Aim :- Implementation of regression analysis for selected data set.

LO :- LO4

In [90]:
```python
import pandas as pd
import matplotlib as plt
import numpy as np
```

In [91]:
```python
from sklearn.linear_model import LinearRegression, LogisticRegression
from sklearn.model_selection import train_test_split
```

In [92]:
```python
df = pd.read_csv("Concrete_Data_Yeh.csv")
```

In [93]:
```python
df
```

Out[93]:

	cement	slag	flyash	water	superplasticizer	coarseaggregate	fineaggregate	a
0	540.0	0.0	0.0	162.0	2.5	1040.0	676.0	
1	540.0	0.0	0.0	162.0	2.5	1055.0	676.0	
2	332.5	142.5	0.0	228.0	0.0	932.0	594.0	2
3	332.5	142.5	0.0	228.0	0.0	932.0	594.0	3
4	198.6	132.4	0.0	192.0	0.0	978.4	825.5	3
...	...	...	...	...	...	...	...	
1025	276.4	116.0	90.3	179.6	8.9	870.1	768.3	
1026	322.2	0.0	115.6	196.0	10.4	817.9	813.4	
1027	148.5	139.4	108.6	192.7	6.1	892.4	780.0	
1028	159.1	186.7	0.0	175.6	11.3	989.6	788.9	
1029	260.9	100.5	78.3	200.6	8.6	864.5	761.5	

1030 rows × 9 columns

In [94]:
```python
X = df.drop(columns=['csMPa'],axis=1)
```

In [95]:
```python
X
```

Out[95]:

	cement	slag	flyash	water	superplasticizer	coarseaggregate	fineaggregate	a
0	540.0	0.0	0.0	162.0	2.5	1040.0	676.0	
1	540.0	0.0	0.0	162.0	2.5	1055.0	676.0	
2	332.5	142.5	0.0	228.0	0.0	932.0	594.0	2
3	332.5	142.5	0.0	228.0	0.0	932.0	594.0	3
4	198.6	132.4	0.0	192.0	0.0	978.4	825.5	3
...	...	...	...	...	...	...	...	
1025	276.4	116.0	90.3	179.6	8.9	870.1	768.3	
1026	322.2	0.0	115.6	196.0	10.4	817.9	813.4	
1027	148.5	139.4	108.6	192.7	6.1	892.4	780.0	
1028	159.1	186.7	0.0	175.6	11.3	989.6	788.9	
1029	260.9	100.5	78.3	200.6	8.6	864.5	761.5	

1030 rows × 8 columns

◀ ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ ▶

In [96]:
```python
y = df['csMPa']
```

In [97]:
```python
y
```

Out[97]:
```
0       79.99
1       61.89
2       40.27
3       41.05
4       44.30
        ...
1025    44.28
1026    31.18
1027    23.70
1028    32.77
1029    32.40
Name: csMPa, Length: 1030, dtype: float64
```

In [105...
```python
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.1,random_sta
X_train.shape, X_test.shape
```

Out[105...
```
((927, 8), (103, 8))
```

In [108...
```python
models = LinearRegression()
models.fit(X_train,y_train)
```

Out[108...
```
▾ LinearRegression ⓘ ⓘ

LinearRegression()
```

In [109...
```python
print("Accuracy",models.score(X_test,y_test)*100)
```

```
Accuracy 61.963128002736454
```

Conclusion:- This experiment applies regression on finding strength of concrete in MPa i.e megapascal