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

Categorical variables are : season, weekday, workingday,
Depended variables are: weathersit, temp, humidity,
windspeed

2. Why is it important to use drop_first=True during dummy variable creation?
drop_first=True is important to use, as it helps in reducing the extra column created during dummy variable creation. Hence it reduces the correlations created among dummy variables.

3. Looking at the pair-plot among the numerical variables, which one has the highestcorrelation with the target variable?
Temp and atemp

4. How did you validate the assumptions of Linear Regression after building the modelon the training set?
Here we have calculated r2_score.

```
from sklearn.metrics import r2_score
r2_score(y_test, y_test_pred)
```

5. Based on the final model, which are the top 3 features contributing significantlytowards explaining the demand of the shared bikes?
Here are the TOP 3 Features with highest VIF

- Spring season : 1.66
- yr : 1.50
- Mist : 1.38