



Chapter 6 - exercise 2: Scores

Cho dữ liệu

Hãy thực hiện các yêu cầu sau:

```
raw_data = {'first_name': ['Jason', 'Molly', 'Tina', 'Jake', 'Amy'], 'last_name': ['Miller', 'Jacobson', 'Ali', 'Milner', 'Cooze'], 'female': [0, 1, 1, 0, 1], 'age': [42, 52, 36, 24, 73], 'preTestScore': [4, 24, 31, 2, 3], 'postTestScore': [25, 94, 57, 62, 70]}
```

1. Tạo DataFrame df từ dữ liệu trên. In df.
2. Câu 2: Tạo một Scatterplot của preTestScore và postTestScore, với kích thước s của mỗi điểm được xác định theo age. Thiết lập xlabel, ylabel và title.
3. Tạo một Scatterplot của preTestScore và postTestScore. Lần này kích thước s sẽ gấp 4.5 lần postTestScore. Màu c được xác định theo giới tính (df.female). Thiết lập xlabel, ylabel và title.

```
In [5]: import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

%matplotlib inline
```

```
In [6]: # Câu 1:
raw_data = {'first_name': ['Jason', 'Molly', 'Tina', 'Jake', 'Amy'],
            'last_name': ['Miller', 'Jacobson', 'Ali', 'Milner', 'Cooze'],
            'female': [0, 1, 1, 0, 1],
            'age': [42, 52, 36, 24, 73],
            'preTestScore': [4, 24, 31, 2, 3],
            'postTestScore': [25, 94, 57, 62, 70]}

df = pd.DataFrame(raw_data, columns = ['first_name', 'last_name', 'age', 'female',
df
```

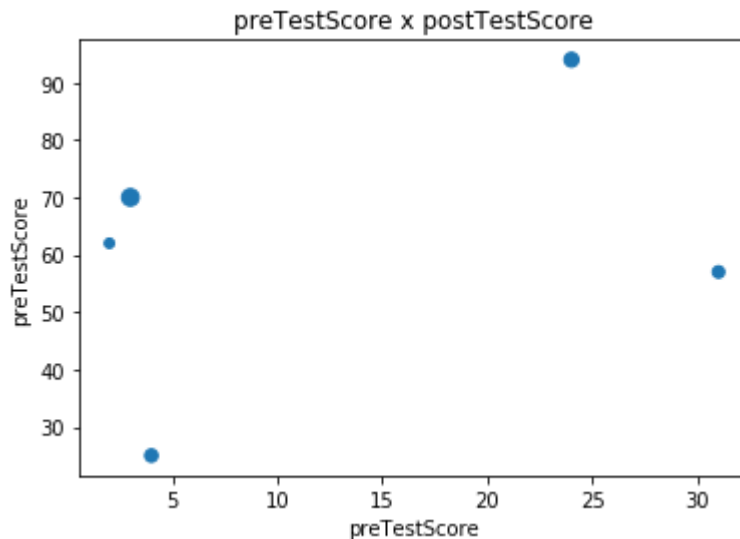
Out[6]:

	first_name	last_name	age	female	preTestScore	postTestScore
0	Jason	Miller	42	0	4	25
1	Molly	Jacobson	52	1	24	94
2	Tina	Ali	36	1	31	57
3	Jake	Milner	24	0	2	62
4	Amy	Cooze	73	1	3	70

```
In [7]: # Câu 2: Tạo một Scatterplot của preTestScore và postTestScore, với kích thước s c
plt.scatter(df.preTestScore, df.postTestScore, s=df.age)

# Thiết lập xlabel, ylabel và title.
plt.title("preTestScore x postTestScore")
plt.xlabel('preTestScore')
plt.ylabel('preTestScore')
```

Out[7]: Text(0,0.5,'preTestScore')



```
In [8]: # Câu 3: Tạo một Scatterplot của preTestScore và postTestScore. Lần này kích thước
# màu c được xác định theo giới tính (df.female)
plt.scatter(df.preTestScore, df.postTestScore, s= df.postTestScore * 4.5, c = df.f

# Thiết lập xlabel, ylabel và title
plt.title("preTestScore x postTestScore")
plt.xlabel('preTestScore')
plt.ylabel('preTestScore')
```

Out[8]: Text(0,0.5,'preTestScore')

