# Première partie

La matrice des distances est la suivante :

1 2 3 4 5

1 87.53 31.30 65.64 56.04

2 87.53 84.22 49.85 62.00

3 31.30 84.22 76.94 66.97

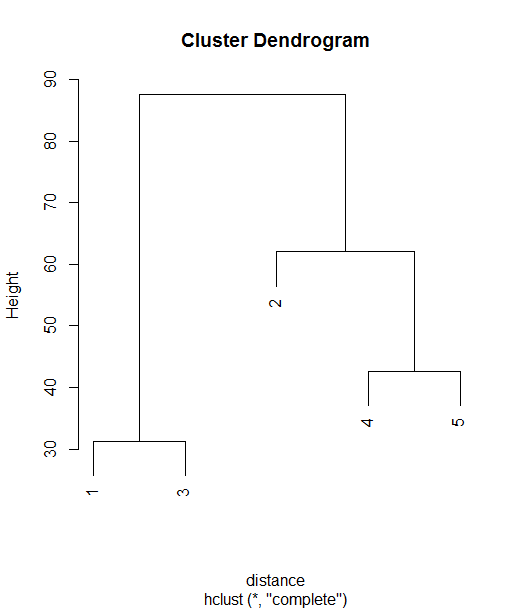
4 65.64 49.85 76.94 42.67

5 56.04 62.00 66.97 42.67

On l’obtient en R avec la commande suivante :

round(dist(data, method= "euclidean", upper=T), 2)

Le dendrogramme en utilisant la méthode du saut maximal est le suivant :

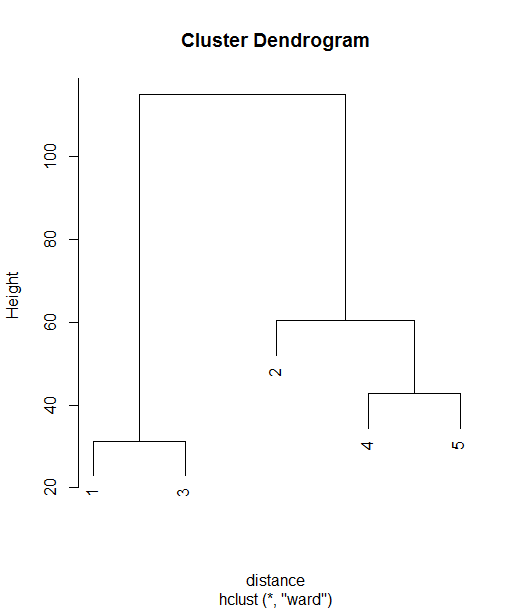


On l’obtient sous R avec les commandes suivantes :

distance.max <- hclust(distance, method="complete")

plot(distance.max)

Le dendrogramme utilisant la méthode du saut minimal (méthode ward) est le suivant :



On l’obtient avec les commandes suivantes :

distance.min <- hclust(distance, method="ward")

plot(distance.min)

# Seconde partie

Les dendrogrammes donnent envie de choisir trois groupes :

* En coupant à hauteur 15 sur le diagramme avec critère de Ward
* En coupant à hauteur 7 sur le diagramme avec critère de lien maximal

En-dessous de cette hauteur, il est difficile de visualiser le moindre groupe.

Les éléments sont répartis selon les proportions suivantes en 3 groupes :

* Groupe 1 : 16
* Groupe 2 : 21
* Groupe 3 : 27

Les statistiques descriptives donnent par groupe:

group: 1

var n mean sd median trimmed mad min max range skew kurtosis se

Var1 1 16 55.20 0.70 54.95 55.12 0.42 54.52 57.13 2.61 1.40 1.14 0.17

Var2 2 16 53.63 0.52 53.66 53.64 0.63 52.79 54.40 1.61 -0.14 -1.42 0.13

Var3 3 16 63.19 0.71 62.95 63.15 0.49 62.32 64.58 2.26 0.65 -1.04 0.18

Var4 4 16 65.22 0.78 65.20 65.24 0.79 63.73 66.38 2.65 -0.30 -1.01 0.19

Var5 5 16 51.23 0.58 51.27 51.26 0.54 50.09 52.07 1.98 -0.31 -0.98 0.14

Var6 6 16 64.59 0.67 64.76 64.63 0.30 62.95 65.75 2.80 -0.87 0.37 0.17

Var7 7 16 67.27 0.70 67.42 67.28 0.77 65.97 68.44 2.47 -0.27 -1.16 0.17

---------------------------------------------------------------------------------

group: 2

var n mean sd median trimmed mad min max range skew kurtosis se

Var1 1 21 54.21 0.87 54.21 54.23 1.13 52.50 55.77 3.27 -0.14 -1.10 0.19

Var2 2 21 53.33 0.78 53.38 53.32 0.99 51.89 54.84 2.95 0.05 -1.11 0.17

Var3 3 21 61.13 0.87 61.28 61.23 0.83 58.76 62.09 3.33 -1.01 0.33 0.19

Var4 4 21 64.62 0.84 64.55 64.64 1.11 63.20 65.76 2.56 0.01 -1.49 0.18

Var5 5 21 52.21 0.69 52.25 52.19 0.65 50.91 54.01 3.10 0.34 0.32 0.15

Var6 6 21 65.14 0.96 65.45 65.27 0.59 62.58 66.61 4.03 -1.10 0.79 0.21

Var7 7 21 68.09 0.77 68.08 68.05 0.67 66.72 70.12 3.40 0.65 0.36 0.17

---------------------------------------------------------------------------------

group: 3

var n mean sd median trimmed mad min max range skew kurtosis se

Var1 1 27 53.67 0.80 53.75 53.70 0.86 51.43 55.23 3.80 -0.57 0.51 0.15

Var2 2 27 52.47 0.72 52.52 52.53 0.39 50.33 53.61 3.28 -0.90 1.03 0.14

Var3 3 27 62.08 0.95 62.22 62.15 0.98 59.77 63.45 3.68 -0.70 -0.44 0.18

Var4 4 27 64.98 1.04 65.04 65.00 1.17 62.99 66.94 3.95 -0.07 -0.97 0.20

Var5 5 27 50.27 0.77 50.38 50.31 0.58 48.13 51.70 3.57 -0.69 0.61 0.15

Var6 6 27 64.87 0.83 64.66 64.82 0.71 63.68 66.58 2.90 0.50 -0.82 0.16

Var7 7 27 66.34 0.69 66.48 66.35 0.59 64.97 67.53 2.56 -0.23 -1.03 0.13

Voici une représentation en boxplot des variables selon les 3 groupes :

