Scores

Introduction:

This time you will create the data.

Exercise based on Chris Albon work, the credits belong to him.

Step 1. Import the necessary libraries

In [1]:
```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns

%matplotlib inline
```

Step 2. Create the DataFrame that should look like the one below.

In [2]:
```
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', 'preTestScore', 'postTestScore'])

df
```

Out[2]:

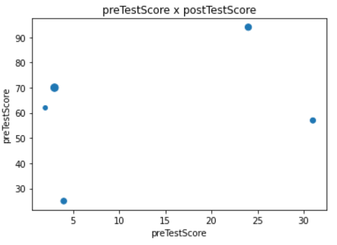
	first_name	last_name	age	female	preTestScore	postTestScore
0	Jason	Miller	42	0	4	25
1	Molly	Jacobson	52	1	24	94
2	Tina	Ali	36	1	31	57
3	Jake	Milner	24	0	2	62
4	Amy	Cooze	73	1	3	70

Step 3. Create a Scatterplot of preTestScore and postTestScore, with the size of each point determined by age

Hint: Don't forget to place the labels

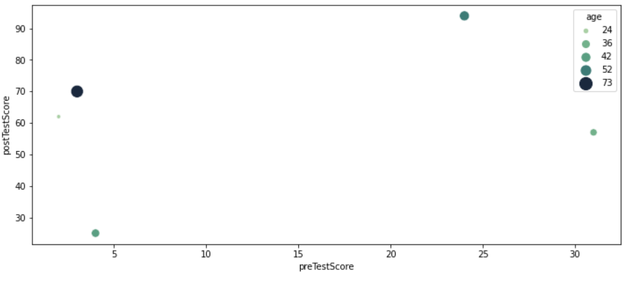
In [3]:
```
plt.scatter(df.preTestScore, df.postTestScore, s=df.age)

#set labels and titles
plt.title("preTestScore x postTestScore")
plt.xlabel('preTestScore')
plt.ylabel('preTestScore')
```

Out[3]:
```
Text(0, 0.5, 'preTestScore')
```



In [4]:
```
plt.figure(figsize=(12,5))
sns.scatterplot(data=df, x="preTestScore", y="postTestScore", hue="age",palette="ch:r=-.5,l=.75",size="age",
    sizes=(20, 200), legend="full")
```

Out[4]:
```
<AxesSubplot:xlabel='preTestScore', ylabel='postTestScore'>
```

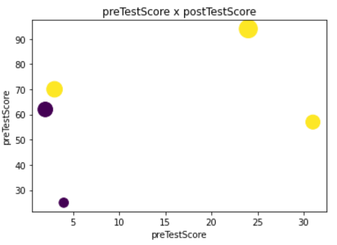


Step 4. Create a Scatterplot of preTestScore and postTestScore.

This time the size should be 4.5 times the postTestScore and the color determined by sex

In [5]:
```
plt.scatter(df.preTestScore, df.postTestScore, s= df.postTestScore * 4.5, c = df.female)

#set labels and titles
plt.title("preTestScore x postTestScore")
plt.xlabel('preTestScore')
plt.ylabel('preTestScore')
```

Out[5]:
```
Text(0, 0.5, 'preTestScore')
```



In [6]:
```
plt.figure(figsize=(12,5))
sns.scatterplot(data=df, x="preTestScore", y="postTestScore", hue="female",palette="ch:r=-.5,l=.75",size="female",
    sizes=(20, 200), legend="full")
```

Out[6]:
```
<AxesSubplot:xlabel='preTestScore', ylabel='postTestScore'>
```