1. From your analysis of the categorical variables from the dataset, what could you infer about their effect on the dependent variable?

There is not much effect from the categorical variables

2. Why is it important to use drop\_first=True during dummy variable creation?

to avoid a statistical issue called multicollinearity when creating dummy variables for categorical features in machine learning.

3. Looking at the pair-plot among the numerical variables, which one has the highest correlation with the target variable?

WorkingDay has the highest correlation

4. How did you validate the assumptions of Linear Regression after building the model on the training set?

By Checking the VIF values i was able to validate the assumptions

5. Based on the final model, which are the top 3 features contributing significantly towards explaining the demand of the shared bikes?

Based on the model results weathersit, weekday and windspeed are the top 3 features