

## Кластерлеу әдісі

Мысал

Студент	Баға (x)	Қатысу (%) (y)
A	90	95
B	85	90
C	40	50
D	45	55

Біз  $k = 2$  деп шештік → екі кластер: C1 және C2

- C1 — кластер 1 орталығы
- C2 — кластер 2 орталығы

Бастапқы центрлерді таңдаймыз (мысалы):

Кластер	Баға (x)	Қатысу (%) (y)
C1	90	95
C2	40	50

Евклид қашықтығының формуласы

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Белгі	Түсіндірме
$x_1$	нүктенің (студенттің) баға мәні
$y_1$	нүктенің (студенттің) қатысу мәні
$x_2$	кластер орталығының баға мәні
$y_2$	кластер орталығының қатысу мәні
$d$	нүкте мен кластер орталығы арасындағы қашықтық

A студенті → C1 орталығы

**Берілгендер:**

- A:  $x_1 = 90, y_1 = 95$
- C1:  $x_2 = 90, y_2 = 95$

Айырмаларды табамыз:

$$x_1 - x_2 = 90 - 90 = 0$$

$$y_1 - y_2 = 95 - 95 = 0$$

Квадраттау:  $0^2 + 0^2 = 0$

Түбір:

$$d = \sqrt{0} = 0$$

А студенті дәл C1 орталығына тең, қашықтық 0

C студенті → C1 орталығы

**Берілгендер:**

- C:  $x_1 = 40, y_1 = 50$
- C1:  $x_2 = 90, y_2 = 95$

Айырмаларды табу:

$$x_1 - x_2 = 40 - 90 = -50$$

$$y_1 - y_2 = 50 - 95 = -45$$

Квадраттау:

$$(-50)^2 + (-45)^2 = 2500 + 2025 = 4525$$

Квадрат түбір алу:

$$d = \sqrt{4525} \approx 67.3$$

C студенті → C2 орталығы

- C2:  $x_2 = 40, y_2 = 50$
- Айырма:  $x_1 - x_2 = 40 - 40 = 0, y_1 - y_2 = 50 - 50 = 0$
- 
- Квадрат:  $0^2 + 0^2 = 0$
- Квадрат түбір:  $\sqrt{0} = 0$
- Сол себепті C студенті **C2 кластеріне жақын.**

Студент	$x_1$	$y_1$	C1( $x_2, y_2$ )	d(C1)	C2( $x_2, y_2$ )	d(C2)	Қай кластерге
A	90	95	90, 95	0	40, 50	67.3	C1
B	85	90	90, 95	7.07	40, 50	63.64	C1
C	40	50	90, 95	67.3	40, 50	0	C2
D	45	55	90, 95	61.64	40, 50	7.07	C2

### Қорытынды түсінік:

1.  $x_1, y_1$  — нақты нүктенің координаттары (студенттің мәндері)
2.  $x_2, y_2$  — кластер орталығының координаттары
3.  $d$  — нүкте мен орталық арасындағы қашықтық
4. Кластер таңдалуы — нүкте қай орталыққа жақын болса, сол кластерге кіретінін білдіреді.

### Аудиториялық тапсырма

Берілген деректерді **K-means** ( $k = 2$ ) арқылы кластерлерге бөл:

Нүкте	x	y
A	2	3
B	3	4
C	8	7
D	9	8