

# Méthodes hiérarchiques

En R

---

Véronique Tremblay

On utilise la librairie `cluster` (Maechler et al. (2019)).

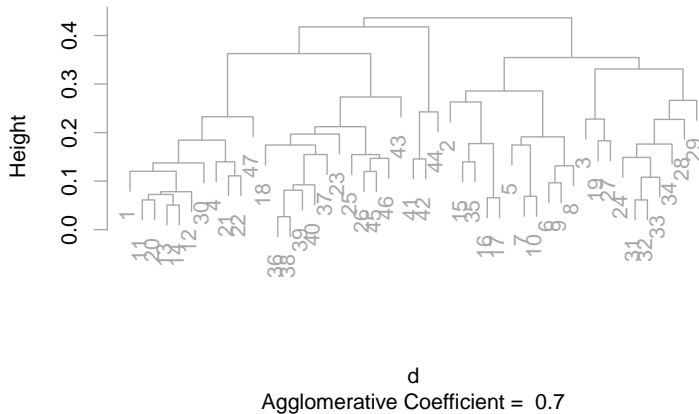
## Exemple: méthode de la moyenne

---

```
m_moy <- agnes(d, # La matrice de dissimilarité  
               method = 'average' # La méthode de la moyenne  
               )
```

```
plot(m_moy, which.plots = 2, col = 'gray65')
```

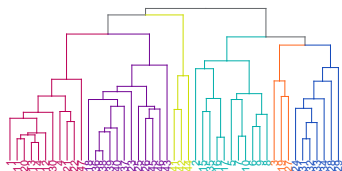
### Dendrogram of `agnes(x = d, method = "average")`



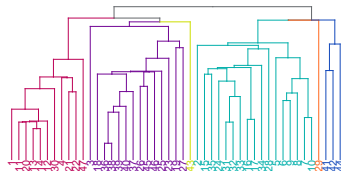
# Pour faire mieux

```
dend <- m_moy %>%  
  as.dendrogram %>%  
  set("branches_k_color", k = 6) %>%  
  set("branches_lwd", 0.2) %>%  
  set("labels_colors", k = 6) %>%  
  set("labels_cex", c(0.4))  
  
ggd1 <- as.ggdend(dend)  
g <- ggplot(ggd1)  
  
gmoy <- g +  
  scale_color_manual(values=pal_Videns_discrete)+ # Palette  
  xlab(' ')+  
  ylab('Dissemblance')+  
  ggtitle('Moyenne')+  
  ylim(-0.05,0.5)
```

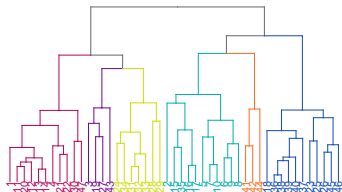
Moyenne



Plus proche voisin



Voisin le plus distant



# Extraire les groupes

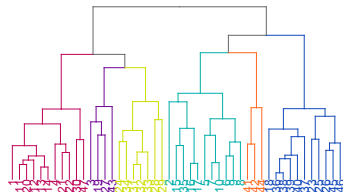
---

```
gr_agnes <- cutree(m_moy, k = 6)

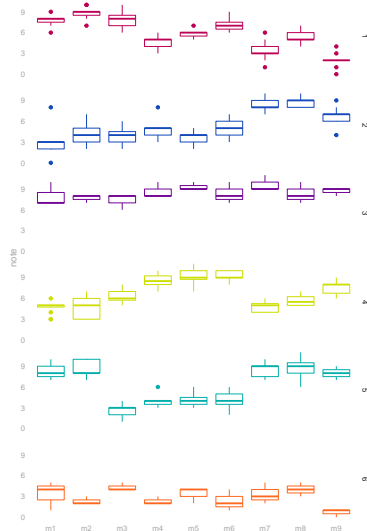
musique_seg <- musique %>%
  mutate(segment = as.factor(gr_agnes))
```

# Décrire les groupes

Voisin le plus distant



k	F	H	Jamais	Parfois	Souvent
1	11	0	2	7	2
2	3	8	7	3	1
3	0	4	1	2	1
4	11	0	0	4	7
5	0	7	0	3	4
6	2	1	0	0	3





- `agnes` de la librairie `cluster`
- dendrogram pour améliorer la visualisation
- Décrire les groupes

Maechler, Martin, Peter Rousseeuw, Anja Struyf, Mia Hubert, and Kurt Hornik. 2019. *Cluster: Cluster Analysis Basics and Extensions*.