**Assignment-based Subjective Questions**

1. **From your analysis of the categorical variables from the dataset, what could you infer about their effect on the dependent variable?**  
   **Answer**: The demand of bike is less in the month of spring when compared with other seasons. The demand bike increased in the year 2019 when compared with year 2018.
2. **Why is it important to use drop first=True during dummy variable creation?  
   Answer:** 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 highest correlation with the target variable?  
   Answer:** The numerical variable 'registered' has the highest correlation with the target variable 'cnt', if we consider all the features.
4. **How did you validate the assumptions of Linear Regression after building the model on the training set?  
   Answer:** By doing residual analysis, comparing actual vs predicted plots.
5. **Based on the final model, which are the top 3 features contributing significantly towards explaining the demand of the shared bikes?  
   Answer:** holiday, temp, hum