

OASIS INFOBYTE INTERNSHIP (Data Science)

Task 5: Sales Prediction Using Python

Name: Palak Parmar

College: Bansal Group Of Institute

Submission Date: 08/11/2025

Introduction:

This beginner-friendly project demonstrates how to predict sales using a simple Machine Learning model. We create a small dataset of advertising spends and their corresponding sales to train a Linear Regression model. The code is simple, self-contained, and ideal for beginners.

Libraries to install (run once in terminal):

pip install pandas

pip install scikit-learn

pip install matplotlib

Code (copy and run in Python):

Beginner-friendly Sales Prediction (no external files required)

```
import pandas as pd
```

```
from sklearn.model_selection import train_test_split
```

```
from sklearn.linear_model import LinearRegression
```

```
from sklearn.metrics import r2_score, mean_absolute_error
```

```
# Sample dataset
```

```
data = {
```

```
    'TV': [230.1, 44.5, 17.2, 151.5, 180.8, 8.7, 57.5, 120.2, 220.5, 75.3],
```

```
    'Radio': [37.8, 39.3, 45.9, 41.3, 10.8, 48.9, 32.8, 19.6, 33.5, 20.5],
```

```
    'Newspaper': [69.2, 45.1, 69.3, 58.5, 58.4, 75.0, 23.5, 11.6, 49.9, 25.4],
```

```
    'Sales': [22.1, 10.4, 9.3, 18.5, 12.9, 7.2, 11.8, 13.2, 19.7, 10.1]
```

```
}
```

```
df = pd.DataFrame(data)
```

```
# Features and target
```

```
X = df[['TV', 'Radio', 'Newspaper']]
```

```
y = df['Sales']
```

```
# Split data
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

```
# Train model
```

```
model = LinearRegression()
```

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

```
# Predict
```

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

```
# Evaluate
```

```
print("R2 Score:", round(r2_score(y_test, y_pred), 3))
```

```
print("Mean Absolute Error:", round(mean_absolute_error(y_test, y_pred), 3))
```

```
print("Actual Sales:", list(y_test.values))
```

```
print("Predicted Sales:", [round(x,2) for x in y_pred])
```

Expected Output (example): - R2 Score: around 0.9 indicates good accuracy - Mean Absolute Error: small value shows less difference between actual and predicted sales - Lists of actual and predicted sales will be printed for comparison. This code uses a built-in sample dataset, so it runs easily for all beginners.

Conclusion: This project explains how to build a simple regression model for predicting sales using Python. By understanding relationships between advertisement spending and sales, this task introduces data analysis and prediction basics.