Exercise 1:

Create a numpy array containing the numbers from 1 to 10, and then reshape it to a 2x5 matrix.

In [69]:
```
import numpy as np
arr = np.arange(1, 11)
n = arr.reshape(2, 5)
print(n)
```

```
[[ 1  2  3  4  5]
 [ 6  7  8  9 10]]
```

Exercise 2:

Create a numpy array containing the numbers from 1 to 20, and then extract the elements between the 5th and 15th index.

In [12]:
```
arr=np.array(list(range(1,21)))
n=arr[5:15]
print(n)
```

```
[ 6  7  8  9 10 11 12 13 14 15]
```

Exercise 3:

Create a Pandas series with the following data: {'apples': 3, 'bananas': 2, 'oranges': 1}. Then, add a new item to the series with the key 'pears' and the value 4.

In [8]:
```
import pandas as pd
data= {'apples': 3, 'bananas': 2, 'oranges': 1}
n=pd.Series( data)
n['pears']=4
print(n)
```

```
apples     3
bananas    2
oranges    1
pears      4
dtype: int64
```

Exercise 4:

Create a dataframe with the following columns: name, age, and gender. The dataframe should have 10 rows of data.

In [9]:
```
import pandas as pd
data={
    'name':['Gopika','Sreekanth','Siya','Pregath','Geethika','bilha','Shidin','Sankeerth','Athul','Abhinav'],
    'age':[31,28,30,26,33,25,35,30,30,20],
    'gender':['Female','Male','Female','Male','Female','Female','Male','Male','Male','Male']
}
df=pd.DataFrame(data)
df
```

Out[9]:

	name	age	gender
0	Gopika	31	Female
1	Sreekanth	28	Male
2	Siya	30	Female
3	Pregath	26	Male
4	Geethika	33	Female
5	bilha	25	Female
6	Shidin	35	Male
7	Sankeerth	30	Male
8	Athul	30	Male
9	Abhinav	20	Male

Exercise 5:

Add a new column to the data frame created in question 1, called occupation. The values for this column should be Programmer, Manager, and Analyst, corresponding to the rows in the dataframe.

In [47]:
```
occupation=['Programmer','Manager','Analyst','Programmer','Manager','Analyst','Programmer','Manager','Analyst','Progr
df['occupation']=occupation
df
```

Out[47]:

	name	age	gender	occupation
0	Gopika	31	Female	Programmer
1	Sreekanth	28	Male	Manager
2	Siya	30	Female	Analyst
3	Pregath	26	Male	Programmer
4	Geethika	33	Female	Manager
5	bilha	25	Female	Analyst
6	Shidin	35	Male	Programmer
7	Sankeerth	30	Male	Manager
8	Athul	30	Male	Analyst
9	Abhinav	20	Male	Programmer

Exercise 6:

Select the rows of the dataframe where the age is greater than or equal to 30.

In [55]:
```
df[df['age']>=30]
```

Out[55]:

	name	age	gender	occupation
0	Gopika	31	Female	Programmer
2	Siya	30	Female	Analyst
4	Geethika	33	Female	Manager
6	Shidin	35	Male	Programmer
7	Sankeerth	30	Male	Manager
8	Athul	30	Male	Analyst

Exercise 7:

Convert this dataframe to a csv file and read that csv file, finally display the contents.

In [66]:
```
file = 'dataframe.csv'
df.to_csv(file,index=False)
read=pd.read_csv(file)
read
```

Out[66]:

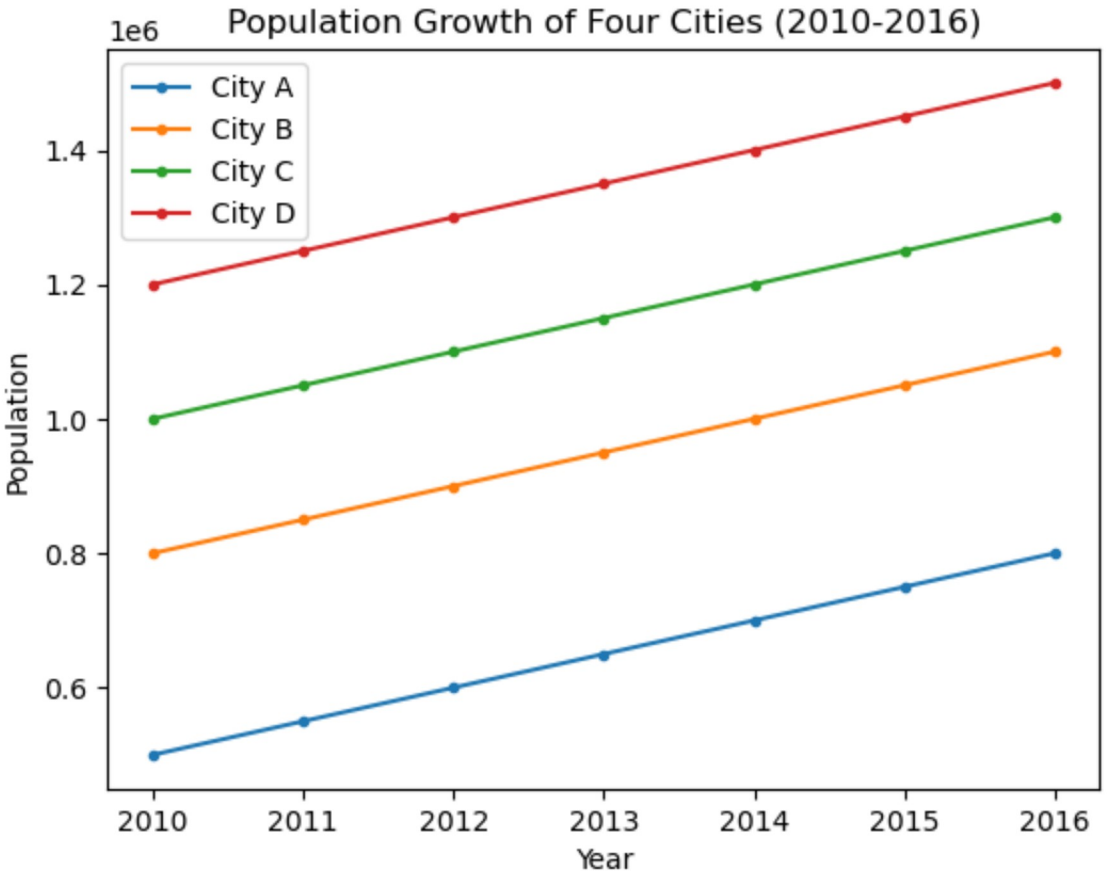
	name	age	gender	occupation
0	Gopika	31	Female	Programmer
1	Sreekanth	28	Male	Manager
2	Siya	30	Female	Analyst
3	Pregath	26	Male	Programmer
4	Geethika	33	Female	Manager
5	bilha	25	Female	Analyst
6	Shidin	35	Male	Programmer
7	Sankeerth	30	Male	Manager
8	Athul	30	Male	Analyst
9	Abhinav	20	Male	Programmer

Exercise 8:

Create a line plot using matplotlib pyplot that displays the population of four different cities over time. Each city should have its own line, and the x-axis should represent years (e.g. 2010, 2011, 2012, etc.) while the y-axis should represent the population. The data for the four cities is provided below:

City A: [500000, 550000, 600000, 650000, 700000, 750000, 800000] City B: [800000, 850000, 900000, 950000, 1000000, 1050000, 1100000] City C: [1000000, 1050000, 1100000, 1150000, 1200000, 1250000, 1300000] City D: [1200000, 1250000, 1300000, 1350000, 1400000, 1450000, 1500000]

In [66]:
```
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
x=['City A','City B','City C','City D']
year=[2010,2011,2012,2013,2014,2015,2016]
data={' A':[500000, 550000, 600000, 650000, 700000, 750000, 800000],
    ' B':[800000, 850000, 900000, 950000, 1000000, 1050000, 1100000],
    ' C':[1000000, 1050000, 1100000, 1150000, 1200000, 1250000, 1300000],
    ' D':[1200000, 1250000, 1300000, 1350000, 1400000, 1450000, 1500000]}
d=pd.DataFrame(data,index=year)
plt.plot(year,d,marker='.',label=x)
plt.xlabel('Year')
plt.ylabel('Population')
plt.title('Population Growth of Four Cities (2010-2016)')
plt.legend()
plt.show()
```

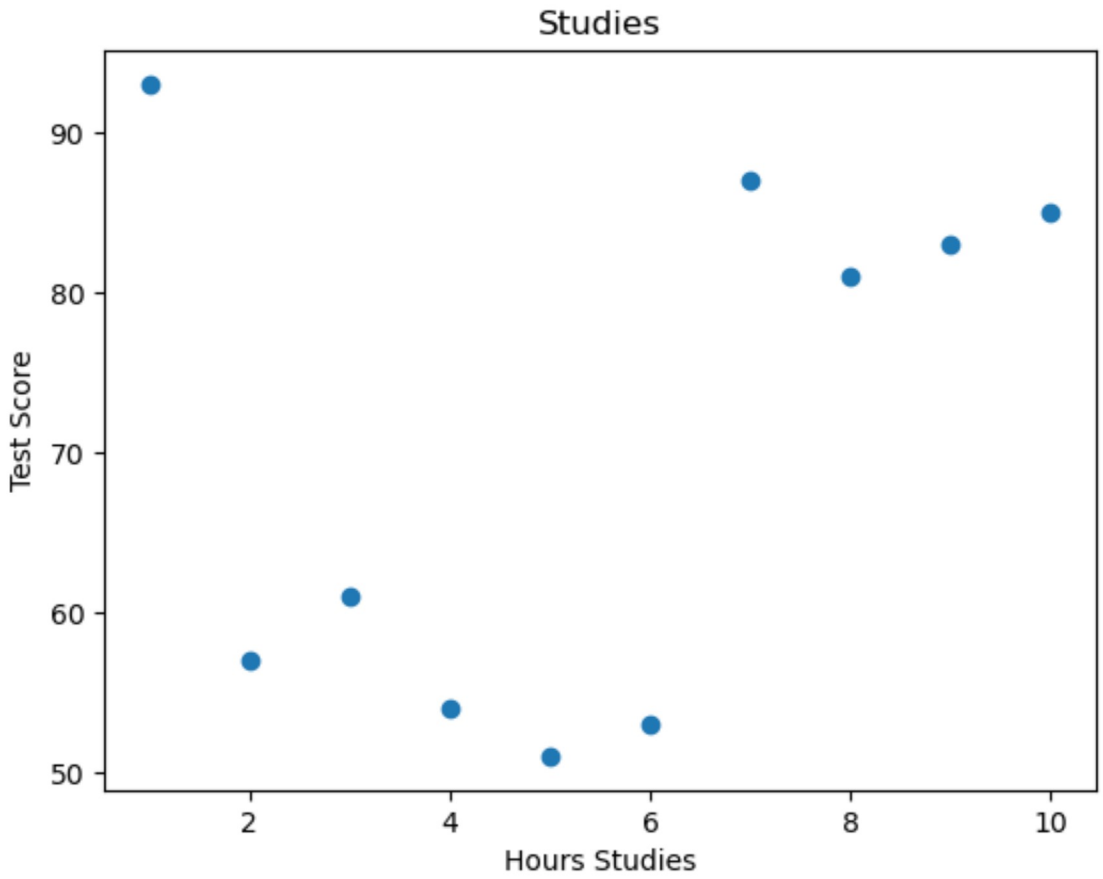


Exercise 9:

Create a scatter plot using seaborn that shows the relationship between the number of hours studied and the test scores obtained by a group of students. Use the following data: Hours Studied: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10] Test Scores: [93, 57, 61, 54, 51, 53, 87, 81, 83, 85]

In [51]:
```
import pandas as pd
import matplotlib.pyplot as plt
s=[1,2,3,4,5,6,7,8,9,10],
e=[93, 57, 61, 54, 51, 53, 87, 81, 83, 85]
plt.scatter(s,e)
plt.title("Studies")
plt.xlabel('Hours Studies')
```

```
plt.ylabel('Test Score')
plt.show()
```

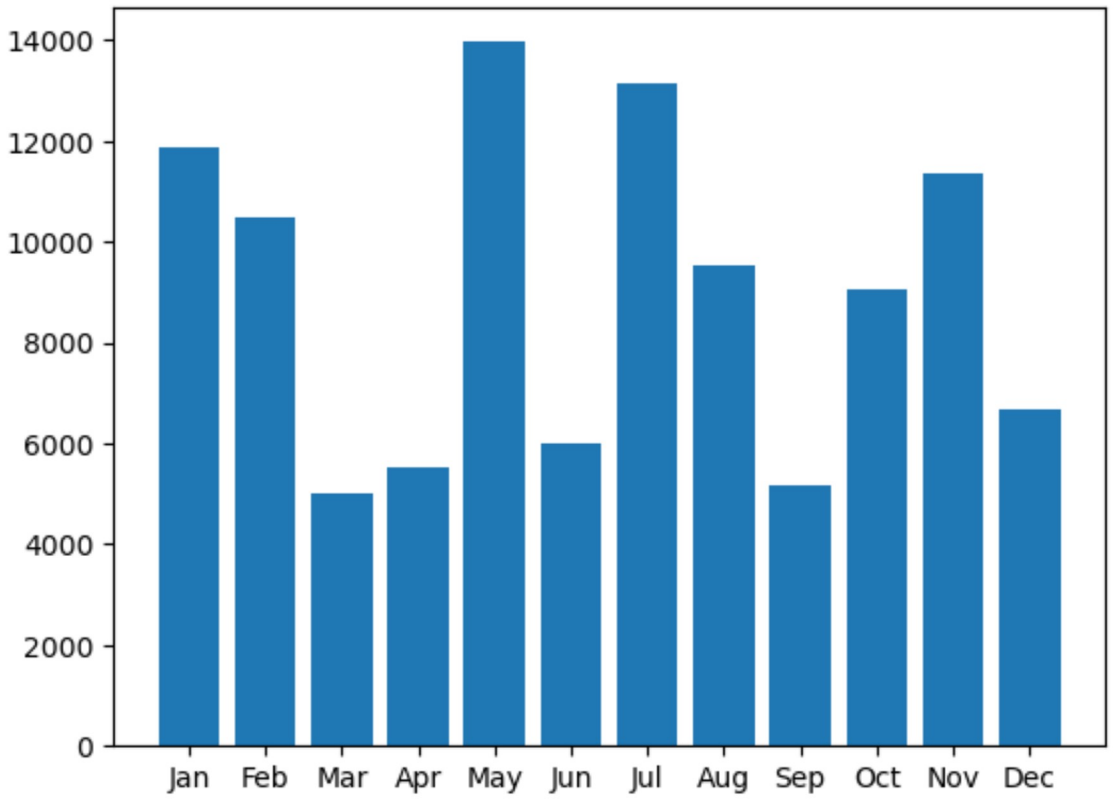


Exercise 10:

Create a bar chart using matplotlib pyplot that shows the total sales for each month of the year. Use the following data: Month: ["Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug", "Sep", "Oct", "Nov", "Dec"] Sales: [11860, 10480, 4997, 5523, 13965, 6011, 13158, 9533, 5158, 9058, 11346, 6675]

In [20]:
```
import matplotlib.pyplot as plt
Month= ["Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug", "Sep", "Oct", "Nov", "Dec"]
Sales= [11860, 10480, 4997, 5523, 13965, 6011, 13158, 9533, 5158, 9058, 11346, 6675]
plt.bar(Month,Sales)
```

```
plt.show()
```


14000
12000
10000
8000
6000
4000
2000
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

In [ ]: