SUMMARY OF CAR SHARING DEMAND PREDICTION PROJECT

1. **Data Loading and Exploration**

   o Loaded the car rental dataset using pandas.

   o Explored the data structure, checked for missing values, and summarized statistics.

2. **Data Preprocessing**

   o Converted date columns to datetime format.

   o Encoded categorical variables (e.g., season, holiday, working day).

   o Extracted new features such as day of week and month.

3. **Exploratory Data Analysis (EDA)**

   o Visualized rental patterns by hour, day of week, temperature, weather, and month.

   o Analyzed the impact of weather and holidays on rentals.

   o Plotted correlation heatmaps to understand feature relationships.

4. **Feature Selection**

   o Selected relevant features including time, weather, and calendar variables for modeling.

5. **Data Splitting and Scaling**

   o Split the data into training and testing sets.

   o Scaled features where necessary (especially for KNN).

6. **Model Training and Evaluation**

   o Trained five regression models: Linear Regression, Random Forest, Gradient Boosting, XGBoost, and KNN.

   o Evaluated models using metrics: RMSE, MAE, MAPE, and R^2.

   o Analyzed feature importance and residuals for each model.

7. **Model Comparison**

   o Compared all models in a summary table.

   o Identified XGBoost and Random Forest as the best performers.

8. **Conclusion**

   o Advanced ensemble models (XGBoost, Random Forest) provide the most accurate predictions for car rental demand.