Dela Paz, Lance Kenneth F.

CoM 232

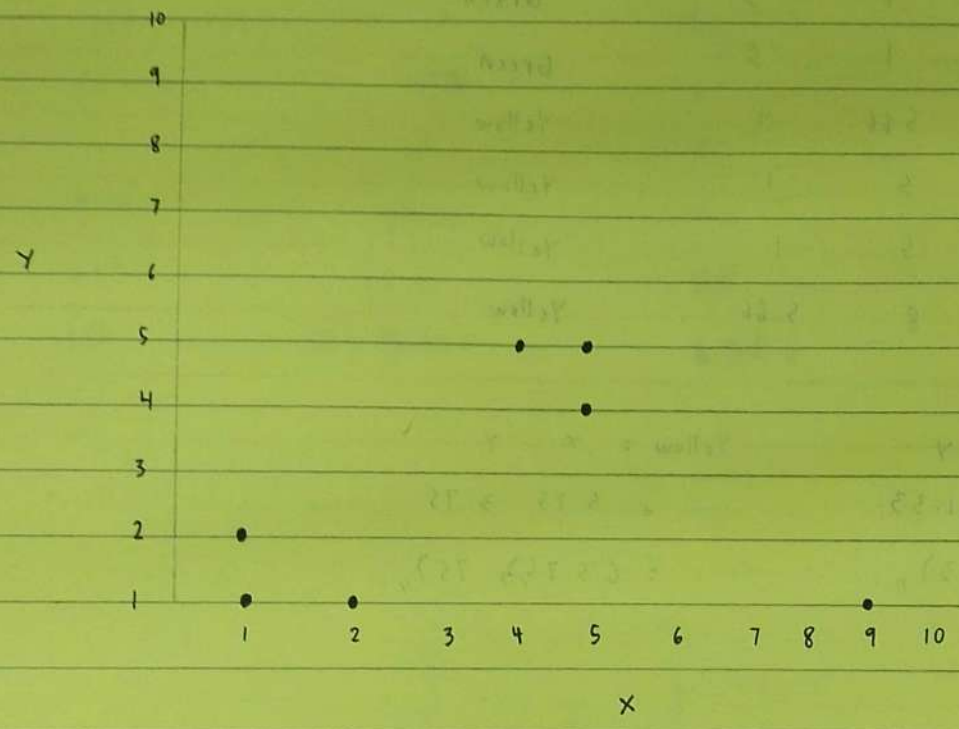
No. _____

Date _____

Midterm Exam

A.

1. Plot the data in a scatter plot



2. Compute all pairwise Euclidean distances between data points and put the distances in a full distance matrix

	A (1,1)	B (1,2)	C (2,1)	D (5,5)	E (4,5)	F (5,4)	G (9,1)
A (1,1)	0	1	1	5.66	5	5	8
B (1,2)	1	0	1.41	5	4.24	4.47	8.06
C (2,1)	1	1.41	0	5	4.47	4.24	7
D (5,5)	5.66	5	5	0	1	1	5.66
E (4,5)	5	4.24	4.47	1	0	1.41	6.40
F (5,4)	5	4.47	4.24	1	1.41	0	5
G (9,1)	8	8.06	7	5.66	6.40	5	0

B.

1.

Data points	x	y	d_1	d_2	Cluster
A	1	1	0	5.66	Green
B	1	2	1	5	Green
C	2	1	1	5	Green
D	5	5	5.66	0	Yellow
E	4	5	5	1	Yellow
F	5	4	5	1	Yellow
G	9	1	8	5.66	Yellow

2. Green = x y

= 1.33 1.33

= (1.33, 1.33)

Yellow = x y

= 5.75 3.75

= (5.75, 3.75)

3.

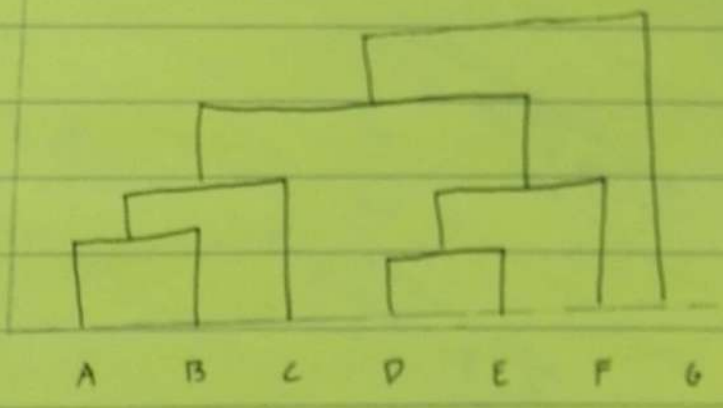
Data Points	x	y	d_1	d_2	Cluster
A	1	1	0.47	5.49	Green
B	1	2	0.75	5.06	Green
C	2	1	0.75	4.65	Green
D	5	5	5.19	1.46	Yellow
E	4	5	4.54	2.15	Yellow
F	5	4	4.54	0.79	Yellow
G	9	1	7.68	4.26	Yellow

No.
Date

6. 1.

Merge order	cluster 1	cluster 2	distance
1st	A	B	1
2nd	AB	C	1
3rd	P	E	1
4th	DE	F	1
5th	ABC	G	4.24
6th	ABCDEF	G	5

2.



D.

eps = 1, min_samples = 2

Data Point	Neighbors	Number of Neighbor	Point type	Cluster
A	B C	2	Core	Green
B	A	1	Non - Core	Green
C	A	1	Non - Core	Green
D	E F	2	Core	Blue
E	D	1	Non - Core	Blue
F	D	1	Non - Core	Blue
G		0	Noise	Red