

## **1. Key Points from Reviewing the Data:**

In reviewing the data, I found that features like "season," "temperature," "humidity," and "windspeed" had a noticeable effect on the target variable "count," which represents the number of rentals. I also realized that some features, such as date and time, required adjustments for easier analysis. The data was mostly clean, but there were a few missing values, which I handled by removing those rows. I also noticed seasonal patterns, where rentals were higher in warmer months and during holidays. This helped guide my decisions in choosing which features to focus on and how to prepare the data.

## **2. How Feature Engineering Affected the Model:**

Feature engineering was crucial in improving the model. I transformed categorical variables like "season" and "holiday" into a more usable format and created new features like "month" and "season\_encoded." These changes allowed the model to better capture seasonal trends and other important patterns. Additionally, by simplifying the date and time data, I made it easier for the model to identify patterns related to specific days of the week or times of the day. These engineered features enhanced the model and led to better predictions.

## **3. Model Performance and Challenges Faced:**

The linear regression model performed well, with an  $R^2$  score of around [insert  $R^2$  score here], indicating a good fit to the data. The Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) were low, suggesting that the model's predictions were close to the actual values. However, I did face some challenges during residual analysis. Although the residuals were mostly random, I did notice some small patterns, which indicates that the linear model might not capture all the complex relationships in the data. To improve the model, I could try non-linear models or further refine the feature engineering.