Random Forest Review

1. Random Forest:

Categorial Features: *member_casual , season*
Contínuos Features: *TMAX*
Target Variable: *rideable_type*
NaN Values: dropped all the rows with null values

Random Forest Accuracy: 0.64
Random Forest Confusion Matrix: [[1650972   10486]
                                          [937518   11376]]

The Random Forest model achieves an accuracy of 64%, indicating moderate effectiveness in predicting the correct rideable_type. However, the confusion matrix reveals a significant imbalance in predictive performance: while true negatives are high at 1,650,972, suggesting good specificity, the model suffers from a high number of false negatives (937,518), indicating poor sensitivity or a strong bias toward the more prevalent class. This results in only 11,376 true positives, which is concerning for applications needing reliable identification of the positive class. The low false positives (10,486) suggest that the model is conservative, potentially at the cost of missing many true positive cases.

Categorial Features: *member_casual , season*
Contínuos Features: *TMAX*
Target Variable: *rideable_type*
NaN Values: Handled NaN Values.

Random Forest Accuracy: 0.53
Random Forest Confusion Matrix: [[749533 916209]
                                          [674585 1075940]]

The Random Forest model shows an accuracy of 53%, which is lower than might be desirable for robust classification tasks. The confusion matrix indicates a nearly balanced but still challenging scenario with considerable false positives

(916,209) and false negatives (674,585), alongside true positives (1,075,940) and true negatives (749,533). This performance suggests that handling null values by imputation rather than dropping them has not necessarily improved the model's predictive accuracy or the balance between sensitivity and specificity. The model's ability to correctly classify the positive class has improved in terms of true positives but at the cost of increasing false positives, highlighting a trade-off between detecting more positives and making more errors in classifying negatives.

Categorial Features: *member_casual , day_of_week*
Contínuos Features: *TMAX*
Target Variable: *rideable_type*

Random Forest Accuracy: 0.54
Random Forest Confusion Matrix: [[843729 822013]
                                                 [746061 1004464]]