Q1 b) The k-means algo converges to a local optimum in finite steps.

$$\mathcal{J} = \sum_{n=1}^{N} \sum_{k=1}^{K} r^{nk} \| x^n - \mu^k \|^2$$

→ distortion ~~funk~~ function monotonically decreases

Proof: show that next iteration is smaller or equal to previous loss funktion.

↳ $X^{(t)} = \sum_{n=1}^{N} \sum_{k=1}^{K} r^{nk} \| x^n - \mu^k \|^2 = \mathcal{J}$ where $X^{(t)}$ is the current partition $X_1^{(t)}, \dots, X_k^{(t)}$ with centroids $\mu_1^{(t)} \dots \mu_k^{(t)}$ and Assignation fct $A(t)$ then,

$$X^{(t)} \geq \sum_{n=1}^{N} \sum_{k=1}^{K} \| x^n - \mu^k_{A^{(t-1)}(x^n)} \|^2$$

→ $A(x_i)$ min quantity $\|x_n - \mu^k\|^2$ over all $k = 1 \dots K$

$$↳ \geq \sum_{n=1}^{N} \sum_{k=1}^{K} \| x^n - \mu^{k(t+1)} \|^2$$

$$\geq X^{(t+1)}$$

⇒ There is no infinite sequence of partitions such that the loss function decreases strictly since there is a finite number of partitons $\binom{n}{k}$.

Hence $X^{(t)}_{t \in \mathbb{N}}$ has a finite number of values.

There ~~exis~~ exists $X^{(t+1)} = X^{(t)}$