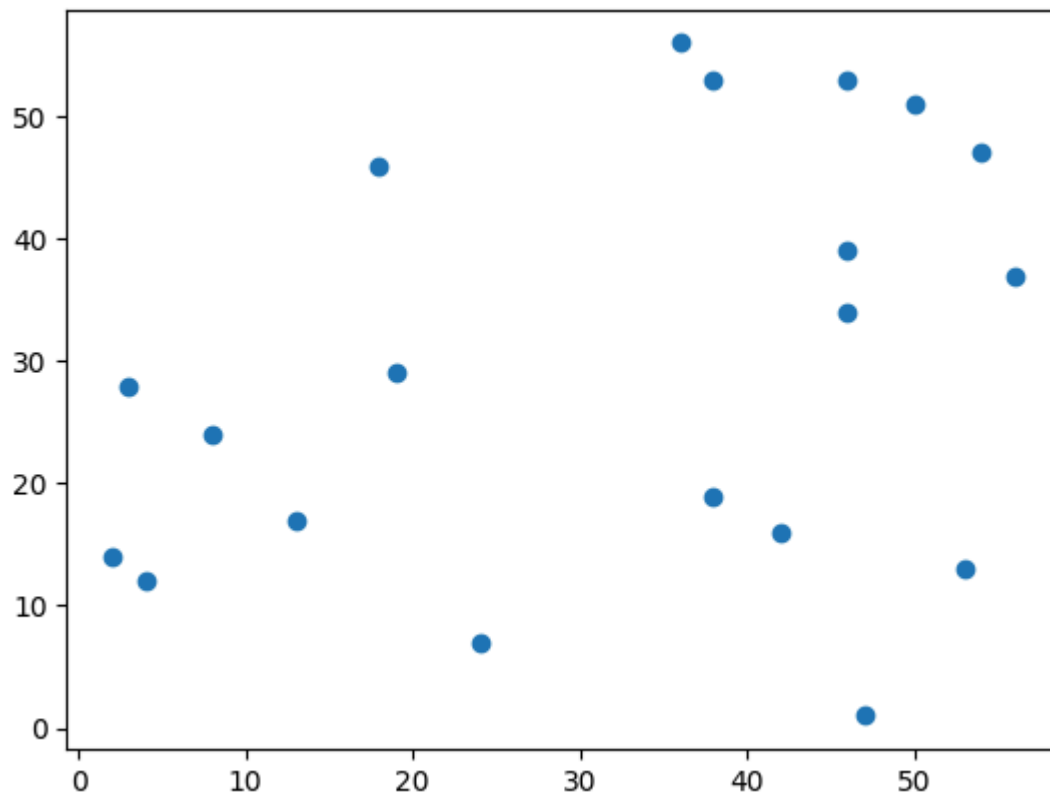


```
In [ ]: import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
from sklearn.cluster import KMeans
```

```
In [ ]: # data
np.random.seed(30)
x = np.random.randint(1,60,20)
y = np.random.randint(1,60,20)
data = list(zip(x,y))
print(data)
plt.scatter(x,y)
```

[(38, 19), (38, 53), (46, 53), (46, 39), (53, 13), (13, 17), (24, 7), (3, 28), (54, 47), (18, 46), (47, 1), (4, 12), (42, 16), (8, 24), (56, 37), (2, 14), (50, 51), (46, 34), (36, 56), (19, 29)]

```
Out[ ]: <matplotlib.collections.PathCollection at 0x1dcd0683b90>
```



```
In [ ]: kmeans = KMeans(n_clusters=4, random_state=0, n_init="auto").fit(data)
```

```
c:\Users\user\AppData\Local\Programs\Python\Python311\Lib\site-packages\joblib\externals\loky\backend\context.py:136: UserWarni
ng: Could not find the number of physical cores for the following reason:
[WinError 2] The system cannot find the file specified
Returning the number of logical cores instead. You can silence this warning by setting LOKY_MAX_CPU_COUNT to the number of core
s you want to use.
  warnings.warn(
    File "c:\Users\user\AppData\Local\Programs\Python\Python311\Lib\site-packages\joblib\externals\loky\backend\context.py", line
257, in _count_physical_cores
        cpu_info = subprocess.run(
            ^^^^^^^^^^^^^^^^^^^
    File "c:\Users\user\AppData\Local\Programs\Python\Python311\Lib\subprocess.py", line 548, in run
        with Popen(*popenargs, **kwargs) as process:
            ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
    File "c:\Users\user\AppData\Local\Programs\Python\Python311\Lib\subprocess.py", line 1024, in __init__
        self._execute_child(args, executable, preexec_fn, close_fds,
    File "c:\Users\user\AppData\Local\Programs\Python\Python311\Lib\subprocess.py", line 1509, in _execute_child
        hp, ht, pid, tid = _winapi.CreateProcess(executable, args,
            ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
```

```
In [ ]: kmeans.labels_
```

```
Out[ ]: array([0, 3, 2, 2, 0, 1, 1, 1, 2, 3, 0, 1, 0, 1, 2, 1, 2, 2, 3, 1])
```

```
In [ ]: centers = kmeans.cluster_centers_
        centers
```

```
Out[ ]: array([[45.          , 12.25         ],
               [10.42857143, 18.71428571],
               [49.66666667, 43.5          ],
               [30.66666667, 51.66666667]])
```

In []:

```
Out[ ]: array([12.25      , 18.71428571, 43.5      , 51.66666667])
```

```
In [ ]: plt.scatter(x,y,c=kmeans.labels_)
plt.scatter(centers[:,0],centers[:,1], marker="*",s=100 , label="Centers")
plt.legend(loc="best")
```

```
Out[ ]: <matplotlib.legend.Legend at 0x1dcd5a3ddd0>
```

