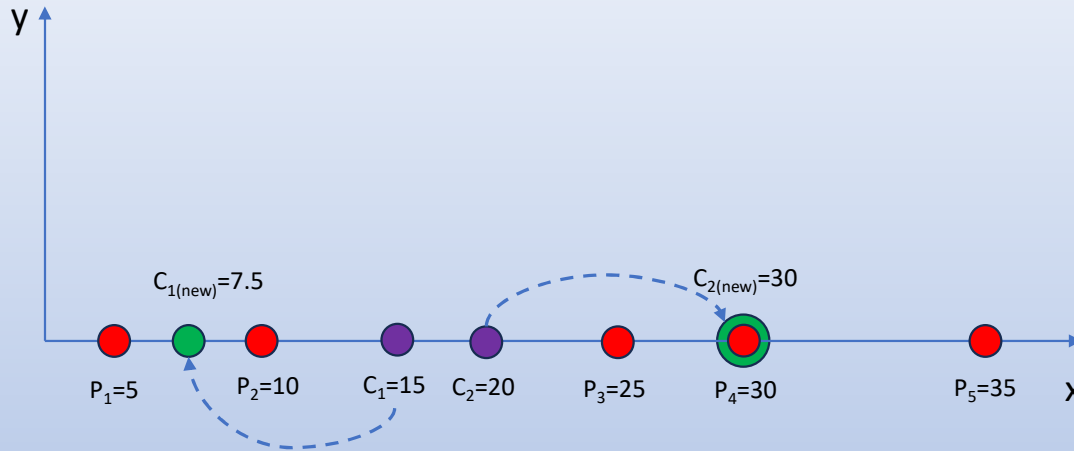


# K-means case study:



**Distances between points and centroids:**

$$\overline{c_j \cdot p_i} = \sqrt{(c_{jx} - p_{ix})^2 + (c_{jy} - p_{iy})^2}$$

**New centroids:**

$$c_{jx(new)} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$c_{jy(new)} = \frac{1}{n} \sum_{i=1}^n y_i$$

i	x	y	$c_1$ X=15, Y=0	$c_2$ X=20, Y=0	cluster
1	5	0	$\overline{c_1 \cdot p_1} = \sqrt{(15-5)^2 + (0-0)^2} = 10$	$\overline{c_2 \cdot p_1} = \sqrt{(20-5)^2 + (0-0)^2} = 15$	$\overline{c_1 \cdot p_1} < \overline{c_2 \cdot p_1} \rightarrow 10 < 15 \rightarrow \mathbf{c_1}$
2	10	0	$\overline{c_1 \cdot p_2} = \sqrt{(15-10)^2 + (0-0)^2} = 5$	$\overline{c_2 \cdot p_2} = \sqrt{(20-10)^2 + (0-0)^2} = 10$	$\overline{c_1 \cdot p_2} < \overline{c_2 \cdot p_2} \rightarrow 5 < 10 \rightarrow \mathbf{c_1}$
3	25	0	$\overline{c_1 \cdot p_3} = \sqrt{(15-25)^2 + (0-0)^2} = 10$	$\overline{c_2 \cdot p_3} = \sqrt{(20-25)^2 + (0-0)^2} = 5$	$\overline{c_1 \cdot p_3} > \overline{c_2 \cdot p_3} \rightarrow 10 > 5 \rightarrow \mathbf{c_2}$
4	30	0	$\overline{c_1 \cdot p_4} = \sqrt{(15-30)^2 + (0-0)^2} = 15$	$\overline{c_2 \cdot p_4} = \sqrt{(20-30)^2 + (0-0)^2} = 10$	$\overline{c_1 \cdot p_4} > \overline{c_2 \cdot p_4} \rightarrow 15 > 10 \rightarrow \mathbf{c_2}$
5	35	0	$\overline{c_1 \cdot p_5} = \sqrt{(15-35)^2 + (0-0)^2} = 20$	$\overline{c_2 \cdot p_5} = \sqrt{(20-35)^2 + (0-0)^2} = 15$	$\overline{c_1 \cdot p_5} > \overline{c_2 \cdot p_5} \rightarrow 20 > 15 \rightarrow \mathbf{c_2}$

**New centroids - calculation:**

$$c_{1x(new)} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{(10 + 5)}{2} = 7.5$$

$$c_{1y(new)} = 0$$

$$c_{2x(new)} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{(25 + 30 + 35)}{3} = 30$$

$$c_{2y(new)} = 0$$