1. Using a graph to illustrate slope and intercept, define basic linear regression.

Answer:-->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.The slope indicates the steepness of a line and the intercept indicates the location where it intersects an axis. The slope and the intercept define the linear relationship between two variables, and can be used to estimate an average rate of change.Linear regression is used to model the relationship between two variables and estimate the value of a response by using a line-of-best-fit. This calculator is built for simple linear regression, where only one predictor variable (X) and one response (Y) are used.

2. In a graph, explain the terms rise, run, and slope.

Answer:--->The vertical change between two points is called the rise, and the horizontal change is called the run. The slope equals the rise divided by the run: Slope =riserun Slope = rise run . You can determine the slope of a line from its graph by looking at the rise and run.The slope of a line measures the steepness of the line. Most of you are probably familiar with associating slope with "rise over run". Rise means how many units you move up or down from point to point. On the graph that would be a change in the y values. Run means how far left or right you move from point to point.The slope of a line is a measure of its steepness. Mathematically, slope is calculated as "rise over run" (change in y divided by change in x).

3. Use a graph to demonstrate slope, linear positive slope, and linear negative slope, as well as the different conditions that contribute to the slope.

Answer:--->If a line has a positive slope (i.e. m > 0), then y always increases when x increases and y always decreases when x decreases. Thus, the graph of the line starts at the bottom left and goes towards the top right.So, a positive slope means the function is increasing while a negative slope means the function is decreasing. Increasing functions have y-values that increase as x-values increase, whereas decreasing functions have y-values that decrease while x-values increase.Graphically, a negative slope means that as the line on the line graph moves from left to right, the line falls. We will learn that “price” and “quantity demanded” have a negative relationship; that is, consumers will purchase less when the price is higher

4. Use a graph to demonstrate curve linear negative slope and curve linear positive slope.

Answer:--->In order to graph a negative slope, it is important to know that a negative number in the ratio will mean to move down along the y-axis or to the left along the x-axis. A positive number in the ratio will mean to move up along the y-axis or to the right along the x-axis.Straight lines that are downward sloping have negative slopes; curves that are downward sloping also have negative slopes. We know, of course, that the slope changes from point to point on a curve, but all of the slopes along these two curves will be negative.A positive slope means y increases as x increases (visually, the line moves up as you go from left to right). A negative slope means y decreases as x increases (visually, the line moves down as you go from left to right). A zero slope means that y is constant and does not change as x changes.

5. Use a graph to show the maximum and low points of curves.

Answer:--->To find the maximum/minimum of a curve you must first differentiate the function and then equate it to zero. This gives you one coordinate. To find the other you must resubstitute the one already found into the original function.If you have the graph, or can draw the graph,

the maximum is just the y value at the vertex of the graph. If you are unable to draw a graph, there are formulas you can use to find the maximum. If you are given the formula y = ax2 + bx + c, then you can find the maximum value using the formula max = c - (b2 / 4a)A global maximum point refers to the point with the largest -value on the graph of a function when a largest -value exists. A global minimum point refers to the point with the smallest -value

6. Use the formulas for a and b to explain ordinary least squares.

Answer:--->Least Squares Regression Formula

a = y-intercept. b = slope of the line.A linear regression line has an equation of the form Y = a + bX, where X is the explanatory variable and Y is the dependent variable. The slope of the line is b, and a is the intercept (the value of y when x = 0).

7. Provide a step-by-step explanation of the OLS algorithm.

Answer:---->The ordinary least squares (OLS) method is a linear regression technique that is used to estimate the unknown parameters in a model. The method relies on minimizing the sum of squared residuals between the actual and predicted values.Ordinary Least Square Method Set a difference between dependent variable and its estimation: Square the difference: Take summation for all data. To get the parameters that make the sum of square difference become minimum, take partial derivative for each parameter and equate it with zero,Step 1: For each (x,y) calculate x2 and xy: Step 2: Sum x, y, x2 and xy (gives us Σx, Σy, Σx2 and Σxy): Step 3: Calculate Slope m: m = N Σ(xy) − Σx Σy N Σ(x2) − (Σx)2 Step 4: Calculate Intercept b: b = Σy − m Σx N. Step 5: Assemble the equation of a line:

8. What is the regression's standard error? To represent the same, make a graph.

Answer:--->The standard error of the regression (S), also known as the standard error of the estimate, represents the average distance that the observed values fall from the regression line. Conveniently, it tells you how wrong the regression model is on average using the units of the response variablethe standard error is the square root of the sum of the squared residuals divided by (n - 2).

9. Provide an example of multiple linear regression.

Answer:--->Multiple Linear Regression is one of the important regression algorithms which models the linear relationship between a single dependent continuous variable and more than one independent variable. Example: Prediction of CO2 emission based on engine size and number of cylinders in a car. For example, scientists might use different amounts of fertilizer and water on different fields and see how it affects crop yield. They might fit a multiple linear regression model using fertilizer and water as the predictor variables and crop yield as the response variable.There are two types of multiple linear regression: ordinary least squares (OLS) and generalized least squares (GLS). The main difference between the two is that OLS assumes there is not a strong correlation between any two independent variables.

10. Describe the regression analysis assumptions and the BLUE principle.

Answer:---->BLUE is an acronym for the following: Best Linear Unbiased Estimator. In this context, the definition of “best” refers to the minimum variance or the narrowest sampling distribution.Under 1 - 6 (the classical linear model assumptions) OLS is BLUE (best linear unbiased estimator), best in the sense of lowest variance. ... Under 1 - 5 (the Gauss-Markov assumptions) OLS is BLUE and efficient (as described above). Under 1 - 4, OLS is unbiased, and consistent.Find the linear estimator that is unbiased and has minimum variance. This leads to Best Linear Unbiased Estimator (BLUE)04

11. Describe two major issues with regression analysis.

Answer:--->Heteroskedasticity: variance of error term is not constant. F-test is unreliable. Standard error underestimated.Nonconstant variance and weighted least squares. Autocorrelation and time series methods. Multicollinearity, which exists when two or more of the predictors in a regression model are moderately or highly correlated with one another. Overfitting. Excluding important predictor variables.Linear regression and logistic regression are two types of regression analysis techniques that are used to solve the regression problem using machine learning. They are the most prominent techniques of regression.Since linear regression assumes a linear relationship between the input and output varaibles, it fails to fit complex datasets properly. In most real life scenarios the relationship between the variables of the dataset isn't linear and hence a straight line doesn't fit the data properly.

12. How can the linear regression model's accuracy be improved?

Answer:--->How to improve the accuracy of a Regression Model Handling Null/Missing Values.

Data Visualization.

Feature Selection and Scaling.

3A. Feature Engineering.

3B. Feature Transformation.

Use of Ensemble and Boosting Algorithms.

Hyperparameter Tuning.

13. Using an example, describe the polynomial regression model in detail.

Answer:--->Polynomial regression is one of the machine learning algorithms used for making predictions. For example, it is widely applied to predict the spread rate of COVID-19 and other infectious diseases Polynomial regression is a form of regression analysis in which the relationship between the independent variable x and the dependent variable y is modelled as an nth degree polynomial in x.

14. Provide a detailed explanation of logistic regression.

Answer:----->Logistic regression is a statistical analysis method to predict a binary outcome, such as yes or no, based on prior observations of a data set. A logistic regression model predicts a dependent data variable by analyzing the relationship between one or more existing independent variables.Logistic regression is a statistical method used to predict the outcome of a dependent variable based on previous observations. It's a type of regression analysis and is a commonly used algorithm for solving binary classification problems.Logistic regression is a Machine Learning classification algorithm that is used to predict the probability of certain classes based on some dependent variables. In short, the logistic regression model computes a sum of the input features (in most cases, there is a bias term), and calculates the logistic of the result.

15. What are the logistic regression assumptions?

Answer:---->Basic assumptions that must be met for logistic regression include independence of errors, linearity in the logit for continuous variables, absence of multicollinearity, and lack of strongly influential outliers.There should be a linear and additive relationship between dependent (response) variable and independent (predictor) variable(s). ... There should be no correlation between the residual (error) terms. ... The independent variables should not be correlated. ... The error terms must have constant variance. Assumption 1: Linear Relationship. Assumption 2: Independence. Assumption 3: Homoscedasticity. Assumption 4: Normality.

16. Go through the details of maximum likelihood estimation.

Answer:--->In statistics, maximum likelihood estimation (MLE) is a method of estimating the parameters of an assumed probability distribution, given some observed data. This is achieved by maximizing a likelihood function so that, under the assumed statistical model, the observed data is most probable.MLE works by calculating the probability of occurrence for each data point (we call this the likelihood) for a model with a given set of parameters. These probabilities are summed for all the data points. We then use an optimizer to change the parameters of the model in order to maximise the sum of the probabilities.