**MLA0201-Fundamentals of Machine Learning**

**Name: Kaviya Kumara K(192425055)**

Experiment 20:

Implement Future Sales Prediction using a suitable machine learning algorithm.

**Code:**

import pandas as pd

import matplotlib.pyplot as plt

import numpy as np

from sklearn.linear\_model import LinearRegression

from sklearn.metrics import mean\_squared\_error

data = pd.read\_csv("sales\_data.csv")

data["Date"] = pd.to\_datetime(data["Date"])

data = data.sort\_values("Date")

data["TimeIndex"] = np.arange(len(data))

X = data[["TimeIndex"]]

y = data["Sales"]

split\_index = int(len(data) \* 0.8)

X\_train = X.iloc[:split\_index]

X\_test = X.iloc[split\_index:]

y\_train = y.iloc[:split\_index]

y\_test = y.iloc[split\_index:]

model = LinearRegression()

model.fit(X\_train, y\_train)

print("Future Sales Prediction model trained successfully")

y\_pred = model.predict(X\_test)

rmse = np.sqrt(mean\_squared\_error(y\_test, y\_pred))

print("RMSE:", rmse)

future\_days = 30

future\_index = pd.DataFrame({

"TimeIndex": np.arange(len(data), len(data) + future\_days)

})

future\_sales = model.predict(future\_index)

plt.figure(figsize=(10, 6))

plt.plot(data["Date"], y, label="Historical Sales")

plt.plot(

data["Date"].iloc[split\_index:],

y\_pred,

label="Test Prediction"

)

future\_dates = pd.date\_range(

start=data["Date"].iloc[-1] + pd.Timedelta(days=1),

periods=future\_days

)

plt.plot(

future\_dates,

future\_sales,

label="Future Forecast"

)

plt.xlabel("Date")

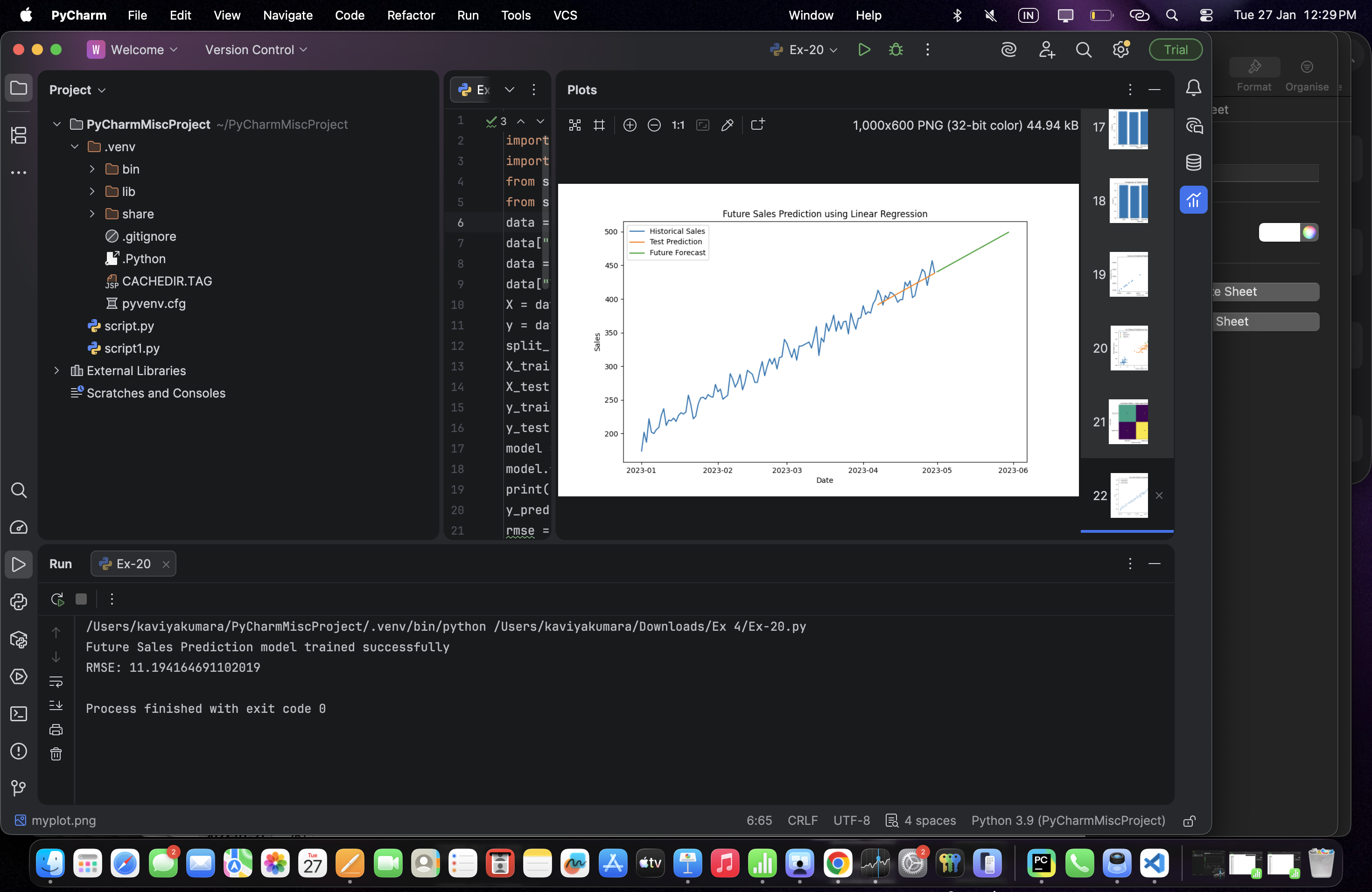
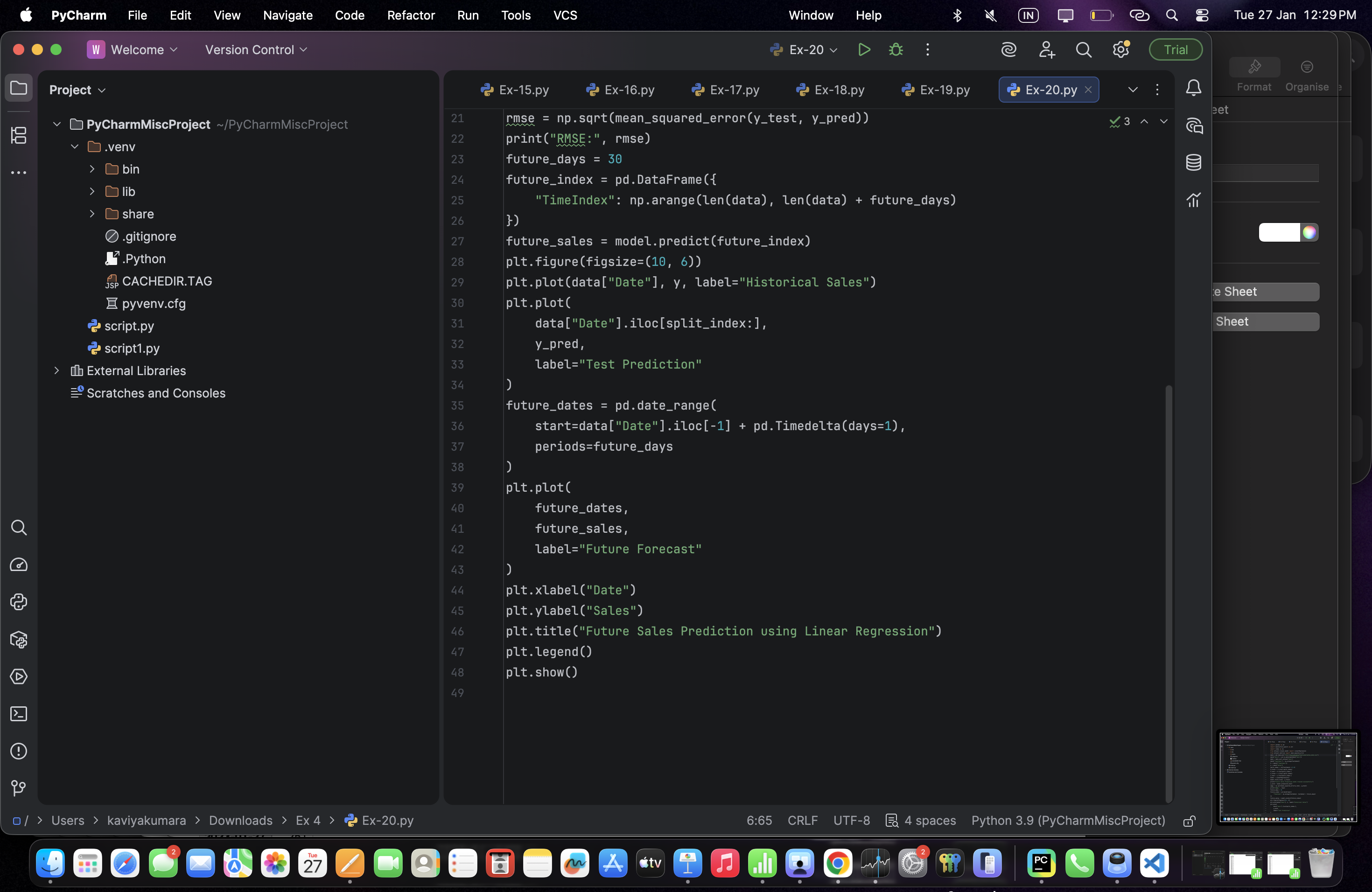
plt.ylabel("Sales")

plt.title("Future Sales Prediction using Linear Regression")

plt.legend()

plt.show()

**Output:**

****