Exercise 16

Consider the following distance matrix of the objects A, B, C, D, and E:

	A	B	C	D	E
A	0	9	7	12	8
B	9	0	4	2	5
C	7	4	0	1	13
D	12	2	1	0	6
E	8	5	13	6	0

(a) Perform an agglomerative clustering of the objects based on **Single**-Linkage with Paper and Pencil. Draw the corresponding Dendrogram (with P&P).

(b) Perform an agglomerative clustering of the objects based on **Complete**-Linkage. Draw the corresponding Dendrogram (with P&P).

(c) Perform an agglomerative clustering of the objects based on **Average**-Linkage. Draw the corresponding Dendrogram (with P&P).

(d) Repeat (a) – (c) in R and compare to your P&P results.

Exercise 17

Consider the data set `USairpollution` in the package `HSAUR2`, but without the variable `SO2` for the moment.

(a) Perform an agglomerative hierarchical clustering (of the cities) with the function `hclust()`. Use the Euclidean distance, and perform the clustering based on Single-, Complete- as well as Average-Linkage. Compare the results of the different linkage methods (e.g. based on Dendrograms). Which are your preferred choices of clusters in each setting?

(b) Choose one clustering from part (a). Compare the different variables in the data set, including `SO2`, in the resulting clusters. What is your conclusion, are there cluster specific properties? How would you describe the individual clusters?