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

An independent variable is the variable that is changed or controlled in a scientific experiment to test the effects on the dependent variable. A dependent variable is the variable being tested and measured in a scientific experiment. The dependent variable is 'dependent' on the independent variable

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

Simple linear regression is used to estimate the relationship between two quantitative variables. You can use simple linear regression when you want to know: How strong the relationship is between two variables (e.g., the relationship between rainfall and soil erosion).

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

The slope represents the change in the dependent variable for each unit change in the independent variable, while the intercept represents the predicted value of the dependent variable when the independent variable is zero.

**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).**

**(x1,y1) = (3,2)**

**(x2, y2) = (2,2)**

**Slope = (y2-y1)/(x2-x1)**

**= (2-2)/(2-3)**

**=(0)/-1**

**=0**

**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).

**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.

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

Multiple linear regression (MLR), also known simply as multiple regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. Multiple regression is an extension of linear (OLS) regression that uses just one explanatory variable.

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

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.

**In a regression equation, what is multicollinearity?**

Multicollinearity is the occurrence of high intercorrelations among two or more independent variables in a multiple regression model.

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

In regression analysis, heteroskedasticity results in an unequal scatter of the residuals, also known as the error term. This means that the observed values deviate from the predicted values ununiformly.

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

Ridge regression is a method of estimating the coefficients of multiple-regression models in scenarios where the independent variables are highly correlated. It has been used in many fields including econometrics, chemistry, and engineering.  
Ridge regression is almost identical to linear regression (sum of squares) except we introduce a small amount of bias. In return, we get a significant drop in variance. In other words, by starting with a slightly worse fit, Ridge Regression can provide better long term predictions

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

n 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 in order to enhance the prediction accuracy and interpretability of the resulting statistical model.

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

Polynomial regression is a form of Linear regression where only due to the Non-linear relationship between dependent and independent variables, we add some polynomial terms to linear regression to convert it into Polynomial regression.

In polynomial regression, the relationship between the dependent variable and the independent variable is modeled as an nth-degree polynomial function. When the polynomial is of degree 2, it is called a quadratic model; when the degree of a polynomial is 3, it is called a cubic model, and so on.

The degree of order which to use is a Hyperparameter, and we need to choose it wisely. But using a high degree of polynomial tries to overfit the data, and for smaller values of degree, the model tries to underfit, so we need to find the optimum value of a degree. Polynomial Regression models are usually fitted with the method of least squares. The least square method minimizes the variance of the coefficients under the Gauss-Markov Theorem.

**15. Describe the basis function.**

Basis functions allow modeling 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.

**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.