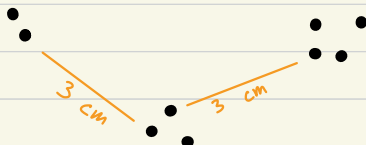# Exercises for hierarchical clustering

1. Sketch (by hand) the dendrograms for the following datasets in the plane :
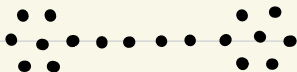
(i)

(ii)

3 cm   3 cm

(iii)

(iv)

(v) The drastic change in the dendrogram from (iv) to (v) above is an example of the problem of "chaining". How might you address this ?

2. On the course GitHub page, see the jupyter notebook "plot_linkage_comparison.ipynb". These are other forms of hierarchical (agglomerative) clustering. Which do you prefer?

Two clusters A,B merge based on the next lowest value of:

$$\min_{a \in A, b \in B} d(a,b)$$

| Single Linkage | Average Linkage | Complete Linkage | Ward Linkage |