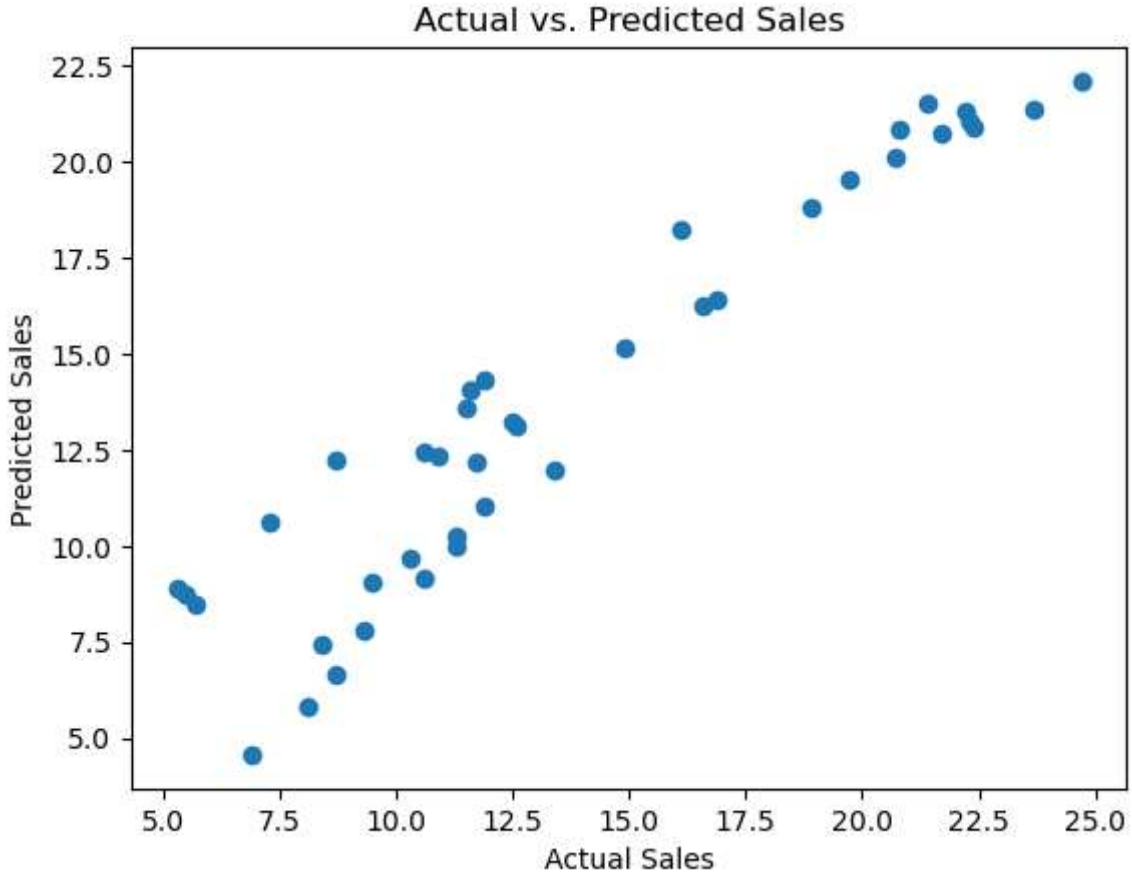
In [1]:
```
# Import necessary libraries
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
import matplotlib.pyplot as plt

# Load the dataset
dataset_path = "C:\\Users\\ARYAN PARIKH\\Desktop\\Oasis Internship\\archive (4
advertising_data = pd.read_csv(dataset_path)

# Display the first few rows of the dataset
print(advertising_data.head())

# Explore the dataset to understand its structure
print(advertising_data.info())

# Select relevant features and target variable
X = advertising_data[['TV', 'Radio', 'Newspaper']]
y = advertising_data['Sales']

# Split the dataset into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, randc

# Train a Linear Regression model
model = LinearRegression()
model.fit(X_train, y_train)

# Make predictions on the test set
y_pred = model.predict(X_test)

# Evaluate the model
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)

print(f'Mean Squared Error: {mse}')
print(f'R-squared Score: {r2}')

# Visualize predicted vs. actual sales
plt.scatter(y_test, y_pred)
plt.xlabel('Actual Sales')
plt.ylabel('Predicted Sales')
plt.title('Actual vs. Predicted Sales')
plt.show()
```

```
C:\Users\ARYAN PARIKH\AppData\Roaming\Python\Python311\site-packages\pandas
\core\arrays\masked.py:60: UserWarning: Pandas requires version '1.3.6' or n
ewer of 'bottleneck' (version '1.3.5' currently installed).
  from pandas.core import (
```

```
   Unnamed: 0     TV  Radio  Newspaper  Sales
0           1  230.1   37.8       69.2   22.1
1           2   44.5   39.3       45.1   10.4
2           3   17.2   45.9       69.3    9.3
3           4  151.5   41.3       58.5   18.5
4           5  180.8   10.8       58.4   12.9
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 5 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   Unnamed: 0  200 non-null    int64
 1   TV          200 non-null    float64
 2   Radio       200 non-null    float64
 3   Newspaper   200 non-null    float64
 4   Sales       200 non-null    float64
dtypes: float64(4), int64(1)
memory usage: 7.9 KB
None
Mean Squared Error: 3.1740973539761046
R-squared Score: 0.899438024100912
```



In [ ]: