```
# Import necessary libraries
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
# Load the dataset
advertising = pd.read_csv('company_data.csv')
# Display the first few rows of the data
print(advertising.head())
#Scatter plot representation of data using seaborn
import matplotlib.pyplot as plt
import seaborn as sns
sns.pairplot(advertising, x_vars=['TV', 'Radio', 'Newspaper'],
y_vars='Sales', height=5, aspect=1, kind='scatter')
plt.show()
# Assuming 'Sales' is the dependent variable and the rest are features
X = advertising.drop(columns=['Sales'])
y = advertising['Sales']
# Split the dataset into training and testing sets (80% training, 20% testing)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
# Create the linear regression model
model = LinearRegression()
# Train the model on the training data
model.fit(X_train, y_train)
# Predict on the test data
y_pred = model.predict(X_test)
# Model evaluation
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
# Display results
print("\nMean Squared Error:", mse)
```

```
print("R-squared:", r2)
# Display the coefficients and intercept
print("\nCoefficients:", model.coef_)
print("Intercept:", model.intercept_)
#To Display Real Values and Predicted Values
y_pred = model.predict(X_test)
for(i,j) in zip(y_test,y_pred):
    if i!=j:
        print("Actual value :",i,"Predicted value :",j)
print("\nNumber of mislabeled points from test data set :", (y_test !=
y_pred).sum())
```

output

==========

```
C:\Users\mlm\PycharmProjects\pythonProject1\.venv\Scripts\python.exe
C:\Users\mlm\PycharmProjects\BIBIN\multiple_linear_regresion.py
     TV  Radio  Newspaper  Sales
0  230.1   37.8       69.2   22.1
1   44.5   39.3       45.1   10.4
2   17.2   45.9       69.3   12.0
3  151.5   41.3       58.5   16.5
4  180.8   10.8       58.4   17.9

Mean Squared Error: 3.562686377102421
R-squared: 0.8484372382624404

Coefficients: [ 0.05323974  0.12099215 -0.00557407]
Intercept: 4.682857108397286
Actual value : 15.0 Predicted value : 17.30978223003696
```

Number of mislabeled points from test data set : 40

Actual value : 18.0 Predicted value : 19.364242727795762

Number of mislabeled points from test data set : 40

Actual value : 17.0 Predicted value : 18.94760357642275

Number of mislabeled points from test data set : 40

Actual value : 11.9 Predicted value : 8.90140256975679

Number of mislabeled points from test data set : 40

Actual value : 11.5 Predicted value : 12.033937588310465

Number of mislabeled points from test data set : 40

Actual value : 22.1 Predicted value : 21.12109798558051

Number of mislabeled points from test data set : 40

Actual value : 11.9 Predicted value : 12.671038929007896

Number of mislabeled points from test data set : 40

Actual value : 11.6 Predicted value : 12.893587797987534

Number of mislabeled points from test data set : 40

Actual value : 1.6 Predicted value : 9.462919576765735

Number of mislabeled points from test data set : 40

Actual value : 21.7 Predicted value : 21.063799751434463

Number of mislabeled points from test data set : 40

Actual value : 7.3 Predicted value : 6.203057368924365

Number of mislabeled points from test data set : 40

Actual value : 12.5 Predicted value : 12.085379481573359

Number of mislabeled points from test data set : 40

Actual value : 7.6 Predicted value : 8.063538892716368

Number of mislabeled points from test data set : 40

Actua