1. In a linear equation, what is the difference between a dependent variable and an independent variable?

>>>The dependent variable (also called the outcome variable) is the variable being predicted or explained in a regression model.

The independent variable (also called the predictor variable) is the variable used to explain or predict the values of the dependent variable.

2. What is the concept of simple linear regression? Give a specific example.

>>>Simple linear regression is a statistical method used to model the relationship between a single independent variable and a dependent variable.

3. In a linear regression, define the slope.

>>>In linear regression, the slope represents the change in the dependent variable for a unit change in the independent variable.

It indicates how much the dependent variable is expected to change for each one-unit change in the independent variable.

4. Determine the graph's slope, where the lower point on the line is represented as (3, 2) and the higher point is represented as (2, 2).

>>>The slope is calculated as (change in y) / (change in x).

Given points (3, 2) and (2, 2), the slope = (2 - 2) / (3 - 2) = 0.

5. In linear regression, what are the conditions for a positive slope?

>>>A positive slope indicates that as the independent variable increases, the dependent variable also tends to increase.

Positive correlation between variables.

6. In linear regression, what are the conditions for a negative slope?

>>>A negative slope indicates that as the independent variable increases, the dependent variable tends to decrease.

Negative correlation between variables.

7. What is multiple linear regression and how does it work?

>>>Multiple linear regression involves predicting a dependent variable using two or more independent variables.

It extends the concept of simple linear regression to cases where multiple factors contribute to the outcome.

Sum of Squares Due to Error (SSE) in Multiple Linear Regression:

SSE represents the sum of squared differences between the observed values and the predicted values in a multiple linear regression model.

Sum of Squares Due to Regression (SSR) in Multiple Linear Regression:

SSR represents the sum of squared differences between the predicted values and the mean of the dependent variable.

Multicollinearity:

Multicollinearity occurs when two or more independent variables in a regression model are highly correlated, making it difficult to distinguish their individual effects.

11. What is heteroskedasticity, and what does it mean?

>>>Heteroskedasticity refers to the presence of non-constant variance in the residuals of a regression model.

12. Describe the concept of ridge regression.

>>>Ridge regression is a regularization technique used to prevent overfitting in regression models.

13. Describe the concept of lasso regression.

>>>Lasso (Least Absolute Shrinkage and Selection Operator) regression is another regularization technique that helps prevent overfitting.

14. What is polynomial regression and how does it work?

>>>Polynomial regression involves fitting a polynomial equation to the data instead of a straight line.

15. Describe the basis function.

>>>A basis function is a mathematical function used in polynomial regression to transform the original input features into a higher-dimensional space.

This allows capturing more complex relationships between variables.

16. Describe how logistic regression works.

>>>Logistic regression is used to model the probability of a binary outcome (1 or 0) based on one or more predictor variables.