Regression Modelling

1) **What is linear Regression?**
   Linear Regression is a linear approach to modelling the relationship between a scalar response and one or more explanatory variables. In the case of one explanatory variable is called simple linear regression for more than one explanatory variables the process is called multiple linear Regression.
2) **What is Simple Linear Regression?**
   A statistical Technique that uses a straight-line relationship to predict a numerical dependent variable Y from a single numerical independent Variable X.
3) **What is Y-intercept?**
   Y-intercept is a straight line where it cross the y-axis of a graph.
4) **What is coefficient?**
   A number used to multiply a variable.
5) **What is the Least-Square Method?**
   The simple linear Regression method that seeks to minimize the sum of the squared differences between the actual values of the dependent variable Y and the predicted values of Y.
6) **What is the Measures of Variation?**
   After a regression model has been fit to a set of data, measures of variation determine how much of the variation in the dependent variable Y can be explained by variation in the independent variable X.
7) **What is SSR?**
   SSR is the sum of squared deviation of predicated values from the mean values.
8) **What is SSE?**
   SSE is the sum of squared deviation of actual values from predicted values.
9) **What is R-square?**
   R-square value designates the total proportion of variance in the dependent variable explained by the independent variable. It is a value between 0 and 1. The value toward 1 indicates a better model fit.
10) **What is adjusted R-squared?**
    Simple R-squared value will keep increasing with addition of independent variable, to fix this issue adjusted R-squared is considered for multivariate regression to understand the explanatory power of the independent variable.
11) **What is The Coefficient of Determination?**
    The ratio of the regression sum of squares to the total sum of squares, represented by the symbol r^2
12) **What is the Coefficient of Correlation?**
    The measure of the strength of the linear relationship between two variables, represented by r.
13) **What is the Standard Error of Estimate?**
    The standard deviation around the fitted line of regression that measures the variability of the actual Y values from the predicted Y, represented by the symbol S_{yx}.

14) **What are the Residuals?**

The difference between the observed and predicted values of the dependent variable Y for a given value of X.

15) **What is RMSE?**

RMSE indicates how close the predicted values are to the actual values; hence the lower RMSE value signifies that the model performance is good.

16) **What is Mean absolute Error?**

This is the mean or Average of absolute value of the errors that is the predicted – Actual.

17) **What is the standard Error or Residuals?**

This is the average observed values from the regression line. The smaller value shows that the models fitting is good.

18) **What is confidence interval?**

Confidence interval is a range of value we are fairly sure our true value lies in.

19) **What is t-value?**

The t-value measures the size of the difference relative to the variation in your sample data. T is simply the calculated difference represented in units of standard errors. The greater the magnitude of t, the greater the evidence against the null hypothesis.

20) **What is P-value?**

The probability of obtaining test statistic equal to or more extreme than the result obtained from the sample data, given that the null hypothesis is true.