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

The categorical variables were season, mnth, holiday, weathersit, yr, workingday and weekday. I visualized these using a boxplot and found the following effect on dependent variable: The boxplot for season showed that 1 had least value for cnt while 3 had max value. The weather 4 seemed to be the least favorable while highest count was in 1 and 2. Reduction was seen during Holiday. Mnth 9 had the highest number of rentals.

2. Why is it important to use drop_first=True during dummy variable creation?
This is done to avoid the dummy variables being correlated

1. Explain the linear regression algorithm in detail
Regression is the most commonly used prediction analysis model. Linear Regression is a supervised ML algorithm. It is based on the equation y = mx + c
A linear relationship is assumed between the dependent and independent variables. In regression, a best fit line is calculated which describes the relationship between these.
The dependent variable should be of continuous data type.
Regression can be divided into simple and multiple linear regression.
Simple linear regression: The dependent variable is predicted using only one independent variable.
Multiple Linear Regression: The dependent variable is predicted using multiple independent variables.

2. Explain the Anscombe’s quartet in detail.
Anscombe's quartet comprises four data sets that have very simiar simple descriptive statistics, yet look very different when plotted on a graph.

3. What is Pearson’s R?
Pearson’s R or Pearson’s correlation coefficient is the summary of the strength of the linear association between the variables. It should be between -1 and +1.

4. What is scaling? Why is scaling performed? What is the difference between normalized scaling and standardized scaling?
It is a method to standardize the range of independent variables of the data. It is performed during the data processing of the data. If it is not done, the algorithm tends to weigh greater values as higher and consider the smaller values as lower.