LAB - 3

Linear Regression using ~~error square method~~ matrix method

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

dataset = pd.read_csv('/content/sales.csv')
print(dataset.head())
weeks = dataset['xi(week)'].values
sales = dataset['yi(sales in thousands)'].values

X = weeks.reshape(-1, 1)
Y = sales.reshape(-1,1)

X_b = np.c_[np.ones((len(X),1)), X]

theta = np.linalg.inv(X_b.T.dot(X_b)).dot(X_b.T).dot(Y)
b = theta[0]
m = theta[1]

print(f"The regression equation is: y = {m[0]:.2f}x
      + {b[0]:.2f}")
predicted_sales_7 = m*7+b
predicted_sales_9 = m*9+b

print(f" Predicted sales for the 7th week: {predicted_sales
      -7[0]:.2f} thousand")
print(f" Predicted sales for the 9th week: {predicted_sales_9
      [0]:.2f} thousand")
```

```
plt.scatter(weeks, sales, color='blue', label='Data points')
plt.plot(weeks, m*weeks+b, color='red',
         label=f'Linear Regression: y = {m[0]:.2f}x +
         {b[0]:.2f}')
plt.xlabel('Weeks')
plt.ylabel('Sales (in thousands)')
plt.title('Sales data and Linear regression')
plt.legend()
plt.show()
```

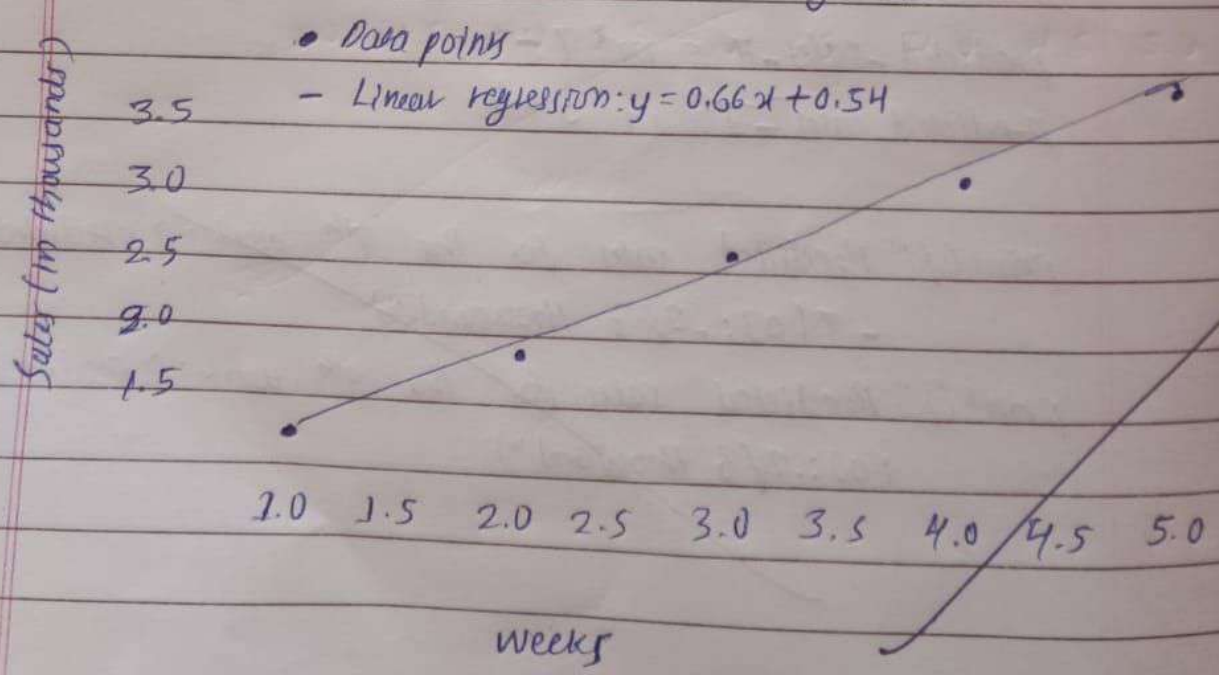
Output:-

	Xi (week)	yi (Sales in thousands)
0	0	1.2
1	1	1.8
2	2	2.6
3	3	3.2
4	4	3.8

The regression equation is: $y = 0.66x + 0.54$

Predicted Sales for the 7th week: 5.16 thousand

Predicted Sales for the 9th week: 6.48 thousand



Linear Regression using error squared method

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

dataset = pd.read_csv('content/sales.csv')
print(dataset.head())

weeks = dataset['X(Weeks)'].values
Sales = dataset['y(Sales in thousands)'].values
n = len(weeks)

sum_x = np.sum(weeks)
sum_y = np.sum(sales)
sum_x2 = np.sum(weeks**2)
sum_xy = np.sum(weeks * sales)

m = (n * sum_xy - sum_x * sum_y) / (n * sum_x2
    - sum_x**2)
b = (sum_y - m * sum_x) / n

print(f"The regression equation is: y = {m:.2f}x +
    {b:.2f}")

print(f"Predicted sales for 7th week: {predicted_sales_7:1.2f}
    thousand")
print(f"Predicted sales for 9th week: {predicted_sales:.2f}
    thousand")
```

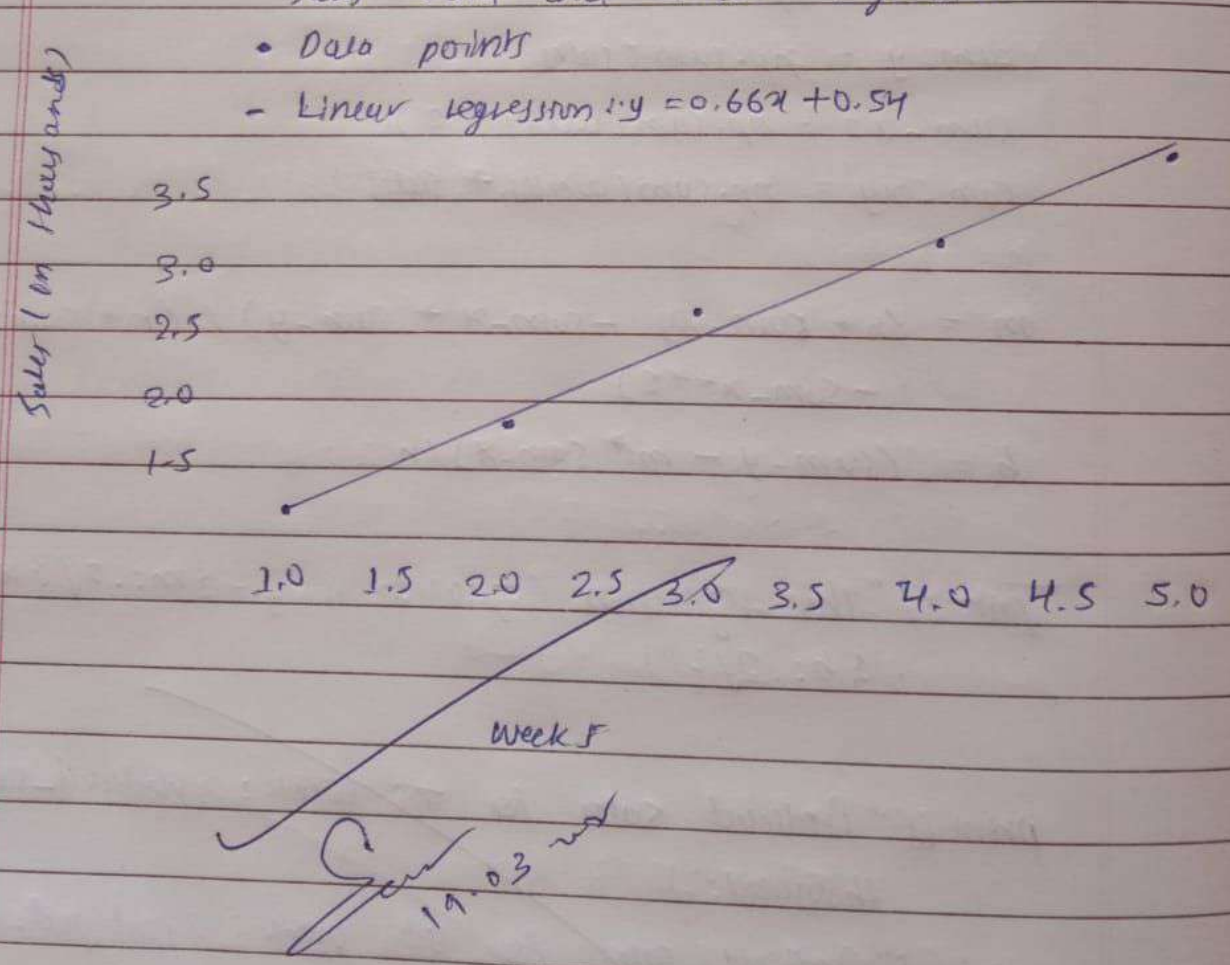
```
plt.scatter(weeks, sales, color='blue', label='Data points')
plt.plot(weeks, m*weeks+b, color='red',
         label=f'Linear Regression: y={m:.2f}x+{b:.2f}')
plt.xlabel('Weeks')
plt.ylabel('Sales (in thousands)')
plt.title('Sales Data and Linear Regression')
plt.legend()
plt.show()
```

Output :-

The regression equation is: $y = 0.66x + 0.54$

Predicted sales for the 7th week : 5.16 thousand

Predicted sales for the 9th week : 6.48 thousand



19.03