Simple Linear Regression using Python

- ## Predicting Salary using Experience data
- ## Total 11_steps

01_step Importing libraries

In [ ]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

02_step Loading Data

In [ ]:
```
df = pd.read_csv("salary_data.csv")
df.head()
```

Out[ ]:

	YearsExperience	Salary
0	1.1	39343
1	1.3	46205
2	1.5	37731
3	2.0	43525
4	2.2	39891

03_step Checcking Numll values

In [ ]:
```
df.isnull().sum()
```

Out[ ]:
```
YearsExperience    0
Salary             0
dtype: int64
```

05_Spliting Data

In [ ]:
```
X = df[["YearsExperience"]]
y = df["Salary"]
```

06_step Importing sciket-learn

In [ ]:
```
pip install scikit-learn
```

In [ ]:
```
import sklearn
```

07_step model selection

In [ ]:
```
from sklearn.model_selection import train_test_split
X_train,X_test,y_train,y_test = train_test_split(X,y, test_size=0.2, random_state=0)
```

08_step Impoting Linearregression

In [ ]:
```
from sklearn.linear_model import LinearRegression
model = LinearRegression().fit(X_train,y_train)
model
```

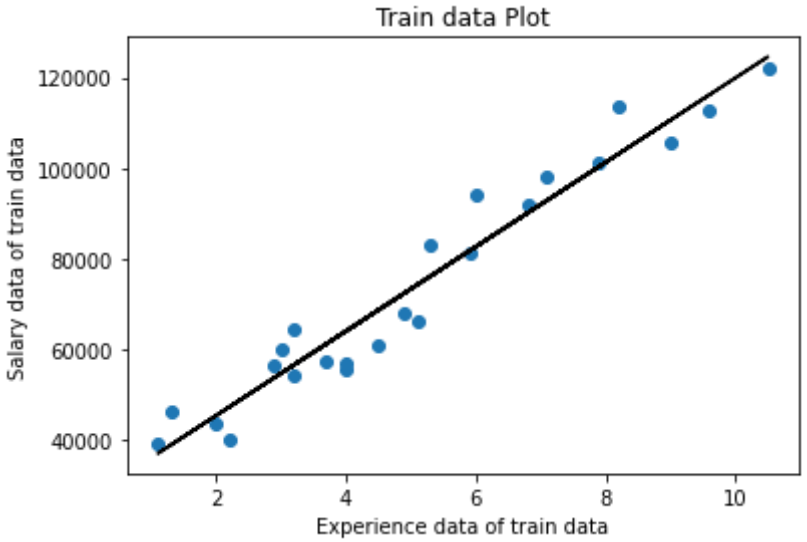
Out[ ]:
```
▾ LinearRegression
LinearRegression()
```

In [ ]:
```
model.predict(X_test)
```

Out[ ]:
```
array([ 40748.96184072, 122699.62295594,  64961.65717022,  63099.14214487,
       115249.56285456, 107799.50275317])
```

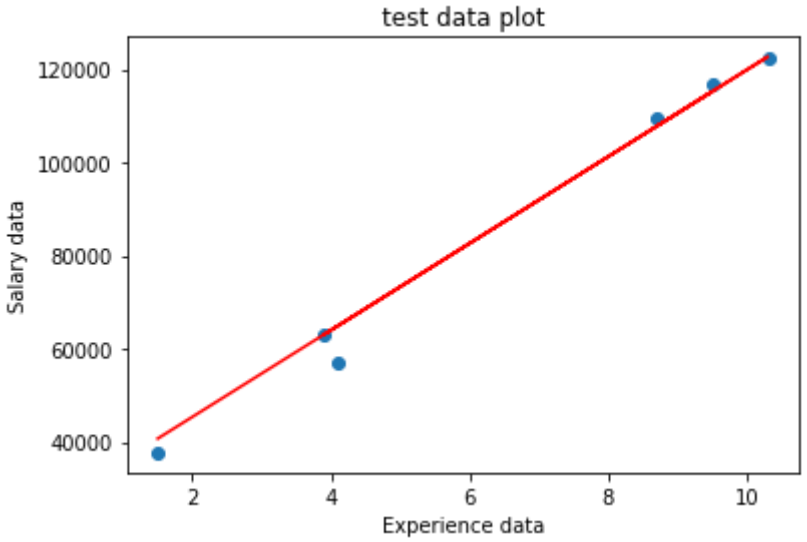
09_step Visualization of train and test data

In [ ]:
```
plt.scatter(X_train, y_train)
plt.plot(X_train, model.predict(X_train), color= "Black")
plt.title("Train data Plot")
plt.xlabel("Experience data of train data")
plt.ylabel("Salary data of train data")
plt.show()
```



In [ ]:
```
from turtle import color


plt.scatter(X_test, y_test)
plt.plot(X_test, model.predict(X_test), color= "Red")
plt.title("test data plot")
plt.xlabel("Experience data")
plt.ylabel("Salary data")
plt.show()
```



10_step Checking fitness of your model

In [ ]:
```
model.score(X_train, y_train)
```

Out[ ]:
```
0.9411949620562126
```

In [ ]:
```
model.score(X_test, y_test)
```

Out[ ]:
```
0.988169515729126
```

11_step Predicting Values using your model

In [ ]:
```
model.predict(X_test)
```

Out[ ]:
```
array([ 40748.96184072, 122699.62295594,  64961.65717022,  63099.14214487,
       115249.56285456, 107799.50275317])
```

In [ ]:
```
model.predict([[3]])
```

```
c:\Users\Salman\AppData\Local\Programs\Python\Python39\lib\site-packages\sklearn\base.py:450: UserWarning: X does not have valid
feature names, but LinearRegression was fitted with feature names
  warnings.warn(
```

Out[ ]:
```
array([54717.82453082])
```