

TD 2 - Classification

1. Les Iris

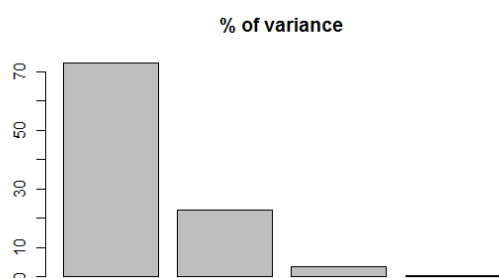
1. Présentation des données et questions

Sur des iris appartenant à 3 espèces différentes (*setosa*, *versicolor*, *virginica*), on a mesuré la longueur et la largeur des pétales et des sépales. 150 fleurs ont été mesurées (50 par espèces).

Nous allons répondre aux questions suivantes :

- Présenter les données
- Interpréter les résultats de l'ACP, en mettant en évidence les différences entre espèces.
- Commenter ensuite les résultats de la CAH. En quoi complètent-ils l'interprétation de l'ACP ?
- Commenter ensuite les résultats de la classification par moyennes mobiles (k-means)
- Concluez avec le tableau croisant les groupes des k-means et les vraies espèces

2. Résultats

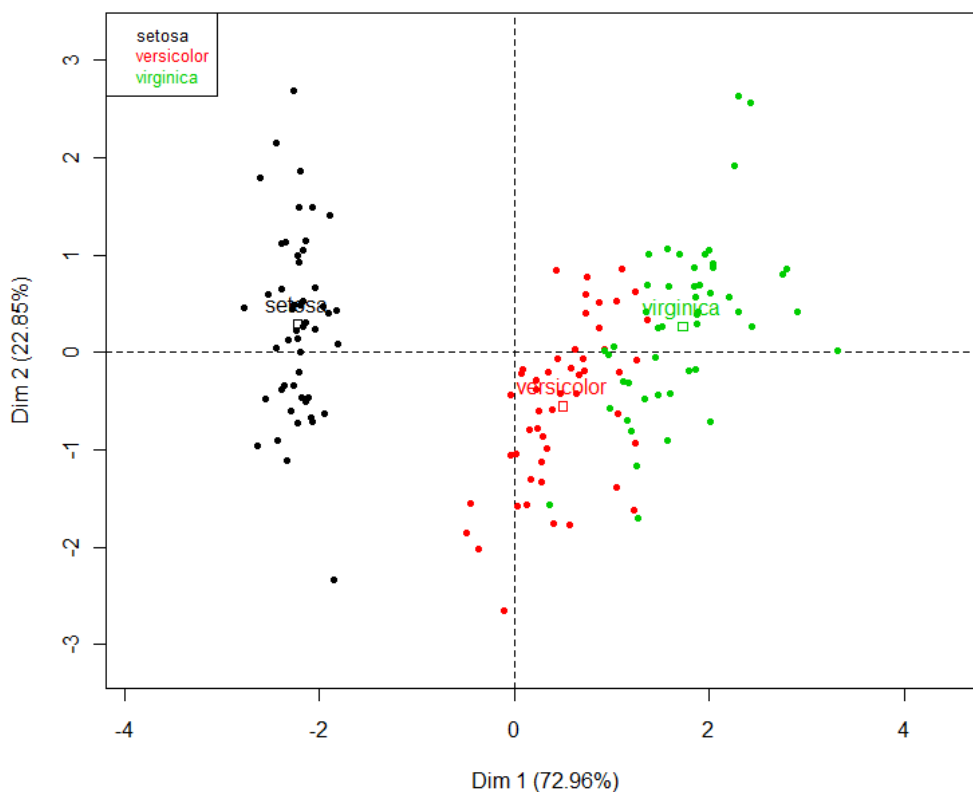


	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
Sepal.Length	1.0000000	-0.1175698	0.8717538	0.8179411
Sepal.Width	-0.1175698	1.0000000	-0.4284401	-0.3661259
Petal.Length	0.8717538	-0.4284401	1.0000000	0.9628654
Petal.Width	0.8179411	-0.3661259	0.9628654	1.0000000

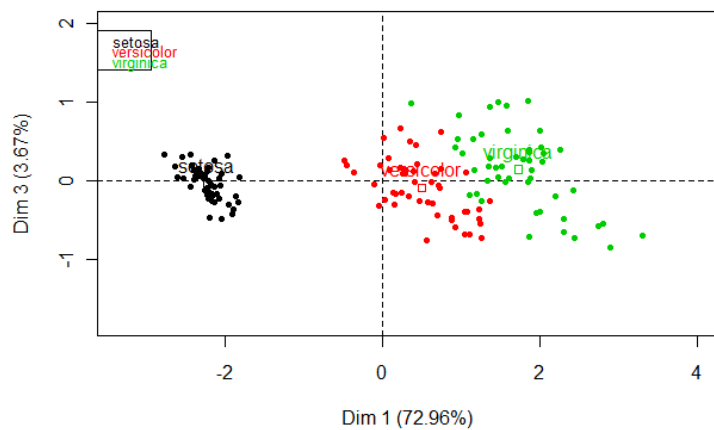
Matrice des corrélations

% de variance expliqué par chaque axe factoriel

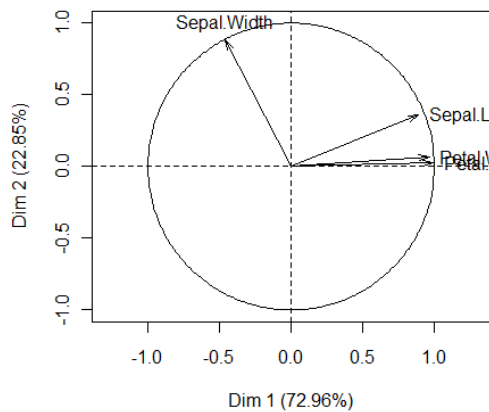
Individuals factor map (PCA)



Individuals factor map (PCA)

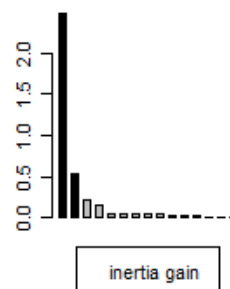
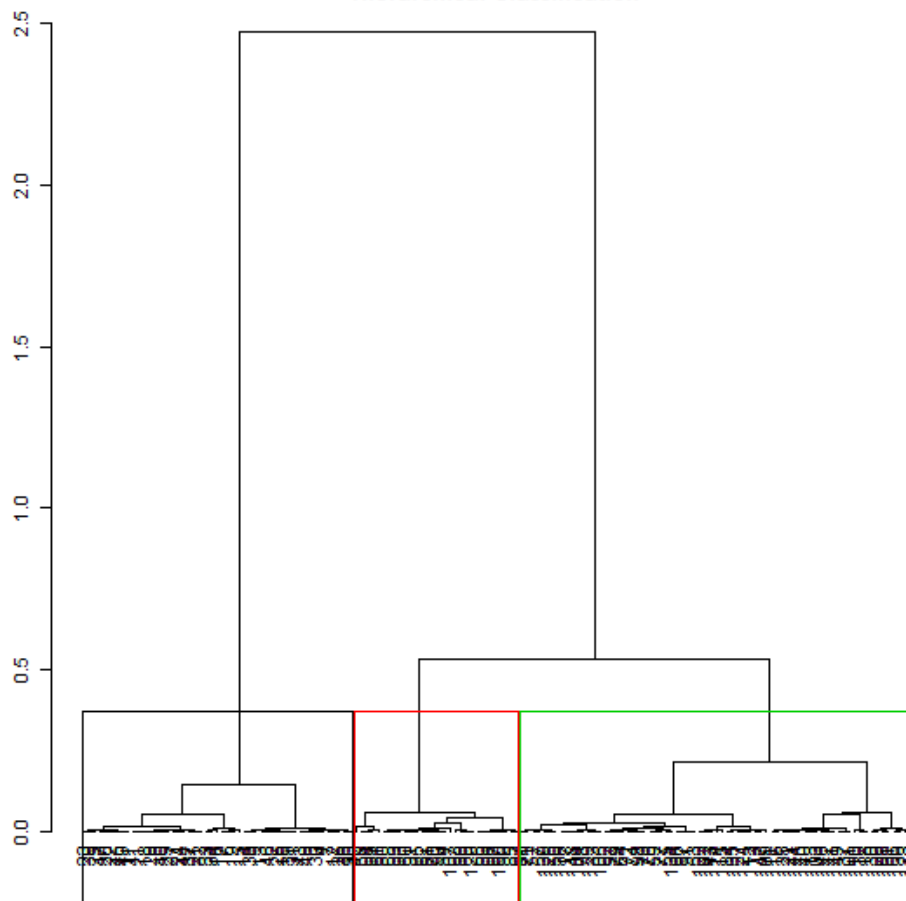


Variables factor map (PCA)

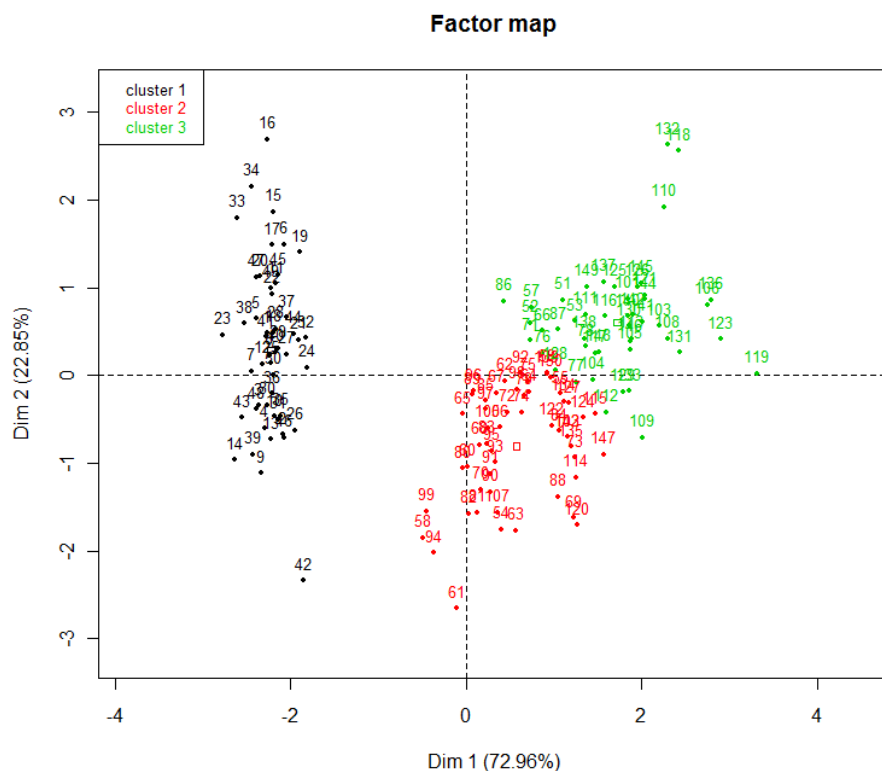


Hierarchical Clustering

Hierarchical Classification



Résultats des k-means :



Description des groupes

	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
Sepal.Width	7.364799	3.428	3.057333	0.3752546	0.4344110	1.774142e-13
Sepal.Length	-8.757174	5.006	5.843333	0.3489470	0.8253013	2.002060e-18
Petal.Width	-10.831410	0.246	1.199333	0.1043264	0.7596926	2.443627e-27
Petal.Length	-11.263787	1.462	3.758000	0.1719186	1.7594041	1.980605e-29

\$quanti\$`2`

	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
Petal.Length	4.221328	4.581132	3.758000	0.8040127	1.7594041	2.428668e-05
Petal.Width	3.503794	1.494340	1.199333	0.3935399	0.7596926	4.586800e-04
Sepal.Length	2.599814	6.081132	5.843333	0.6475080	0.8253013	9.327438e-03
Sepal.Width	-4.835449	2.824528	3.057333	0.3474466	0.4344110	1.328456e-06

\$quanti\$`3`

	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
Petal.Width	7.397121	1.880851	1.199333	0.3552873	0.7596926	1.391692e-13
Petal.Length	7.097107	5.272340	3.758000	0.6734169	1.7594041	1.273951e-12
Sepal.Length	6.220666	6.465957	5.843333	0.6116864	0.8253013	4.950500e-10
Sepal.Width	-2.501743	2.925532	3.057333	0.3028012	0.4344110	1.235835e-02

Tableau groupe kmeans/vraie espèce :

	grp	1	2	3
setosa		50	0	0
versicolor		0	41	9
virginica		0	12	38