# Take-home messages

1. **Unsupervised** vs supervised learning: No ground truth (labels) accessible. Consequently, no split in training and test sets.
2. Scanning for optimal k:
    - *k*-means & GMM: for every k → run the algorithm→ evaluate.
    - Hierarchical: run the algorithm once → scan the dendrogram→ evaluate.
3. There are still hyper-parameters to fix, e.g.:
    - *k*-means: distance metric + other design choices (handling empty clusters).
    - Hierarchical: distance metric & linkage methods.
    - Probabilistic: probability distribution.

# Take-home messages

4. *k*-means and GMM can be sensitive to initialisation.
5. For empty clusters issue in *k*-means, a design choice needs to be made:
    - Continue running the algorithm and return $k' < k$.
    - Do multiple runs with different random initialisations.
    - Choose the best k that corresponds to the clustering with minimum within-cluster distance (monotonically decreasing with k).
6. Coding tip: Sometimes, it's useful to fix the random seed.