4.Write a python programme to implement linear Regression

```
import pandas as pd
```

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

```
#definging target(y) and feature(x)
y = salary['Salary']
X = salary[['Experience Years']]
```

```
#train test split
from sklearn.model_selection import train_test_split
X_train,X_test,y_train,y_test= train_test_split(X,y,train_size =  0.7,random_state =42)
```

```
#Check shape of train and test sample
print(X_train.shape , X_test.shape , y_train.shape , y_test.shape)
```

```
#select model
from sklearn.linear_model import LinearRegression
model = LinearRegression()
```

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

```
print(model.intercept_)
```

```
26767.86524944281
```

```
print(model.coef_)
```

```
[9459.35953483]
```

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

```
print(y_pred)
```

```
[ 69334.98315618  64605.30338877  64605.30338877  83524.02245843
  45686.5843191   61767.49552832 117577.71678382  85415.89436539
 126091.14036516  50416.26408652  82578.08650494  57037.8157609 ]
```

```
#model accuracy
from sklearn.metrics import mean_absolute_error,mean_absolute_percentage_error,
mean_squared_error
```

```
print("MSE:",mean_absolute_error(y_test,y_pred))
```

```
MSE: 5211.981371554558
```

```
print("MAPE:",mean_absolute_percentage_error(y_test,y_pred))
```

```
MAPE: 0.0761963556601618
```

```
print("MSE:",mean_squared_error(y_test,y_pred))
```

```
MSE: 35344480.17477033
```

```
y_pred_line = model.predict(X)
y_pred_test = model.predict(X_test)
```

```
# Scatter plot of actual vs predicted
import matplotlib.pyplot as plt
import seaborn as sns

# Plot actual data
plt.figure(figsize=(8, 6))
plt.scatter(X, y, marker = '*', label='Actual Data')                    # all real data
plt.plot(X, y_pred_line, marker = 'o', label='Regression Line')         # regression
curve
plt.scatter(X_test, y_pred_test, color= 'green', label='Predicted Points')  # predicted
points on test data

plt.xlabel("X")
plt.ylabel("y")
plt.title("Linear Regression: Actual vs Predicted with Regression Line" )
plt.legend()
plt.grid(True)
plt.show()
```