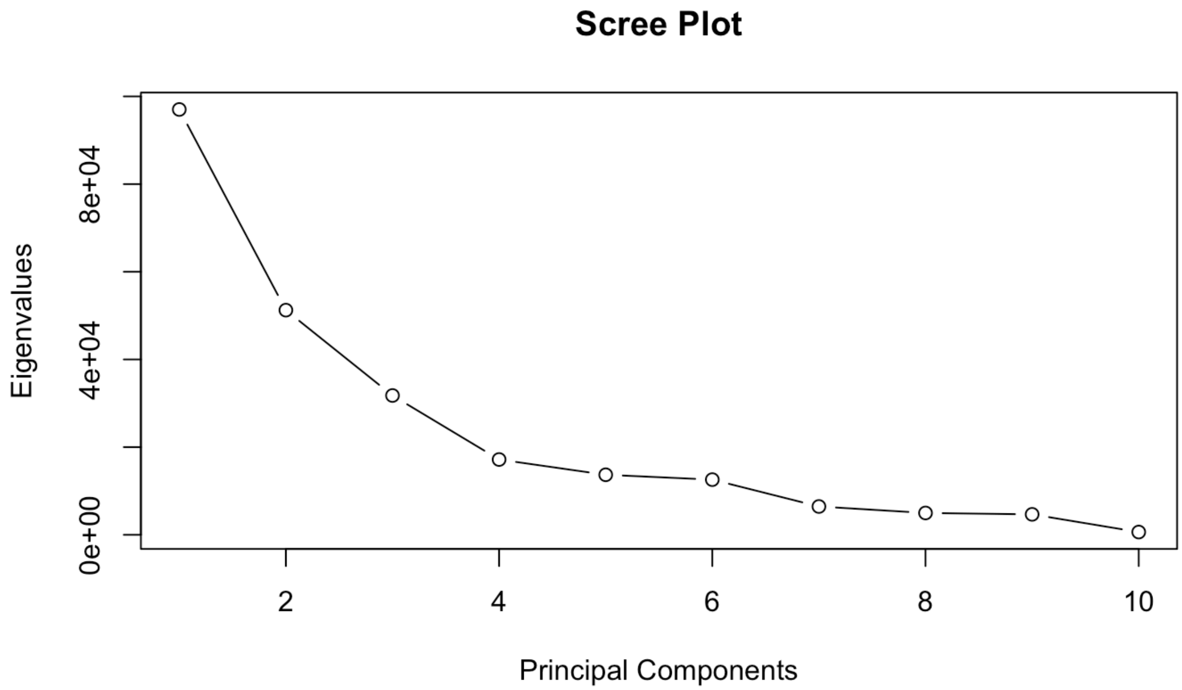

Figure 5: Scree Plot

Additionally, the cumulative variance explained plot shows the rapid increase in explained variance within the first few components, supporting the selection of a limited number of principal components to represent the data effectively.

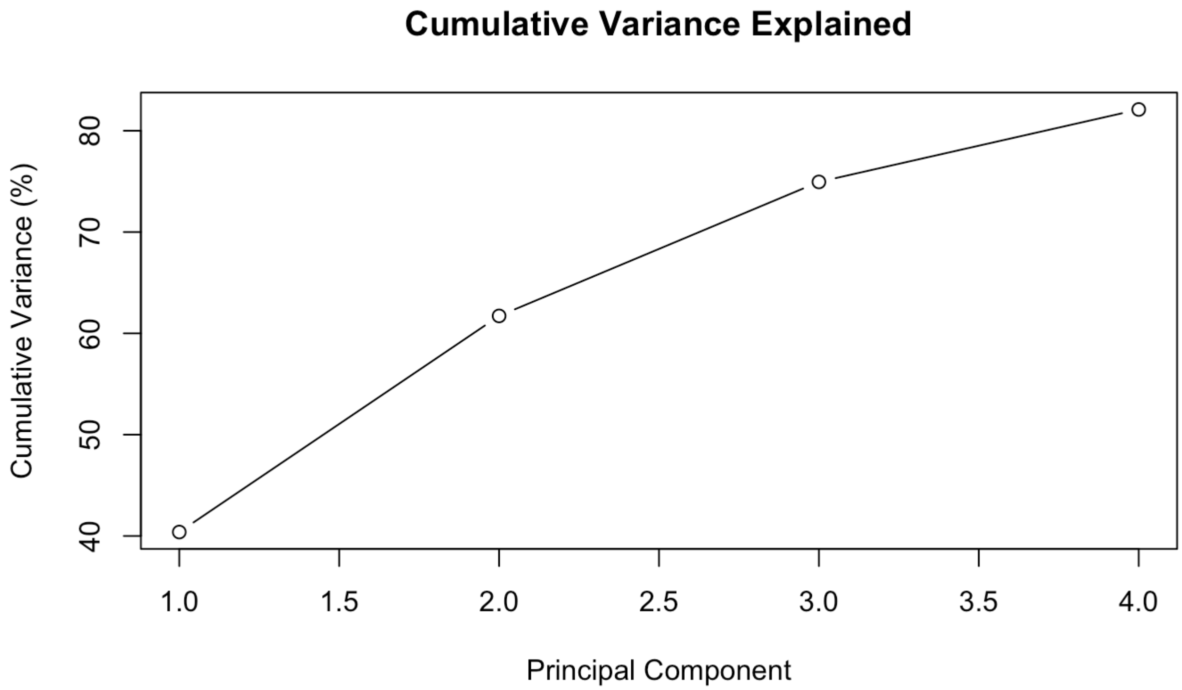

Figure 6: Cumulative Variance

Finally, the plot of the first principal component highlights its role in capturing the main rental cycle, with peaks during the day and lower values in late-night hours, aligning with commuter and leisure bike rental patterns.

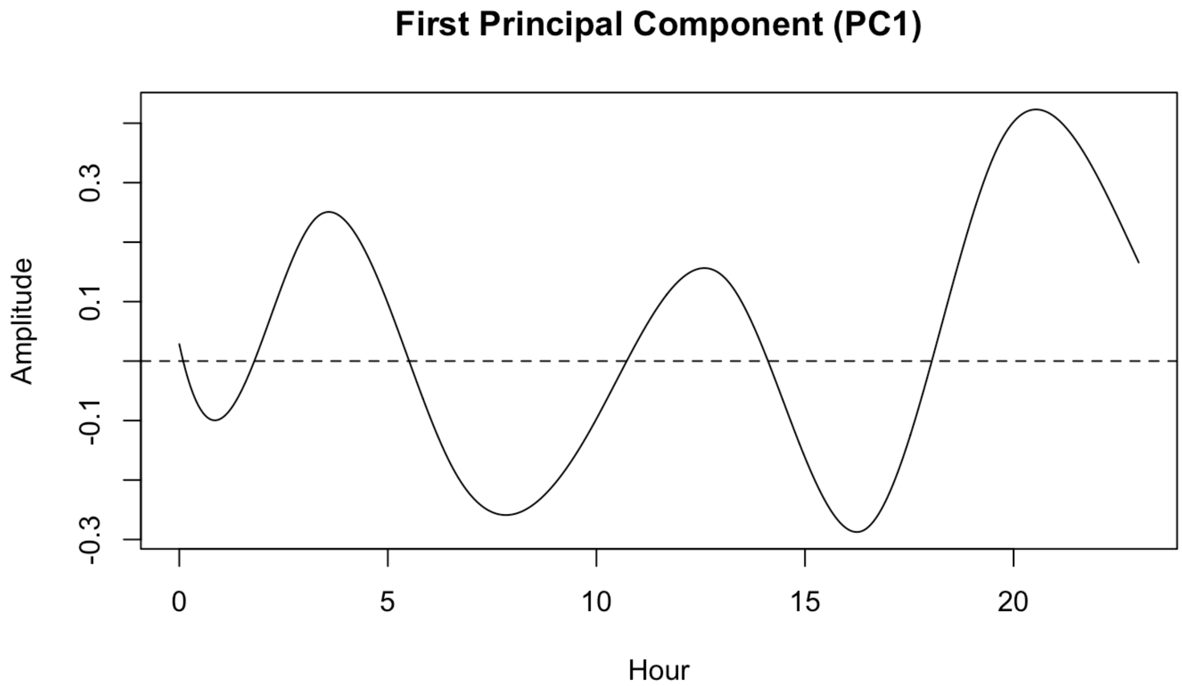

Figure 7: First Principal Component

5.5 Data Registration

To standardize the data further, data registration was performed by aligning individual rental curves with the mean rental curve. This adjustment corrects for minor temporal shifts in rental peaks, making the curves more comparable. Post-registration, the aligned curves display clearer patterns, particularly around typical peak hours, confirming the consistency of these peaks across days.

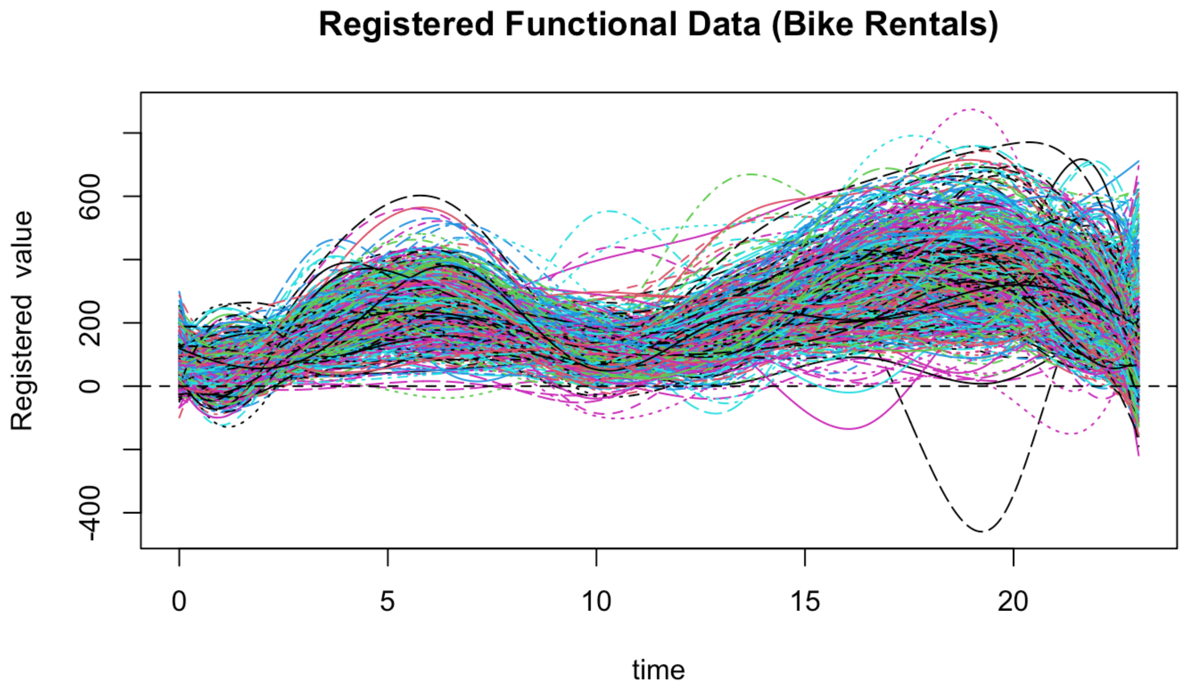

Figure 8: Registered Functional Data

For comparison, the original, unregistered curves are also displayed, showing the slight misalignment that data registration effectively corrected.

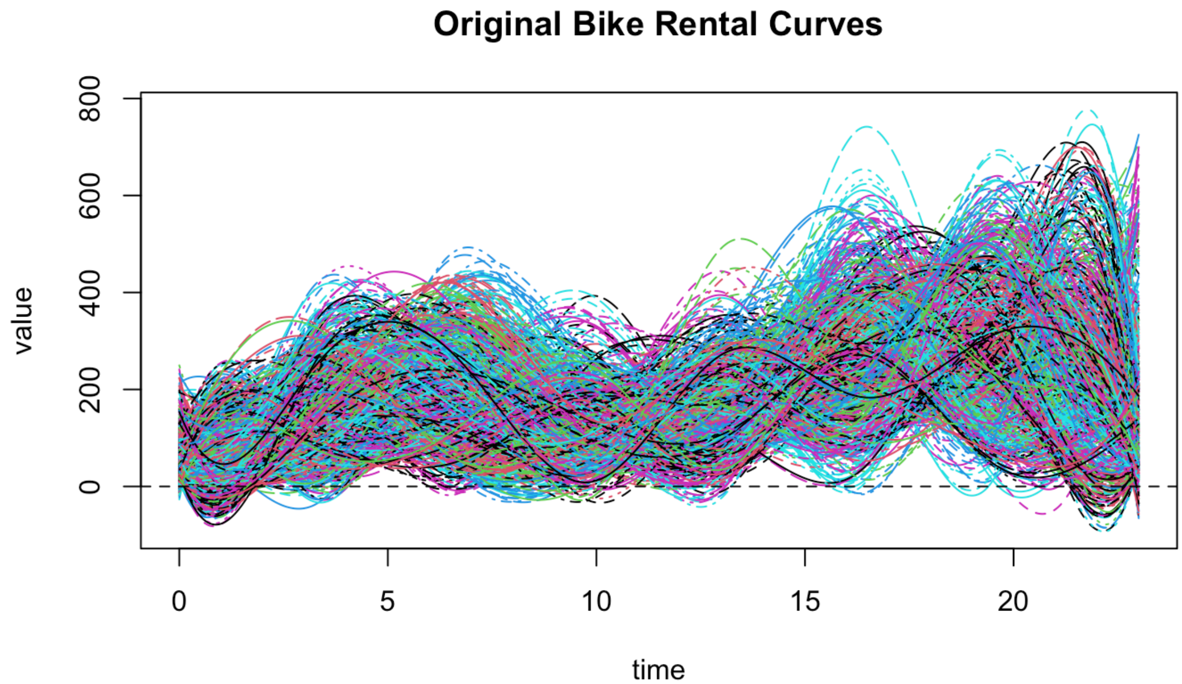

Figure 9: Original Bike Rental Curves

6 Functional Regression Analysis

This study uses Functional Regression Analysis to explore the relationship between hourly bike rentals and temperature, with rental counts as the response variable and temperature as the predictor. Functional regression allows for a continuous assessment of how temperature influences rental patterns throughout the day, capturing both the timing and magnitude of this effect.

6.1 Regression Model Setup and Estimated Coefficients

A functional regression model was created with temperature as the predictor and hourly bike rentals as the response. The model used B-spline basis functions to represent smooth temperature and rental curves over a 24-hour period. This smoothing process captures natural variations while minimizing noise to highlight significant underlying patterns.

The functional regression model showed a positive association between temperature and bike rentals, particularly during midday and afternoon hours, indicating that warmer temperatures lead to increased rental activity, especially during peak hours. The estimated regression coefficient for temperature, visualized over a 24-hour period, supports this trend, with higher coefficients during peak times reflecting a stronger temperature influence on rentals.

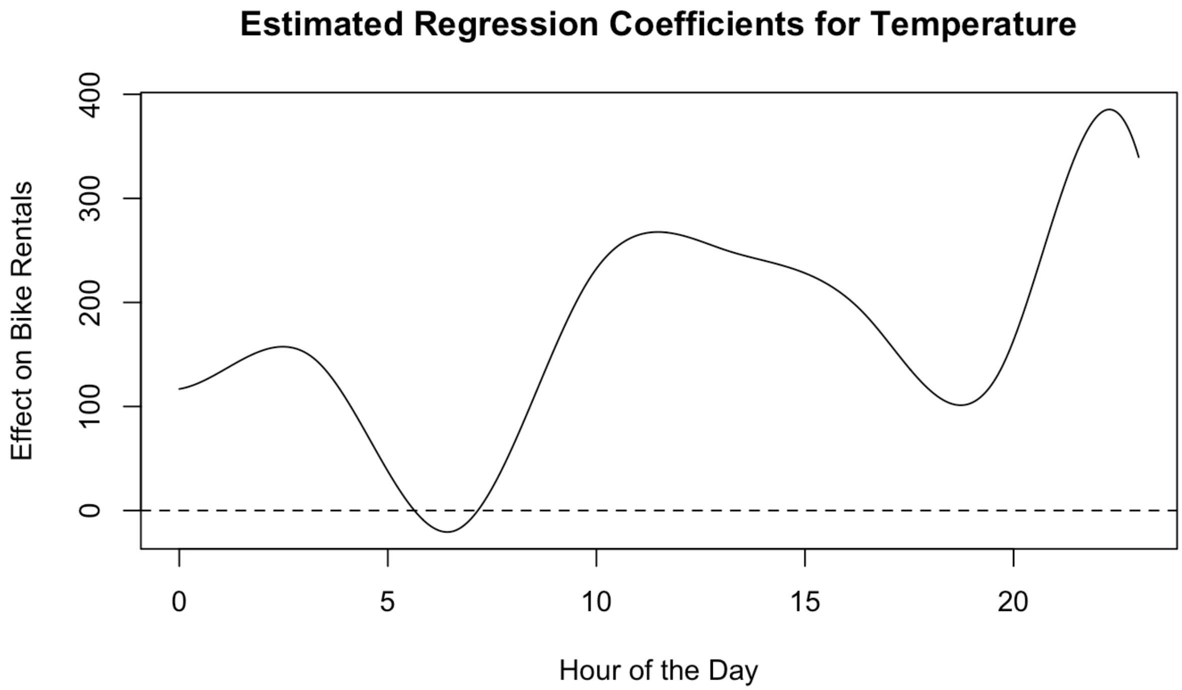

Figure 10: Estimated Regression Coefficients for Temperature

6.2 Model Evaluation: R-squared and F-statistic Analysis

The performance of the model was evaluated using two key metrics: the *R-squared* value and the F-ratio.

The *R-squared* value of approximately 0.2706 indicates that about 27% of the variability in hourly bike rentals can be explained by changes in temperature. Although this value suggests that temperature significantly influences bike rentals, it also indicates that other factors, such as day of the week, holiday status, and weather conditions, likely contribute to the demand patterns observed. The moderate *R-squared* value highlights temperature as a critical, but not exclusive, determinant of bike rental behavior.

The model's *F-ratio*, calculated to be 1288.6, demonstrates that the model fit is statistically significant. A high *F-ratio* reinforces that temperature is a meaningful predictor of hourly rental counts, providing evidence that temperature should be included in any model seeking to explain bike rental demand variations.

The positive relationship between temperature and bike rentals suggests that warmer periods encourage higher bike demand, particularly in the midday and afternoon. This observation aligns with typical outdoor activity patterns, where users prefer cycling during warmer times, particularly during peak hours.

This Functional Regression Analysis confirms that temperature significantly influences bike rental demand patterns. The midday peak effect of temperature on bike rentals suggests potential scheduling adjustments and resource allocation strategies, especially on warmer days.

6.3 Functional Regression Analysis with Humidity and Windspeed

In this section, the functional regression model is extended by incorporating humidity and windspeed as additional predictors alongside temperature. These environmental factors are known to influence outdoor activities and, consequently, bike rental patterns. By examining all three variables within a single model, a more comprehensive understanding is achieved regarding how weather conditions impact rental demand throughout the day.

To model the diurnal patterns of humidity and windspeed, functional representations for each variable were created over a 24-hour period, following the same approach applied to temperature. Each variable was smoothed using a B-spline basis with 10 basis functions and a moderate smoothing penalty (λ = 1e-2). This process captured general trends and minimized noise, ensuring that the functional data objects for humidity, windspeed, and temperature are consistent in form and suitable for comparison within a unified regression model.

Model Specification

The extended functional regression model is specified as follows:

$$\begin{aligned} Bike\ Rentals\ &= \beta\text{^}0 + \beta\_temp(t) * temp(t) + \beta\_hum(t) * hum(t) \\ &+ \beta\_windspeed(t) * windspeed(t) + \varepsilon(t) \end{aligned}$$

where:

- $temp(t), hum(t), and\ windspeed(t)$ are the hourly functional data representations of temperature, humidity, and windspeed, respectively.
- $\beta\_temp(t)\ , \beta\_hum(t)\ , and\ \beta\_windspeed(t)$ are the functional coefficients estimated for each predictor, describing their respective influence on bike rentals over the course of a day.
- $\varepsilon(t)$ is the residual error term.

The model estimated separate functional coefficients for **temperature**, **humidity**, and **windspeed**, revealing how the influence of each variable on bike rentals changes throughout the day. Key insights from these coefficients are summarized below:

- **Temperature**: The functional coefficient for temperature demonstrates a positive relationship with bike rentals, especially during peak hours. This finding suggests that as temperature increases, bike rental counts tend to rise, particularly in the late morning and early afternoon. This pattern aligns with expectations, as warmer temperatures generally encourage outdoor activities.

- **Humidity**: The functional coefficient for humidity displays a variable relationship with bike rentals throughout the day. Higher humidity levels tend to have a slight negative effect on rentals, especially during midday, likely reflecting the discomfort associated with high humidity levels. The dampening effect of humidity on bike rentals may be more pronounced during hours when people are typically outdoors, such as late morning and early afternoon.
- **Windspeed**: The functional coefficient for windspeed suggests a complex relationship with bike rentals. At certain times of day, such as early morning and late evening, higher windspeed appears to discourage rentals, possibly due to the discomfort or difficulty of cycling in windy conditions. However, the effect of windspeed fluctuates throughout the day, indicating that other factors might interact with windspeed to influence its impact on rental demand.

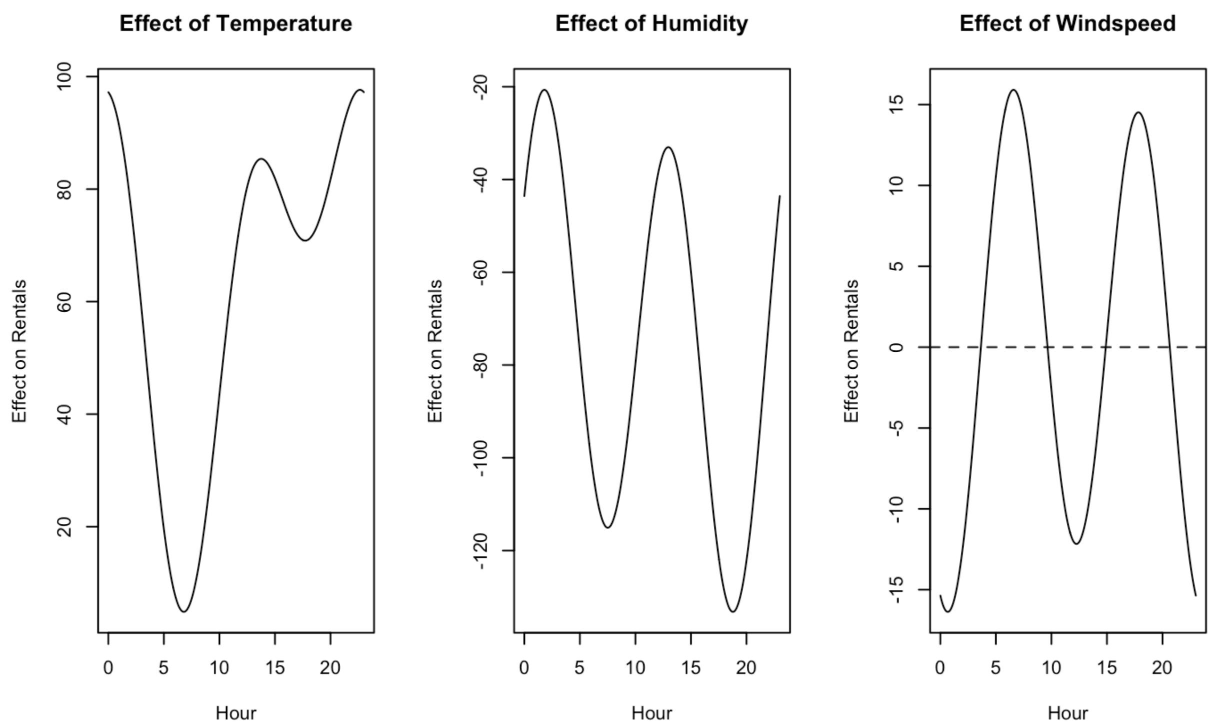

Figure 11: Functional Regression Analysis with Humidity and Windspeed

The inclusion of humidity and windspeed in the functional regression model provides a richer understanding of how environmental conditions shape bike rental behavior. While temperature remains a strong positive predictor, humidity and windspeed contribute additional nuance, particularly during peak outdoor activity hours. These findings suggest that bike rental services may benefit from adjusting availability or offering weather-based incentives during periods of high humidity or windspeed.

7 Functional ANOVA (FANOVA)

To capture seasonal effects on bike rental patterns, Functional ANOVA (FANOVA) was performed to compare variations in bike rentals across different seasons. FANOVA is well-suited for this analysis as it allows continuous, smooth rental functions to be examined, identifying significant differences in bike rentals across seasons.

$Y_i(t) = \mu(t) + \alpha_j(t) + \varepsilon_i(t)$

Explanation: The FANOVA model explains bike rentals $Y_i(t)$ for day i within season j as:

- $\mu(t)$: the overall mean rental pattern over time t,
- $\alpha_j(t)$: the effect of season j on rentals over the day, and
- $\varepsilon_i(t)$: the residual error for day i.

This FANOVA setup helps in identifying how seasonal factors influence rental patterns throughout the day.

7.1 Seasonal Basis Functions and Model Setup

Seasonal basis functions were constructed using a Fourier basis representation to capture seasonal variations smoothly over a 24-hour period. This approach enabled the construction of a design matrix representing the four seasons: Winter, Spring, Summer, and Fall. Each season's unique pattern was analyzed using FANOVA to evaluate differences across these seasonal groups.

The FANOVA model reveals clear variations in bike rental behaviors across seasons, as shown in the seasonal effect curves. These curves demonstrate distinct rental patterns in Summer and Fall, where demand is notably higher compared to Winter and Spring, likely due to more favorable weather conditions.

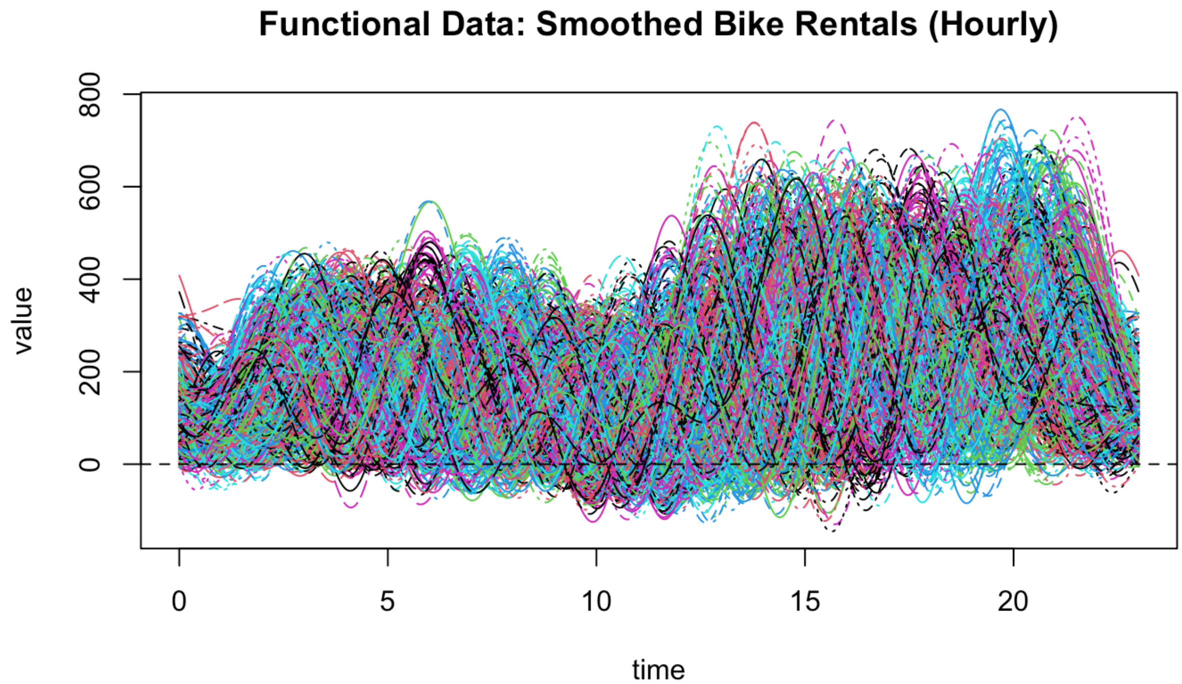

Figure 12: Functional Data: Smoothed Hourly Bike Rentals

7.2 Seasonal Effects and Interpretation

The estimated functional coefficients from the FANOVA model reveal the specific effect each season has on bike rental patterns over the 24-hour cycle:

Summer demonstrates a pronounced effect on bike rentals, especially during midday and evening hours. The curve for Summer has the highest rental counts, suggesting that favorable weather during this season results in higher bike usage.

Like Summer, **Fall** also shows high rental counts, although they are generally lower than Summer. This season demonstrates steady usage throughout the day, with peaks during typical commuting hours.

In **Spring**, rental counts are moderate, showing a consistent but lower influence across the day. This pattern suggests that, while bike usage is higher than in Winter, it does not reach Summer and Fall levels, possibly due to variable weather.

Winter exhibits the lowest rental counts across the day. The seasonal effect curve shows reduced demand during colder months, with lower usage during all hours, likely reflecting the impact of colder temperatures on outdoor cycling.

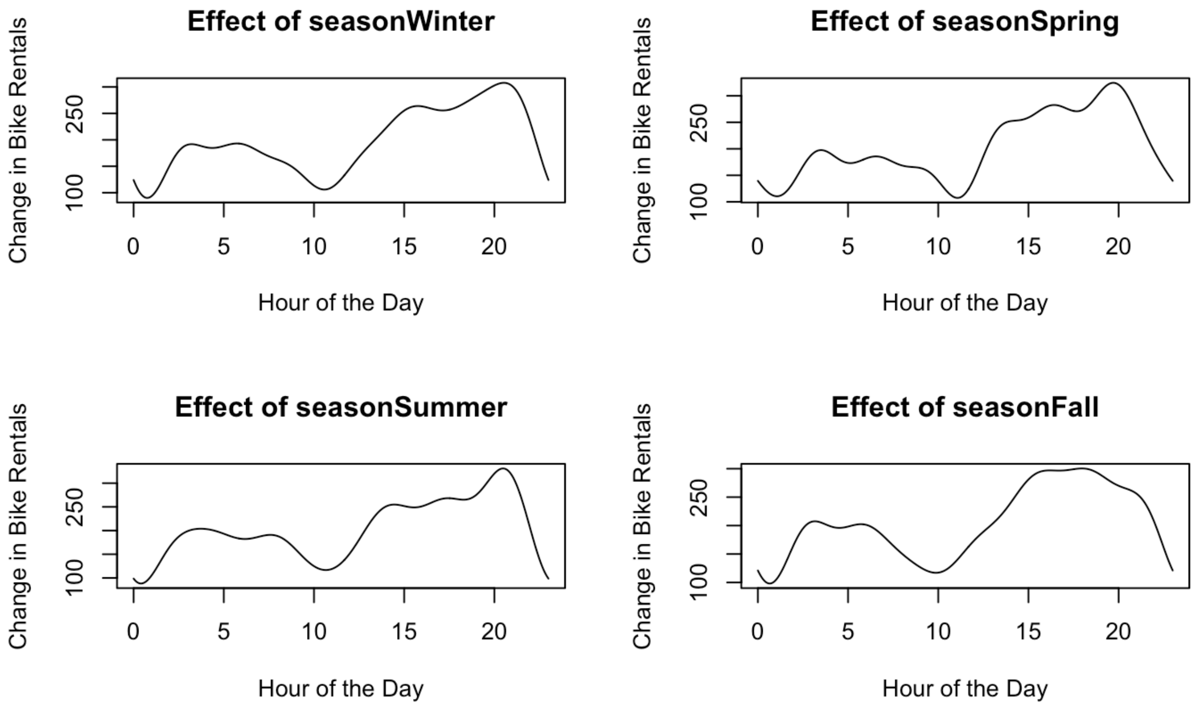

Figure 13: Seasonal Effects

7.3 Interpretation of FANOVA Results

The FANOVA results illustrate that seasonal factors significantly impact bike rental patterns, with Summer and Fall seeing the highest usage, particularly during midday and evening hours. In contrast, Winter shows consistently lower rental activity throughout the day.

These seasonal patterns provide actionable insights for bike-sharing systems, indicating periods of high demand during warmer months and potential reductions in usage during winter. This information can help bike-sharing companies adjust bike availability and maintenance schedules, as well as inform strategies for increasing usage during lower-demand seasons.

8 Depth Measures and Outlier Detection

This section explores the analysis of depth measures, focusing on the **Modified Band Depth (MBD)** to assess daily rental patterns and identify outliers. Depth measures such as MBD quantify the centrality of each day's rental curve within the dataset, distinguishing typical rental days from those with unusual patterns. This analysis helps identify demand consistency and variability, providing valuable insights for operational planning.

8.1 Functional Depth Measures with MBD

MBD was used to compute depth values for each day's rental pattern, evaluating how closely each day aligns with the central trend. Higher depth values indicate days with rental patterns that closely follow the central trend in the dataset, while lower values suggest deviations. The functional boxplot generated using MBD illustrates the central 50% of daily rental curves, with the median curve representing the most typical rental day.

$$MBD(X_{\mathrm{i}}) = \left(\frac{1}{C(n,2)}\right) \sum_{1 \leq j < k \leq n} I\left(X_{j(t)} \leq X_{\mathrm{i}}(t) \leq X_{k}(t)\right)$$

Explanation: Modified Band Depth (MBD) is a measure of how "central" or typical a function $X_i(t)$ is within a set of daily rental functions. Here:

- $C(n, 2)$ denotes the number of unique pairs of functions, n being the total number of days.
- $I\left(X_{j(t)} \leq X_{\mathrm{i}}(t) \leq X_{k}(t)\right)$ is an indicator function counting how often $X_i(t)$ lies between any pair of other curves $X_{j(t)}$ and $X_k(t)$ over time.

The functional boxplot shows that most rental days adhere closely to this median pattern, with clear peaks during commuting hours—morning and evening—consistent with typical daily travel patterns in urban bike-sharing systems. Variation around the median reflects natural fluctuations in demand across different hours of the day.

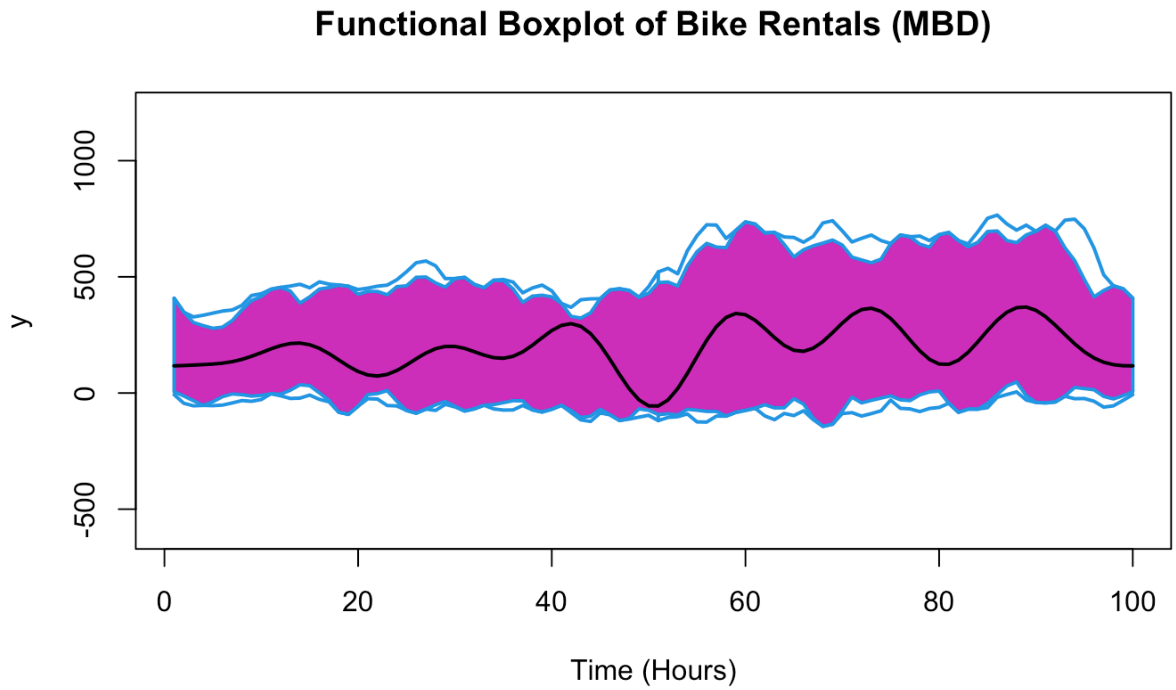

Figure 14: Functional Boxplot of Bike Rentals (MBD)

8.2 Outlier Detection Using Depth Measures

Outliers, which are days with the lowest 5% depth values, exhibit unusual rental patterns and can indicate the need for special operational strategies. Days with high rental counts may require an increased number of available bikes to meet demand, while low-demand days might suggest opportunities for maintenance or targeted promotions to encourage usage.

These outlier insights are critical for effective resource allocation, as they help bike-sharing operators prepare for extreme usage scenarios and minimize potential service disruptions during high-demand periods.

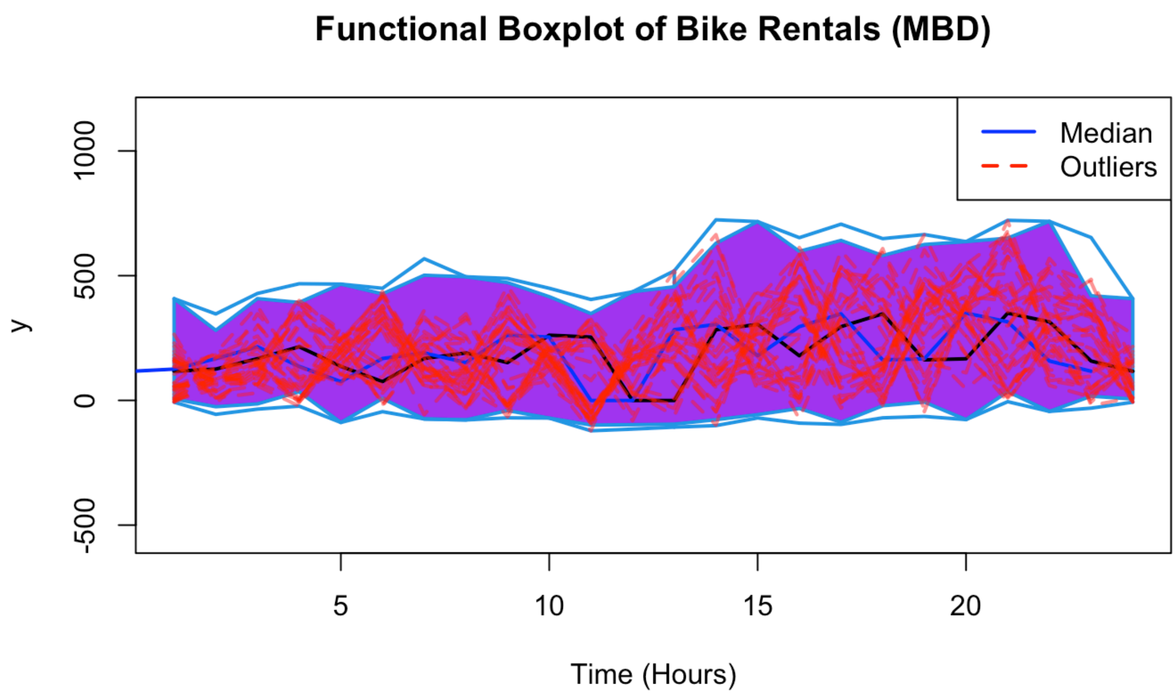

Figure 15: Outlier Detection Using Depth Measures

8.3 Comparison of Weekday and Weekend Depth Measures

To assess whether rental behaviors differ between weekdays and weekends, a Wilcoxon rank-sum test was conducted on the depth values for each group. This non-parametric test evaluates whether a statistically significant difference exists in the central tendency of weekday versus weekend depth values.

The test returned a *p-value* of 0.5945, indicating no significant difference between weekday and weekend depth values. This finding suggests that bike-sharing demand patterns are generally consistent throughout the week. Despite potential intra-day differences driven by commuting hours, the central rental structure remains stable across the entire week. This insight is valuable for resource allocation, as it implies that similar bike availability strategies can be employed on both weekdays and weekends without major adjustments, streamlining operational logistics.

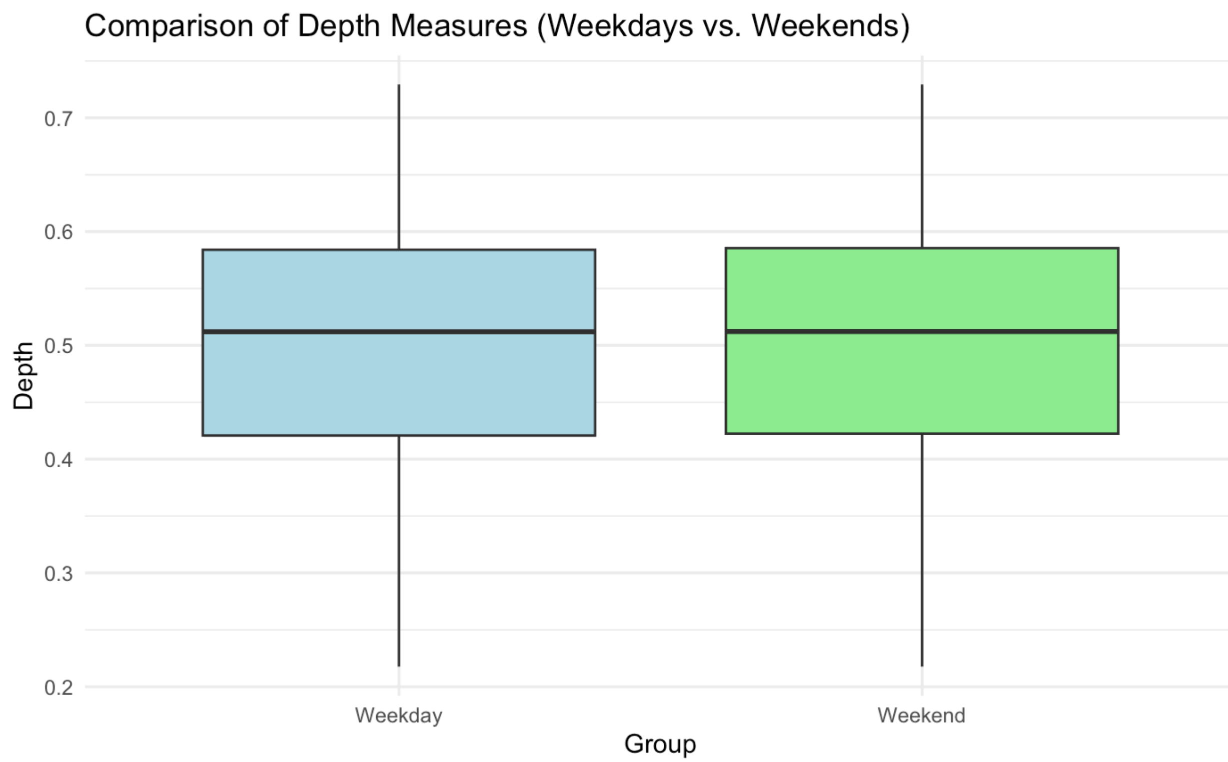

Figure 16: Comparison of Depth Measures (Weekdays vs. Weekends)

9 Additional Functional Analysis for Casual and Registered Users

This section extends the functional data analysis to examine patterns among two distinct user types in the bike-sharing dataset: casual users and registered users. Analyzing the differing rental patterns between these user types provides insights into user behavior and can inform service adjustments to better meet the needs of each group.

Functional Data Construction for Casual and Registered Rentals

Hourly bike rental counts for casual and registered users were extracted and organized into matrices where each row corresponds to the 24-hour rental activity for a single day. Smoothing functions were then applied to these matrices to create functional data objects for each user type, denoted as $〚\boldsymbol{fd}〛\_casual\ \ and\ \ 〚\boldsymbol{fd}〛\_registered.$

The smoothed data for both user types reveal contrasting patterns. Casual users generally show more variable rental patterns across days, with increased activity outside typical commuting hours, suggesting that this group primarily engages in recreational cycling. In contrast, registered users exhibit regular peaks aligned with commuting hours, indicating structured, work-related usage.

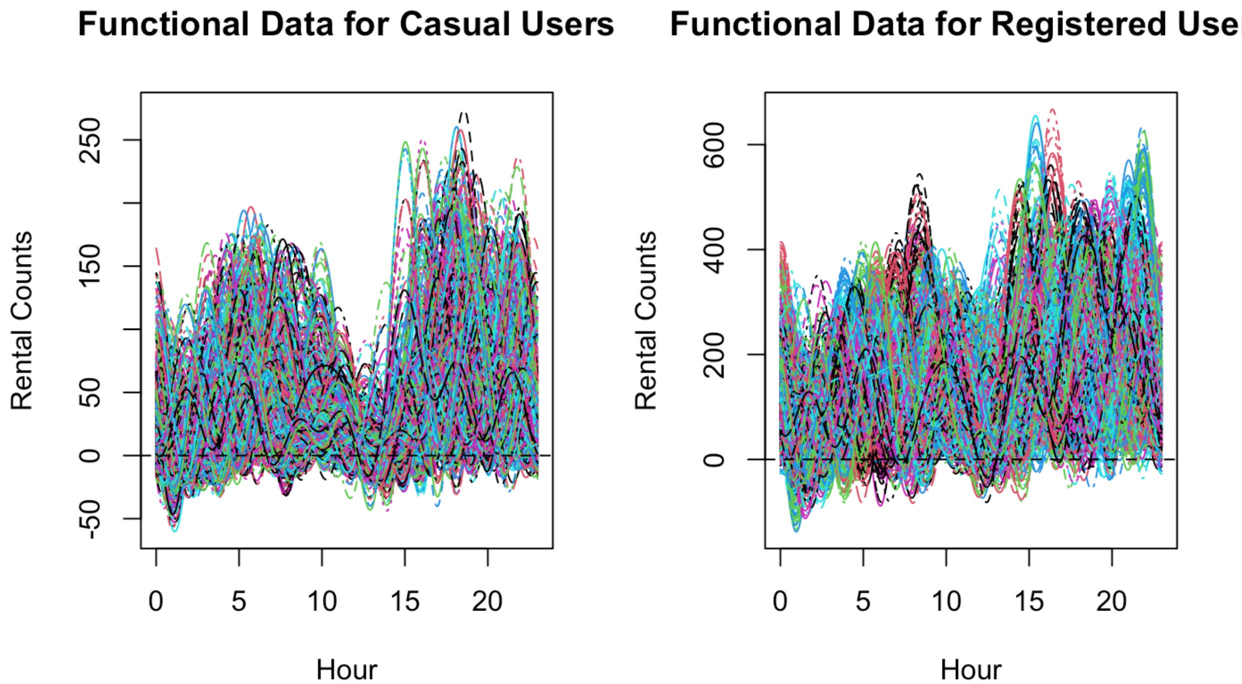

Figure 17: Functional Analysis of Casual vs Registered Users

To further examine these differences, **Functional Principal Component Analysis (FPCA)** was applied to each group. For casual users, the first principal component (PC1) emphasizes peak rental hours that typically occur during weekends and outside standard commuting times. By contrast, for registered users, PC1 aligns with peak weekday commuting periods. This distinction reflects the flexible nature of casual usage versus the structured patterns of registered users.

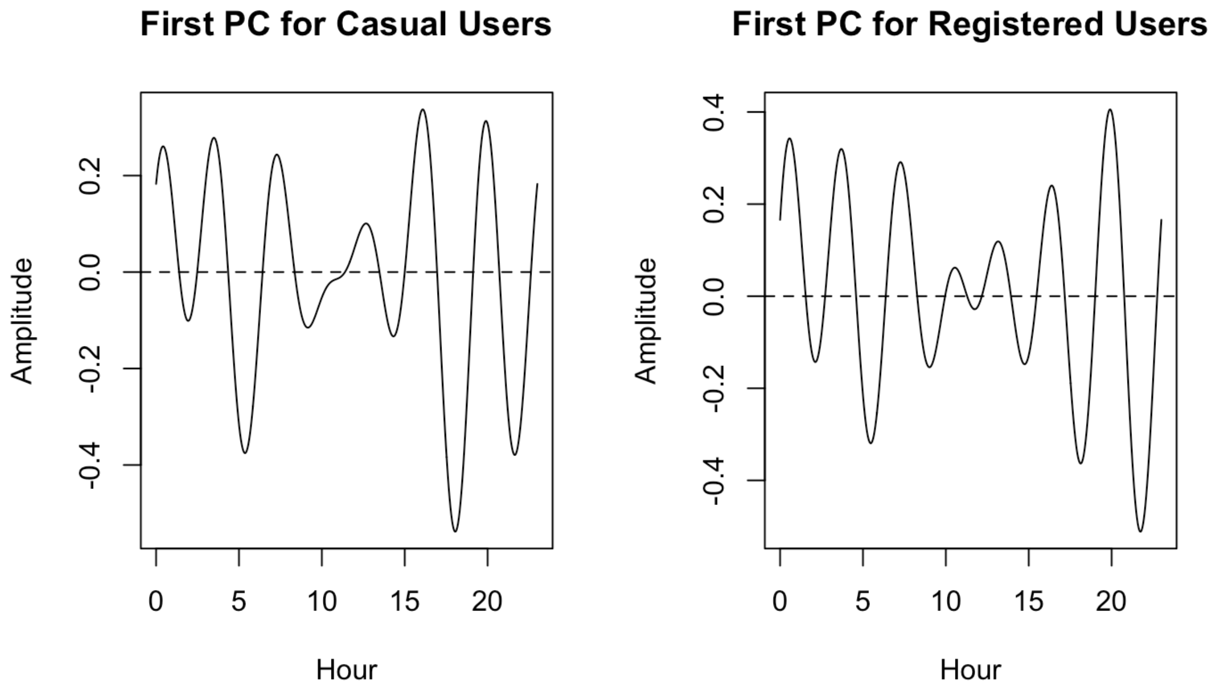

Figure 18: Functional Principal Component Analysis of Casual vs Registered Users

In summary, casual users exhibit flexible rental patterns associated with recreational usage, while registered users display consistent peaks during commuting hours. These insights highlight the value of adapting bike-sharing resources to accommodate the unique behaviors of each user type, potentially optimizing bike availability during distinct peak periods for casual and registered riders.

10 Conclusion

This project utilized Functional Data Analysis (FDA) to examine bike-sharing rental data, uncovering key demand patterns over time and under varying conditions. By applying techniques such as smoothing and basis functions, discrete hourly rental counts were converted into smooth, continuous curves, which allowed for clearer visualization of demand peaks and troughs. Through Functional Principal Component Analysis (FPCA), primary structures within rental data were captured, while Functional Regression and Functional ANOVA (FANOVA) assessed the impacts of temperature and seasonality on rentals. Depth measures and statistical tests identified stable demand patterns across weekdays and weekends, supporting reliable operational strategies.

The main insights from this analysis include:

- **Hourly Demand Peaks**: Clear peaks in demand during morning and evening hours indicate high rental activity, corresponding with commuting times.
- **Temperature Influence**: Warmer temperatures have a positive effect on bike rentals, indicating higher demand during mild and warm periods.
- **Seasonal Variation**: Rental demand shows strong seasonal patterns, with increased rentals in warmer months and lower counts in winter.
- **Weekday Stability**: Depth measures suggest consistent demand patterns between weekdays and weekends, supporting similar management strategies for both periods.

These findings equip bike-sharing services with actionable insights into user demand, enabling them to optimize bike availability and resource allocation effectively.

Based on these findings, the following recommendations are proposed to enhance operational efficiency and meet user demand:

- **Increase Bike Availability During Peak Hours**: To accommodate high demand during weekday commute hours, allocate additional bikes in the morning and evening. Slight reductions during off-peak hours may optimize overall utilization.
- **Seasonal Adjustments to Fleet Size**: Given increased rental activity during warmer months, consider expanding fleet capacity, particularly in summer. This adjustment can help meet demand and reduce potential shortages.
- **Encourage Winter Usage**: Seasonal discounts or promotional offers during winter could boost rentals, helping balance demand across all seasons.
- **Consistent Resource Allocation Throughout the Week**: The analysis indicates that weekday and weekend demands are largely similar, so a uniform bike distribution strategy across the week may streamline operations with minimal weekend adjustments required.

These recommendations, grounded in FDA insights, support a strategic approach to managing bike-sharing resources, improving service reliability, and enhancing user satisfaction year-round.

11 References

Ramsay, J. O., & Silverman, B. W. (2005). Functional Data Analysis (2nd ed.). Springer.

Ramsay, J. O., Hooker, G., & Graves, S. (2009). Functional Data Analysis with R and MATLAB. Springer.

Febrero-Bande, M., & Oviedo de la Fuente, M. (2012). Statistical Computing in Functional Data Analysis: The R Package fda.usc. Journal of Statistical Software, 51(4), 1-28.

Yao, F., Müller, H.-G., & Wang, J.-L. (2005). Functional Data Analysis for Sparse Longitudinal Data. Journal of the American Statistical Association, 100(470), 577-590. (Basis for fdapace)

Cuevas, A., Febrero, M., & Fraiman, R. (2007). Robust Estimation and Classification for Functional Data via Projection-Based Depth Notions. Computational Statistics, 22(3), 481-496. (Basis for depthTools)

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … & Yutani, H. (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Goldsmith, J., Scheipl, F., Huang, L., Wrobel, J., Di, C., Gellar, J., & Crainiceanu, C. M. (2011). Refund: Regression with Functional Data.