PART III GENERAL QUESTION

1. What are the assumptions of linear regression?
Linear regression assumes a linear relationship between the input and output variables. It also assumes that the residuals (errors) are normally distributed, have constant variance , the data points are independent, and that the input features are not highly correlated.

2. When should you use logistic regression instead of linear regression?
Logistic regression should be used when the target variable is categorical, especially for binary classification (like 0 or 1, yes or no). Linear regression is not suitable for classification because it predicts continuous values, not probabilities.

3. What is the interpretation of coefficients in logistic regression?
Each coefficient in logistic regression shows how a one-unit change in a feature affects the log-odds of the positive outcome. A positive coefficient increases the probability of the positive class, while a negative one decreases it.

4. What is the difference between sigmoid and softmax functions?
The sigmoid function is used in binary classification and gives a probability between 0 and 1. The softmax function is used for multiclass classification and gives a probability distribution across all possible classes.

5. Why is R-squared not suitable for evaluating logistic regression models?
R-squared measures how well a model explains variance in a continuous outcome, which works for linear regression. Logistic regression predicts categories, so evaluation metrics like accuracy, precision, recall, and ROC-AUC are more appropriate.