Yulu Business Case

Finding the variables that are significant in predicting the demand for shared electric cycles in the Indian market

About Yulu

Yulu is India's leading micro-mobility service provider, which offers unique vehicles for the daily commute. Starting off as a mission to eliminate traffic congestion in India, Yulu provides the safest commute solution through a user-friendly mobile app to enable shared, solo and sustainable commuting.

Yulu zones are located at all the appropriate locations (including metro stations, bus stands, office spaces, residential areas, corporate offices, etc) to make those first and last miles smooth, affordable, and convenient!

Yulu has recently suffered considerable dips in its revenues. They have contracted a consulting company to understand the factors on which the demand for these shared electric cycles depends. Specifically, they want to understand the factors affecting the demand for these shared electric cycles in the Indian market.

The company wants to know:

- Which variables are significant in predicting the demand for shared electric cycles in the Indian market?
- How well those variables describe the electric cycle demands

Dataset link:

https://d2beiqkhq929f0.cloudfront.net/public_assets/assets/000/001/428/original/bike_sharing.csv?1642089089

Colab link:

https://colab.research.google.com/drive/1cDqdii3A5q6DvA29AoNSL-3XJR2_ltTj?usp=sharing

Exploratory Data Analysis (EDA)

1. Structure & Characteristics of the Dataset

The dataset contains 10,886 records and 12 columns. The columns and their data types are as follows:

- `datetime`: object
- `season`: int64
- `holiday`: int64
- `workingday`: int64
- `weather`: int64
- `temp`: float64
- `atemp`: float64
- `humidity`: int64
- `windspeed`: float64
- `casual`: int64
- `registered`: int64
- `count`: int64

There are no missing values in the dataset.

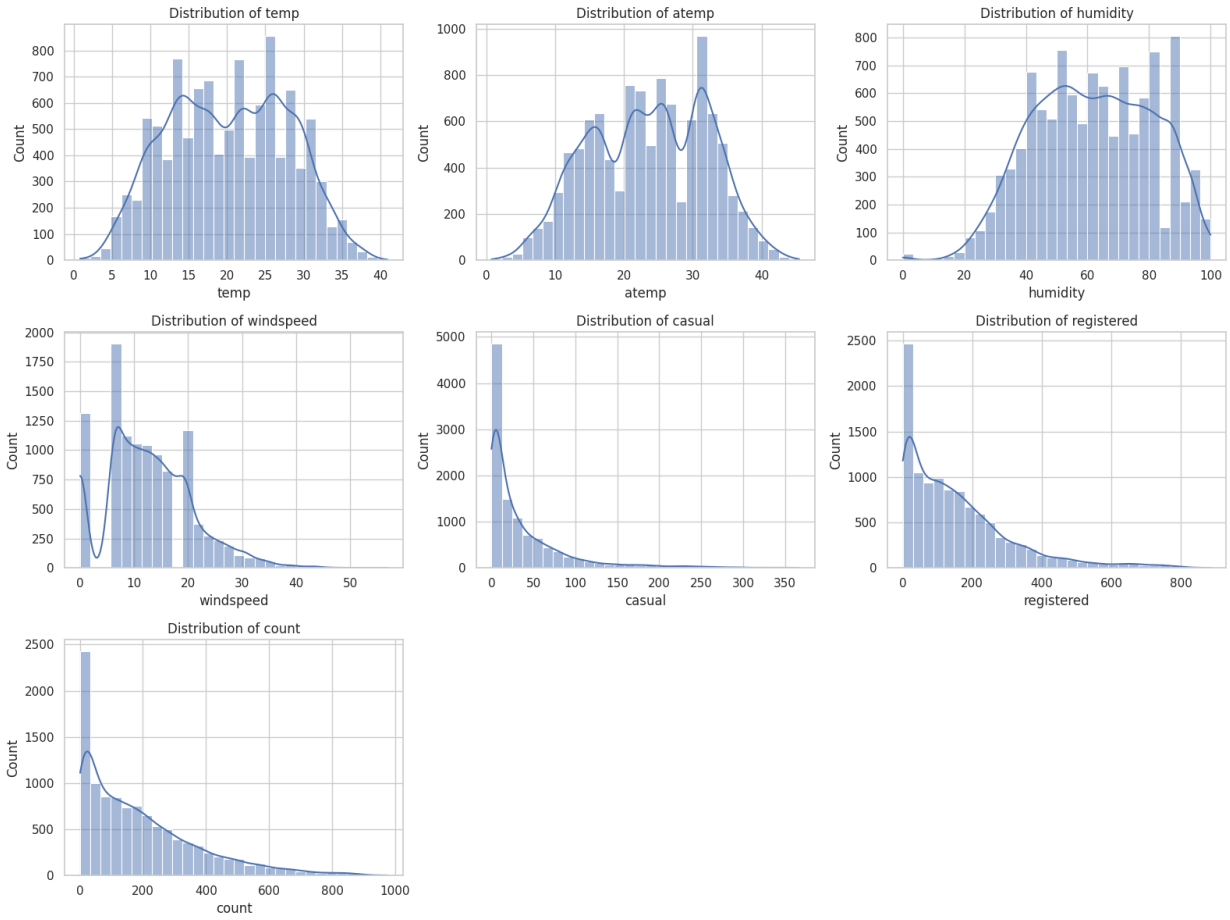
2. Data Conversion and Cleaning

- Convert `datetime` to datetime type
- Extract hour, day, month, and year from `datetime` for further analysis
- Convert categorical variables (`season`, `holiday`, `workingday`, `weather`) to the 'category' data type

Univariate Analysis

Continuous Variables

We'll create distribution plots for the continuous variables: temp, atemp, humidity, windspeed, casual, registered, and count.



Insights:

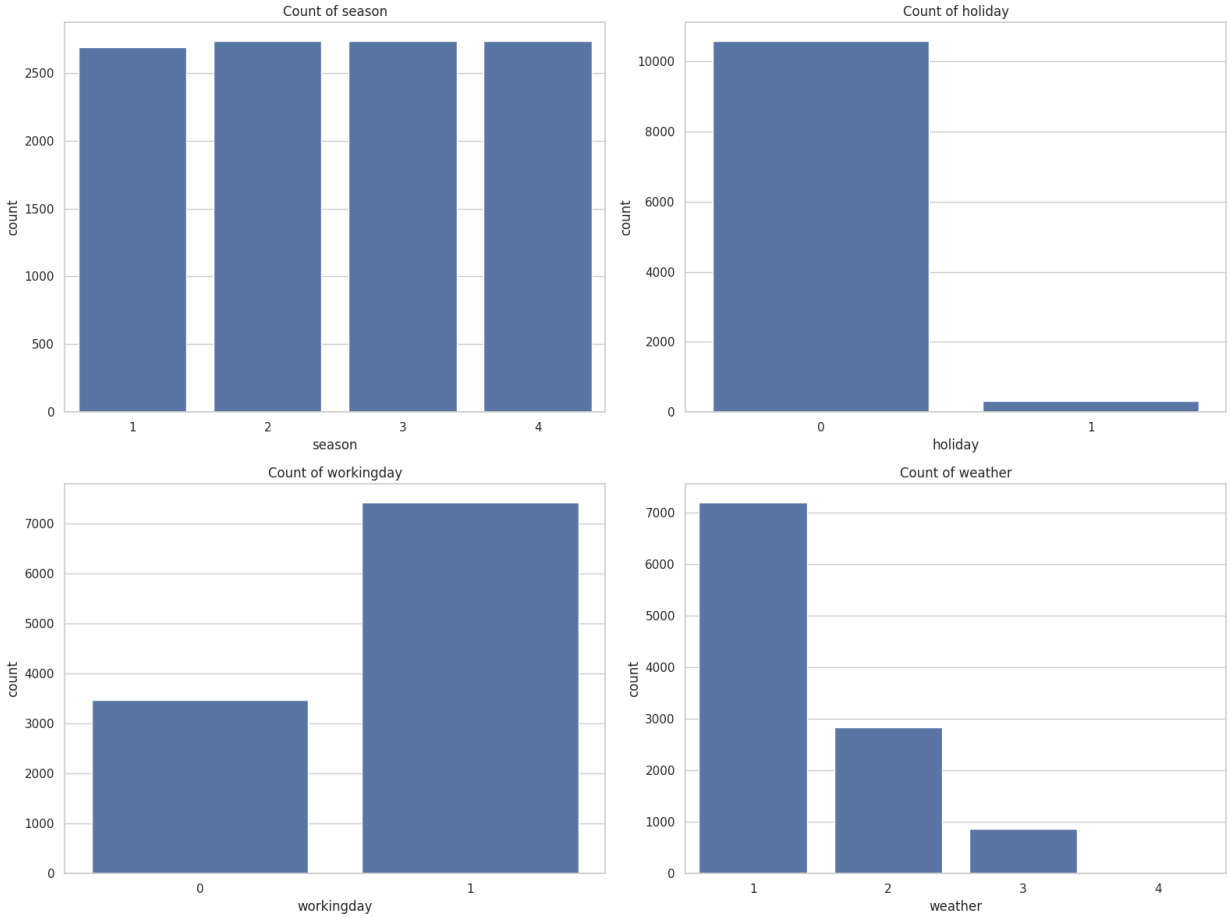
1. **Temperature (`temp`) and Apparent Temperature (`atemp`)**:
    a. Both variables exhibit a roughly normal distribution.
    b. The majority of the temperature readings are between 10°C and 30°C.
2. **Humidity**:
    a. Humidity is widely spread with peaks around 50-60% and 80-90%.
3. **Windspeed**:
    a. Wind speed is heavily right-skewed, with most values being below 20 km/h.
4. **Casual and Registered Users**:

a. The distribution of casual users is heavily skewed to the right, indicating most values are low.
b. Registered users also show a right-skew but have a wider distribution compared to casual users.

5. **Count of Total Users**:
    a. The count of total users (sum of casual and registered) is right-skewed, with most values between 0 and 200.

Categorical Variables

We'll create bar plots for the categorical variables: season, holiday, workingday, and weather.



Insights:

1. **Season**:
    a. Data is fairly evenly distributed across all four seasons.

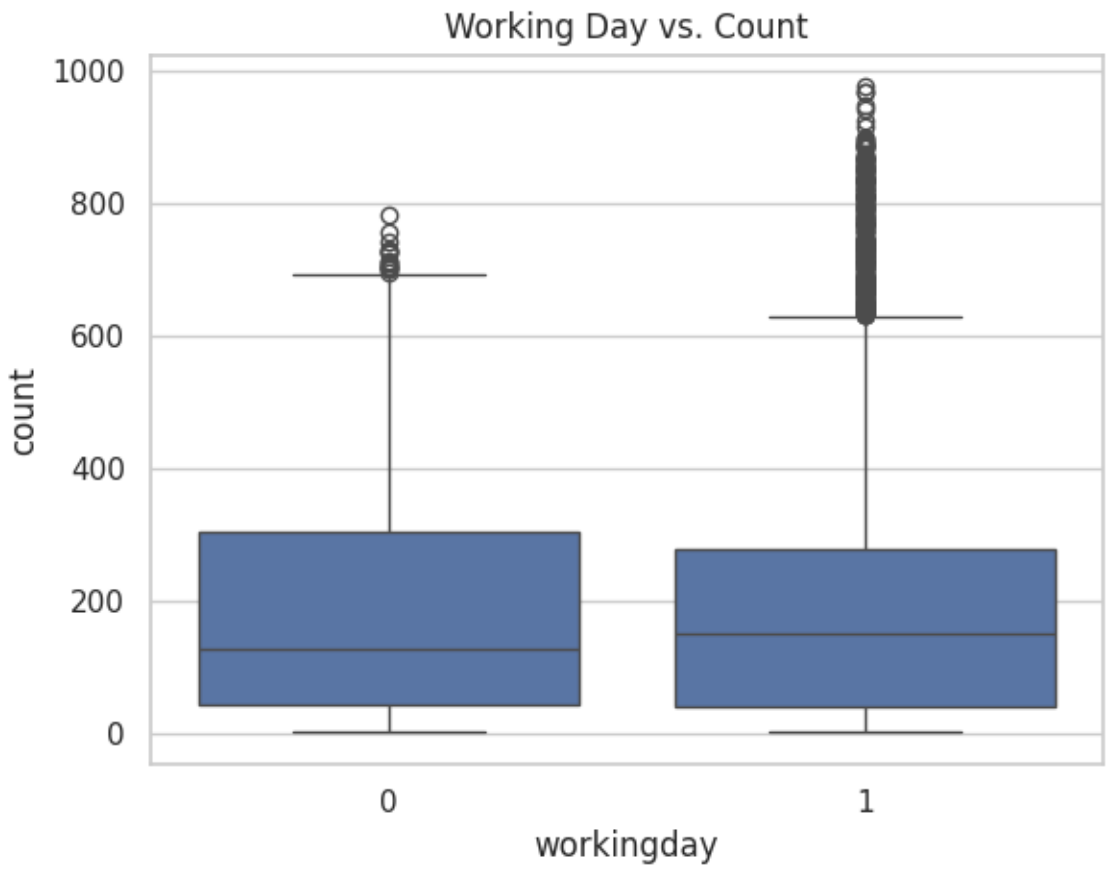
2. **Holiday**:
    a. There are significantly fewer holidays than non-holidays.
3. **Working Day**:
    a. There are more working days than non-working days.
4. **Weather**:
    a. Most of the weather conditions are either clear or slightly cloudy.

Bivariate Analysis

We'll analyze the relationships between key variables such as `workingday`, `season`, `weather`, and the `count` of rented electric cycles.

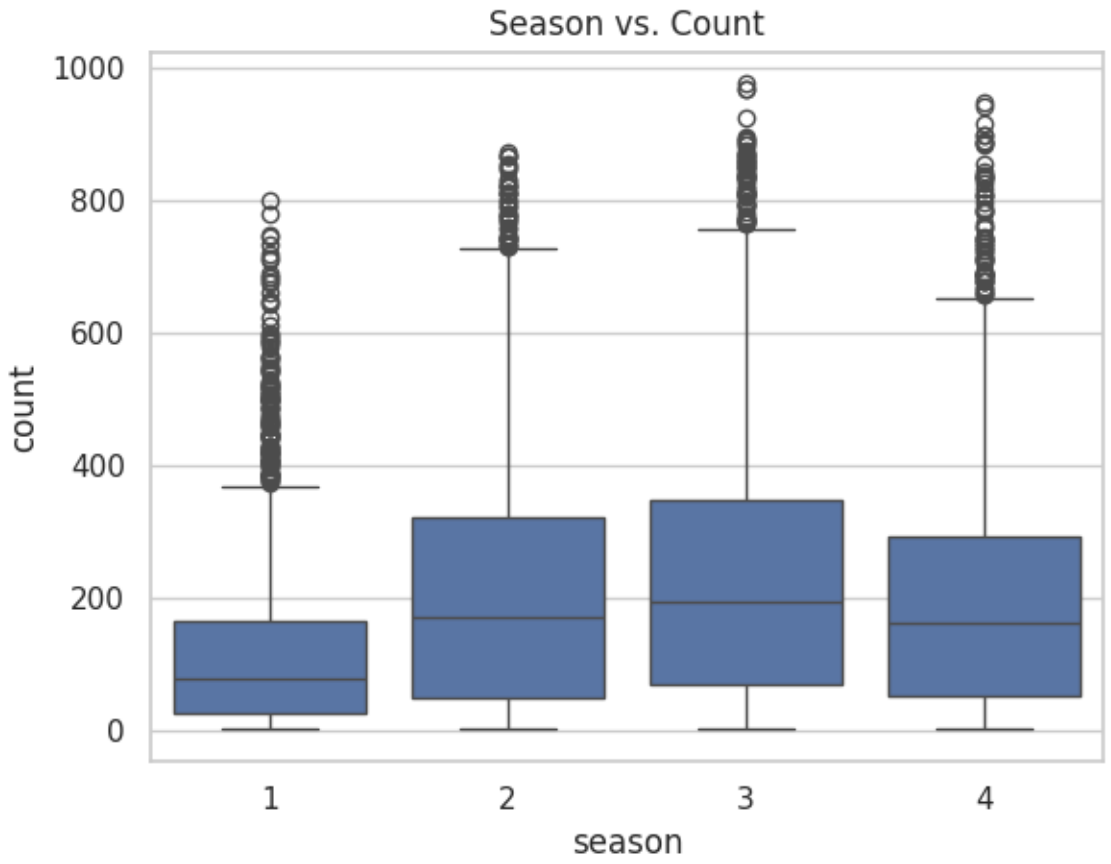
Relationships to Explore:

Working day vs. Count



- **Observation**:
    - The median count of rentals is slightly higher on non-working days compared to working days.
    - There is a similar spread in the rental counts for both working and non-working days, but with a slightly higher upper whisker for non-working days.
- **Insight**:
    - Working days have a moderate effect on the rental counts, with a slight increase in rentals on non-working days. This could be due to more people having free time to use the cycles on weekends or holidays.

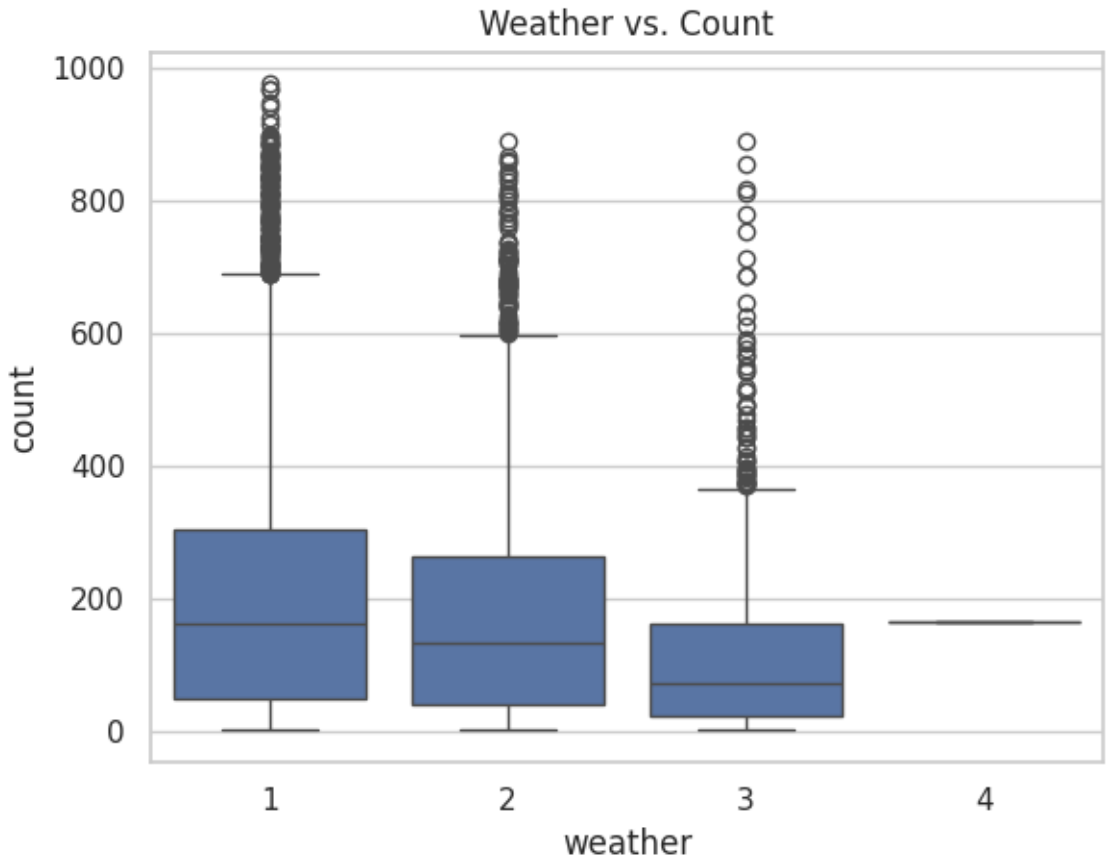
Season vs. Count



- **Observation**:
    - The median rental count is higher in seasons 2 (summer) and 3 (fall) compared to seasons 1 (spring) and 4 (winter).
    - There is more variability in rental counts during summer and fall, as indicated by the larger interquartile range.

- **Insight**:
    - The season has a significant effect on rental counts, with more rentals occurring during summer and fall. This may be due to more favorable weather conditions and longer daylight hours during these seasons.

Weather vs. Count



- **Observation**:
    - The count of rentals tends to be higher in weather conditions 1 and 2, which correspond to clear, few clouds, partly cloudy, and misty conditions.
    - The median count is lower in weather conditions 3 and 4, indicating that people rent fewer cycles during light snow, light rain, and heavy rain.
- **Insight**:
    - Weather has a significant impact on the number of rentals. Adverse weather conditions reduce the number of rentals.

These insights suggest that Yulu's demand for shared electric cycles is significantly influenced by weather and season, with a moderate impact from the distinction between working and non-working days.

Hypothesis Testing

1. Effect of Working Day on Rentals

- **Null Hypothesis (H0):** Working day has no effect on the number of electric cycles rented.
- **Alternate Hypothesis (H1):** Working day has an effect on the number of electric cycles rented.

Since the p-value (0.2264) is greater than the commonly used significance level (alpha = 0.05), we **fail to reject the null hypothesis**. This means there is not enough evidence to suggest that the number of electric cycles rented significantly differs between working days and non-working days.

2. Effect of Season on Rentals

- **Null Hypothesis (H0):** The number of cycles rented is similar across different seasons.
- **Alternate Hypothesis (H1):** The number of cycles rented differs across different seasons.

Since the p-value (6.16e-149) is much smaller than the commonly used significance level (alpha = 0.05), we **reject the null hypothesis**. This means there is strong evidence to suggest that the number of electric cycles rented differs significantly across different seasons.

3. Effect of Weather on Rentals

- **Null Hypothesis (H0):** The number of cycles rented is similar under different weather conditions.
- **Alternate Hypothesis (H1):** The number of cycles rented differs under different weather conditions.

Since the p-value (5.48e-42) is much smaller than the commonly used significance level (alpha = 0.05), we **reject the null hypothesis**. This means there is strong evidence to suggest

that the number of electric cycles rented differs significantly across different weather conditions.

4. Dependence of Weather on Season

- **Null Hypothesis (H0):** Weather is independent of the season.
- **Alternate Hypothesis (H1):** Weather is dependent on the season.

Since the p-value (1.55e-07) is much smaller than the commonly used significance level (alpha = 0.05), we **reject the null hypothesis**. This means there is strong evidence to suggest that weather conditions are not independent of the seasons.

Overall Insights:

- **Weather and Season**: Both significantly affect the demand for shared electric cycles. Yulu should consider these factors when planning and managing their fleet.
- **Working Days**: Have a moderate effect but are not statistically significant in this context.

Recommendations:

1. **Seasonal Demand Management:**
    a. **Increase Fleet During High-Demand Seasons**: As the number of rentals significantly varies across seasons, Yulu should consider increasing the availability of electric cycles during high-demand seasons (spring, summer, and fall).
    b. **Promotional Campaigns in Low-Demand Seasons**: Implement promotional campaigns or discounts during the low-demand season (winter) to attract more users.
2. **Weather-Dependent Strategies:**
    a. **Dynamic Pricing Based on Weather Conditions**: Introduce dynamic pricing where rental rates can be adjusted based on weather conditions. For instance, offer lower prices during less favorable weather to encourage usage.
    b. **Weather Forecast Integration**: Integrate weather forecasts into the Yulu app to provide users with real-time information on weather conditions and suggest optimal times for riding.
3. **Targeted Marketing and Communication:**

a. **Segmented Marketing Campaigns**: Develop targeted marketing campaigns based on user segments. For example, promote more to casual users during weekends and holidays, while focusing on registered users during working days.
b. **Seasonal Promotions**: Highlight specific promotions and benefits during different seasons to encourage consistent usage throughout the year.

4. **Infrastructure and Accessibility Improvements:**
    a. **Expand Yulu Zones**: Increase the number of Yulu zones in areas with high demand during peak seasons and favorable weather conditions, such as near parks, recreational areas, and tourist spots.
    b. **Improved Weather Protection**: Provide weather protection features, such as covered parking areas for the cycles and options for users to rent raincoats or ponchos during the rainy season.
5. **User Experience Enhancements:**
    a. **Feedback Mechanism**: Implement a robust feedback mechanism to collect user insights about their experience during different seasons and weather conditions. Use this feedback to improve service offerings.
    b. **In-App Notifications**: Send in-app notifications to users about weather changes, seasonal promotions, and best times to ride, enhancing their overall experience and encouraging usage.
6. **Data-Driven Decision Making:**
    a. **Continuous Data Monitoring**: Regularly monitor and analyze rental data to identify trends and patterns in user behavior related to weather and seasons. Use these insights to make informed decisions.
    b. **Predictive Analytics**: Utilize predictive analytics to forecast demand based on historical data, weather forecasts, and seasonal trends, ensuring optimal fleet management and resource allocation.