Q1. What is y = f(x) equation called

It is called mapping function. It maps input variables to output variable

Q2. What should the be relation between input and output variable to start with regression

x and Y should be linearly associated

Q3. What are Y and X called as in the y = mx+c

Y- is called output variable

X- is called input variable

Q4. How to evaluate relation between two quantitative variables

Using scatter diagram

Q5. Can we plot gender and salary of people in scatter diagram?

No

Q6. How to quantify the strength of the association between variables

Using Correlation coefficient

Q7. What is the range of correlation coefficient (r)

-1 to +1

Q8. What is the thumb rule for r for a relation to be called as strong relation?

r>= 0.8

or r<= -0.8

Q9. What is the equation of a straight line

y=mx+c

Q10. What does OLS mean

Ordinary least squares method

Q1. What does scatter diagram tell us

1. Linearity
2. Direction of linear relation
3. Strength of the relation (in a subjective way)

Q2. What does correlation coefficient measure?

Strength of the linear association between two numerical variables

Q3. How to evaluate the overall linear regression model’s significance

P value of the F statistic of the model

Q4. How to gauge the linear regression model’s accuracy

Coefficient of determination R2 value helps us to gauge the model’s accuracy

Q5. What is the formula to calculate residual?

Residual= Actual value- fitted value

Q6. How to increase the accuracy of the linear model

Feature engineering/ Transformations

Q7. What is the range of R^2 value

0 to 1

Q8. How to handle data which has non-linear relation between input and output

Transformations

Q9. What should be the relation between parameters and output variable in linear regression?

Linear

Q10. What is OLS technique objective

To find the best fit straight line for the given data which gives least errors (mean of the error should be 0)

Q1. When do we resort to linear regression?

When y is continuous variable.

Q2. Logistic regression is used to predict the output which is?

Categorical- Binary

Q4. In Logistic regression the predicted output will be??

Probability of class assigned to 1

Q5. What is the function used to calculate probability in logistic regression?

p= e^y/1+e^y

Q6. Assumptions of logistic regression

Out come is binary and mutually exclusive.

Out come of one event is independent of other

Q7. What is Forward feature selection

Adding one variable at a time to the model, evaluate its importance and proceed further till we add all the variables. If the variable is not important can be removed from the model

Q8. The default cut-off value used in logistic regression

0.5

Q10. What does odds mean

odds is the ration of probability of event occurring vs probability of event not occurring