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

variable?

A linear equation in two variables can be described as a linear relationship between x and y, that is, two variables in which the value of one of them (usually y) depends on the value of the other one (usually x). In this case, x is the independent variable, and y depends on it, so y is called the dependent variable

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

We could use the equation to predict weight if we knew an individual's height. In this example, if an individual was 70 inches tall, we would predict his weight to be: Weight = 80 + 2 x (70) = 220 lbs. In this simple linear regression, we are examining the impact of one independent variable on the outcome

3. In a linear regression, define the slope.

In a regression line passing through a set of data points in data sets Argument1 and Argument2, the slope is the vertical distance divided by the horizontal distance between any two points on the line. This ratio is also known as the rate of change along the line.

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

In summary, if the slope is positive, y increases as x increases, and the function runs "uphill" (going left to right). If the slope is negative, y decreases as x increases and the function runs downhill. If the slope is zero, y does not change, thus is constant—a horizontal line.

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

If the slope is negative, y decreases as x increases and the function runs downhill. If the slope is zero, y does not change, thus is constant—a horizontal line

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

Multiple linear regression refers to a statistical technique that uses two or more independent variables to predict the outcome of a dependent variable. The technique enables analysts to determine the variation of the model and the relative contribution of each independent variable in the total variance

8. In multiple linear regression, define the number of squares due to error.

Sum of squares (SS) is a statistical tool that is used to identify the dispersion of data as well as how well the data can fit the model in regression analysis. The sum of squares got its name because it is calculated by finding the sum of the squared differences.

9. In multiple linear regression, define the number of squares due to regression.

Regression sum of squares (also known as the sum of squares due to regression or explained sum of squares) The regression sum of squares describes how well a regression model represents the modeled data. A higher regression sum of squares indicates that the model does not fit the data well

10. In a regression equation, what is multicollinearity?

Multicollinearity exists whenever an independent variable is highly correlated with one or more of the other independent variables in a multiple regression equation. Multicollinearity is a problem because it will make the statistical inferences less reliable

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

In statistics, heteroskedasticity (or heteroscedasticity) happens when the standard deviations of a predicted variable, monitored over different values of an independent variable or as related to prior time periods, are non-constant

12. Describe the concept of ridge regression.

Ridge regression is a model tuning method that is used to analyse any data that suffers from multicollinearity. This method performs L2 regularization. When the issue of multicollinearity occurs, least-squares are unbiased, and variances are large, this results in predicted values being far away from the actual values.

13. Describe the concept of lasso regression.

Lasso regression algorithm is defined as a regularization algorithm that assists in the elimination of irrelevant parameters, thus helping in the concentration of selection and regularizes the models. Lasso models can be evaluated using various metrics such as RMSE and R-Square.

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

A polynomial regression model is a machine learning model that can capture non-linear relationships between variables by fitting a non-linear regression line, which may not be possible with simple linear regression. It is used when linear regression models may not adequately capture the complexity of the relationship

15. Describe the basis function.

In mathematics, a basis function is an element of a particular basis for a function space. Every function in the function space can be represented as a linear combination of basis functions, just as every vector in a vector space can be represented as a linear combination of basis vectors.

16. Describe how logistic regression works.

Logistic regression is a data analysis technique that uses mathematics to find the relationships between two data factors. It then uses this relationship to predict the value of one of those factors based on the other. The prediction usually has a finite number of outcomes, like yes or no.