In [1]:
```python
#import packages
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

#read dataset
dataset = pd.read_csv("Kmeans_dataset.csv")
print(dataset)
```

```
      X     Y
0  0.10  0.60
1  0.15  0.71
2  0.08  0.90
3  0.16  0.85
4  0.20  0.30
5  0.25  0.50
6  0.24  0.10
7  0.30  0.20
```

In [32]:
```python
f1 = dataset['X'].values
f2 = dataset['Y'].values
```

In [35]:
```python
f2
```

Out[35]:
```
array([0.6 , 0.71, 0.9 , 0.85, 0.3 , 0.5 , 0.1 , 0.2 ])
```

In [33]:
```python
f11 = dataset[['X','Y']].values
f11
```

Out[33]:
```
array([[0.1 , 0.6 ],
       [0.15, 0.71],
       [0.08, 0.9 ],
       [0.16, 0.85],
       [0.2 , 0.3 ],
       [0.25, 0.5 ],
       [0.24, 0.1 ],
       [0.3 , 0.2 ]])
```

In [36]:
```python
# initial centroid points
centers = np.array([[0.1,0.6],[0.3,0.2]])
print(centers)
```

```
[[0.1 0.6]
 [0.3 0.2]]
```

In [37]:
```python
# Apply K-Means Clustering
from sklearn.cluster import KMeans
model = KMeans(n_clusters=2, init=centers, n_init=1)
# n_clusters = number of clusters
# init = initial centroids
# n_init = number of init parameter

#train the algorithm
model.fit(f11)
```

```
# print labels
print(model.labels_)


# output : [0 0 0 0 1 0 1 1]
# 0 --> cluster no. 1 and 1 --> cluster no. 2
```

```
[0 0 0 0 1 0 1 1]
```

In [38]:
```
# population around cluster 2
print(np.count_nonzero(model.labels_ == 1))
```
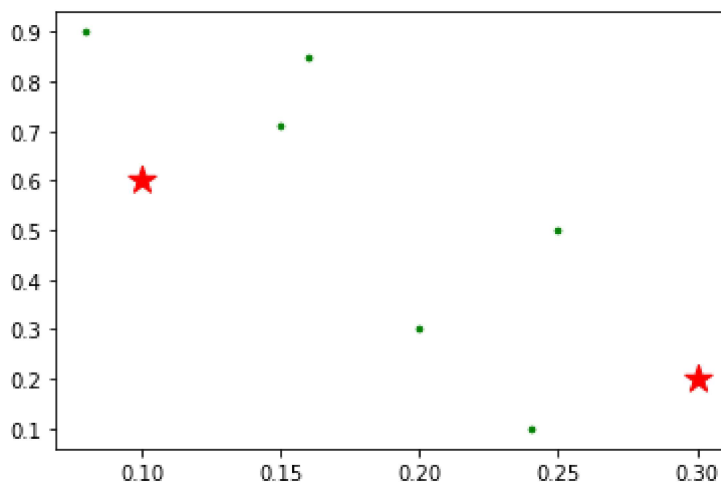
```
3
```

In [39]:
```
# new centroids
print(model.cluster_centers_)
```

```
[[0.148      0.712     ]
 [0.24666667 0.2       ]]
```
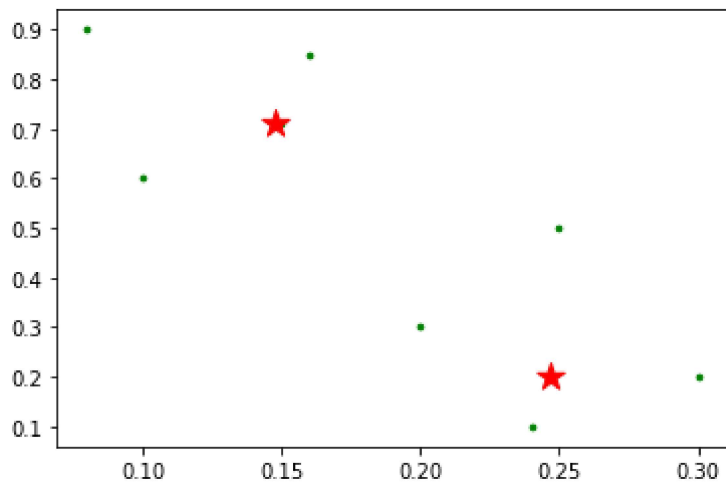
In [40]:
```
# Cluster with initial centroids
C_x = np.array([0.1, 0.3])
C_y = np.array([0.6, 0.2])
plt.scatter(f1, f2, c='green', s=7)
plt.scatter(C_x, C_y, marker='*', s=200, c='r')
```

Out[40]: `<matplotlib.collections.PathCollection at 0x1c9c379e6b0>`

In [41]:
```
# cluster with new centroids
N_x = np.array([model.cluster_centers_[0][0], model.cluster_centers_[
N_y = np.array([model.cluster_centers_[0][1],model.cluster_centers_[1
plt.scatter(f1, f2, c='green', s=7)
plt.scatter(N_x, N_y, marker='*', s=200, c='r')
```

Out[41]: `<matplotlib.collections.PathCollection at 0x1c9c37f5c90>`

In [ ]: