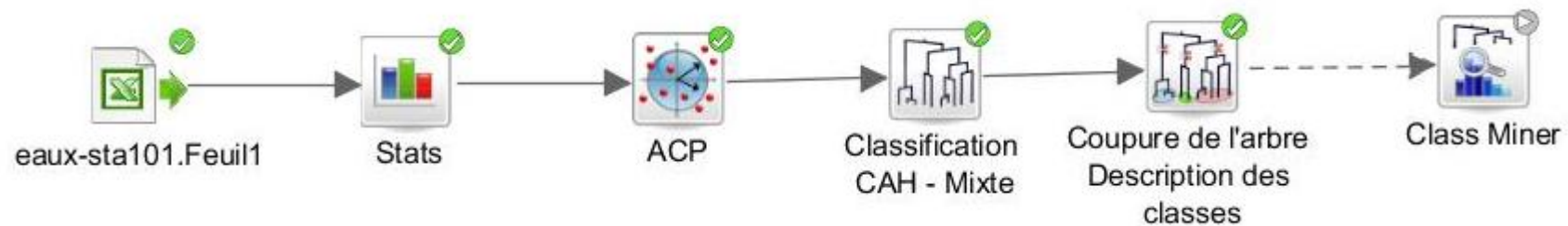


STA101 – ED7

CLASSIFICATION PAR LA METHODE DE WARD : EAUX

Cette feuille d'ED fait suite à la feuille d'ED 5 en poursuivant l'étude du fichier eaux. La chaîne de traitements réalisée avec SPAD est la suivante :



Nous étudions la partition en 3 classes.

- 1) Quels autres choix seraient possibles au niveau du nombre de classes de la partition retenue ? Pourquoi la partition en 3 classes semble t'elle préférable ?
- 2) Expliquer en quoi consiste l'étape de consolidation présentée dans le tableau 1.
- 3) Interpréter brièvement les classes à l'aide du tableau 4.
- 4) Quelle information peut-on tirer du tableau 6 ?
- 5) Interpréter en détail les résultats de la classification en commentant les tableaux 7, 8, 9 et 10.

TABLEAU 1 CLASSIFICATION HIERARCHIQUE (VOISINS RECIPROQUES) SUR LES 7 PREMIERS AXES FACTORIELS

DESCRIPTION DES		50 NOEUDS D'INDICES LES PLUS ELEVES					HISTOGRAMME DES INDICES DE NIVEAU
NUM.	AINE	BENJ	EFF.	POIDS	INDICE		
64	61	5	3	3.00	0.00004	*	
65	28	7	2	2.00	0.00006	*	
66	15	1	2	2.00	0.00007	*	
67	30	41	2	2.00	0.00007	*	
68	63	46	3	3.00	0.00010	*	
69	25	65	3	3.00	0.00013	*	
70	58	17	3	3.00	0.00022	*	
71	22	26	2	2.00	0.00023	*	
72	67	24	3	3.00	0.00025	*	
73	59	57	3	3.00	0.00025	*	
74	8	56	2	2.00	0.00039	*	
75	23	14	2	2.00	0.00040	*	
76	29	21	2	2.00	0.00046	*	
77	64	36	4	4.00	0.00051	*	
78	72	76	5	5.00	0.00054	*	
79	74	34	3	3.00	0.00061	*	
80	3	70	4	4.00	0.00074	*	
81	77	69	7	7.00	0.00077	*	
82	44	75	3	3.00	0.00094	*	
83	50	4	2	2.00	0.00099	*	
84	51	83	3	3.00	0.00108	*	
85	73	78	8	8.00	0.00175	*	
86	68	33	4	4.00	0.00178	*	
87	52	66	3	3.00	0.00187	*	
88	85	60	10	10.00	0.00210	*	
89	45	48	2	2.00	0.00217	*	
90	84	87	6	6.00	0.00344	*	
91	80	82	7	7.00	0.00362	*	
92	81	79	10	10.00	0.00411	*	
93	10	27	2	2.00	0.00444	*	
94	62	53	3	3.00	0.00653	*	
95	71	90	8	8.00	0.00661	*	
96	88	86	14	14.00	0.00669	*	
97	89	16	3	3.00	0.00699	*	
98	19	43	2	2.00	0.01640	*	
99	18	55	2	2.00	0.01708	*	
100	92	96	24	24.00	0.01722	*	
101	100	91	31	31.00	0.03868	**	
102	94	54	4	4.00	0.04009	**	
103	20	95	9	9.00	0.04419	**	
104	6	98	3	3.00	0.06951	***	
105	99	103	11	11.00	0.07178	***	
106	42	11	2	2.00	0.11012	****	
107	101	105	42	42.00	0.14469	*****	
108	104	102	7	7.00	0.22796	*****	
109	97	93	5	5.00	0.23382	*****	
110	106	108	9	9.00	0.72147	*****	
111	109	31	6	6.00	0.82130	*****	
112	110	111	15	15.00	1.67760	*****	
113	112	107	57	57.00	2.68709	*****	
SOMME DES INDICES DE NIVEAU =						7.00000	

Classification hiérarchique directe (sur facteurs)

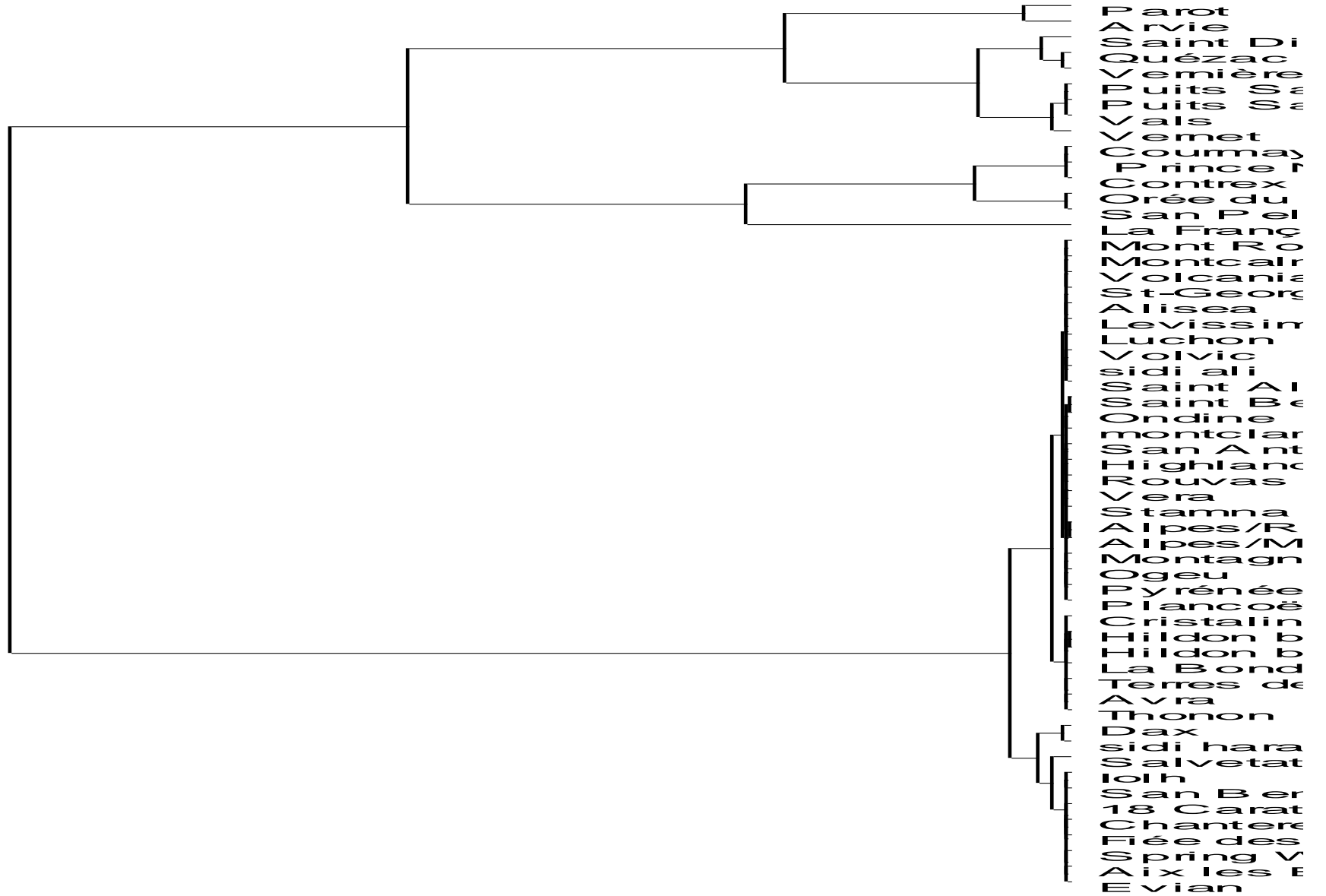


TABLEAU 2 : PARTITION PAR COUPURE D'UN ARBRE HIERARCHIQUE Coupure de l'arbre en 3 classes

FORMATION DES CLASSES (INDIVIDUS ACTIFS)
DESCRIPTION SOMMAIRE

CLASSE	EFFECTIF	POIDS	CONTENU
aa1a	42	42.00	1 A 42
aa2a	6	6.00	43 A 48
aa3a	9	9.00	49 A 57

COORDONNEES ET VALEURS-TEST AVANT CONSOLIDATION
AXES 1 A 5

CLASSES			VALEURS-TEST					COORDONNEES					
IDEN - LIBELLE	EFF.	P.ABS	1	2	3	4	5	1	2	3	4	5	DISTO.
Coupure 'a' de l'arbre en 3 classes													
aa1a - Classe 1 / 3	42	42.00	-6.2	-1.2	1.5	-1.0	-1.0	-0.96	-0.14	0.10	-0.03	-0.02	0.96
aa2a - Classe 2 / 3	6	6.00	3.0	6.1	0.2	0.8	0.8	2.28	3.43	0.07	0.14	0.09	17.03
aa3a - Classe 3 / 3	9	9.00	5.0	-3.7	-1.9	0.5	0.6	2.97	-1.64	-0.53	0.07	0.05	11.81

CONSOLIDATION DE LA PARTITION

AUTOUR DES 3 CENTRES DE CLASSES, REALISEE PAR 10 ITERATIONS A CENTRES MOBILES

PROGRESSION DE L'INERTIE INTER-CLASSES

ITERATION	I.TOTALE	I.INTER	QUOTIENT
0	7.00000	4.36469	0.62353
1	7.00000	4.47544	0.63935
2	7.00000	4.47544	0.63935
3	7.00000	4.47544	0.63935

ARRET APRES L'ITERATION 3 L'ACCROISSEMENT DE L'INERTIE INTER-CLASSES
PAR RAPPORT A L'ITERATION PRECEDENTE N'EST QUE DE 0.000 %.

TABLEAU 3 : DECOMPOSITION DE L'INERTIE CALCULEE SUR 7 AXES.

INERTIES	INERTIES		EFFECTIFS		POIDS		DISTANCES	
	AVANT	APRES	AVANT	APRES	AVANT	APRES	AVANT	APRES
INTER-CLASSES	4.3647	4.4754						
INTRA-CLASSE								
CLASSE 1 / 3	0.3745	0.4228	42	43	42.00	43.00	0.9597	0.9122
CLASSE 2 / 3	1.0687	1.0687	6	6	6.00	6.00	17.0276	17.0276
CLASSE 3 / 3	1.1921	1.0331	9	8	9.00	8.00	11.8128	14.2138
TOTALE	7.0000	7.0000						

QUOTIENT (INERTIE INTER / INERTIE TOTALE) : AVANT ... 0.6235
 APRES ... 0.6393

TABLEAU 4 : COORDONNEES ET VALEURS-TEST APRES CONSOLIDATION

IDEN - LIBELLE	CLASSES			VALEURS-TEST					COORDONNEES					DISTO.
	EFF.	P.ABS		1	2	3	4	5	1	2	3	4	5	
Coupure 'a' de l'arbre en 3 classes														
aa1a - Classe 1 / 3	43	43.00		-6.3	-1.5	1.4	-0.9	-1.3	-0.94	-0.16	0.10	-0.03	-0.03	0.91
aa2a - Classe 2 / 3	6	6.00		3.0	6.1	0.2	0.8	0.8	2.28	3.43	0.07	0.14	0.09	17.03
aa3a - Classe 3 / 3	8	8.00		5.2	-3.6	-1.9	0.4	0.9	3.32	-1.70	-0.56	0.06	0.10	14.21

TABLEAU 5 : COMPOSITION DE: Coupure 'a' de l'arbre en 3 classes

Classe 1 / 3
 Evia Mont Cris Fiée Volc Luch Volv Alpe Alpe Ondi Thon Aix La B Dax Salv Stam Iolh Avra Rouv Alis San Levi
 Vera San Sain Plan Sain St-G Hild Hild Mont Ogeu High Terr Pyrè Mont Chan 18 C Spri Vern sidi sidi mont
 Classe 2 / 3
 Orée Cont San La F Cour Pri
 Classe 3 / 3
 Sain Arvi Quéz Puit Paro Vern Puit Vals

TABLEAU 6 PARANGONS

CLASSE 1/ 3
EFFECTIF: 43

IRK	DISTANCE	IDENT.
1	0.03227	Pyrénées
2	0.04371	Stamna
3	0.04960	Ogeu
4	0.05899	Alpes/Moulettes
5	0.05899	Alpes/Roche des Ecrins
6	0.06676	Montagne des Pyrénées
7	0.07649	Cristaline-St-Cyr
8	0.09536	Vera

CLASSE 2/ 3
EFFECTIF: 6

IRK	DISTANCE	IDENT.
1	3.15144	Contrex
2	4.09072	Orée du bois
3	4.38146	Prince Noir
4	4.59600	Courmayeur
5	5.68514	San Pellegrino
6	39.01183	La Française

CLASSE 3/ 3
EFFECTIF: 8

IRK	DISTANCE	IDENT.
1	2.48698	Saint Diéry
2	3.98037	Quézac
3	4.57666	Vals
4	4.61803	Vernière
5	4.92649	Puits Saint Georges/Monoprix
6	5.01151	Puits Saint Georges/Casino
7	15.10218	Arvie
8	18.18308	Parot

MATRICE DES DISTANCES ENTRE CLASSES

	1	2	3
1	0.000		
2	4.830	0.000	
3	4.571	5.275	0.000

TABLEAU 7 : CARACTERISATION DE LA PARTITION PAR LES MODALITES DES VARIABLES QUALITATIVES

(TOUS LES KHI-2 ONT 2 DEGRES DE LIBERTE)

V.TEST	PROBA	MODALITE	QUESTION	IDEN	KHI-2	POIDS
3.56	0.000	G	PG	m1	17.15	16.00

TABLEAU 8 : CARACTERISATION DES CLASSES PAR LES MODALITES DES VARIABLES QUALITATIVES

Classe 1 / 3

V.TEST	PROBA	POURCENTAGES			MODALITES		IDEN	POIDS
		CLA/MOD	MOD/CLA	GLOBAL	CARACTERISTIQUES	DES VARIABLES		
				75.44	Classe 1 / 3		aa1a	43
3.03	0.001	87.80	83.72	71.93	P	PG	m2	41
3.02	0.001	100.00	44.19	33.33	S	TYPE	m2	19
-3.02	0.001	63.16	55.81	66.67	M	TYPE	m1	38
-3.03	0.001	43.75	16.28	28.07	G	PG	m1	16

Classe 2 / 3

V.TEST	PROBA	POURCENTAGES			MODALITES		IDEN	POIDS
		CLA/MOD	MOD/CLA	GLOBAL	CARACTERISTIQUES	DES VARIABLES		

Classe 3 / 3

V.TEST	PROBA	POURCENTAGES			MODALITES		IDEN	POIDS
		CLA/MOD	MOD/CLA	GLOBAL	CARACTERISTIQUES	DES VARIABLES		
				14.04	Classe 3 / 3		aa3a	8
4.32	0.000	50.00	100.00	28.07	G	PG	m1	16
-4.32	0.000	0.00	0.00	71.93	P	PG	m2	41

TABLEAU 9 : CARACTERISATION PAR LES VARIABLES CONTINUES

DE Coupure 'a' de l'arbre en 3 classes
(TOUTES LES STATISTIQUES DE FISHER ONT 2 DEGRES DE LIBERTE AU NUMERATEUR)

V.TEST	PROBA.	NUM . VARIABLE	IDEN	DEG.LIB.DEN	FISHER
0.15	0.442	9 . NO3	NO3	54	0.82
2.39	0.008	11 . CL	CL	54	5.18
6.83	0.000	6 . NA	NA	54	43.95
6.84	0.000	7 . K	K	54	44.19
7.73	0.000	4 . CA	CA	54	63.83
8.20	0.000	5 . MG	MG	54	77.65
8.41	0.000	10 . HCO3	HCO3	54	84.81
10.21	0.000	8 . SUL	SUL	54	182.58

TABLEAU 10 STATISTIQUES PAR CLASSE POUR LES VARIABLES CONTINUES

EDITION PAR PROBABILITES CROISSANTES
(TOUTES LES STATISTIQUES DE FISHER ONT 2 DEGRES DE LIBERTE AU NUMERATEUR)

9 . NO3 - NO3

CLASSES	EFFECTIF	POIDS	MOYENNE	ECART TYPE	MINIMUM	MAXIMUM
aa1a - Classe 1 / 3	43	43.00	4.233	7.326	0.000	35.000
aa2a - Classe 2 / 3	6	6.00	0.833	0.898	0.000	2.000
aa3a - Classe 3 / 3	8	8.00	2.500	3.202	0.000	8.000
ENSEMBLE	57	57.00	3.632	6.579	0.000	35.000

FISHER = 0.82 / 54 DEGRES DE LIBERTE AU DENOMINATEUR PROBA (FISHER > 0.82) = 0.442 / VALEUR-TEST = 0.15

11 . CL - CL

CLASSES	EFFECTIF	POIDS	MOYENNE	ECART TYPE	MINIMUM	MAXIMUM
aa1a - Classe 1 / 3	43	43.00	21.512	39.763	0.000	220.000
aa2a - Classe 2 / 3	6	6.00	189.333	355.600	1.000	982.000
aa3a - Classe 3 / 3	8	8.00	115.125	131.457	18.000	387.000
ENSEMBLE	57	57.00	52.316	142.037	0.000	982.000

FISHER = 5.18 / 54 DEGRES DE LIBERTE AU DENOMINATEUR PROBA (FISHER > 5.18) = 0.008 / VALEUR-TEST = 2.39

6 . NA - NA

CLASSES	EFFECTIF	POIDS	MOYENNE	ECART TYPE	MINIMUM	MAXIMUM
aa1a - Classe 1 / 3	43	43.00	19.791	36.939	0.000	192.000
aa2a - Classe 2 / 3	6	6.00	126.333	236.118	1.000	653.000
aa3a - Classe 3 / 3	8	8.00	466.125	233.660	154.000	968.000
ENSEMBLE	57	57.00	93.649	195.603	0.000	968.000

FISHER = 43.95 / 54 DEGRES DE LIBERTE AU DENOMINATEUR PROBA (FISHER > 43.95) = 0.000 / VALEUR-TEST = 6.83

7 . K - K

CLASSES	EFFECTIF	POIDS	MOYENNE	ECART TYPE	MINIMUM	MAXIMUM
aa1a - Classe 1 / 3	43	43.00	2.581	5.096	0.000	28.000
aa2a - Classe 2 / 3	6	6.00	6.833	7.198	2.000	22.000
aa3a - Classe 3 / 3	8	8.00	58.000	37.577	18.000	130.000
ENSEMBLE	57	57.00	10.807	24.259	0.000	130.000

FISHER = 44.19 / 54 DEGRES DE LIBERTE AU DENOMINATEUR PROBA (FISHER > 44.19) = 0.000 / VALEUR-TEST = 6.84

4 . CA - CA

CLASSES	EFFECTIF	POIDS	MOYENNE	ECART TYPE	MINIMUM	MAXIMUM
aa1a - Classe 1 / 3	43	43.00	60.116	47.402	1.000	253.000
aa2a - Classe 2 / 3	6	6.00	387.833	131.096	208.000	528.000
aa3a - Classe 3 / 3	8	8.00	115.250	70.802	45.000	241.000
ENSEMBLE	57	57.00	102.351	118.978	1.000	528.000

FISHER = 63.83 / 54 DEGRES DE LIBERTE AU DENOMINATEUR PROBA (FISHER > 63.83) = 0.000 / VALEUR-TEST = 7.73

5 . MG - MG

CLASSES	EFFECTIF	POIDS	MOYENNE	ECART TYPE	MINIMUM	MAXIMUM
aa1a - Classe 1 / 3	43	43.00	11.860	10.449	0.000	40.000
aa2a - Classe 2 / 3	6	6.00	72.833	10.123	55.000	84.000
aa3a - Classe 3 / 3	8	8.00	64.375	28.173	21.000	95.000
ENSEMBLE	57	57.00	25.649	28.157	0.000	95.000

FISHER = 77.65 / 54 DEGRES DE LIBERTE AU DENOMINATEUR PROBA (FISHER > 77.65) = 0.000 / VALEUR-TEST = 8.20

10 . HCO3 - HCO3

CLASSES	EFFECTIF	POIDS	MOYENNE	ECART TYPE	MINIMUM	MAXIMUM
aa1a - Classe 1 / 3	43	43.00	223.907	167.635	4.000	820.000
aa2a - Classe 2 / 3	6	6.00	272.667	78.147	168.000	403.000
aa3a - Classe 3 / 3	8	8.00	1741.125	685.528	1170.000	3380.000
ENSEMBLE	57	57.00	441.982	602.971	4.000	3380.000

FISHER = 84.81 / 54 DEGRES DE LIBERTE AU DENOMINATEUR PROBA (FISHER > 84.81) = 0.000 / VALEUR-TEST = 8.41

8 . SUL - SUL

CLASSES	EFFECTIF	POIDS	MOYENNE	ECART TYPE	MINIMUM	MAXIMUM
aa1a - Classe 1 / 3	43	43.00	26.791	55.109	1.000	365.000
aa2a - Classe 2 / 3	6	6.00	1023.167	323.008	549.000	1371.000
aa3a - Classe 3 / 3	8	8.00	54.125	56.492	10.000	158.000
ENSEMBLE	57	57.00	135.509	326.356	1.000	1371.000

FISHER = 182.58 / 54 DEGRES DE LIBERTE AU DENOMINATEUR PROBA (FISHER > 182.58) = 0.000 / VALEUR-TEST = 10.21

TABLEAU 11 : CARACTERISATION PAR LES VARIABLES CONTINUES DES CLASSES

Classe 1 / 3

V.TEST	PROBA	MOYENNES		ECARTS TYPES		VARIABLES CARACTERISTIQUES		IDEN
		CLASSE GENERALE	CLASSE GENERALE	CLASSE GENERAL	CLASSE GENERAL	NUM.LIBELLE		
		Classe 1 / 3		(POIDS = 43.00		EFFECTIF = 43)		aa1a
1.20	0.115	4.23	3.63	7.33	6.58	9.NO3		NO3
-2.84	0.002	21.51	52.32	39.76	142.04	11.CL		CL
-4.37	0.000	26.79	135.51	55.11	326.36	8.SUL		SUL
-4.45	0.000	2.58	10.81	5.10	24.26	7.K		K
-4.66	0.000	60.12	102.35	47.40	118.98	4.CA		CA
-4.74	0.000	223.91	441.98	167.63	602.97	10.HCO3		HCO3
-4.95	0.000	19.79	93.65	36.94	195.60	6.NA		NA
-6.42	0.000	11.86	25.65	10.45	28.16	5.MG		MG

Classe 2 / 3

V.TEST	PROBA	MOYENNES		ECARTS TYPES		VARIABLES CARACTERISTIQUES		IDEN
		CLASSE GENERALE		CLASSE GENERAL		NUM.LIBELLE		
		Classe 2 / 3		(POIDS = 6.00		EFFECTIF = 6)		aa2a
6.98	0.000	1023.17	135.51	323.01	326.36	8.SUL	SUL	
6.16	0.000	387.83	102.35	131.10	118.98	4.CA	CA	
4.30	0.000	72.83	25.65	10.12	28.16	5.MG	MG	
2.48	0.007	189.33	52.32	355.60	142.04	11.CL	CL	
0.43	0.334	126.33	93.65	236.12	195.60	6.NA	NA	
-0.42	0.337	6.83	10.81	7.20	24.26	7.K	K	
-0.72	0.236	272.67	441.98	78.15	602.97	10.HCO3	HCO3	
-1.09	0.137	0.83	3.63	0.90	6.58	9.NO3	NO3	

Classe 3 / 3

V.TEST	PROBA	MOYENNES		ECARTS TYPES		VARIABLES CARACTERISTIQUES		IDEN
		CLASSE GENERALE		CLASSE GENERAL		NUM.LIBELLE		
		Classe 3 / 3		(POIDS = 8.00		EFFECTIF = 8)		aa3a
6.51	0.000	1741.12	441.98	685.53	602.97	10.HCO3	HCO3	
5.88	0.000	58.00	10.81	37.58	24.26	7.K	K	
5.76	0.000	466.13	93.65	233.66	195.60	6.NA	NA	
4.16	0.000	64.38	25.65	28.17	28.16	5.MG	MG	
1.34	0.091	115.13	52.32	131.46	142.04	11.CL	CL	
0.33	0.372	115.25	102.35	70.80	118.98	4.CA	CA	
-0.52	0.302	2.50	3.63	3.20	6.58	9.NO3	NO3	
-0.75	0.225	54.13	135.51	56.49	326.36	8.SUL	SUL	

Variable: Coupure 'a' de l'arbre en 3 classes - Valeurs-test

