Case Study 1 – Machine Learning Using Python

Problem Statement:

You work in XYZ Company as a Python developer. The company officials want you to build a data science model.

Tasks To Be Performed:

1. Using sklearn import the wine dataset
2. Split the data into train and test set
3. Train the model
4. Make Predictions
5. Check the performance of the model using r2_score

In [1]:
```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import *

df = pd.read_csv(r"csv files\wine.csv")
df.head()
```

Out[1]:

	Wine	Alcohol	Malic.acid	Ash	Acl	Mg	Phenols	Flavanoids	Nonflavanoid.phenols	Proanth	Color.int	Hue	OD	Proline
0	1	14.23	1.71	2.43	15.6	127	2.80	3.06	0.28	2.29	5.64	1.04	3.92	1065
1	1	13.20	1.78	2.14	11.2	100	2.65	2.76	0.26	1.28	4.38	1.05	3.40	1050
2	1	13.16	2.36	2.67	18.6	101	2.80	3.24	0.30	2.81	5.68	1.03	3.17	1185
3	1	14.37	1.95	2.50	16.8	113	3.85	3.49	0.24	2.18	7.80	0.86	3.45	1480
4	1	13.24	2.59	2.87	21.0	118	2.80	2.69	0.39	1.82	4.32	1.04	2.93	735

In [2]:
```
X = df.drop(columns=['Proline'])
y = df['Proline']
```

In [3]:
```
X_train, X_test, y_train, y_test = train_test_split(X,y,train_size=0.8,random_state=0)
```

In [4]:
```
lr = LinearRegression()
lr.fit(X_train,y_train)
```

Out[4]:
```
LinearRegression
LinearRegression()
```

In [5]:
```
y_pred = lr.predict(X_test)
```

In [6]:
```
r2_score(y_test,y_pred)
```

Out[6]: 0.737580212561169

In [ ]: