MSSV: 21114201 Họ và Tên: Trần Ngọc Phát ca TH: 2

```
import pandas as pd
import numpy as np
```

✓ 0.0s Python

Load data vào sales df

```
sale_df = pd.read_csv('../data/advertising_and_sales_clean.csv')
sale_df
```

✓ 0.0s Python

	tv	radio	social_media	influencer	sales
0	16000.0	6566.23	2907.98	Mega	54732.76
1	13000.0	9237.76	2409.57	Mega	46677.90
2	41000.0	15886.45	2913.41	Mega	150177.83
3	83000.0	30020.03	6922.30	Mega	298246.34
4	15000.0	8437.41	1406.00	Micro	56594.18
...	...	...	...	...	...
4541	26000.0	4472.36	717.09	Micro	94685.87
4542	71000.0	20610.69	6545.57	Nano	249101.92
4543	44000.0	19800.07	5096.19	Micro	163631.46
4544	71000.0	17534.64	1940.87	Macro	253610.41
4545	42000.0	15966.69	5046.55	Micro	148202.41

4546 rows × 5 columns

- Create x, an array of the values from the sales_df DataFrame's "radio" column.
- Create y, an array of the values from the sales df DataFrame's "sales" column.

```
[118] ✓ 0.0s
x = np.array(sale_df['radio'])
y = np.array(sale_df['sales'])
x
```
Python

```
array([ 6566.23,  9237.76, 15886.45, ..., 19800.07, 17534.64, 15966.69])
```

Reshape x into a two-dimensional NumPy array.

```
[119] ✓ 0.0s
x_reshape = x.reshape(-1, 1)
x_reshape
```
Python

```
array([[ 6566.23],
       [ 9237.76],
       [15886.45],
       ...,
       [19800.07],
       [17534.64],
       [15966.69]])
```

Print the shape of x and y.

```
[120] ✓ 0.0s
print(x.shape)
print(y.shape)
```
Python

```
(4546,)
(4546,)
```

Import LinearRegression.

```
[121] ✓ 0.0s
from sklearn.linear_model import LinearRegression as lr
from sklearn.model_selection import train_test_split
```
Python

Instantiate a linear regression model.

```
[122] ✓ 0.0s
model = lr()
```
Python

Predict sales values using x, storing as predictions.

```
[123] ✓ 0.0s
x_train, x_test, y_train, y_test = train_test_split(x_reshape, y, test_size=0.2)
model.fit(x_train, y_train)
predictions = model.predict(x_test)
predictions[:5]
```
Python

```
array([ 73529.96083895, 237860.32287776, 172914.31889386, 137228.90744772,
       133162.81134904])
```

Import matplotlib.pyplot as plt.

```
import matplotlib.pyplot as plt
```

✓ 0.0s Python

- Create a scatter plot visualizing y against x, with observations in blue.
- Draw a red line plot displaying the predictions against x.
- Display the plot.

```
plt.title('Radio vs Sales')
plt.scatter(x_test, y_test, color='blue')
plt.plot(x_test, predictions, color='red')
plt.show()
```

✓ 0.0s Python