

Name: Shubham Shigwan  
Email:shubhamshigwan2k@gmail.com

In [ ]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [ ]:

```
salary = pd.read_csv('https://github.com/YBI-Foundation/Dataset/raw/main/Salary%20Data.csv')
```

In [ ]:

```
salary.head()
```

Out[ ]:

	Experience	Years	Salary
0	1.1	39343	
1	1.2	42774	
2	1.3	46205	
3	1.5	37731	
4	2.0	43525	

In [ ]:

```
salary.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 40 entries, 0 to 39
Data columns (total 2 columns):
 #   Column           Non-Null Count  Dtype  
 ---  --  
 0   Experience     40 non-null      float64 
 1   Salary          40 non-null      int64  
dtypes: float64(1), int64(1)
memory usage: 768.0 bytes
```

In [ ]:

```
salary.describe()
```

Out[ ]:

	Experience Years	Salary
count	40.000000	40.000000
mean	5.152500	74743.625000
std	2.663715	25947.122885
min	1.100000	37731.000000
25%	3.200000	56878.250000
50%	4.600000	64472.500000
75%	6.875000	95023.250000
max	10.500000	122391.000000

## Define x and y

In [ ]:

```
y = salary['Salary']
```

In [ ]:

```
X = salary[['Experience Years']]
```

In [ ]:

```
y.shape
```

Out[ ]:

```
(40, )
```

In [ ]:

```
X.shape
```

Out[ ]:

```
(40, 1)
```

## Split Data

In [ ]:

```
from sklearn.model_selection import train_test_split
```

In [ ]:

```
X_train, X_test, y_train, y_test = train_test_split(X,y,train_size=0.8,random_state=2529)
```

In [ ]:

```
X_train.shape, X_test.shape, y_train.shape, y_test.shape
```

Out[ ]:

```
((32, 1), (8, 1), (32,), (8,))
```

In [ ]:

```
from sklearn.linear_model import LinearRegression
```

In [ ]:

```
model = LinearRegression()
```

In [ ]:

```
model.fit(X_train,y_train)
```

Out[ ]:

```
LinearRegression()
```

In [ ]:

```
model.coef_
```

Out[ ]:

```
array([9398.19785815])
```

In [ ]:

```
model.intercept_
```

Out[ ]:

```
26344.85810217697
```

In [ ]:

```
y_pred = model.predict(X_test)
```

In [ ]:

```
y_pred
```

Out[ ]:

```
array([ 90252.60353757,  59238.55060569, 106229.53989642,  63937.649
53476,       68636.74846383, 123146.29604108,  84613.68482268,  62997.829
74895])
```

In [ ]:

```
from sklearn.metrics import mean_absolute_percentage_error
```

In [ ]:

```
mean_absolute_percentage_error(y_test,y_pred)
```

Out[ ]:

```
0.050339688879149136
```