

In-Class Assignment 1

1. Set up your Python Development Environment. Anaconda is recommended. You may also want to follow the introduction to python lab in Chapter 2 of the ISLP (the course textbook)
2. Answer the questions below.
3. Export your completed and executed notebook as a pdf.
4. Upload your pdf to Canvas under In-Class Assignment 1.

Question 1

Print your name with the `print()` function.

```
In [25]: print("Amisha Farhana Shaik")
```

Amisha Farhana Shaik

Question 2

Import the `advertising.csv` dataset as `advertising_df` using pandas `read_csv()`.

```
In [ ]:
```

```
In [14]: # uncomment this line if you need to install pandas
         # !pip install -q pandas
```

[notice] A new release of pip is available: 23.2.1 -> 24.2

[notice] To update, run: `pip install --upgrade pip`

```
In [9]: import pandas as pd
        df = pd.read_csv('advertising.csv')
```

```
In [11]: df
```

```
Out[11]:
```

	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
...
195	196	38.2	3.7	13.8	7.6
196	197	94.2	4.9	8.1	9.7
197	198	177.0	9.3	6.4	12.8
198	199	283.6	42.0	66.2	25.5
199	200	232.1	8.6	8.7	13.4

200 rows × 5 columns

Question 3

Display the first 10 rows of `advertising_df`.

```
In [15]: print(df.head(10))
```

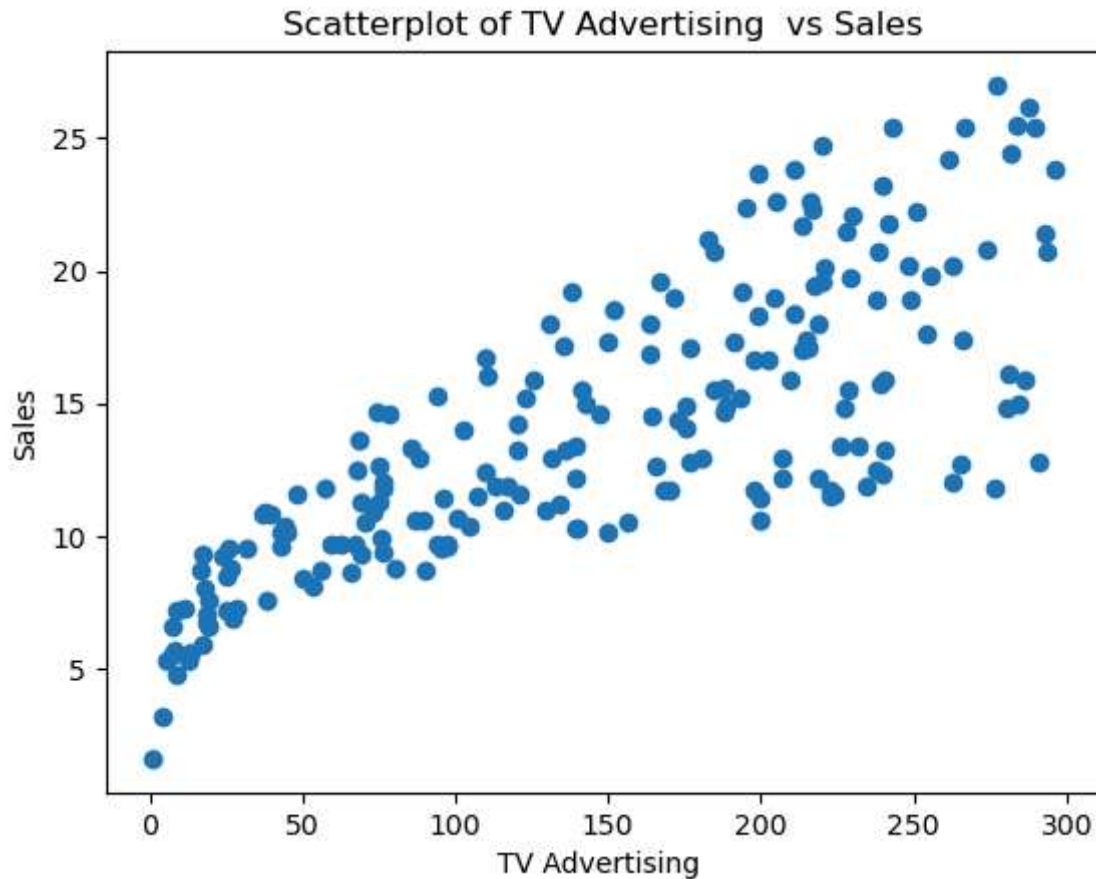
	Unnamed: 0	TV	radio	newspaper	sales
0	1	230.1	37.8	69.2	22.1
1	2	44.5	39.3	45.1	10.4
2	3	17.2	45.9	69.3	9.3
3	4	151.5	41.3	58.5	18.5
4	5	180.8	10.8	58.4	12.9
5	6	8.7	48.9	75.0	7.2
6	7	57.5	32.8	23.5	11.8
7	8	120.2	19.6	11.6	13.2
8	9	8.6	2.1	1.0	4.8
9	10	199.8	2.6	21.2	10.6

Question 4

Display a scatterplot with `TV` on the x-axis and `sales` on the y-axis.

```
In [19]: import matplotlib.pyplot as plt
plt.scatter(df['TV'],df['sales'])
plt.xlabel('TV Advertising ')
plt.ylabel('Sales')
plt.title('Scatterplot of TV Advertising vs Sales')
```

```
Out[19]: Text(0.5, 1.0, 'Scatterplot of TV Advertising vs Sales')
```



Question 5

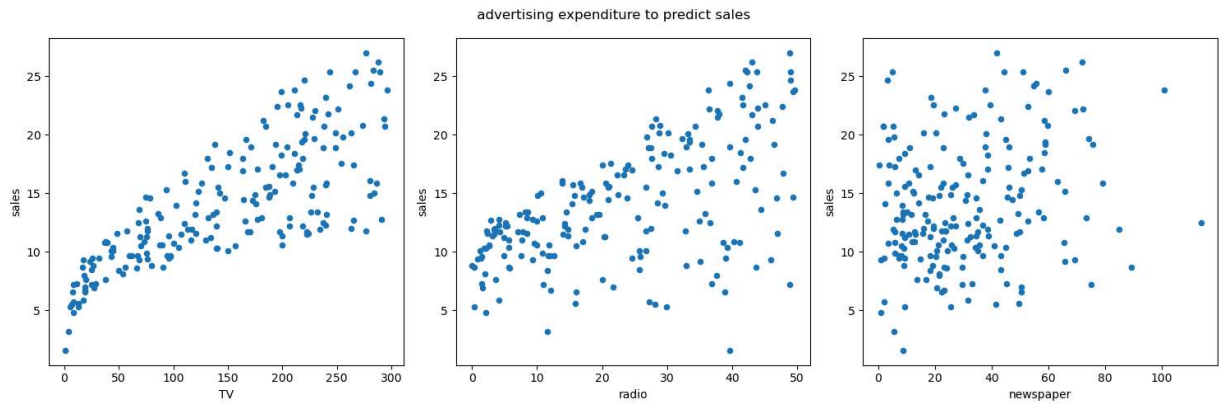
Display a set of three scatterplots in the same figure with each of the predictor variables TV, radio, and newspaper on the x-axes and sales on the y-axes.

```
In [ ]: # uncomment this line if you need to install matplotlib
        # !pip install -q matplotlib
```

[notice] A new release of pip is available: 23.2.1 -> 24.2

[notice] To update, run: `pip install --upgrade pip`

```
In [23]: fig,ax=plt.subplots(1,3,figsize=(15,5))
        for(i,col)in enumerate(['TV','radio','newspaper']):
            df.plot.scatter(x=col,y='sales',ax=ax[i])
        fig.suptitle(' advertising expenditure to predict sales ')
        plt.tight_layout()
```



Question 6

If we want to maximize sales, and have unlimited budget for a single medium, which medium would you allocate your budget to? Why? Does your answer change if budget is limited?

Answer: To maximize sales, I would most probably allocate my budget to TV medium if I have unlimited budget as it is evident from the plot that as we allocate more budget the sales have been higher. If the budget is limited I would probably allocate it to newspaper medium as there have been major sales with even budget as less as few 10s of dollars.

Question 7

From these scatterplots, is it possible to tell if there is an *interaction* between the advertising mediums and sales ?

Answer: I believe it is not possible to tell that there is an interaction between the advertising medium and sales with just the available information alone, maybe if and when we have some more variables we can find an interaction between the mediums and sales.