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

Solution - 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.**

Solution - Simple linear regression is a regression model that estimates the relationship between one independent variable and one dependent variable using a straight line. Both variables should be quantitative.

Total number of sales, Agricultural scientists use linear regression to estimate the effect of fertilizer on the total crops yielded, the effect of drug dosage on blood pressure.

3. **In a linear regression, define the slope.**

Solution - 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.

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).**

Solution – y=2 : Slope = infinity

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

Solution - if the slope is positive, y increases as x increases, and the function runs "uphill"

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

Solution - If the slope is negative, y decreases as x increases and the function runs downhill.

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

Solution - 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.**

Solution - The Sum of Squared Error is the difference between the observed value and the predicted value.

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

Solution - The regression sum of squares is used to denote the relationship between the modeled data and a regression model. A regression model establishes whether there is a relationship between one or multiple variables. Having a low regression sum of squares indicates a better fit with the data.

In a regression equation, what is multicollinearity?

Solution - Multicollinearity occurs when two or more independent variables in a data frame have a high correlation with one another in a regression model. This means that one independent variable can be predicted from another in a regression model.

10. **What is heteroskedasticity, and what does it mean?**

Solution - Heteroskedastic refers to a condition in which the variance of the residual term, or error term, in a regression model varies widely.

11. **Describe the concept of ridge regression.**

Solution - Ridge regression is a method of estimating the coefficients of multiple-regression models in scenarios where the independent variables are highly correlated

12. **Describe the concept of lasso regression.**

Solution - In statistics and machine learning, lasso (least absolute shrinkage and selection operator; also Lasso or LASSO) is a regression analysis method that performs both variable selection and regularization to enhance the prediction accuracy and interpretability of the resulting statistical model.

13. **What is polynomial regression and how does it work?**

Solution - 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.

The main steps involved in Polynomial Regression are given below:

- Data Pre-processing
- Build a Linear Regression model and fit it to the dataset
- Build a Polynomial Regression model and fit it to the dataset
- Visualize the result for Linear Regression and Polynomial Regression model.
- Predicting the output.

14. **Describe the basis function.**

Solution - Basis functions allow modelling non linearity in the data while keeping linearity in parameters, which greatly simplifies the analysis of these models. Using linear combination of different basis function, we can construct complex functions and still use linear regression.

15. **Describe how logistic regression works.**

Solution - In Logistic regression, instead of fitting a regression line, we fit an "S" shaped logistic function, which predicts two maximum values (0 or 1).The curve from the logistic function indicates the likelihood of something such as whether the cells are cancerous or not, a mouse is obese or not based on its weight, etc.