Group Average – Proximity between two clusters is the average pairwise distance between all pairs of points in the two clusters

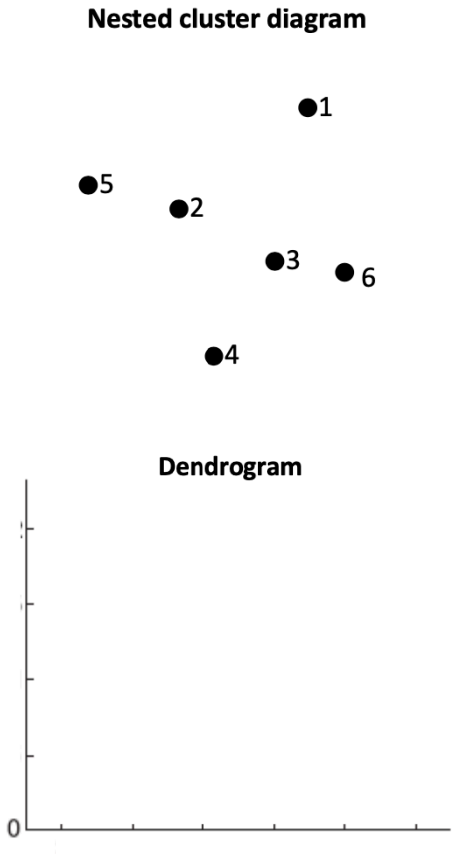
- Given six points:

Point	x Coordinate	y Coordinate
p1	0.40	0.53
p2	0.22	0.38
p3	0.35	0.32
p4	0.26	0.19
p5	0.08	0.41
p6	0.45	0.30

- And their pairwise distances:

	p1	p2	p3	p4	p5	p6
p1	0.00	0.24	0.22	0.37	0.34	0.23
p2	0.24	0.00	0.15	0.20	0.14	0.25
p3	0.22	0.15	0.00	0.15	0.28	0.11
p4	0.37	0.20	0.15	0.00	0.29	0.22
p5	0.34	0.14	0.28	0.29	0.00	0.39
p6	0.23	0.25	0.11	0.22	0.39	0.00

Draw the nested cluster diagram and dendrogram using the group average as your proximity measure



Group Average

Proximity Matrix

	p1	p2	p3	p4	p5	p6
p1	0	0.24	0.22	0.37	0.34	0.23
p2	0.24	0	0.15	0.20	0.14	0.25
p3	0.22	0.15	0	0.15	0.28	0.11
p4	0.37	0.20	0.15	0	0.29	0.22
p5	0.34	0.14	0.28	0.29	0	0.39
p6	0.23	0.25	0.11	0.22	0.39	0

*dist({p2,p5}, {p3,p6}) = (0.15 + 0.25 + 0.28 + 0.39) / (2*2) = 0.2675*
dist({p2,p5}, p1) = (0.24 + 0.34) / 2 = 0.29
dist({p2,p5}, p4) = (0.2 + 0.29) / 2 = 0.245

Round 1

	p1	p2	{p3,p6}	p4	p5
p1	0	0.24	**0.225**	0.37	0.34
p2	0.24	0	**0.2**	0.20	0.14
{p3,p6}	**0.225**	**0.2**	0	**0.185**	**0.335**
p4	0.37	0.20	**0.185**	0	0.29
p5	0.34	0.14	**0.335**	0.29	0

dist(p1, {p3,p6} = (0.22 + 0.23) / 2 = 0.225
dist(p2 {p3,p6} = (0.15 + 0.25) / 2 = 0.2
dist(p4, {p3,p6} = (0.15 + 0.22) / 2 = 0.185
dist(p5, {p3,p6} = (0.28 + 0.39) / 2 = 0.335

Round 2

	p1	{p2,p5}	{p3,p6}	p4
p1	0	**0.29**	0.225	0.37
{p2,p5}	**0.29**	0	**0.2675**	**0.245**
{p3,p6}	0.225	**0.2675**	0	0.185
p4	0.37	**0.245**	0.185	0

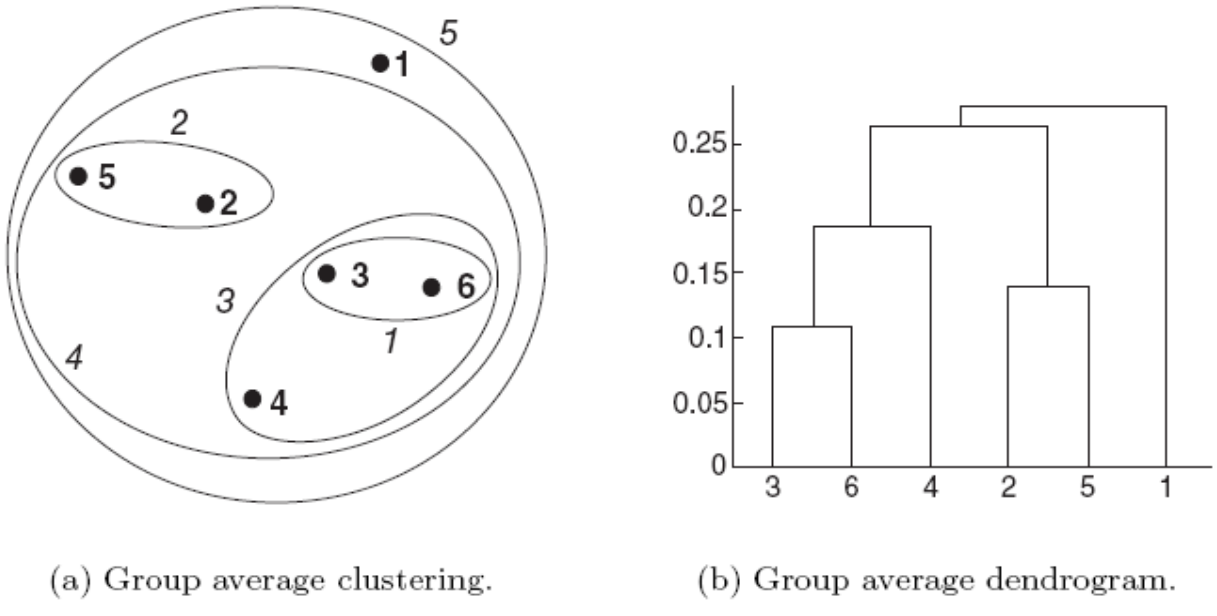
	p1	p2	p3	p4	p5	p6
p1	0	0.24	0.22	0.37	0.34	0.23
p2	0.24	0	0.15	0.20	0.14	0.25
p3	0.22	0.15	0	0.15	0.28	0.11
p4	0.37	0.20	0.15	0	0.29	0.22
p5	0.34	0.14	0.28	0.29	0	0.39
p6	0.23	0.25	0.11	0.22	0.39	0

Round 3

	p1	{p2,p5}	{p3,p6,p4}
p1	0	0.29	**0.273**
{p2,p5}	0.29	0	**0.26**
{p3,p6,p4}	**0.273**	**0.26**	0

*dist({p3,p6,p4},p1) = (0.22 + 0.23 + 0.37)/(3*1) = 0.273*

*dist({p3,p6,p4},{p2,p5}) = (0.15 + 0.25 + 0.2 + 0.28 + 0.39 + 0.29) / (2*3) = 0.26*



(a) Group average clustering.

(b) Group average dendrogram.

Figure 8.18. Group average clustering of the six points shown in Figure 8.15.