

Algorithm K-MEDIANS

Input Data $\mathcal{D} = \{x_i\}_{i=1:n}$, number clusters K

1. Calculate $d_{ij} = ||x_i - x_j||$
2. Solve Linear Program

$$\begin{array}{ll}\min_{u,y} & \sum_{ij} d_{ij} u_{ij} \\ \text{s.t.} & \sum_j u_{ij} = 1 \\ & \sum_j y_j \leq k \\ & u_{ij} \leq y_j \\ & u_{ij} \geq 0 \\ & y_j \leq 1\end{array}$$

obtain fractionary “centers” $y_{1:n}$ and “assignments” $u_{1:n,1:n}$

3. Sample K centers $\mu_1 \dots \mu_K$ by
 - $P[\mu_k = \text{point } j] \propto y_j$ (without replacement)
4. Assign points to centers (deterministically)

$$k(i) = \underset{k}{\operatorname{argmin}} ||x_i - \mu_k||$$