Name: Ankit Sharma

Email: kumarankitx022@gmail.com

Mob no.: +91-7677241423

ASSIGNMENT – 05

ASSUMPTIONS OF LINEAR REGRESSION:

1. **Linearity**: The relationship between X and the mean of Y is linear.
2. **Lack of multicollinearity**: a state of very high inter-correlations or inter-associations among the independent variables.
3. **Homoscedasticity**: The variance of residual is the same for any value of X and if there is a specific pattern, the data is **Heteroscedastic.**
4. **Independence of error (Independence)**: Observations are independent of each other.
5. **Multivariate normality (normality)**: For any fixed value of X, Y is normally distributed.

MAE, MSE, RMSE, R-Square of multiple linear regression:

In [14]:
```
from sklearn.metrics import r2_score, mean_squared_error

print("Mean absolute error: %.2f" % np.mean(np.absolute(pred_y - y_test)))
print("Mean square error: %.2f" %np.mean((pred_y - y_test))**2)
print("R2-score: %.2f" % r2_score(pred_y , y_test) )
print('Variance score: %.2f' % reg.score(X_test, y_test))
print('RMSE: %.2f' %np.sqrt(np.mean((pred_y - y_test) ** 2)))
#print('RMSE: %.2f' %np.sqrt(mean_squared_error(y_test, pred_y)))
```

```
Mean absolute error: 5701.10
Mean square error: 951660.65
R2-score: 0.95
Variance score: 0.95
RMSE: 7516.47
```