

### Task 1: Exploring the Linear Regression model through “SKLearn” library.

Solution:

```
import numpy
import matplotlib.pyplot as plot
import pandas
from sklearn import linear_model
from sklearn.model_selection import train_test_split
reg = linear_model.LinearRegression()

# Import the dataset
dataset = pandas.read_csv('Salary.csv')
x = dataset.iloc[:, :-1].values
y = dataset.iloc[:, 1].values

# Split the dataset into the training set and test set
# We're splitting the data in 1/3, so out of 30 rows, 20 rows will go into the training set,
# and 10 rows will go into the testing set.
xTrain, xTest, yTrain, yTest = train_test_split(x, y, test_size = 1/3, random_state = 0)

reg.fit(xTrain, yTrain)
print("co-efficients->")
print(reg.coef_)
print("intercept->")
print(reg.intercept_)
print("prediction")
print(reg.predict(xTest))
```

```
# plotting items with respect to training set
plot.scatter(xTrain, yTrain, color = 'red')
plot.plot(xTrain, reg.predict(xTrain), color = 'blue')
plot.title('Salary vs Experience (Training set)')
plot.xlabel('Years of Experience')
plot.ylabel('Salary')
plot.show()

# plotting items with respect to test set
plot.scatter(xTest, yTest, color = 'magenta')
plot.plot(xTrain, reg.predict(xTrain), color = 'blue')
plot.title('Salary vs Experience (Test set)')
plot.xlabel('Years of Experience')
plot.ylabel('Salary')
plot.show()
```



