1. **What is the Linear Regression?**

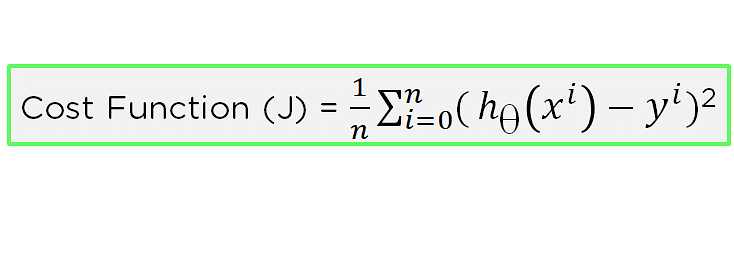
Simple linear regression is **a regression model that estimates the relationship between one independent variable and one dependent variable using a straight line**. Both variables should be quantitative.

Linear regression is **the most basic and commonly used predictive analysis**. Regression estimates are used to describe data and to explain the relationship.

1. **How we can calculate an error in Linear Regression?**

Difference between predicated value and real value is called **error.**

By using MSE (Mean Squared Error) – Cost Function, we can calculate the error



1. **Difference between loss and cost function?**

The loss function is to capture the difference between the actual and predicted values for a single record where as cost functions aggregate the difference for the entire training dataset.

The Most commonly used loss functions are Mean-squared error and Hinge loss.

1. **Explain how gradient descent works in linear regression**

Gradient Descent is **an algorithm that finds the best-fit line for a given training dataset in a smaller number of iterations**. For some combination of m and c, we will get the least Error (MSE). That combination of m and c will give us our best fit line

1. **What is intercept value refers**

The intercept (sometimes called the “constant”) in a regression model **represents the mean value of the response variable when all of the predictor variables in the model are equal to zero**

1. **Write all the assumptions of Linear Regression?**

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.

* Linear relationship.
* Multivariate normality.
* No or little multicollinearity.
* No auto-correlation.
* Homoscedasticity.
* There are four assumptions associated with a linear regression model: 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.

1. **How is hypothesis testing using in linear regression?**

While training linear regression models, hypothesis testing is done to determine whether the relationship between the response and each of the predictor variables is statistically significant or otherwise.

It consists of collecting the sample data and making some conclusions about population data by using some experiments.

1. **How would decide the importance of variable for the multivariate variable?**

**When we fit a multiple regression model,** we use the p-value in the ANOVA table to determine whether the model, as a whole**,** is significant..

1. **R squared and adjusted R squared**

R-squared = (TSS-RSS)/TSS

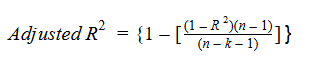
= Explained variation/ Total variation

   = 1 – Unexplained variation/ Total variation

Adjusted R-squared statistic

The Adjusted R-squared takes into account the number of independent variables used for predicting the target variable. In doing so, we can determine whether adding new variables to the model actually increases the model fit.

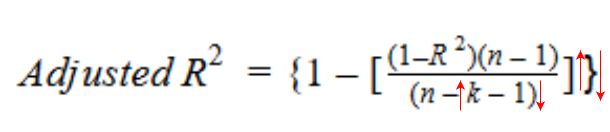
Let’s have a look at the formula for adjusted R-squared to better understand its working.



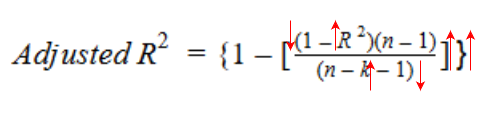
Here,

* **n** represents the number of data points in our dataset
* **k** represents the number of independent variables, and
* **R** represents the R-squared values determined by the model.

So, if R-squared does not increase significantly on the addition of a new independent variable, then the value of Adjusted R-squared will actually decrease.



On the other hand, if on adding the new independent variable we see a significant increase in R-squared value, then the Adjusted R-squared value will also increase.



We can see the difference between R-squared and Adjusted R-squared values if we add a random independent variable to our model.

