Q4

August 13, 2025

[1]:
```
import matplotlib.pyplot as plt
from sklearn.datasets import make_blobs
from sklearn.cluster import KMeans
```

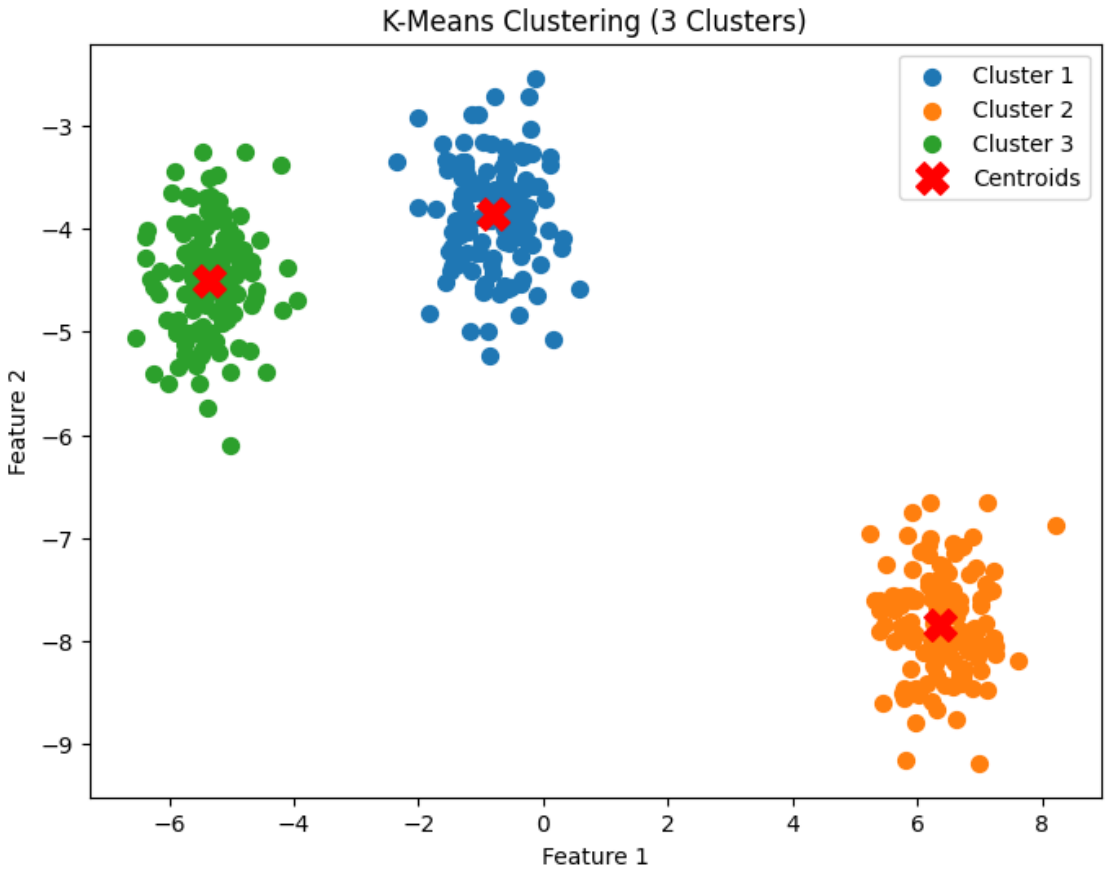
[2]:
```
X, y_true = make_blobs(
    n_samples=400,        # total number of points
    centers=3,            # number of clusters to generate
    cluster_std=0.50,     # standard deviation of clusters
    random_state=35       # reproducibility
)
```

[3]:
```
kmeans = KMeans(n_clusters=3, random_state=42)
kmeans.fit(X)
y_kmeans = kmeans.predict(X)
```

[4]:
```
plt.figure(figsize=(8, 6))
for cluster in range(3):
    plt.scatter(
        X[y_kmeans == cluster, 0],
        X[y_kmeans == cluster, 1],
        s=50,
        label=f"Cluster {cluster+1}"
    )

centers = kmeans.cluster_centers_
plt.scatter(
    centers[:, 0], centers[:, 1],
    c='red', s=200, marker='X', label='Centroids'
)

# Titles and labels
plt.title("K-Means Clustering (3 Clusters)")
plt.xlabel("Feature 1")
plt.ylabel("Feature 2")
plt.legend()
plt.show()
```


K-Means Clustering (3 Clusters)
Cluster 1
Cluster 2
Cluster 3
Centroids
Feature 1
Feature 2
−6
−4
−2
0
2
4
6
8
−3
−4
−5
−6
−7
−8
−9

[ ]:

[ ]:

[ ]:

[ ]: