

```
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
sales_data = pd.read_csv('sales data.csv')
```

```
# Data Exploration
```

```
sales_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 4 columns):
#   Column      Non-Null Count  Dtype
---  -
0    TV          200 non-null    float64
1    Radio       200 non-null    float64
2    Newspaper   200 non-null    float64
3    Sales       200 non-null    float64
dtypes: float64(4)
memory usage: 6.4 KB
```



```
# Data Exploration
```

```
sales_data.describe()
```

	TV	Radio	Newspaper	Sales	
count	200.000000	200.000000	200.000000	200.000000	
mean	147.042500	23.264000	30.554000	15.130500	
std	85.854236	14.846809	21.778621	5.283892	
min	0.700000	0.000000	0.300000	1.600000	
25%	74.375000	9.975000	12.750000	11.000000	
50%	149.750000	22.900000	25.750000	16.000000	
75%	218.825000	36.525000	45.100000	19.050000	
max	296.400000	49.600000	114.000000	27.000000	

```
# Data Exploration
```

```
sales_data.head()
```



	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	

Next steps:

 [View recommended plots](#)

```
# Data Exploration
```

```
sales_data.tail()
```

	TV	Radio	Newspaper	Sales	
195	38.2	3.7	13.8	7.6	
196	94.2	4.9	8.1	14.0	
197	177.0	9.3	6.4	14.8	
198	283.6	42.0	66.2	25.5	
199	232.1	8.6	8.7	18.4	

```
# Feature Selection
X = sales_data[['TV', 'Radio', 'Newspaper']] # Features
y = sales_data['Sales'] # Target variable

# Model Selection
model = LinearRegression()

# Model Training
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
model.fit(X_train, y_train)
```

```
▼ LinearRegression
LinearRegression()
```

```
# Model Evaluation
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
r_squared = r2_score(y_test, y_pred)
print("Mean Squared Error:", mse)
print("R-squared:", r_squared)

Mean Squared Error: 2.9077569102710896
R-squared: 0.9059011844150826
```