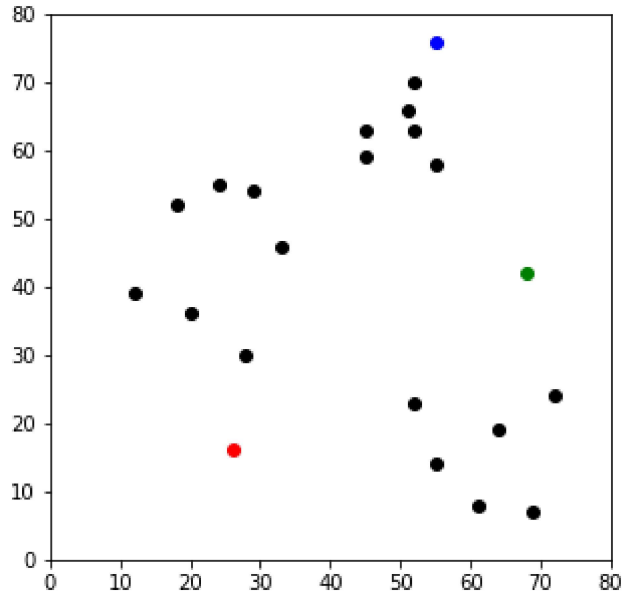


```
In [18]: import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
```

```
In [19]: df = pd.DataFrame({
    'x':[12,20,28,18,29,33,24,45,45,52,51,52,55,52,55,61,64,69,72],
    'y':[39,36,30,52,54,46,55,59,63,70,66,63,58,23,14,8,19,7,24]
})
```

```
In [20]: np.random.seed(200)
k=3
centroids = {
    i+1:[np.random.randint(0,80),np.random.randint(0,80)]
    for i in range(k)
}
```

```
In [21]: fig=plt.figure(figsize=[5,5])
plt.scatter(df['x'],df['y'],color='k')
colmap={1:'r',2:'g',3:'b'}
for i in centroids.keys():
    plt.scatter(*centroids[i],color=colmap[i])
plt.xlim(0,80)
plt.ylim(0,80)
plt.show()
```



```
In [22]: from sklearn.cluster import KMeans
```

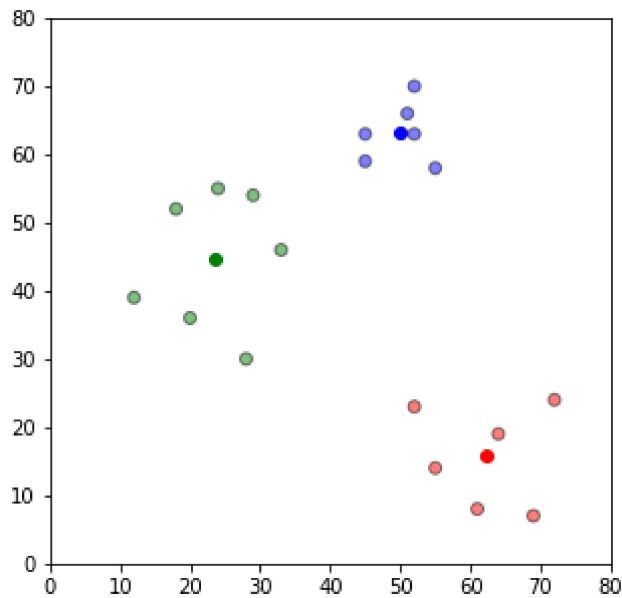
```
In [23]: kmeans=KMeans(n_clusters=3)
kmeans.fit(df)
```

```
Out[23]: KMeans(n_clusters=3)
```

```
In [24]: labels = kmeans.predict(df)
```

```
In [25]: centroids = kmeans.cluster_centers_
```

```
In [26]: fig = plt.figure(figsize=(5,5))
colors = map(lambda x: colormap[x+1],labels)
colors1 = list(colors)
plt.scatter(df['x'],df['y'],color=colors1,alpha=0.5,edgecolor='k')
for idx, centroid in enumerate(centroids):
    plt.scatter(*centroid,color=colormap[idx+1])
plt.xlim(0,80)
plt.ylim(0,80)
plt.show()
```



```
In [ ]:
```