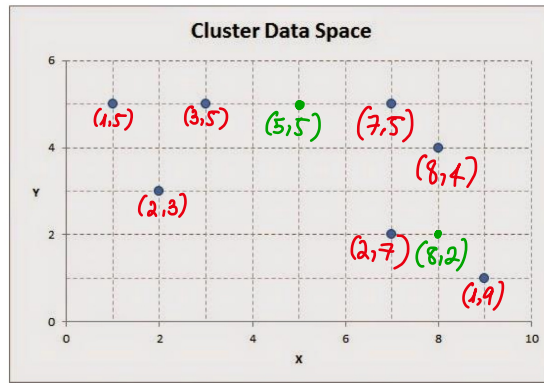


2. (15 נקודות) וקטורי התכונות הבאים (נקודות ב- $R^2$ ), מייצגים מדדים של שגבוב (aardvark) שלא היה ידוע לחוקרים. הזואולוגים טוענים כי קיימים שני סוגים של שגבוב. כדי להבחין בין הסוגים מפעילים את אלגוריתם k-means עבור  $k=2$ . הצנטרואידים ההתחלתיים הוגרלו בנקודות  $C_1=(5,5)$ ,  $C_2=(8,2)$ .



- א. סמנו את הצנטרואידים ההתחלתיים על התרשים.  
 ב. בצעו את צעד השייך ההתחלתי, רשמו את הנוסחה באמצעותה מתבצע השייך כתבו את החישובים שאתם מבצעים. באיזה כלל השתמשתם?

### צד השייך ההתחלתי :

$$X = [(1, 5), (2, 3), (3, 5), (7, 2), (7, 5), (8, 4), (9, 1)]$$

$$d(x_1, x_2) = \sqrt{x_1^2 + x_2^2}$$

$$C_1 = (5, 5)$$

$$\|(1, 5) - (5, 5)\| = \|(-4, 0)\| = \sqrt{16}$$

$$\|(2, 3) - (5, 5)\| = \|(-3, -2)\| = \sqrt{13}$$

$$\|(3, 5) - (5, 5)\| = \|(-2, 0)\| = \sqrt{4}$$

$$\|(7, 2) - (5, 5)\| = \|(2, -3)\| = \sqrt{13}$$

$$\|(7, 5) - (5, 5)\| = \|(2, 0)\| = \sqrt{4}$$

$$\|(8, 4) - (5, 5)\| = \|(3, -1)\| = \sqrt{10}$$

$$\|(9, 1) - (5, 5)\| = \|(4, -4)\| = \sqrt{32}$$

$$C_2 = (8, 2)$$

$$\|(1, 5) - (8, 2)\| = \|(-7, 3)\| = \sqrt{58}$$

$$\|(2, 3) - (8, 2)\| = \|(-6, 1)\| = \sqrt{37}$$

$$\|(3, 5) - (8, 2)\| = \|(-5, 3)\| = \sqrt{34}$$

$$\|(7, 2) - (8, 2)\| = \|(-1, 0)\| = \sqrt{1}$$

$$\|(7, 5) - (8, 2)\| = \|(-1, 3)\| = \sqrt{10}$$

$$\|(8, 4) - (8, 2)\| = \|(0, 2)\| = \sqrt{4}$$

$$\|(9, 1) - (8, 2)\| = \|(1, -1)\| = \sqrt{2}$$

ג. חשבו את הצנטרואידים של הצעד השני, ורשמו את הנוסחה בה השתמשתם.

$$C_1 = \frac{1}{4} \cdot ((1, 5) + (2, 3) + (3, 5) + (7, 5)) \\ = \frac{1}{4} \cdot (13, 18) = (3\frac{1}{4}, 4\frac{1}{2})$$

$$C_2 = \frac{1}{3} \cdot ((7, 2) + (8, 4) + (9, 1)) \\ = \frac{1}{3} \cdot (24, 7) = (8, 2\frac{1}{3})$$

4. (10 נקודות) תהי  $X$  קבוצת הלימוד הבאה:

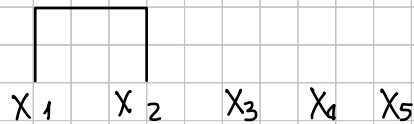
$$X = \{x_i, i = 1, 2, 3, 4, 5\}$$

$$x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, x_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, x_3 = \begin{pmatrix} 5 \\ 4 \end{pmatrix}, x_4 = \begin{pmatrix} 6 \\ 5 \end{pmatrix}, x_5 = \begin{pmatrix} 6.5 \\ 6 \end{pmatrix}$$

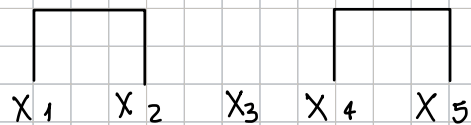
$$D(x) = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 5 & 4 \\ 6 & 5 \\ 6.5 & 6 \end{pmatrix}$$

השתמשו במרחק אוקלידי כדי ליצור את מטריצת המרחק וציירו את הדנדרוגרמה לאחר הפעלת hierarchical clustering עם מרחק מינימלי.

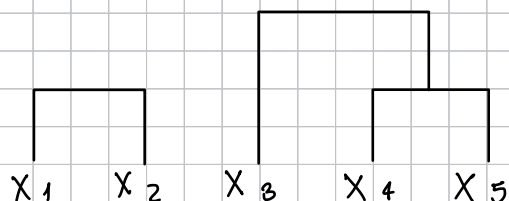
	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
$x_1$	0	1	5	6.4	7.4
$x_2$	1	0	4.2	5.6	6.7
$x_3$	5	4.2	0	1.4	2.5
$x_4$	6.4	5.6	1.4	0	1.1
$x_5$	7.4	6.7	2.5	1.1	0



	$x_1$ $x_2$	$x_3$	$x_4$	$x_5$
$x_1$ $x_2$	0	4.2	5.6	6.7
$x_3$	4.2	0	1.4	2.5
$x_4$	5.6	1.4	0	1.1
$x_5$	6.7	2.5	1.1	0



	$x_1$ $x_2$	$x_3$	$x_4$ $x_5$
$x_1$ $x_2$	0	4.2	5.6
$x_3$	4.2	0	1.4
$x_4$ $x_5$	5.6	1.4	0



	$x_1$ $x_2$	$x_3$ $x_4$ $x_5$
$x_1$ $x_2$	0	4.2
$x_3$ $x_4$ $x_5$	4.2	0

