Exploring Urban Mobility Patterns: A Data-Driven Analysis of Bike Rental Services

By: Aman Kumar
Instructor: Dr. Neeraj Kumar Sharma

Motivation

Emerging Industry:

• Motorbike rental services have become increasingly popular, especially in urban areas and tourist destinations.
• Analyzing this industry can provide insights into the trends, challenges, and opportunities within the market.

Societal Impact:

• Understanding motorbike rental patterns can have broader societal implications, such as promoting sustainable transportation and reducing traffic congestion.

Business Optimization:

• Businesses offering motorbike rental services can benefit from data-driven insights to optimize operations, pricing, and customer experiences.

Tourism and Travel:

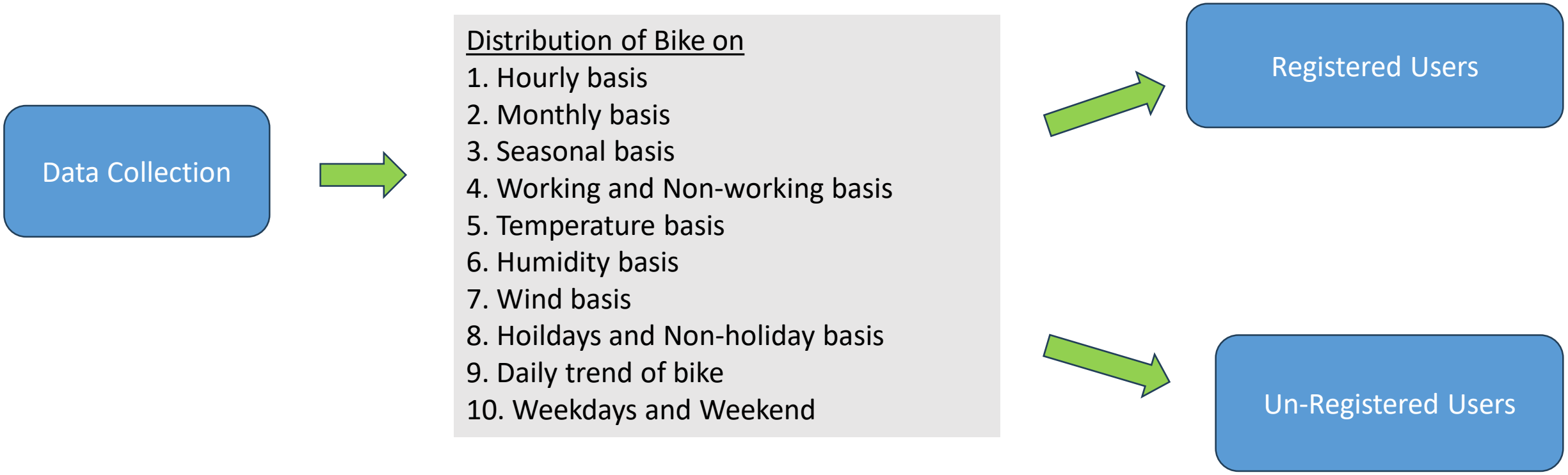
• In tourist-heavy areas, motorbike rentals are a popular mode of transportation for visitors.
•Analysis can shed light on seasonal trends, preferences, and the impact of tourism on motorbike rental demand.

Personal Connection!!

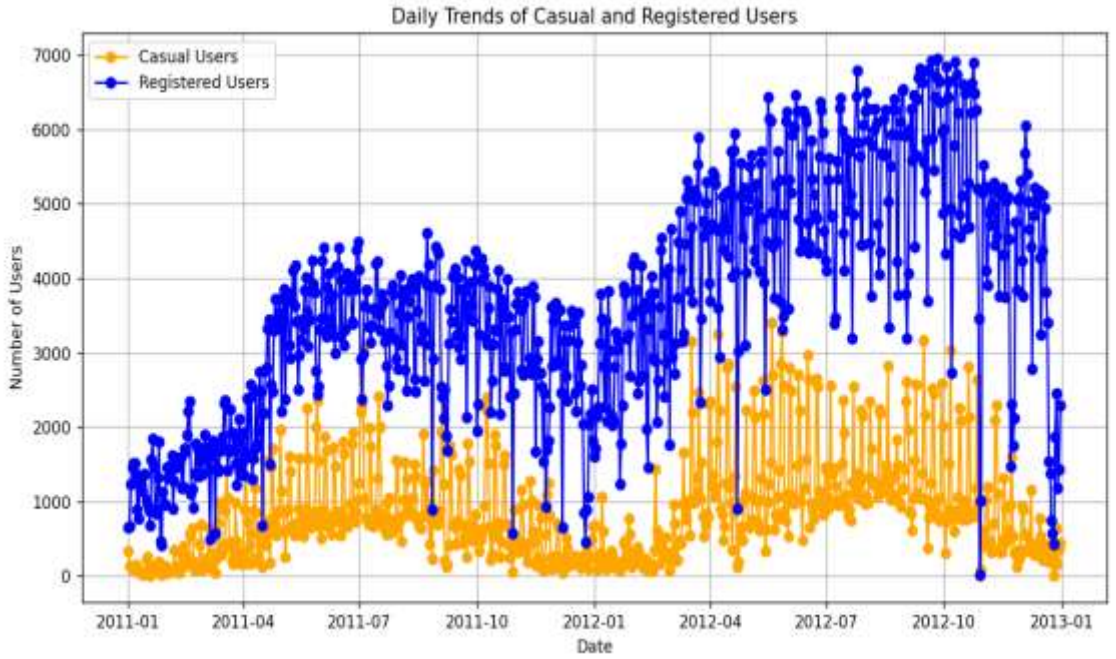
Growing up with a love for motorbikes, I have always been intrigued by the freedom and flexibility they offer.

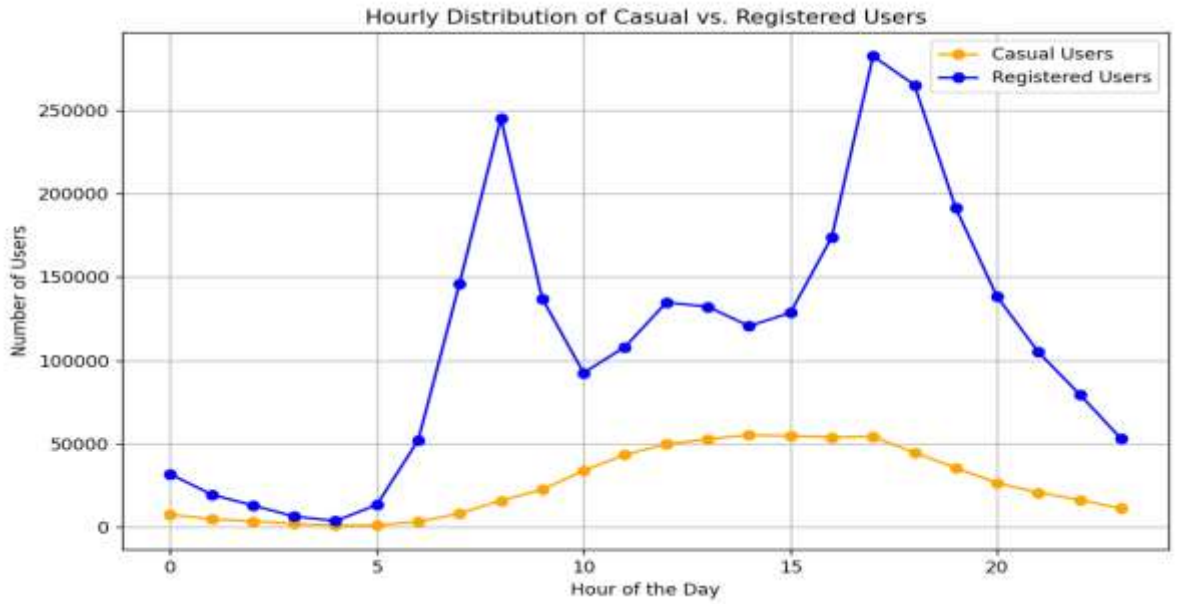
Block Diagram

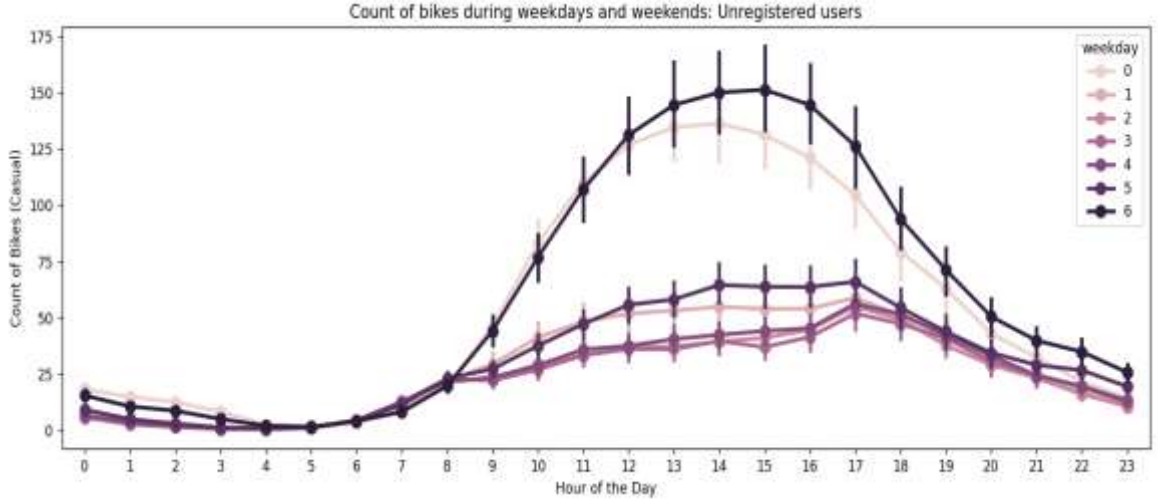
Problem statement: To find number of Bikes rent by Registered/Un-Registered users based on attributes like Seasons, Weekends, Weekday, Weather, Months, Different days of the week ,Temperature, Humidity, Wind Speed etc .

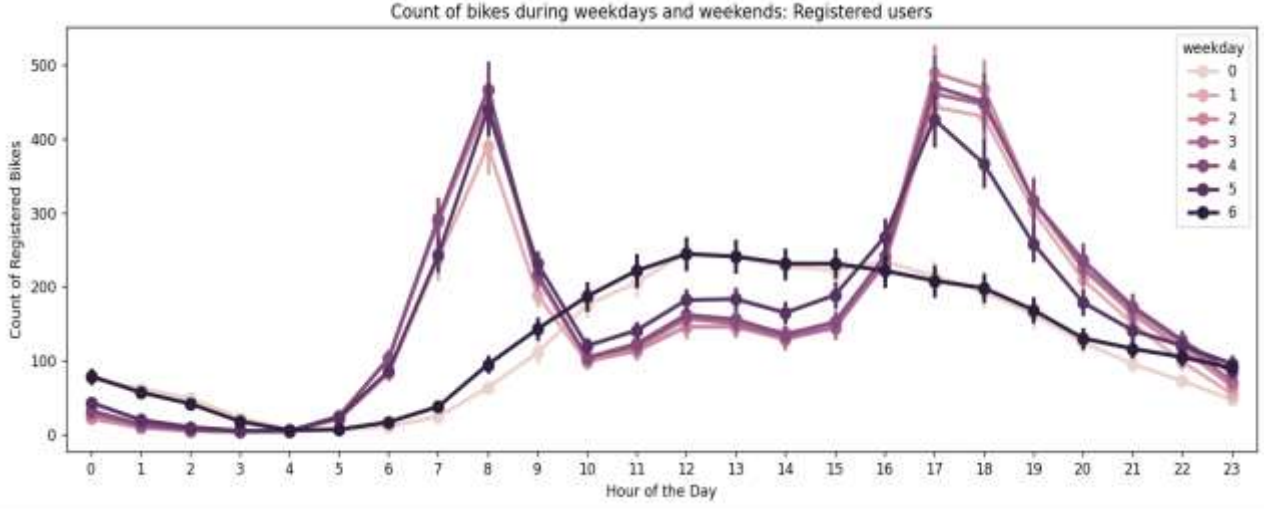


Registered Users And Un-Registered Users

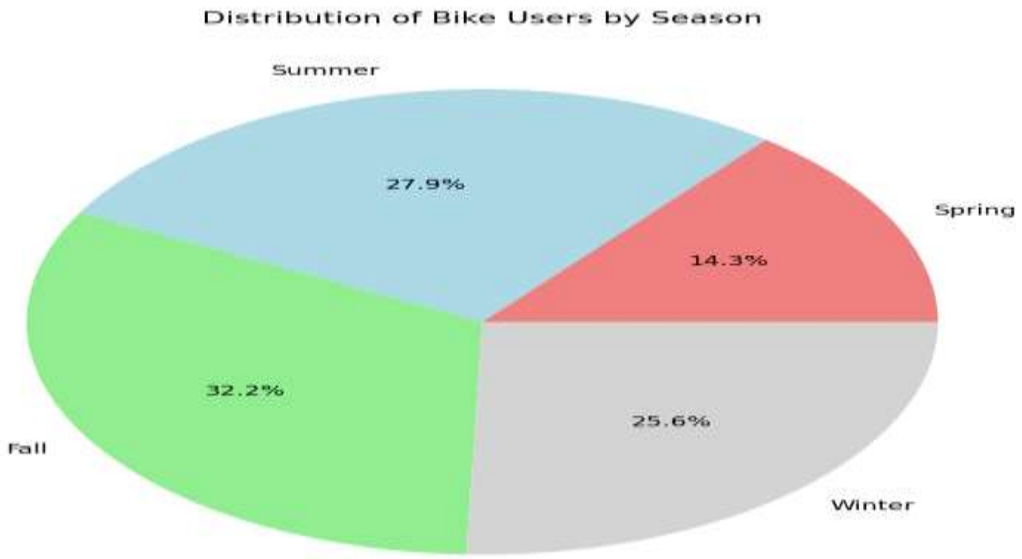


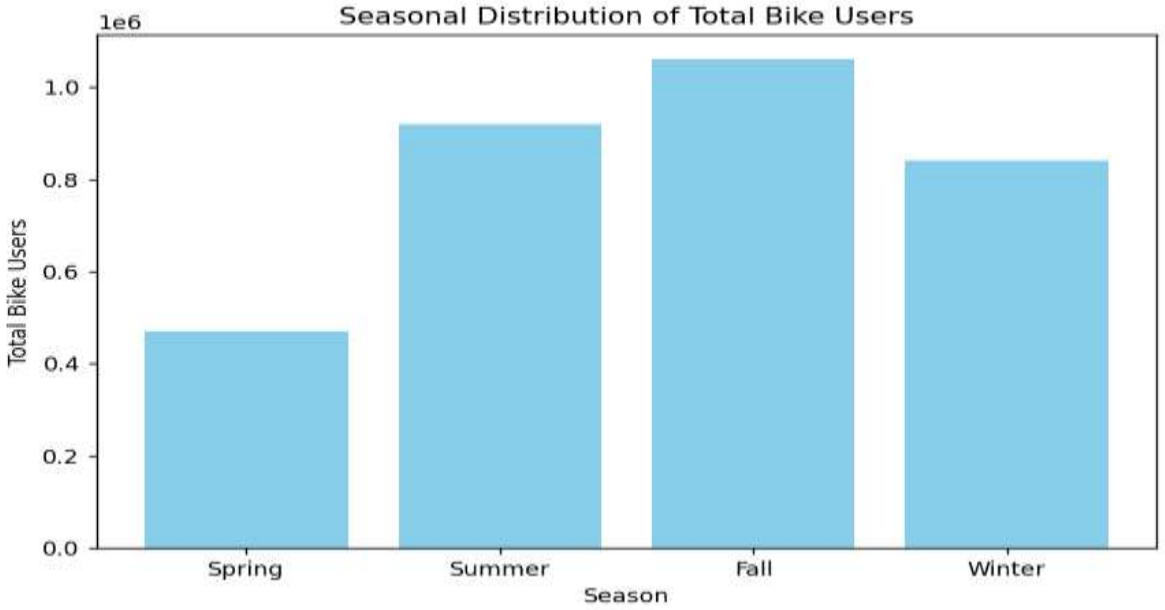




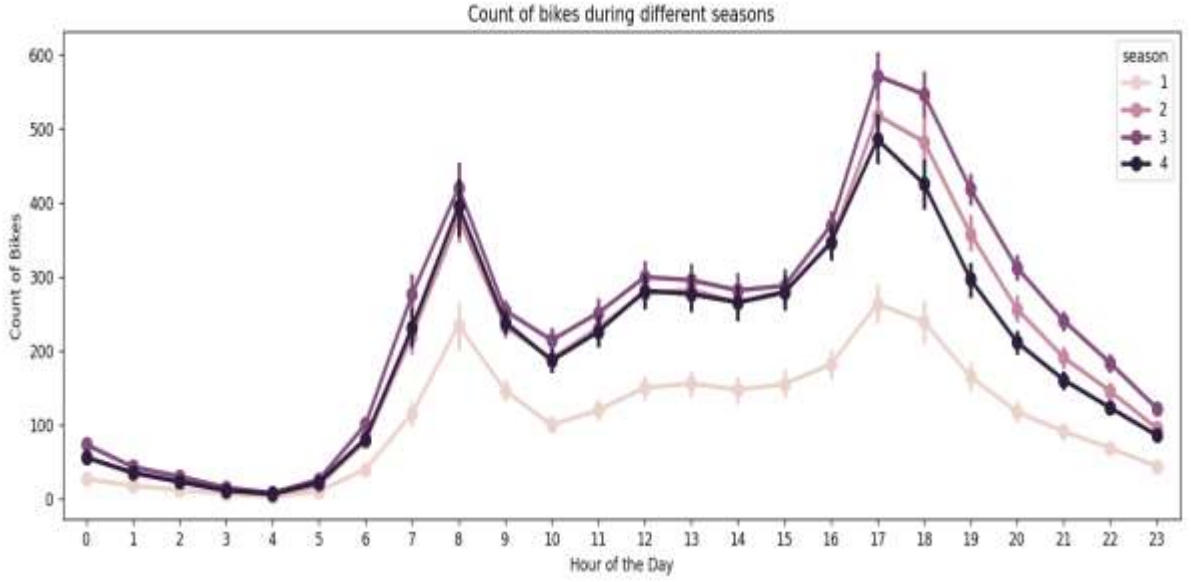


Season

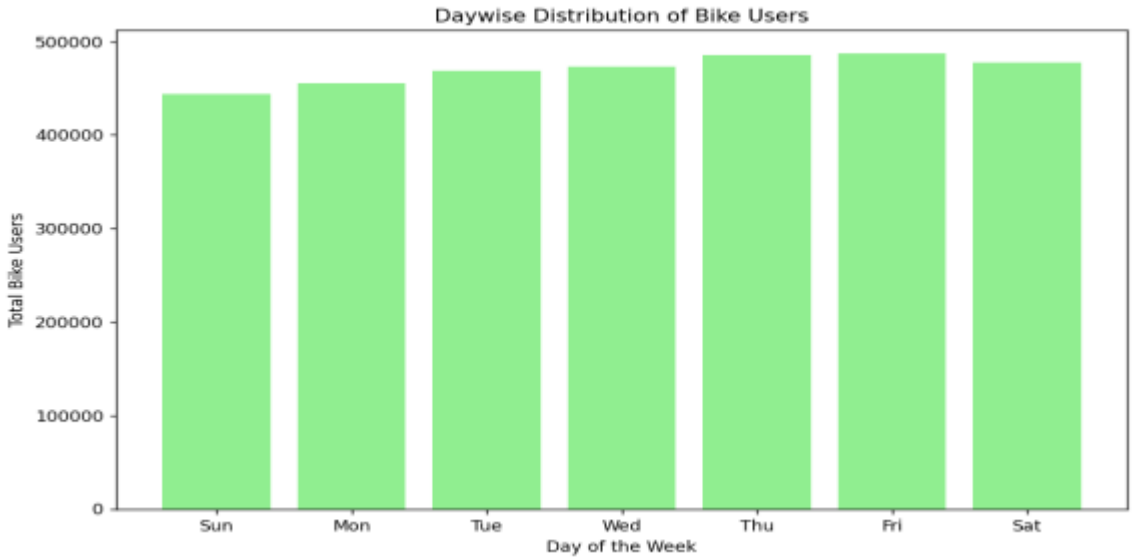


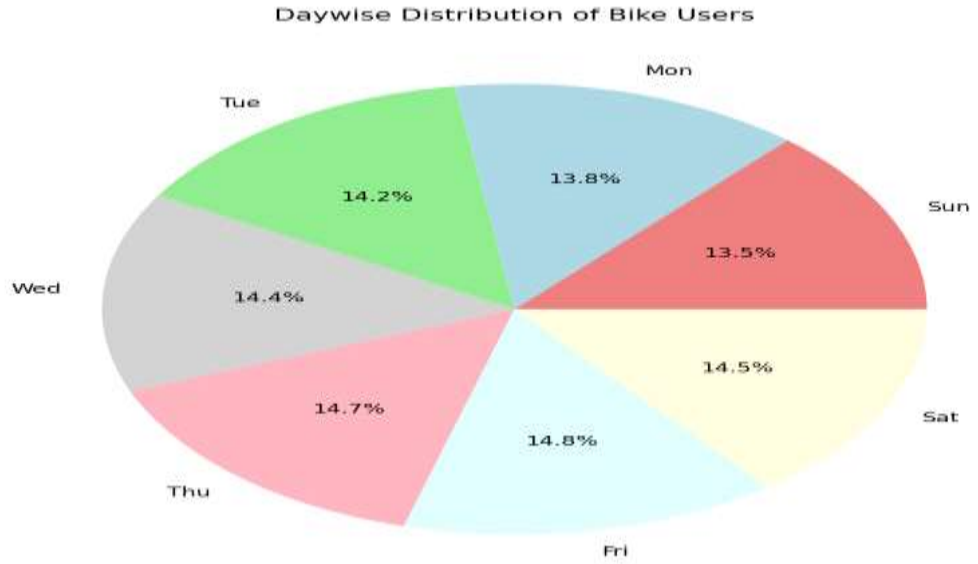



```python
def seasonal_counts(df):
    seasonal_counts = df.groupby('season')['cnt'].sum()

    plt.figure(figsize=(8, 5))
    plt.bar(seasonal_counts.index, seasonal_counts, color='skyblue')
    plt.title('Seasonal Distribution of Total Bike Users')
    plt.xlabel('Season')
    plt.ylabel('Total Bike Users')
    plt.xticks(seasonal_counts.index, ['Spring', 'Summer', 'Fall', 'Winter'])
    plt.show()
```

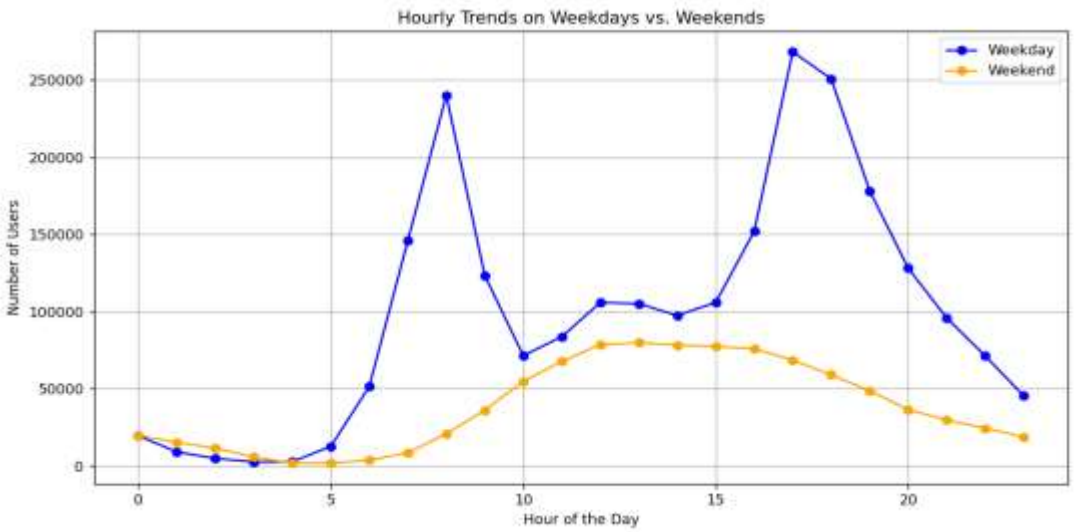


Days Of Week + Weekdays And Weekend

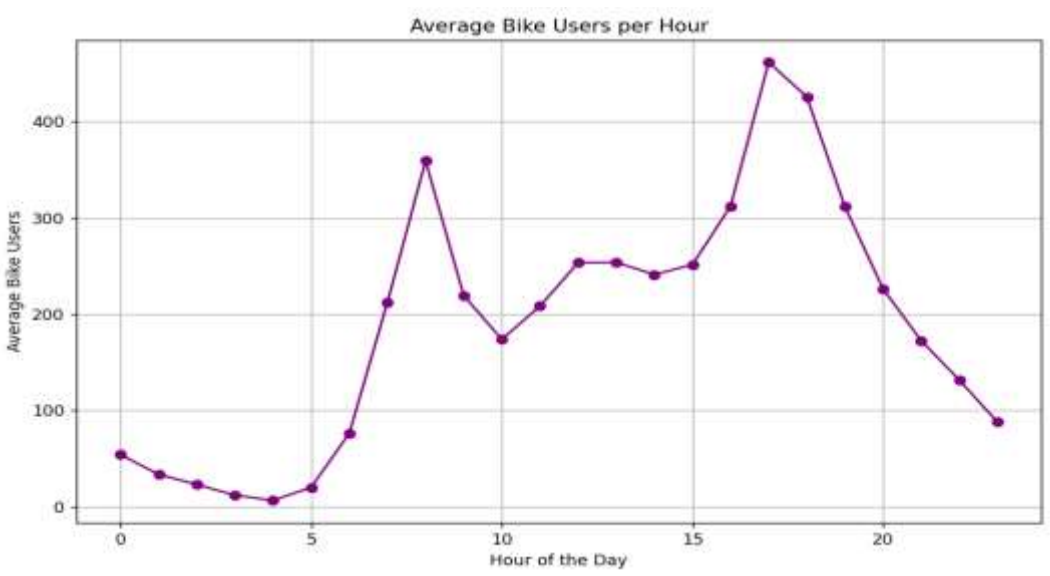


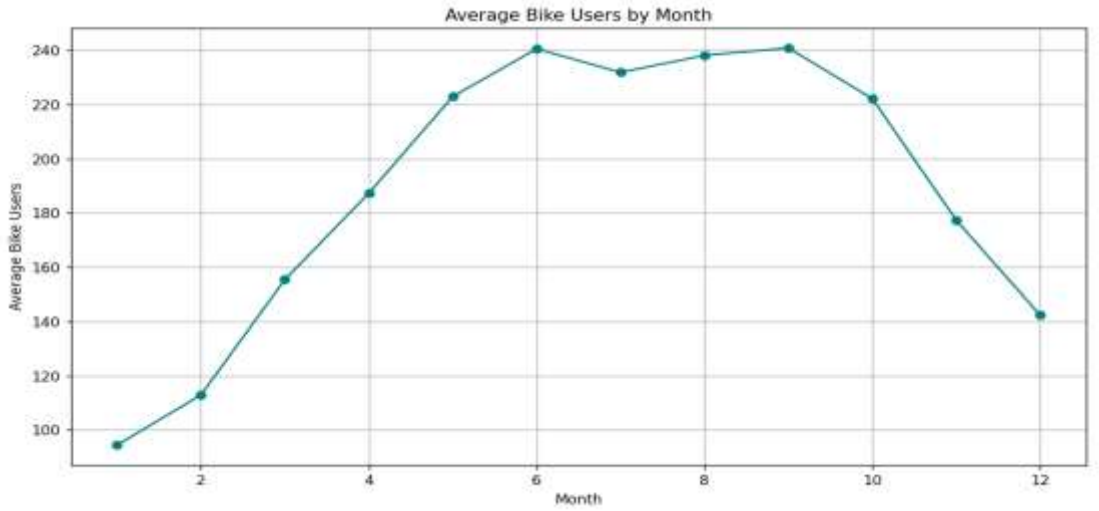


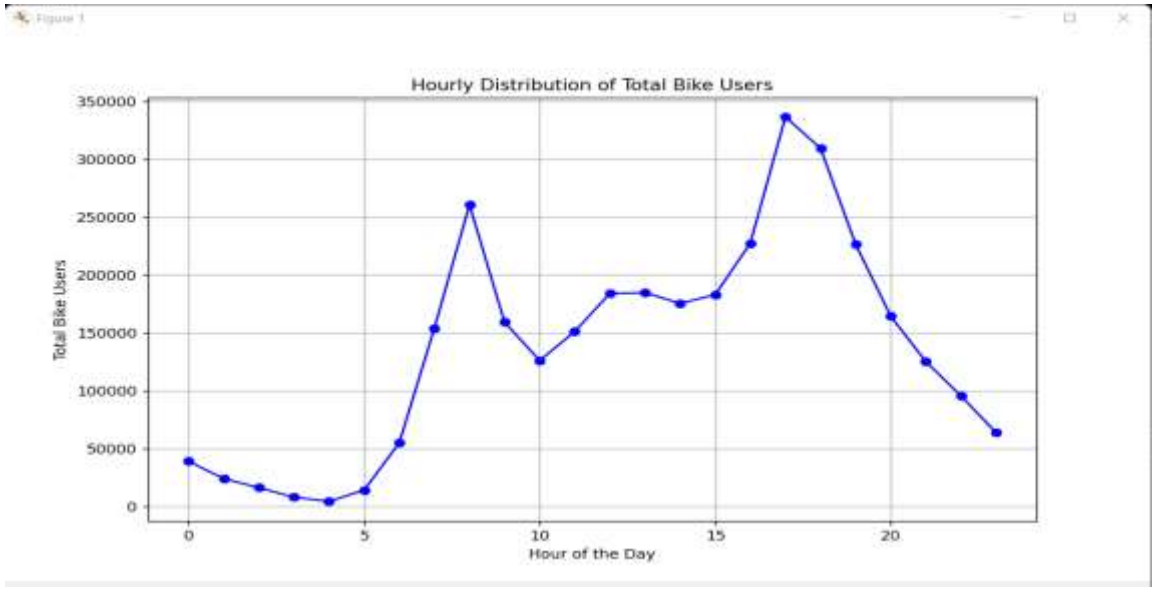

```python
def temp19(df):
    daywise_counts = df.groupby('weekday')['cnt'].sum()

    plt.figure(figsize=(10, 6))
    plt.bar(daywise_counts.index, daywise_counts, color='lightgreen')
    plt.title('Daywise Distribution of Bike Users')
    plt.xlabel('Day of the Week')
    plt.ylabel('Total Bike Users')
    plt.xticks(daywise_counts.index, ['Sun', 'Mon', 'Tue', 'Wed', 'Thu', 'Fri', 'Sat'])
    plt.show()
```

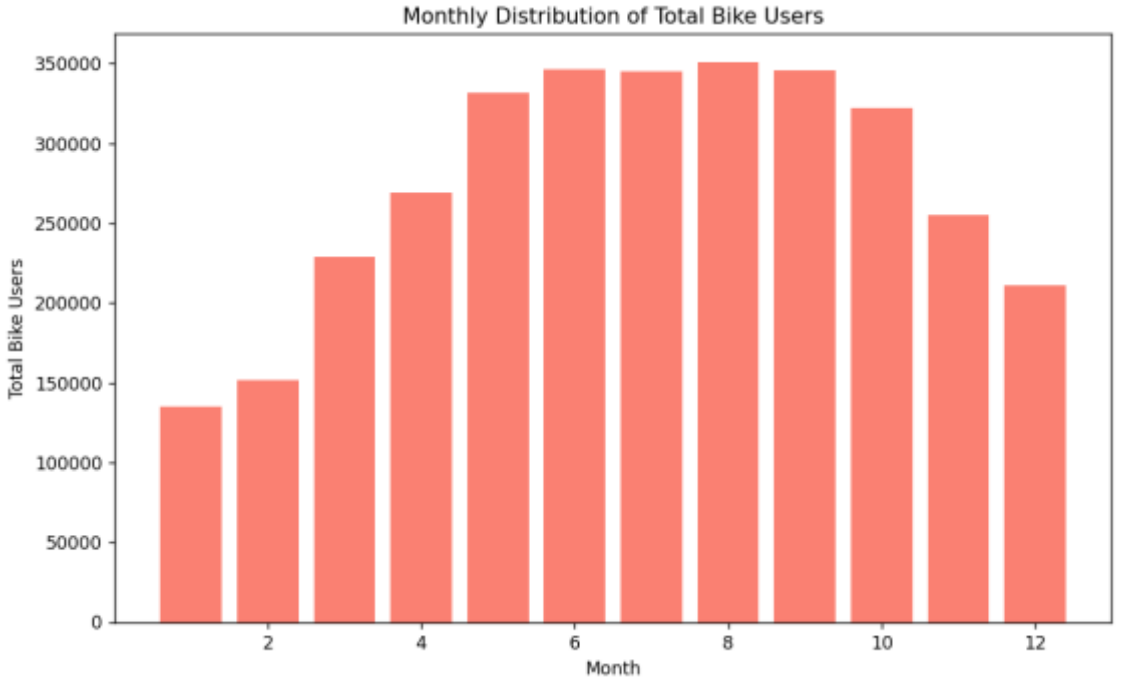


Hourly And Monthly

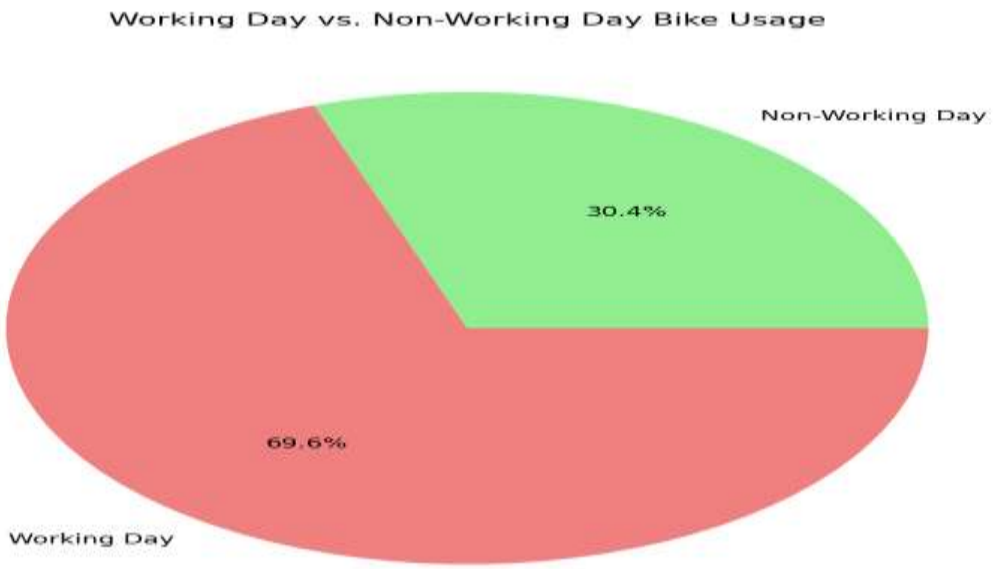


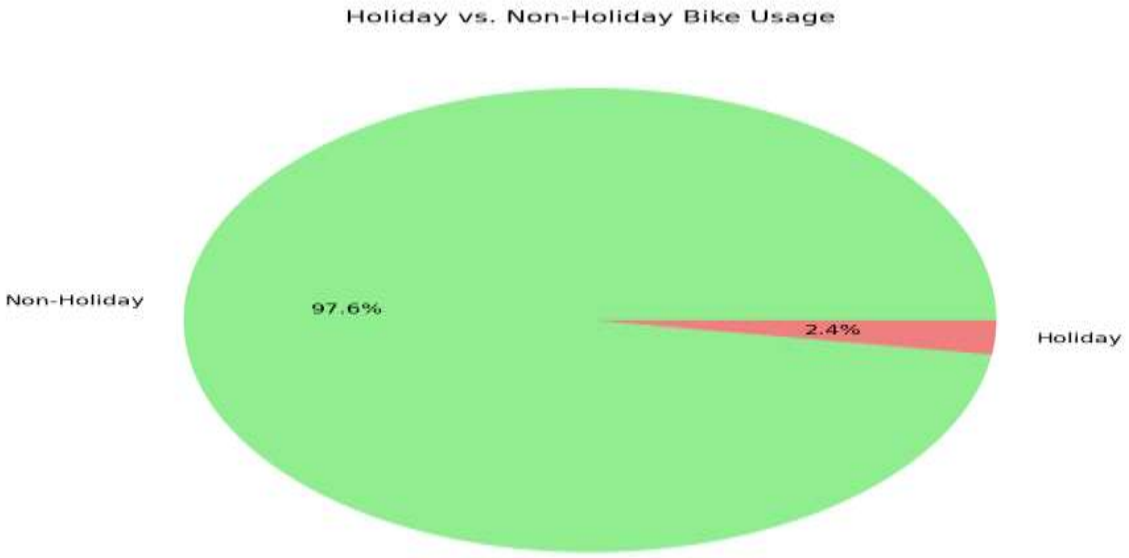


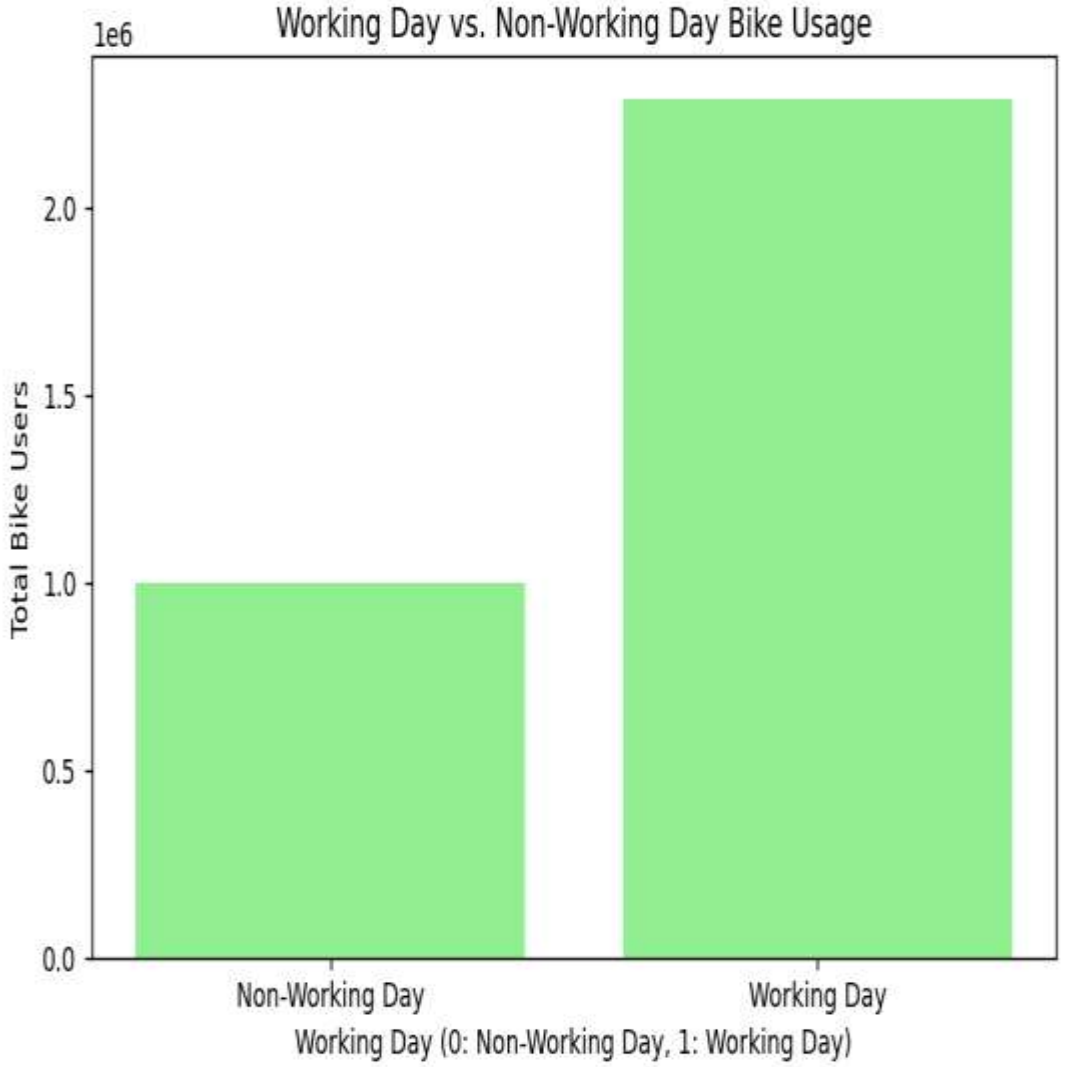




Working And Non-Working Days

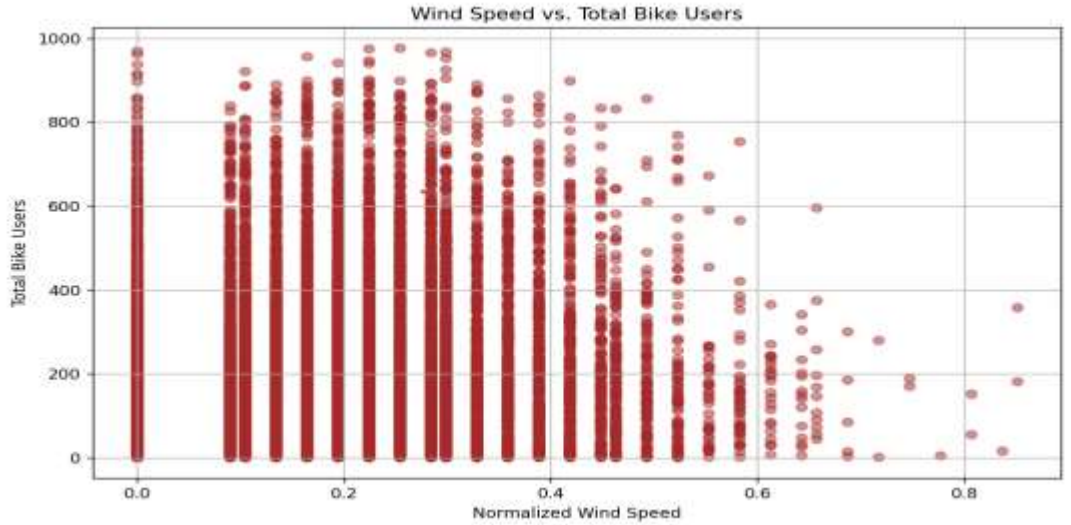


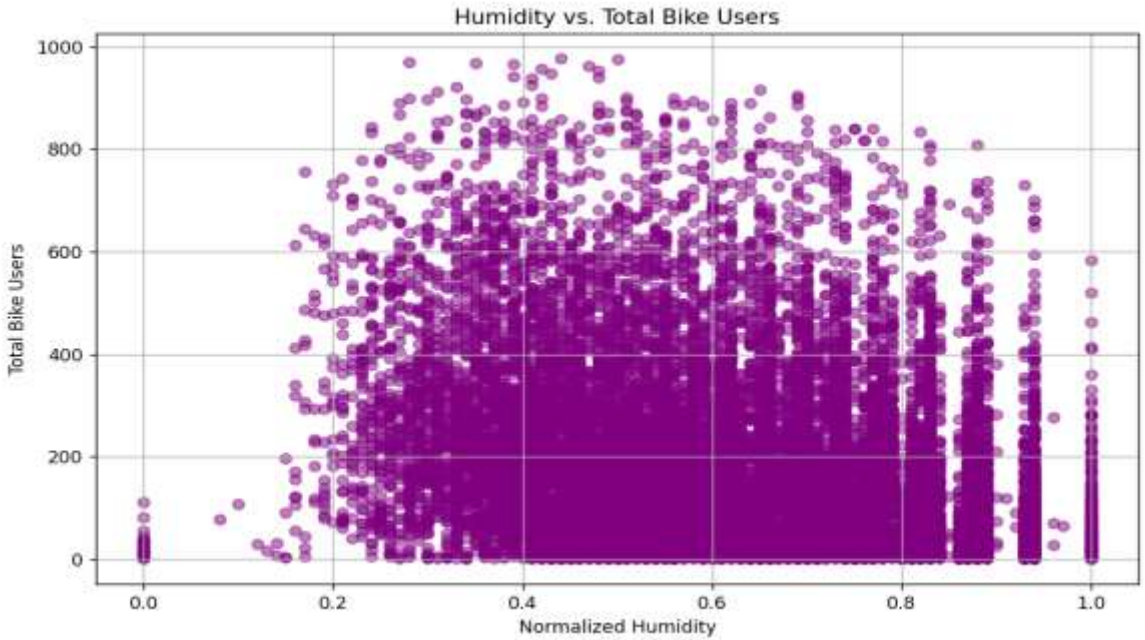


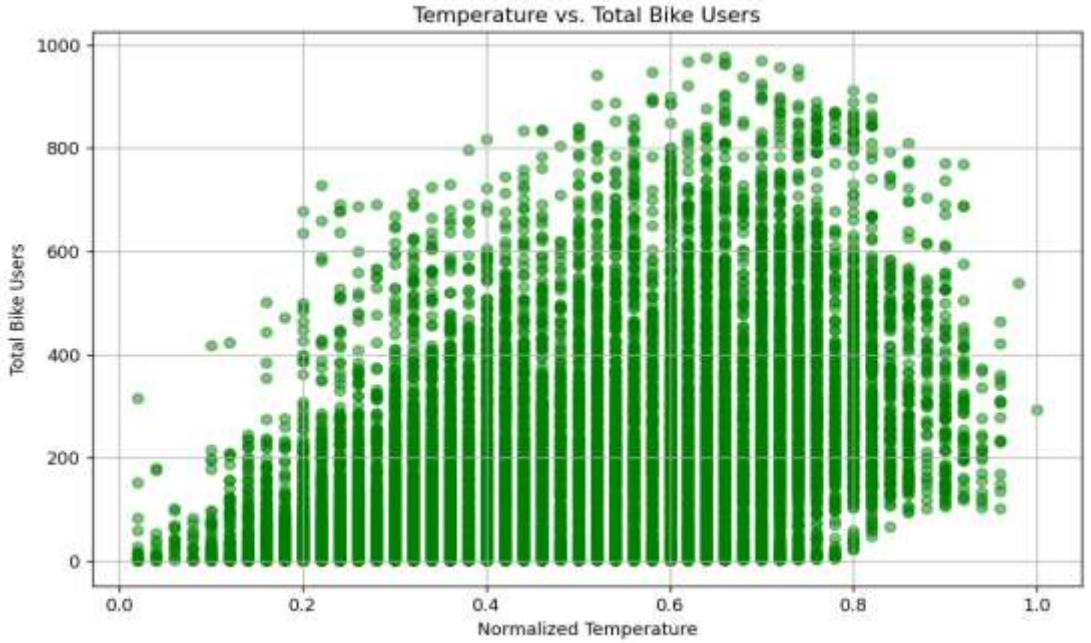


Temperature, Humidity, Wind Speed

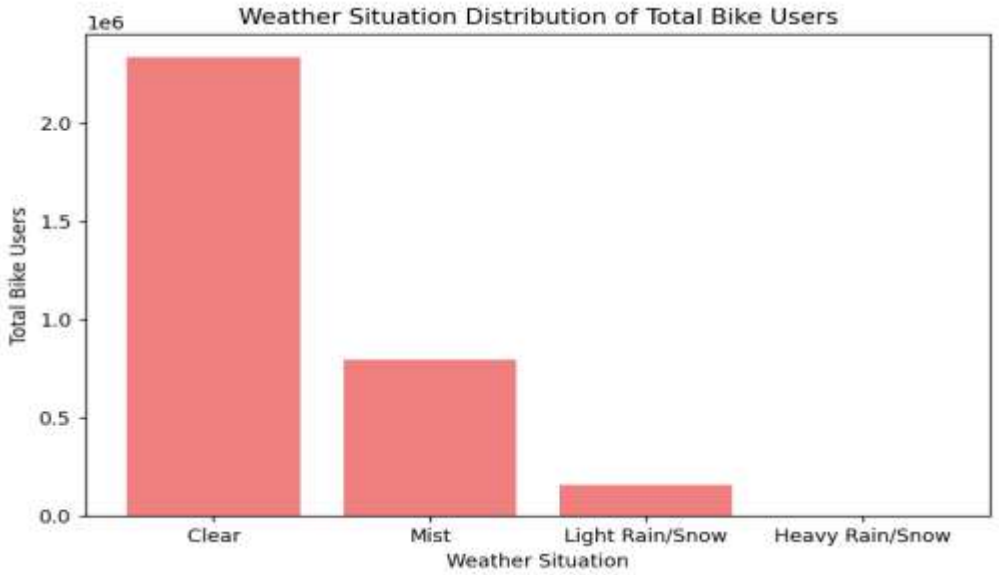
```python
def temp18(df):
    plt.figure(figsize=(12, 6))
    sns.scatterplot(x='hum', y='cnt', hue='season', data=df, palette='coolwarm')
    plt.title('Impact of Humidity on Bike Usage During Different Seasons')
    plt.xlabel('Normalized Humidity')
    plt.ylabel('Total Bike Users')
    plt.show()
```

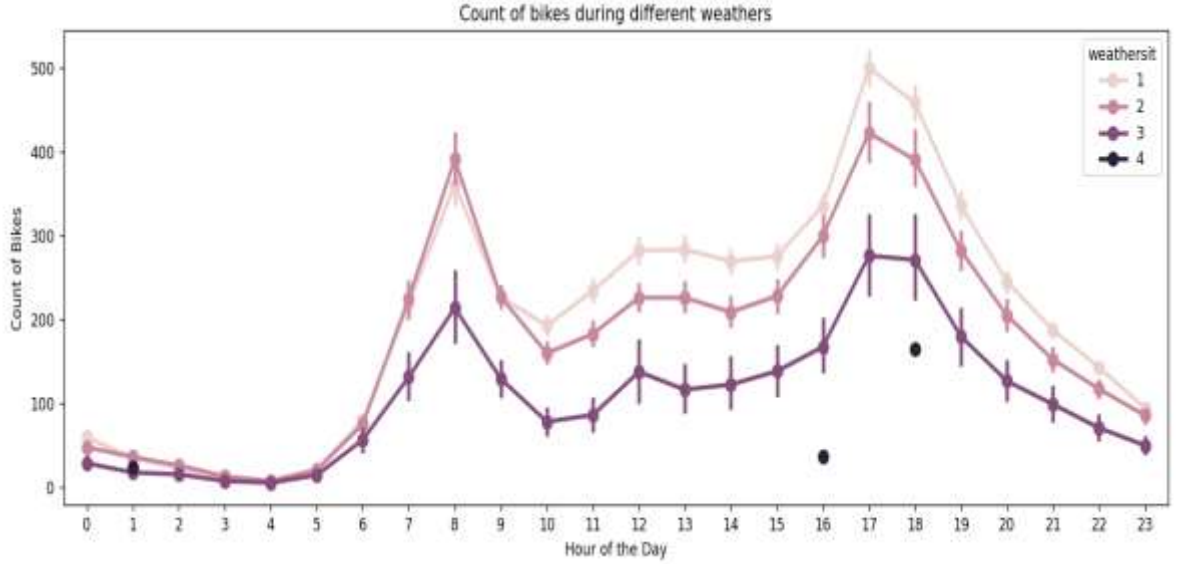


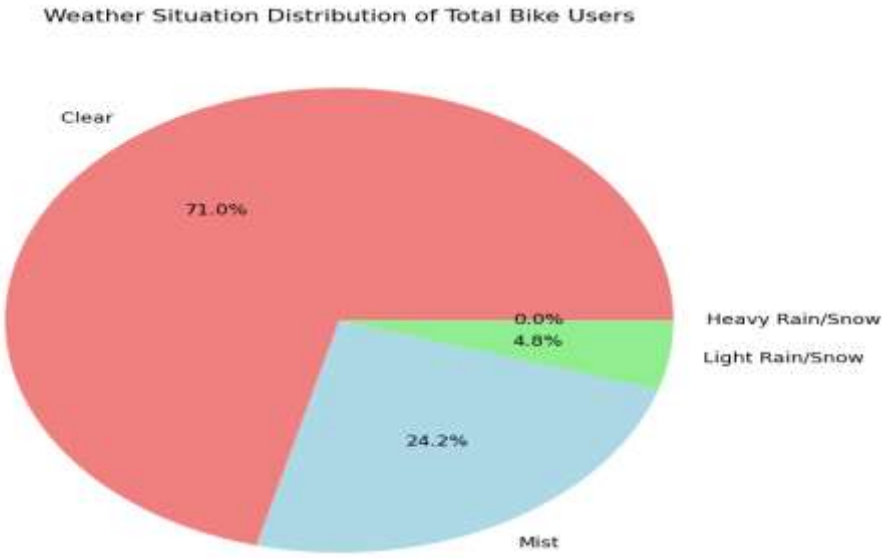


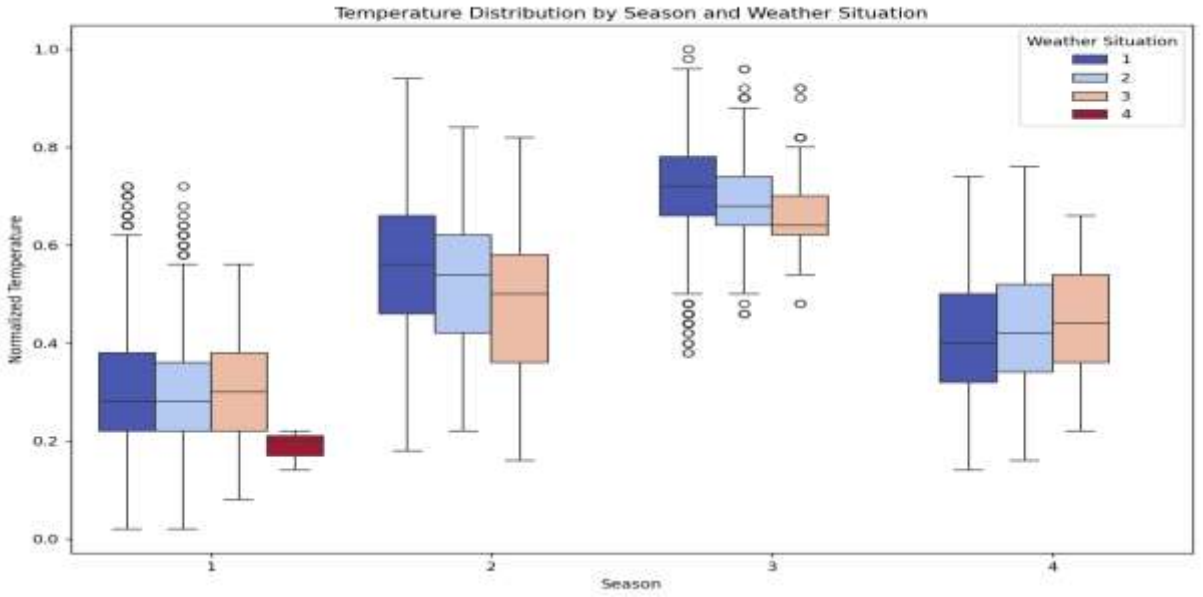


Weather

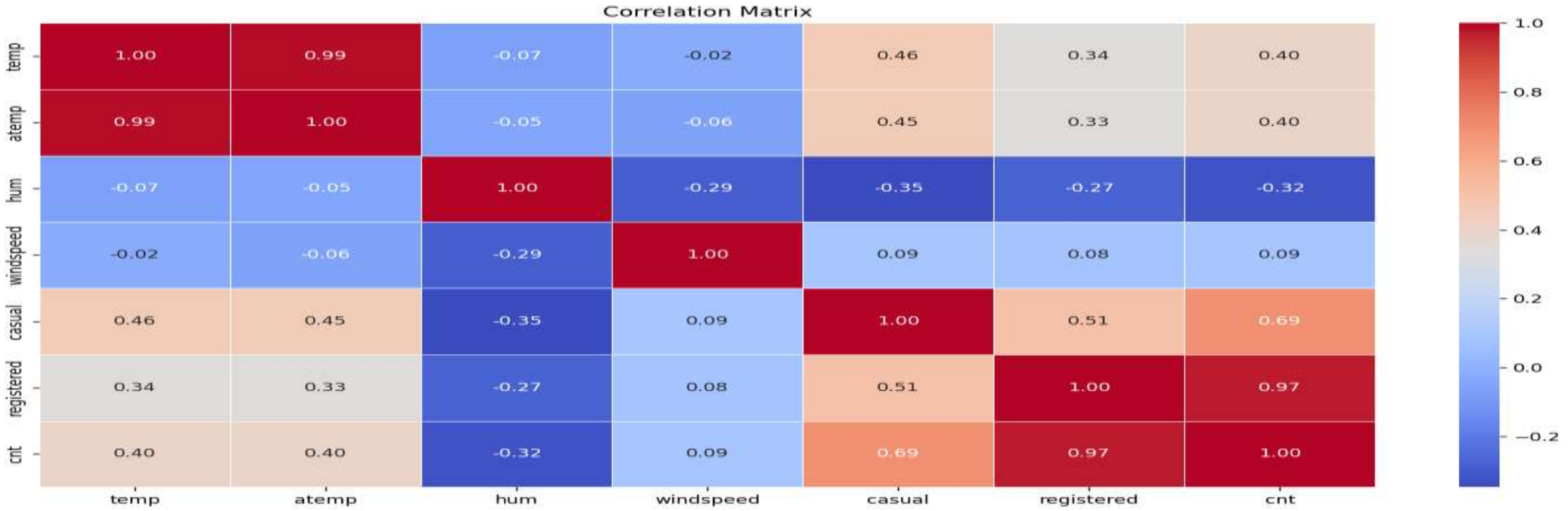

Weather Situation Distribution of Total Bike Users
1e6
Total Bike Users
2.0
1.5
1.0
0.5
0.0
Clear
Mist
Light Rain/Snow
Heavy Rain/Snow
Weather Situation


Count of bikes during different weathers
weathersit
1
2
3
4
Count of Bikes
500
400
300
200
100
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Hour of the Day


Weather Situation Distribution of Total Bike Users
Clear
71.0%
0.0%
4.8%
24.2%
Heavy Rain/Snow
Light Rain/Snow
Mist


Temperature Distribution by Season and Weather Situation
Weather Situation
1
2
3
4
Normalized Temperature
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
4
Season

Correlation Matrix




```python
import seaborn as sns
corr_matrix = df[['temp', 'atemp', 'hum', 'windspeed', 'casual', 'registered', 'cnt']].corr()

plt.figure(figsize=(10, 8))
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm', fmt='.2f', linewidths=0.5)
plt.title('Correlation Matrix')
plt.show()
```

Pseudo Code

```
# Import required libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

# Define the main function
def main():
    # Load the dataset
    df = load_dataset()

    # Perform data analysis and visualization
    # visualize_hourly_distribution(df)
    # visualize_monthly_distribution(df)
    # visualize_seasonal_distribution(df)
    # visualize_workingday_distribution(df)
    # visualize_weather_distribution(df)
    # visualize_hourly_casual_registered(df)
    # visualize_temperature_vs_bike_users(df)
    # visualize_humidity_vs_bike_users(df)
    # visualize_windspeed_vs_bike_users(df)
    # visualize_holiday_distribution(df)
    # visualize_daily_trends(df)
    # visualize_box_plot_season(df)
    # visualize_weekday_vs_weekend_bike_counts(df)
    # visualize_yearly_usage(df)
    # visualize_temperature_impact(df)
    # visualize_hourly_trends_weekdays_weekends(df)
    # visualize_correlation_matrix(df)
    # visualize_holiday_season_distribution(df)
    # visualize_box_plot_weather(df)
    # visualize_monthly_distribution_pie_chart(df)
    # visualize_seasonal_distribution_pie_chart(df)
    # visualize_weekday_distribution_pie_chart(df)
    # visualize_workingday_distribution_pie_chart(df)
    # visualize_bike_usage_comparison_workingdays(df)
    # visualize_average_bike_users_per_hour(df)
    # visualize_pairwise_scatter_plot(df)
    # visualize_average_bike_users_by_month(df)
    # visualize_bike_usage_comparison_weekdays_weekends_season(df)
    # visualize_registered_vs_casual_users_by_temperature(df)
    # visualize_bike_counts_weekdays_weekends(df)
    # visualize_bike_counts_unregistered_users(df)
    # visualize_bike_counts_registered_users(df)
    # visualize_bike_counts_weather_situations(df)
    # visualize_bike_counts_seasons(df)
    # visualize_correlation_matrix_selected_features(df)
    # visualize_windspeed_impact_on_bike_usage(df)
    # visualize_daily_trends_casual_registered(df)
    # visualize_temperature_vs_bike_users_scatter_plot(df)
    # visualize_windspeed_vs_bike_users_scatter_plot(df)
    # visualize_box_plot_total_bike_users_by_season(df)

# Define functions for each type of visualization
# (Function definitions should follow)

# Execute the main function when the script is run
if __name__ == "__main__":
    main()
```

Observation

1. **Hourly Distribution:** There is a clear hourly trend, with peak bike usage during certain hours of the day, indicating potential commuting patterns.
2. **Monthly Distribution:** The bike usage shows variations over the months, with potential seasonality trends. More people may rent bikes during specific months, likely influenced by weather conditions.
3. **Seasonal Distribution:** Different seasons impact bike rental patterns. For example, usage might increase during warmer seasons and decrease during colder ones.
4. **Working Day vs. Non-Working Day:** Bike usage patterns differ between working days and non-working days. It's common to see increased usage on working days, possibly due to commuting purposes.
5. **Weather Situation Distribution:** Weather conditions affect bike rental demand. Clear weather might attract more riders, while adverse conditions like rain or snow may reduce usage.
6. **Casual vs. Registered Users:** Distinguishing between casual and registered users helps understand user behavior. Casual users might contribute more to peak usage, while registered users show consistent patterns.
7. **Temperature, Humidity, and Windspeed Impact:** Temperature and weather-related features have a significant impact on bike usage. There might be an optimal temperature range for bike rentals.
8. **Holiday vs. Non-Holiday Usage:** Holidays may influence bike rental patterns, with potential increased usage for recreational purposes or decreased usage due to travel.
9. **Daily Trends:** Analyzing daily trends provides insights into user preferences and habits. Weekdays may show more commuting-related usage, while weekends may exhibit more leisure-oriented patterns.
10. **Correlation Matrix:** Examining correlations between various features, such as temperature, humidity, and windspeed, helps identify relationships and potential predictors of bike usage.
11. **Bike Usage Over Years:** Understanding how bike usage evolves over the years helps identify overall growth or decline in popularity.