1. Get Euclidean distance:

	1	2	3	4	5	6	7
1	0						
2	0.72	0					
3	0.66	0.12	0				
4	0.96	0.71	0.6	0			
5	0.75	0.82	0.7	0.37	0		
6	1.03	0.49	0.44	0.38	0.7	0	
7	0.25	0.67	0.57	0.73	0.5	0.86	0

a. Single:

$d_{2\to3}$ is the smallest 0.12. We group these two. Compute all min distance from other values to (2,3). The next one is (1,7) with dist at 0.25. Repeat same steps for this also

	1	2,3	4	5	6	7
1	0					
2,3	0.66	0				
4	0.96	0.6	0			
5	0.75	0.7	0.37	0		
6	1.03	0.44	0.38	0.7	0	
7	0.25	0.57	0.73	0.5	0.86	0

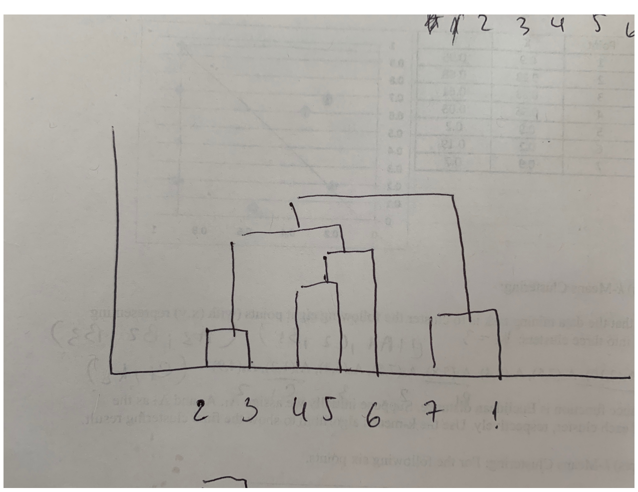
	1,7	2,3	4	5	6
1,7	0				
2,3	0.57	0			
4	0.73	0.6	0		
5	0.5	0.7	0.37	0	
6	0.86	0.44	0.38	0.7	0

Next, (4,5) has the smallest distance. Repeat steps. Then (4,5),6 with min distance 0.38. Then we group (2,3) with (4,5),6

	1,7	2,3	4,5	6
1,7	0			
2,3	0.57	0		
4,5	0.5	0.6	0	
6	0.86	0.44	0.38	0

	1,7	2,3	(4,5),6
1,7	0		
2,3	0.57	0	
(4,5),6	0.5	0.44	0

	1,7	(2,3),((4,5),6
1,7	0	0.5
(2,3),((4,5),6)	0.5	0



b. Complete: Do the same but use max distance as the measurement, we have (2,3) first , but after that we choose max distance

	1	2,3	4	5	6	7
1	0					
2,3	0.72	0				
4	0.96	0.71	0			
5	0.75	0.7	0.82	0		
6	1.03	0.49	0.38	0.7	0	
7	0.25	0.67	0.73	0.5	0.86	0

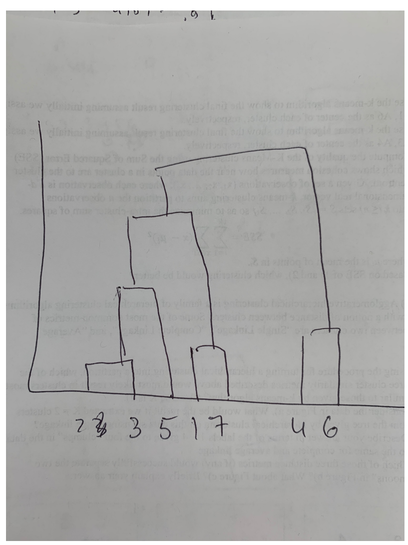
Next, we see min distance is 0.25 of (1,7). Then 4,6 at .38, the (2,3),5 at .7, the 17235 at .75

	1,7	2,3	4	5	6
1,7	0				
2,3	0.72	0			
4	0.96	0.71	0		
5	0.75	0.7	0.82	0	
6	1.03	0.49	0.38	0.7	0

	1,7	2,3	4,6	5
1,7	0			
2,3	0.72	0		
4,6	1.03	0.71	0	
5	0.75	0.7	0.82	0

	1,7	(2,3),5	4,6
1,7	0		
(2,3),5	0.75	0	
4,6	1.03	0.82	0

	((1,7),((2,3),5	4,6
((1,7),((2,3),5)	0	
4,6	1.03	0



2. The final three clusters are: (1) {A1 , C2 , B1 }, (2) {A3 , B2 , B3 }, (3) {C1 ,A2 }

3. Seed 1

	X	Y	A1(1,2)	A6(5,4)
A1	1.00	2.00	0.000	4.472
A2	1.00	4.00	2.000	4.000
A3	3.00	1.00	2.236	3.606
A4	3.00	5.00	3.606	2.236
A5	5.00	2.00	4.000	2.000
A6	5.00	4.00	4.472	0.000

	New X	New Y
C1 = {A1, A2, A3}.	1.667	2.333
C2 = {A4, A5, A6}.	4.333	2.667

	X	Y	C1(1.667, 2.333)	C2(4.333, 3.667)
A1	1.00	2.00	0.746	3.727
A2	1.00	4.00	1.795	3.350
A3	3.00	1.00	1.885	2.982
A4	3.00	5.00	2.982	1.885
A5	5.00	2.00	3.350	1.795
A6	5.00	4.00	3.727	0.746

	New X	New Y
C1 = {A1, A2, A3}.	1.667	2.333
C2 = {A4, A5, A6}.	4.333	2.667

SSE1= 14.666668

Seed 2

	X	Y	A3(3,1)	A4(3,5)
A1	1.00	2.00	2.236	3.606
A2	1.00	4.00	3.606	2.236
A3	3.00	1.00	0.000	4.000
A4	3.00	5.00	4.000	0.000
A5	5.00	2.00	2.236	3.606
A6	5.00	4.00	3.606	2.236

	New X	New Y
C1 = {A1, A3, A5}.	3	1.667
C2 = {A2, A4, A6}.	3	4.33

	X	Y	C1(3,1.667)	C2(3,4.333)
A1	1.00	2.00	2.028	3.073
A2	1.00	4.00	3.073	2.028
A3	3.00	1.00	0.667	3.333
A4	3.00	5.00	3.333	0.667
A5	5.00	2.00	2.028	3.073
A6	5.00	4.00	3.073	2.028

	New X	New Y
C1 = {A1, A3, A5}.	3	1.667
C2 = {A2, A4, A6}.	3	4.33

SSE2 = 17.333334

SSE1<SSE2 so the first one is better

4. Answers:
   a. Average linkage
   b. Average and complete linkage would assign clumps 1 and 3 to the first clusters and 2 and 4 to the second one. Single linkage assigns 1&2 to one and 3&4 to the other
   c. Single linkage would be able to separate two crescents in (b) while complete and average wouldn't. None of the methods would work for (c)