1. a)

K-Means Cluster Centers:

[[3.34338961 28.91328961]

 [33.73770667 -5.85638333]

 [-30.41868947 -30.82280526]

 [-16.8944127   29.74264762]

 [35.10477273 19.25536364]]

b)

GMM Means:

[[3.34338961 28.91328961]

 [33.73770667 -5.85638333]

 [-30.41868947 -30.82280526]

 [-16.8944127   29.74264762]

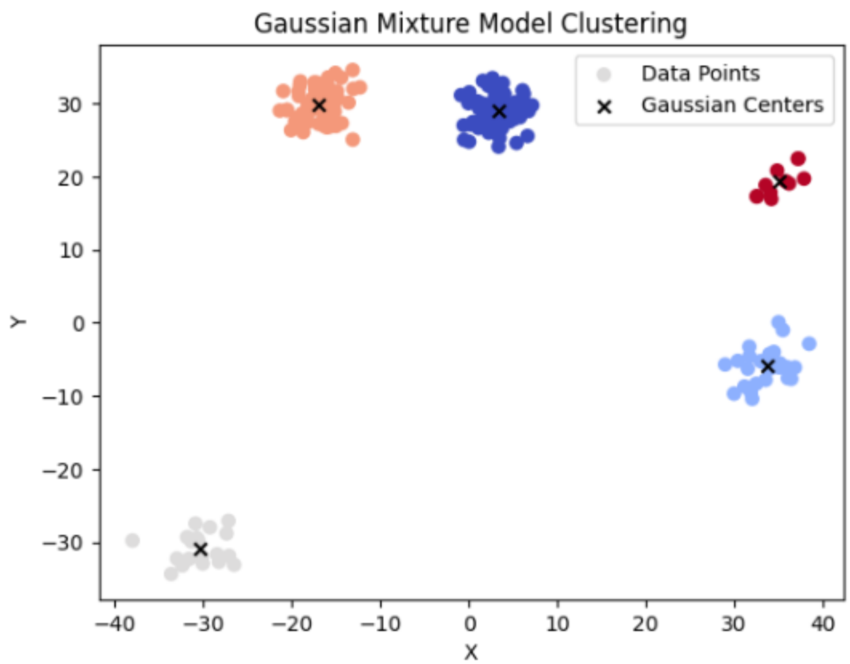
 [35.10477273 19.25536364]]

GMM Covariances:

```
[[[3.64869209 0.32408908]
  [0.32408908 4.20562373]]

 [[4.73082307 1.69561065]
  [1.69561065 5.27469359]]

 [[7.32130237 0.57280891]
  [0.57280891 4.37376756]]

 [[4.22140252 0.91010659]
  [0.91010659 4.90107045]]

 [[3.25874584 2.59189484]
  [2.59189484 3.50090514]]]
```

GMM Weights:

```
[0.385 0.15  0.095 0.315 0.055]
```



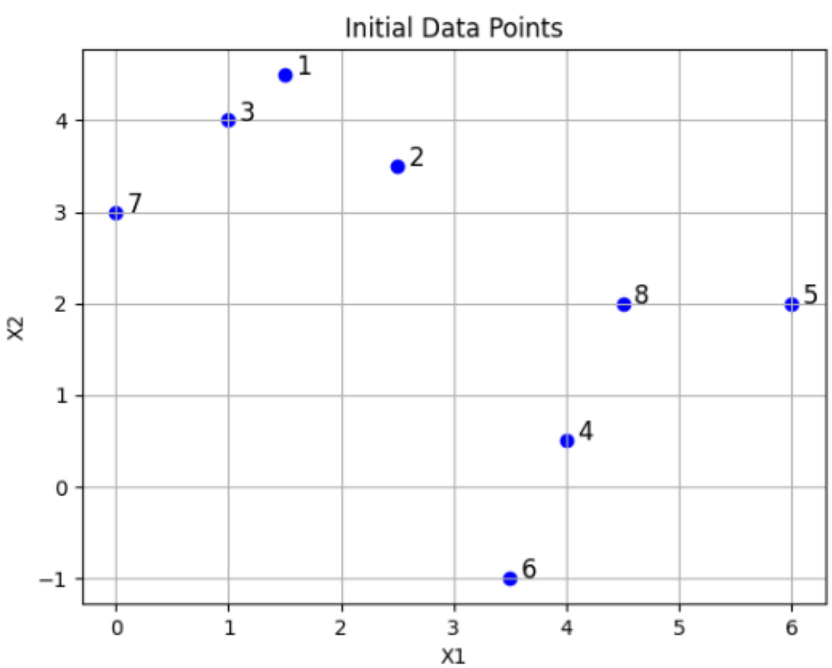
Code :

```
import pandas as pd

import numpy as np

from sklearn.cluster import KMeans

from sklearn.mixture import GaussianMixture

import matplotlib.pyplot as plt


# Load the data

file_path = 'Q1-Data.xlsx'

data = pd.read_excel(file_path)


# Extract the X and Y columns

X = data[['X', 'Y']].values


# K-means clustering
```

```
kmeans = KMeans(n_clusters=5, random_state=0).fit(X)
kmeans_centers = kmeans.cluster_centers_
kmeans_labels = kmeans.labels_

print("K-Means Cluster Centers:")
print(kmeans_centers)

# Plotting K-Means results
plt.scatter(X[:, 0], X[:, 1], c=kmeans_labels, cmap='viridis', marker='o', label='Data Points')
plt.scatter(kmeans_centers[:, 0], kmeans_centers[:, 1], c='red', marker='x', label='Cluster Centers')
plt.title("K-Means Clustering")
plt.xlabel("X")
plt.ylabel("Y")
plt.legend()
plt.show()

# Gaussian Mixture Model
gmm = GaussianMixture(n_components=5, random_state=0).fit(X)
gmm_labels = gmm.predict(X)

print("\nGMM Means:")
print(gmm.means_)
print("\nGMM Covariances:")
print(gmm.covariances_)
print("\nGMM Weights:")
```

```
print(gmm.weights_)

# Plotting GMM results

plt.scatter(X[:, 0], X[:, 1], c=gmm_labels, cmap='coolwarm', marker='o', label='Data Points')

plt.scatter(gmm.means_[:, 0], gmm.means_[:, 1], c='black', marker='x', label='Gaussian Centers')

plt.title("Gaussian Mixture Model Clustering")

plt.xlabel("X")

plt.ylabel("Y")

plt.legend()

plt.show()
```

2. a)



~~f~~)

② b)

cluster 1

	x_1	x_2	Cluster label
1	1.5	4.5	1
2	2.5	3.5	1
3	1	4	1
7	0	3	1

cluster 1

	x_1	x_2	Cluster label
4	4	0.5	2
5	6	2	2
6	3.5	-1	2
8	4.5	2	2

c) Cluster 1 →

$$\begin{bmatrix} \dfrac{1.5+2.5+1+0}{4} \\ \\ \dfrac{4.5+3.5+4+3}{4} \end{bmatrix} = \begin{bmatrix} 5/4 \\ 15/4 \end{bmatrix}$$

$$= \begin{bmatrix} 1.25 \\ 3.75 \end{bmatrix}$$

$$\text{cluster } 2 \rightarrow \begin{bmatrix} \dfrac{4+6+3.5+4.5}{4} \\ \\ \dfrac{0.5+2-1+2}{4} \end{bmatrix} = \begin{bmatrix} 18/4 \\ 3.5/4 \end{bmatrix} = \begin{bmatrix} 4.5 \\ 0.875 \end{bmatrix}$$

d)

$$1 \quad (1.5, 4.5) \quad \sqrt{(1.5-1.25)^2 + (4.5-3.75)^2} = 0.79$$

$$2 \quad (2.5, 3.5) \quad \sqrt{(2.5-1.25)^2 + (3.5-3.75)^2} = 1.27$$

$$3 \quad (1, 4) \quad \sqrt{(1-1.25)^2 + (4-3.75)^2} = 0.35$$

$$7 \quad (0, 3) \quad \sqrt{(0-1.25)^2 + (3-3.75)^2} = 1.45$$

$$1 \quad (1.5, 4.5) \quad \sqrt{(1.5-4.5)^2 + (4.5-0.875)^2} = 4.7$$

$$2 \quad (2.5, 3.5) \quad \sqrt{(2.5-4.5)^2 + (3.5-0.875)^2} = 3.3$$

$$3 \quad (1, 4) \quad \sqrt{(1-4.5)^2 + (4-0.875)^2} = 4.69$$

$$7 \quad (0, 3) \quad \sqrt{(0-4.5)^2 + (3-0.875)^2} = 4.9$$

$\therefore$ distances between 1, 2, 3, 7 observations with cluster 1 center are the minimum.

$4 \quad (4, 0.5) = \sqrt{(4-1.25)^2 + (0.5-3.75)^2} = 4.25$

$5 \quad (6, 2) = \sqrt{(6-1.25)^2 + (2-3.75)^2} = 5.06$

$6 \quad (3.5, -1) = \sqrt{(3.5-1.25)^2 + (-1-3.75)^2} = 5.2$

$8 \quad (4.5, 2) = \sqrt{(4.5-1.25)^2 + (2-3.75)^2} = 3.69$

$4 \quad (4, 0.5) = \sqrt{(4-4.5)^2 + (0.5-0.875)^2} = 0.625$

$5 \quad (6, 2) = \sqrt{(6-4.5)^2 + (2-0.875)^2} = 1.875$

$6 \quad (3.5, -1) = \sqrt{(3.5-4.5)^2 + (-1-0.875)^2} = 2.125$

$8 \quad (4.5, 2) = \sqrt{(4.5-4.5)^2 + (2-0.875)^2} = 1.125$

distances between 4, 5, 6, 8 observations with center 2 are minimum.

$\therefore$ 1, 2, 3, 7 $\rightarrow$ cluster 1 $\rightarrow$ same centroid

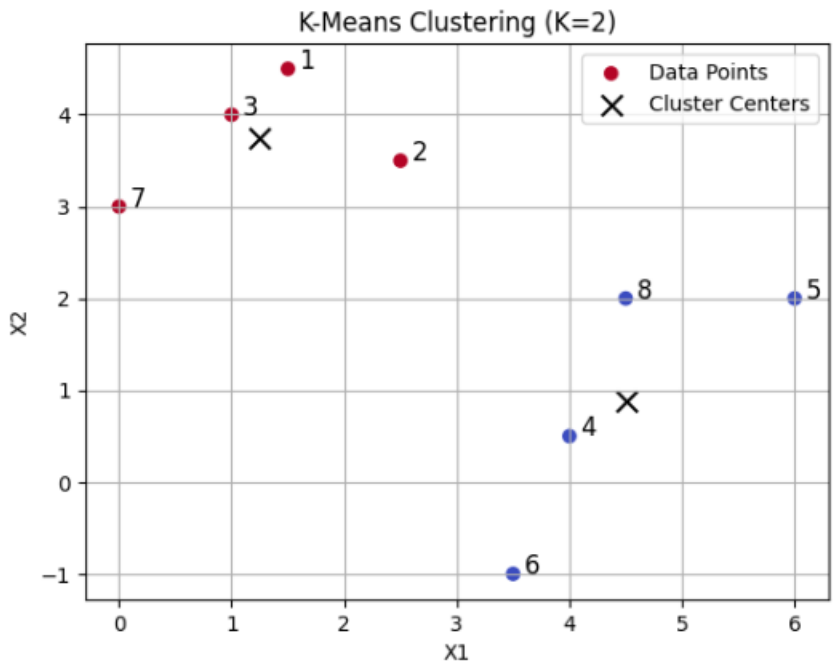
4, 5, 6, 8 $\rightarrow$ cluster 2 $\rightarrow$ same centroid.

e) center 1 $= \begin{bmatrix} \dfrac{1.5 + 2.5 + 1 + 0}{4} \\[2ex] \dfrac{4.5 + 3.5 + 4 + 3}{4} \end{bmatrix} = \begin{bmatrix} 1.25 \\ 3.75 \end{bmatrix}$

$$\text{center } 2 = \begin{bmatrix} \dfrac{4+6+3.5+4.5}{4} \\ \dfrac{0.5+2+(-1)+2}{4} \end{bmatrix} = \begin{bmatrix} 4.5 \\ 0.875 \end{bmatrix}$$

Since we are getting the same centroids and same clusters like in our initial values, we can stop the iterations here and accept these as final values. //

f)




```
Cluster Centers:
[[4.5   0.875]
 [1.25  3.75 ]]
```