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

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. The greater the magnitude of the slope, the steeper the line and the greater the rate of change.

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

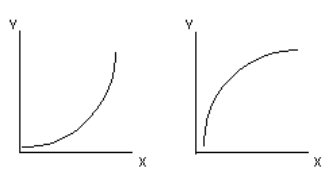
The vertical change between two points is called the rise, and the horizontal change is called the run. The slope equals the rise divided by the run: Slope =riserun Slope = rise run . You can determine the slope of a line from its graph by looking at the rise and run.

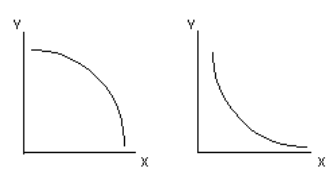
3. Use a graph to demonstrate slope, linear positive slope, and linear negative slope, as well as the different conditions that contribute to the slope.

As with upward sloping straight lines, we can say that generally the slope of the curve is positive. While the slope will differ at each point on the curve, it will always be positive.

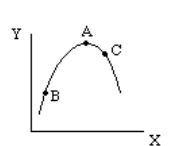
Straight lines that are downward sloping have negative slopes; curves that are downward sloping also have negative slopes.

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





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



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

The least square method is the process of finding the best-fitting curve or line of best fit for a set of data points by reducing the sum of the squares of the offsets (residual part) of the points from the curve. During the process of finding the relation between two variables, the trend of outcomes are estimated quantitatively. This process is termed as regression analysis. The method of curve fitting is an approach to regression analysis. This method of fitting equations which approximates the curves to given raw data is the least squares.

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

The simple linear regression is a model with a single regressor (independent variable) x that has a relationship with a response (dependent or target) y that is a

y = β0 + β1 x + ε — — — — — — — — — — (1)

Where β0: intercept

β1: slope (unknown constant)

ε: random error component

This is a line where y is the dependent variable we want to predict, x is the independent variable, and β0 and β1 are the coefficients that we need to estimate.

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

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.

9. Provide an example of multiple linear regression.

Prediction of CO2 emission based on engine size and number of cylinders in a car.

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

Linearity: The relationship between X and the mean of Y is linear. Homoscedasticity: The variance of residual is the same for any value of X. Independence: Observations are independent of each other. Normality: For any fixed value of X, Y is normally distributed.

The Gauss Markov theorem says that, under certain conditions, the ordinary least squares (OLS) estimator of the coefficients of a linear regression model is the best linear unbiased estimator (BLUE), that is, the estimator that has the smallest variance among those that are unbiased and linear in the observed output

11. Describe two major issues with regression analysis.

* Non-Linearity of the response-predictor relationships
* Correlation of error terms

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

Handling Null/Missing Values.

Data Visualization.

Feature Selection and Scaling.

3A. Feature Engineering.

3B. Feature Transformation.

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

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.

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.

15. What are the logistic regression assumptions?

Assumption #1: The Response Variable is Binary.

Assumption #2: The Observations are Independent.

Assumption #3: There is No Multicollinearity Among Explanatory Variables.

16. Go through the details of maximum likelihood estimation.

maximum likelihood estimation (MLE) is a method of estimating the parameters of an assumed probability distribution, given some observed data