1. Exercise K-means.

- Jhon and Hanna will be in Group-1
- Micheal and Lily will be in Group-2

(Initial clusters [k=2])

Cluster	Centroid		
	Age	Math Score	Student
k1	20	10	Jhon
k2	15	7	Micheal

Distance from cluster 1 (k1) and cluster 2 (k2):

Cluster	Centroid		
	Age	Math Score	Student
k1	0	5.83	Jhon
k2	5.83	0	Micheal

Dataset	Euclidean Distance		
(19,9)	k1	k2	Student
	1.414	4.472	Jhon

Dataset	Euclidean Distance		
(13,6)	k1	k2	Student
	7.382	2.236	Micheal

Final Clusters

Student	Age	Math Score	Cluster
Jhon	20	10	Group-1
Hanna	19	9	
Micheal	15	7	Group-2
Lily	13	6	

2. **Knn exercise**

X1	X2	Y	Distance	Rank	Closest neighbor?	value of Y classification
7	7	Out of range	3.15	4	No	
7	4	Out of range	2.73	2	Yes	
5	6	within range	2	1	Yes	
2.5	4.5	within range	2.81	3	Yes	
4	5	-	-	k=1	(5,6)	in range
				k=2	(5,6),(7,4)	Undefined
				k=3	(5,6),(7,4),(2.5,4.5)	in range

- For k=3, the nearest neighbors for X1=4 and X2=5 are (5,6),(7,4) and (2.5,4.5) and it is within range.

3. **Naïve Bayes**

i. Probability that a randomly selected person will use an iPhone
= (number of people using iphone)/(Total people)
= 5/10 = 0.5

ii. Probability that a person has a given iphone using a Mac laptop
= (number of people using iphone intersect Mac laptop)/
(total count of people)
=4/10 = 0.4

iii. probability that a random person uses a Mac
=(number of people using Mac)/(Total Count)
=6/10 = 0.6

iv. probability that someone uses an iPhone, since that person uses a Mac
= number of people using iphone / number of people using Mac
= 0.4/0.6 = 0.667