Task 2 (Solution)

How many Red colored top Honda SUV cars do you think will be sold in India in 2022 ?

Questions to consider:

1. What factors do you think will impact sales ?

Soln )

* Previous Years Sales

## Economic conditions

## Car prices

## Petro/Diesel Prices

## Technology

## Availability of quality raw materials and labor

## Marketing

## Product features

## Vehicle Performance

## After-sales services

## Parts availability

## Level of competition in the industry and markets

2. Assume you have all the data you need, what statistical methodology or algorithm will you use to make this sales forecast? Please give a brief explanation of why you choose this model.

Soln )

I will go with Linear Regression Algorithm because it is best fit for prediction purpose since we will be having all of the previous year's data. It would help us predict the sales in future.

Linear Regression is used to predict or forecast a continuous (not limited) value, such as the sales made on a day or predict temperature of a city, etc. (basically predict any continuous amount).

Linear Regression can be used to create a predictive model. If additional values get added, the model will make a prediction of a specified target variable.

3. How would you evaluate your model or determine its accuracy?

Soln )

Step 1)

Split your data into a *X-array* that contains all the features and a *Y-array* with the target variable, which we are trying to predict (e.g., car sales)

Step 2)

Split your data into a training set for the model and a testing set in order to test the model.

Step 3)

Create and test your model

Using:

* + - *from sklearn.linear\_model import LinearRegression*
    - *lm = LinearRegression()*
    - *lm.fit(X\_train,y\_train)*

Step 4)

We can calculate accuracy of a model using R2 score

* from sklearn.metrics import r2\_score
* r2\_score(Y\_test,lm.predict(X\_test))