Data **Visualization** in Python

Overview

This project explores the fundamentals of creating interactive plots using Plotly. The dataset collected from Seoul's bike-sharing data to create bar plots, scatter plots, and line plots using Plotly. Along the way, we'll explore how weather patterns in Seoul impact bike-sharing trends.

Problem Statement

The project aims to understand how various factors, including weather patterns, time of day, seasonality, and holidays, influence bike rental trends in Seoul's bike-sharing system.

Goals

Visualize Bike Rental Trends

Create interactive plots to analyze bike rentals over time, daily usage patterns, and variations across seasons.

Explore Weather Impact

Investigate the relationship between weather conditions (temperature, wind speed, etc.) and bike rentals.

Identify Usage Patterns

Understand typical daily usage patterns and how they differ based on season and holidays.

Dataset & Source

The dataset '**seoul_bike_data_renamed.csv**' used is Seoul's bike-sharing data, retrieved from the UCI Machine Learning Repository (Source).
This data provides hourly information on bike rentals, including weather conditions, timing, and holiday status.

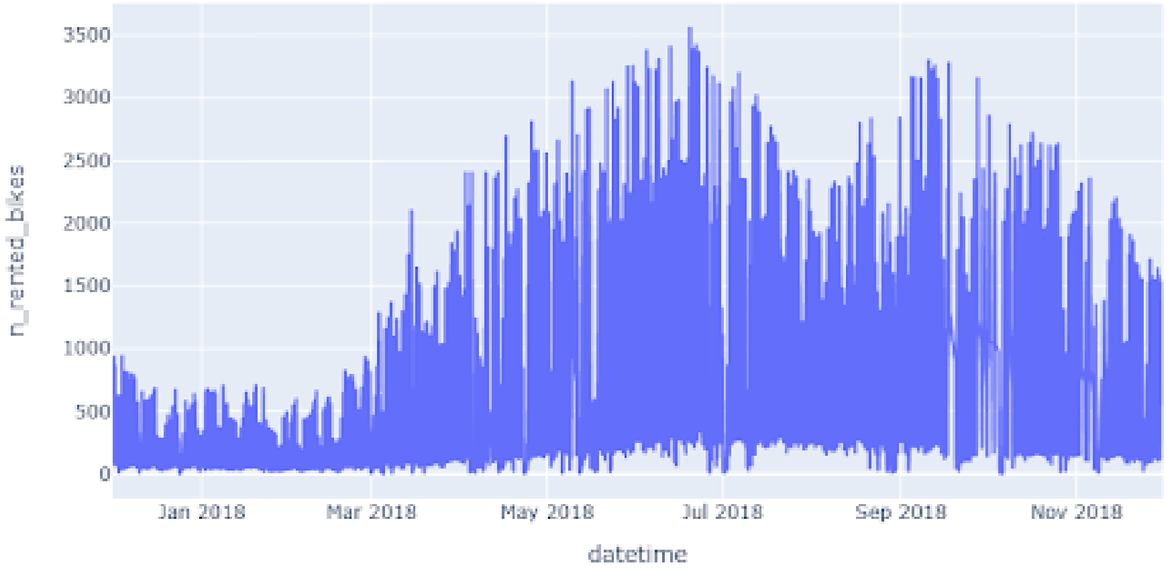
Data Cleaning Steps (Code Snippets [1, 2])

- Load the CSV data using pandas.
- Convert the "date" column to datetime format.
- Create a "datetime" column combining date and hour.
- Convert "is_holiday" and "is_functioning" columns to boolean values (True/False) using a map function.
- Filter the data to include only observations where the bike-sharing system is functioning.

Data Findings (Code Snippets [3-11])

- **Temporal Trends:** Bike rentals show fluctuations over time, with potentially higher activity during specific periods. Daily rentals also exhibit variations.
- **Seasonal Variations:** Bike rentals appear to be seasonal, with potentially higher usage during warmer months (Spring, Summer).
- **Weather Impact:** There might be a positive correlation between temperature and bike rentals, suggesting more rentals on warmer days. The impact of other weather factors like wind speed, humidity, or rainfall needs further investigation.
- **Daily Usage Patterns:** Bike rentals likely follow a daily pattern with peak usage during commute times (mornings and evenings).
- **Holiday Impact:** New Year's Eve might exhibit a unique usage pattern compared to a typical winter day.

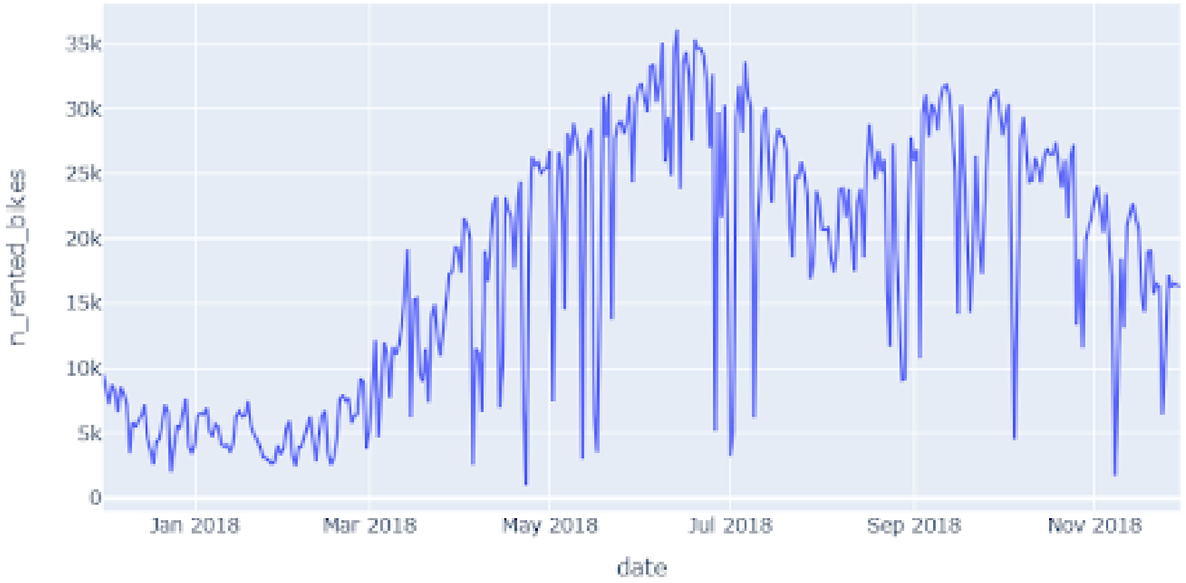
Line plot of rented bikes over time



The line plot visually represents the number of rented bikes over time in Seoul's bike-sharing system.
Here are some key observations:

- **Seasonality:** The plot clearly demonstrates a seasonal pattern in bike rentals. There's a noticeable increase in rentals during the warmer months (spring and summer) and a decrease during colder months (fall and winter). This aligns with the general trend of increased outdoor activity during warmer seasons.
- **Daily Fluctuations:** Within each season, there are daily fluctuations in rental numbers. This likely reflects the daily routines of bike-sharing users, with peaks during commuting hours (morning and evening) and lower activity during off-peak times.
- **Specific Events:** The plot might also reveal spikes or dips associated with specific events, such as holidays, festivals, or adverse weather conditions.

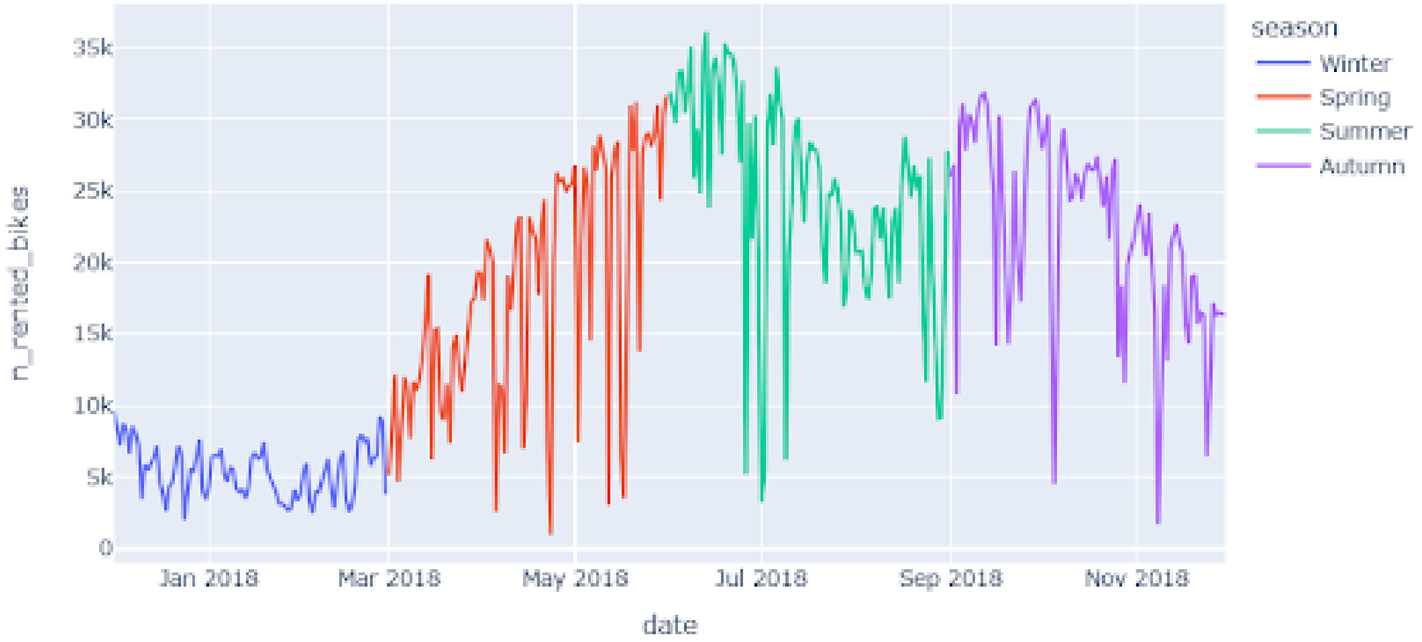
Line plot showing total number of bikes per day over time



The line plot visually represents the number of rented bikes over time in Seoul's bike-sharing system.
Here are some key observations:

- **Seasonality:** The plot clearly demonstrates a seasonal pattern in bike rentals. There's a noticeable increase in rentals during the warmer months (spring and summer) and a decrease during colder months (fall and winter). This aligns with the general trend of increased outdoor activity during warmer seasons.
- **Daily Fluctuations:** Within each season, there are daily fluctuations in rental numbers. This likely reflects the daily routines of bike-sharing users, with peaks during commuting hours (morning and evening) and lower activity during off-peak times.
- **Specific Events:** The plot might also reveal spikes or dips associated with specific events, such as holidays, festivals, or adverse weather conditions.

Line plot showing total number of bikes per day over time and map season to color

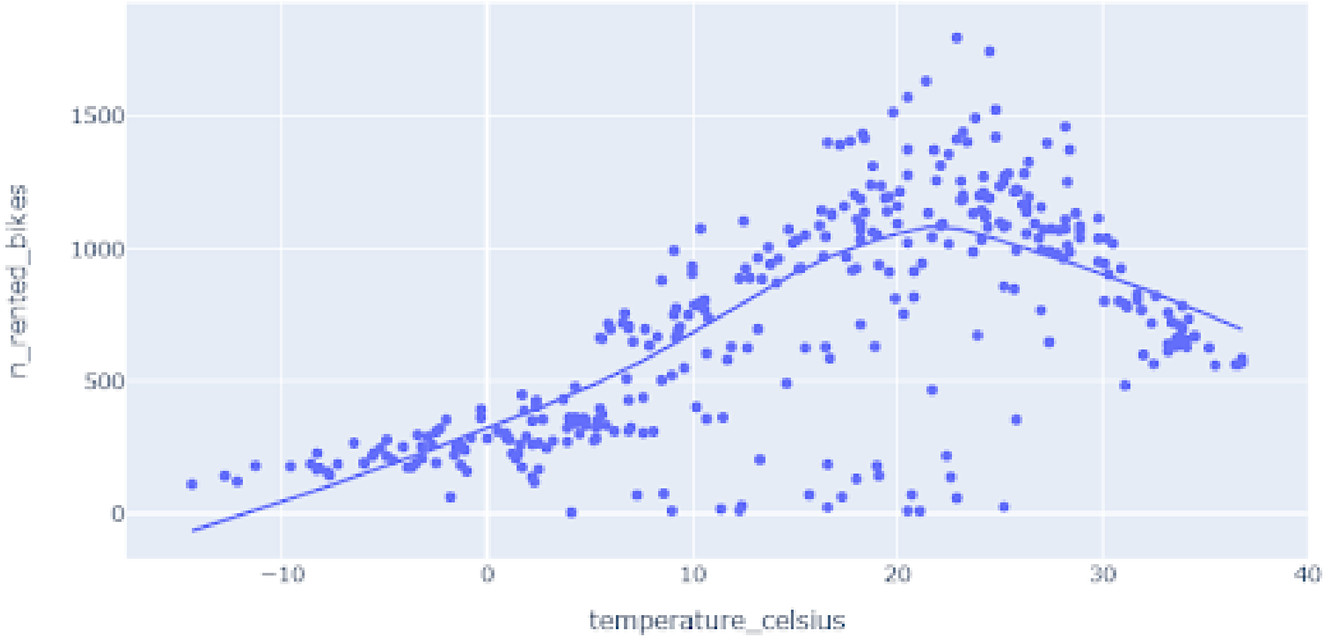


The provided line plot visualizes the number of rented bikes over time in Seoul's bike-sharing system, with different lines representing each season (Winter, Spring, Summer, Autumn).

Here are some key observations:

- **Seasonal Patterns:**
  - Winter: The winter line shows a lower number of rented bikes compared to other seasons. This aligns with the general trend of decreased outdoor activity during colder months.
  - Spring and Summer: Both lines exhibit a significant increase in bike rentals compared to winter. This is likely due to warmer weather and increased outdoor activities.
  - Autumn: The autumn line shows a slight decrease in rentals compared to summer but remains higher than winter.
- **Daily Fluctuations:**
  - Within each season, there are daily fluctuations in rental numbers. This likely reflects the daily routines of bike-sharing users, with peaks during commuting hours (morning and evening) and lower activity during off-peak times.

Scatter plot showing temperature against number of rented bikes

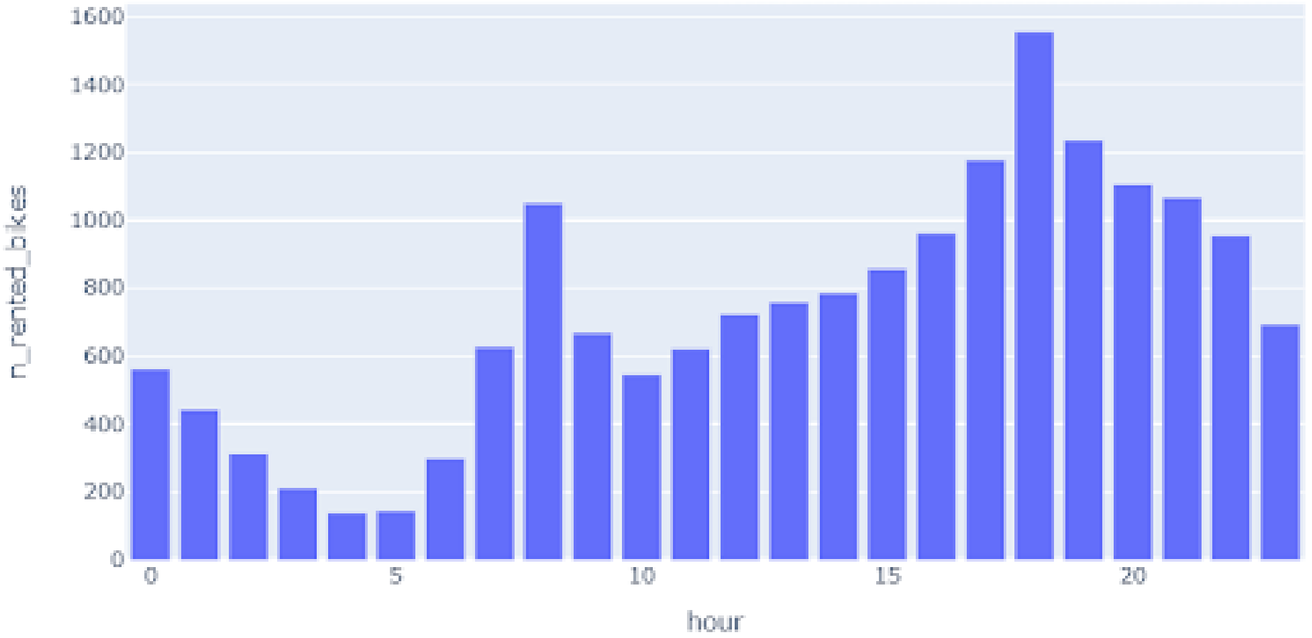


This scatter plot with a trendline shows the relationship between temperature and the number of rented bikes in Seoul's bike-sharing system.

Key Observations:

- **Positive Correlation:** There's a clear positive correlation between temperature and bike rentals. As the temperature increases, the number of rented bikes also tends to increase.
- **Non-Linear Relationship:** The trendline suggests that the relationship is not strictly linear. There might be a point of diminishing returns, where further increases in temperature lead to smaller increases in bike rentals.
- **Data Spread:** The data points are scattered around the trendline, indicating that temperature is not the sole factor influencing bike rentals. Other factors like weather conditions, holidays, and events can also play a role.

Bar chart showing average number of rented bikes per hour

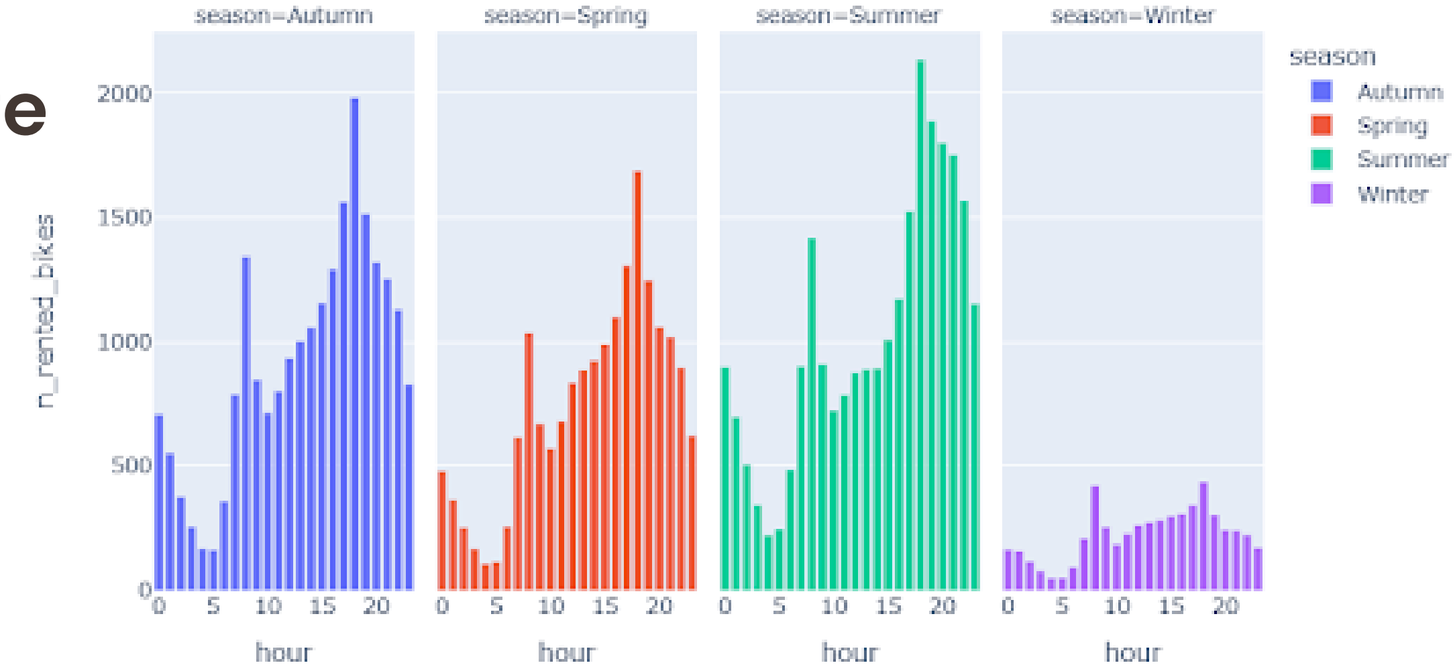


This bar chart visualizes the average number of rented bikes per hour in Seoul's bike-sharing system.

Key Observations:

- **Peak Usage Hours:** There are two distinct peaks in bike rentals: one in the morning (around 7-9 AM) and another in the evening (around 5-7 PM). These peaks likely correspond to commuting hours when people use bikes to get to work or school.
- **Off-Peak Hours:** During the early morning hours (before 6 AM) and late night hours (after 10 PM), bike rentals are significantly lower. This indicates that the system is primarily used for commuting purposes.
- **Consistent Pattern:** The overall pattern of bike rentals throughout the day remains relatively consistent, suggesting that people's commuting habits are fairly stable.

Bar chart showing average number of rented bikes per hour to show usage pattern per season

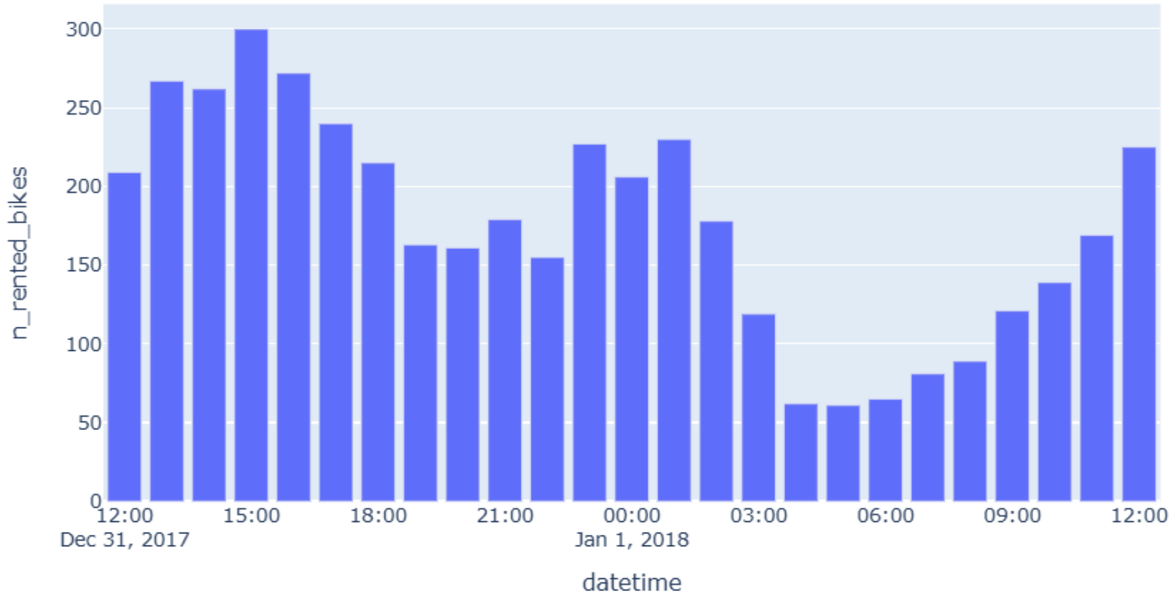


This plot visualizes the average number of rented bikes per hour for each season in Seoul's bike-sharing system. It provides a more detailed breakdown of usage patterns across different seasons.

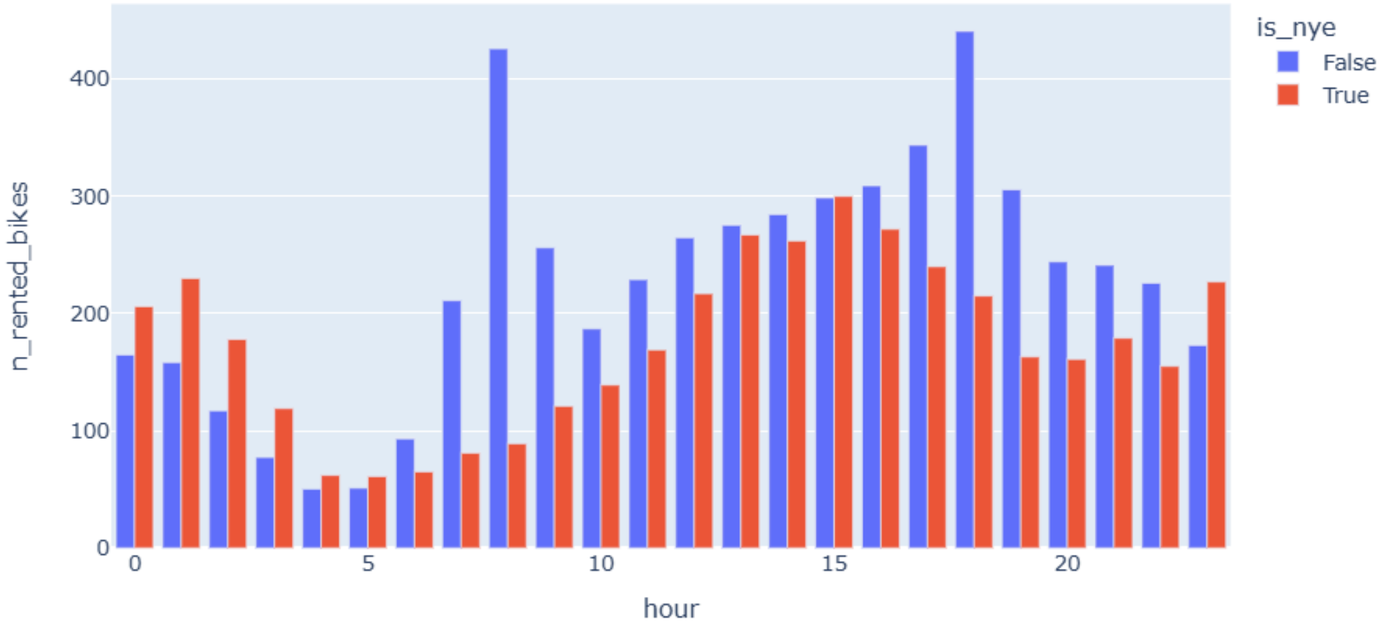
Key Observations:

- **Seasonal Differences:** The plots clearly show distinct seasonal patterns in bike rental behavior.
  - **Summer:** Summer has the highest overall bike rental activity, with peaks in the morning and evening, likely due to commuting and leisure activities.
  - **Spring and Autumn:** These seasons have similar patterns, with moderate levels of bike rentals throughout the day. There are still peaks during commuting hours.
  - **Winter:** Winter shows the lowest overall bike rental activity, with less pronounced peaks during commuting hours. This is likely due to colder weather conditions and reduced outdoor activities.
- **Peak Usage Hours:** Across all seasons, there are two distinct peaks in bike rentals:
  - **Morning Peak:** Around 7-9 AM, likely corresponding to the morning commute.
  - **Evening Peak:** Around 5-7 PM, likely corresponding to the evening commute.

Bar chart showing usage pattern on New Year's Eve



Bar plot that compares New Year's usage with standard winter usage



Insights

- **Understanding bike rental trends** can help optimize bike availability in different locations and at different times.
- **Analyzing weather impact** can inform strategies for managing the bike-sharing system during specific weather conditions.
- **Identifying daily and seasonal usage patterns** can facilitate resource allocation and maintenance scheduling.