Analyse Factorielle des Correspondances
(TP SAS)

Bilan des méthodes

Donnée	Quantitative	Qualitative
Tableau	Individus × Variables	Variable1 × Variable2
		Individus × Variables
Exemples	forsy	CSP, Pommes
Analyse	ACP	AFC, ACM
	Classification	
Programmes		CSP.sas, Pommes.sas

Fichiers de données : CSP.dat

	dro	sci	let	med	pha	iut
age	4814	8941	8719	1583	959	558
oua	463	1118	1224	153	33	83
ind	3610	2608	3365	1634	728	110
art	2870	5091	5966	1583	562	261
gco	2777	2608	3212	1736	794	136
pco	4444	6209	7954	2655	1125	313
lib	9813	7823	9637	8985	3143	120
ing	4444	8195	7648	3880	1307	151
adm	14071	13535	18050	7147	2134	235
mes	833	993	1071	817	182	21
mso	555	621	765	357	66	16
tec	2129	3974	4436	1021	298	235
cam	7776	10058	13308	3931	1125	407
ins	1666	5464	6425	1481	579	141
emb	5184	8071	10861	2399	529	0
con	1481	2856	3059	613	132	177
ous	2314	4098	4436	766	182	516
min	278	869	918	153	17	37
mar	185	373	306	51	0	21
man	741	1863	1989	255	33	83
ser	741	993	1377	204	33	63
arm	2592	3104	3671	817	198	136

Lecture du fichier

```
Data test;
 infile '../Pgm-s/csp.dat' firstobs=2 expandtabs;
 input origine$ dro sci let med pha iut;
run;


Data CSP;
 set test;
 length etude $4.;
 array tmp{*} dro--iut;
    do i=1 to dim(tmp);
       poids=tmp{i};
       call vname(tmp{i},etude);
       output;
    end;
 keep origine etude poids;
run;
```

Analyse préliminaire

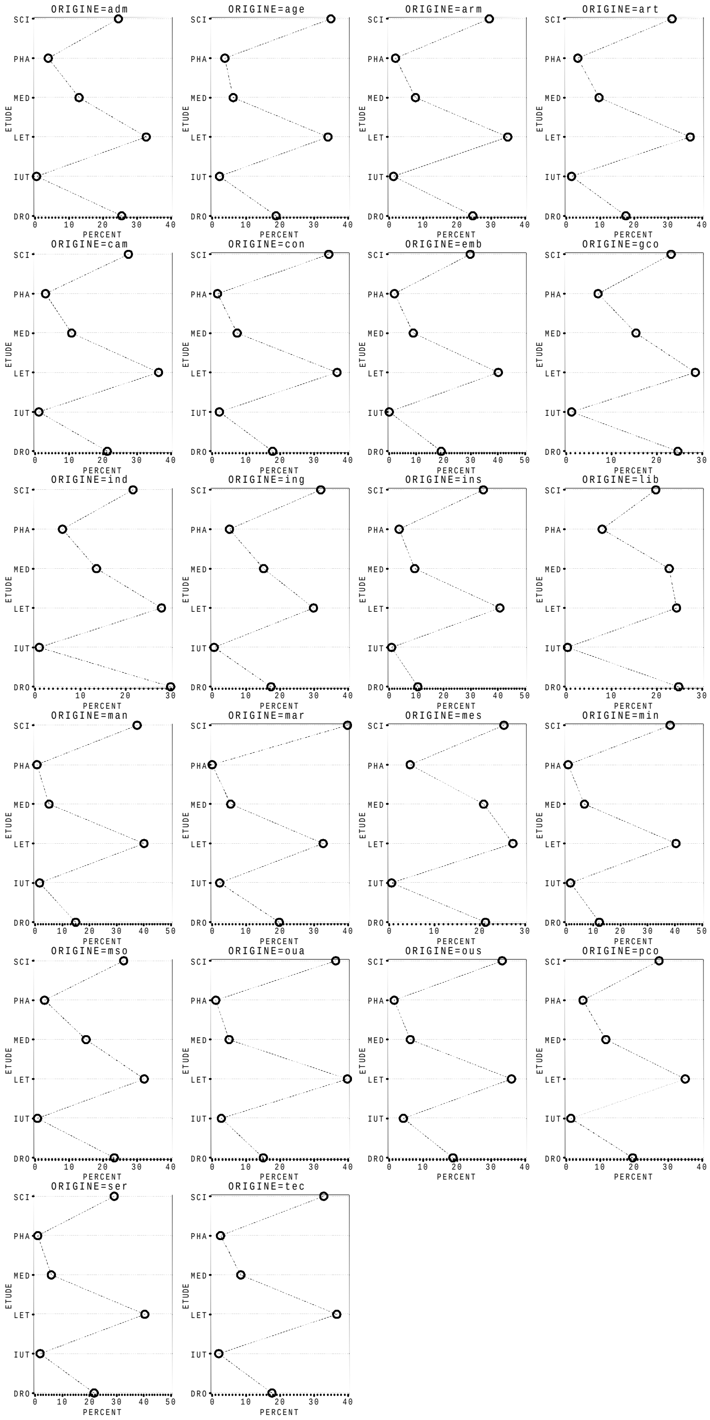
```
Proc Freq data=CSP order=data;
 table origine*etude;
 weight poids;
run;

TABLE OF ORIGINE BY ETUDE

ORIGINE     ETUDE

Frequency|
Percent  |
Row Pct  |
Col Pct  |DRO     |SCI     |LET     |MED     |PHA     |IUT     | Total
---------+--------+--------+--------+--------+--------+--------+
age      |   4814 |   8941 |   8719 |   1583 |    959 |    558 | 25574
         |   1.37 |   2.54 |   2.48 |   0.45 |   0.27 |   0.16 |  7.27
         |  18.82 |  34.96 |  34.09 |   6.19 |   3.75 |   2.18 |
         |   6.52 |   8.99 |   7.36 |   3.75 |   6.77 |  14.61 |
---------+--------+--------+--------+--------+--------+--------+
oua      |    463 |   1118 |   1224 |    153 |     33 |     83 |  3074
         |   0.13 |   0.32 |   0.35 |   0.04 |   0.01 |   0.02 |  0.87
         |  15.06 |  36.37 |  39.82 |   4.98 |   1.07 |   2.70 |
         |   0.63 |   1.12 |   1.03 |   0.36 |   0.23 |   2.17 |
---------+--------+--------+--------+--------+--------+--------+
ind      |   3610 |   2608 |   3365 |   1634 |    728 |    110 | 12055
         |   1.03 |   0.74 |   0.96 |   0.46 |   0.21 |   0.03 |  3.43
         |  29.95 |  21.63 |  27.91 |  13.55 |   6.04 |   0.91 |
         |   4.89 |   2.62 |   2.84 |   3.87 |   5.14 |   2.88 |
...
---------+--------+--------+--------+--------+--------+--------+
ing      |   4444 |   8195 |   7648 |   3880 |   1307 |    151 | 25625
         |   1.26 |   2.33 |   2.17 |   1.10 |   0.37 |   0.04 |  7.28
         |  17.34 |  31.98 |  29.85 |  15.14 |   5.10 |   0.59 |
         |   6.02 |   8.24 |   6.46 |   9.19 |   9.23 |   3.95 |
---------+--------+--------+--------+--------+--------+--------+
Total       73781    99465   118397    42221    14159     3820  351843
            20.97    28.27    33.65    12.00     4.02     1.09  100.00
```

```
Proc Freq data=CSP order=data;
 table origine*etude /noprint chisq;
 weight poids;
run;
```

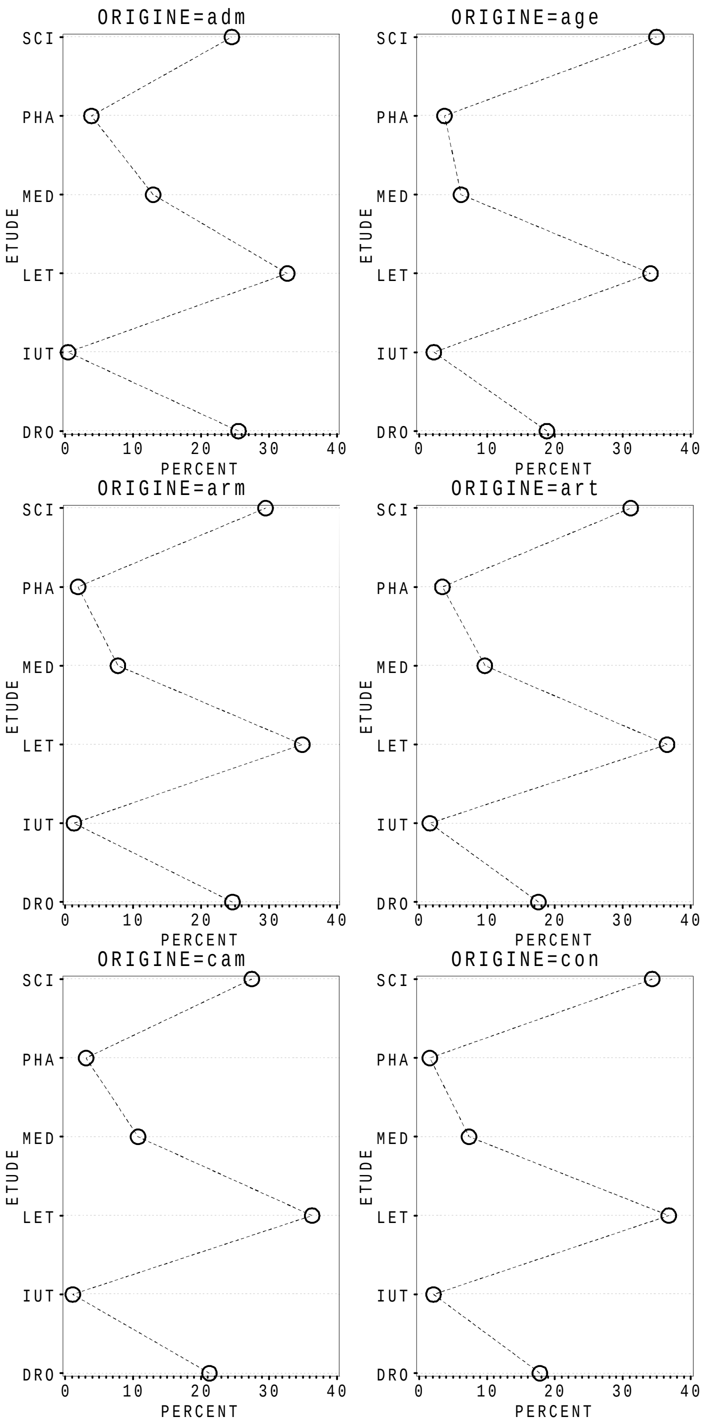
```
STATISTICS FOR TABLE OF ORIGINE BY ETUDE

Statistic                     DF     Value        Prob
------------------------------------------------------
Chi-Square                   105 21181.127       0.001
Likelihood Ratio Chi-Square  105 20904.564       0.001
Mantel-Haenszel Chi-Square     1   284.429       0.001
Phi Coefficient                      0.245
Contingency Coefficient              0.238
Cramer's V                           0.110

Sample Size = 351843
```

AFC

```
Proc Corresp data=CSP outf=tabfreq noprint;
  tables origine,etude;
  weight poids;
run;

goptions htext=2.5 hby=3 nodisplay;

Proc Datasets library=work;
  delete graphe gout/ memtype=catalog;
run;
quit;

Proc sort data=tabfreq;
 by origine;
run;

Proc Gplot data=tabfreq(where=(_type_ ='RP')) gout=graphe;
 symbol1 v=circle i=join l=3 h=5;
 axis1 label=(a=90);
 plot etude*percent /frame autovref lvref=33 vaxis=axis1;
 by origine;
run;
quit;

Goptions display;

%inc 'panels.sas';

%panels(rows=6,cols=4,plots=22,order=DOWN,
        equate=Y,gin=graphe,gout=graphe);
```


ORIGINE=adm
ORIGINE=age
ORIGINE=arm
ORIGINE=art
ORIGINE=cam
ORIGINE=con
ORIGINE=emb
ORIGINE=gco
ORIGINE=ind
ORIGINE=ing
ORIGINE=ins
ORIGINE=lib
ORIGINE=man
ORIGINE=mar
ORIGINE=mes
ORIGINE=min
ORIGINE=mso
ORIGINE=oua
ORIGINE=ous
ORIGINE=pco
ORIGINE=ser
ORIGINE=tec
ETUDE
SCI
PHA
MED
LET
IUT
DRO
PERCENT

```
Proc sort data=tabFreq;
 by etude;
run;

Goptions nodisplay;

Proc Gplot data=TabFreq(where=(_type_ ='CP')) gout=graphe;
 symbol1 v=circle i=join l=3 h=5;
 axis1 label=(a=90);
 plot origine*percent /frame autovref lvref=33 vaxis=axis1;
 by etude;
run;
quit;

Goptions display;

%panels(rows=3,cols=2,plots=6,order=DOWN,
        equate=Y,gin=graphe,gout=graphe);

quit;
```


ORIGINE=adm
ORIGINE=age
ORIGINE=arm
ORIGINE=art
ORIGINE=cam
ORIGINE=con
ETUDE
SCI
PHA
MED
LET
IUT
DRO
PERCENT
0 10 20 30 40

```
Proc Corresp data=CSP outc=graphe dim=3;
 table origine,etude;
 weight poids;
run;
```

```
Inertia and Chi-Square Decomposition

Singular  Principal Chi-
Values    Inertias  Squares Percents   15   30   45   60   75
                                    ----+----+----+----+----+---
0.21290   0.04533     15948  75.29% *************************
0.08322   0.00693   2436.52  11.50% ****
0.07551   0.00570    2006.1   9.47% ***
0.03454   0.00119    419.73   1.98% *
0.03246   0.00105   370.738   1.75% *
          -------   -------
          0.06020   21181.1 (Degrees of Freedom = 105)
```

Row Coordinates

	Dim1	Dim2	Dim3
adm	0.093103	-.097952	-.037983
age	-.204953	0.024170	0.096188
arm	-.119086	-.134009	0.009826
art	-.127387	0.042609	0.011147

Summary Statistics for the Row Points

	Quality	Mass	Inertia
adm	0.993601	0.156809	0.051659
age	0.901919	0.072686	0.069400
arm	0.943464	0.029894	0.016967
art	0.974484	0.046421	0.014376

Partial Contributions to Inertia for the Row Points

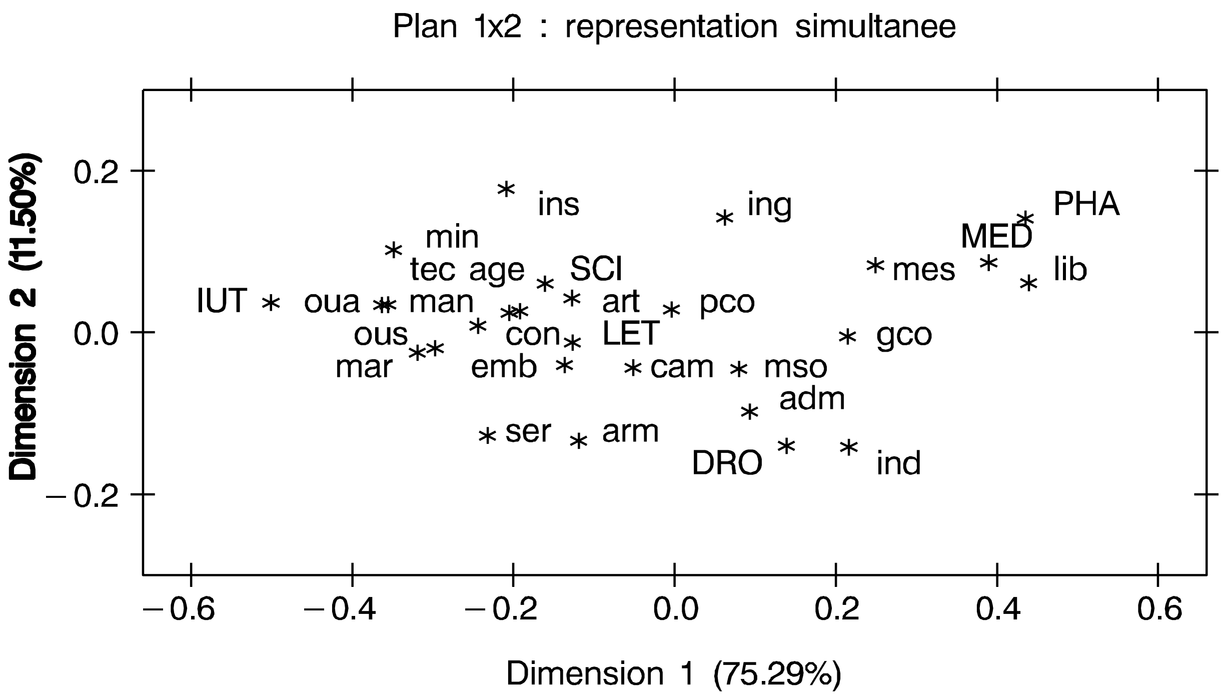
	Dim1	Dim2	Dim3
adm	0.029987	0.217257	0.039678
age	0.067359	0.006132	0.117947
arm	0.009353	0.077523	0.000506
art	0.016619	0.012170	0.001012

Squared Cosines for the Row Points

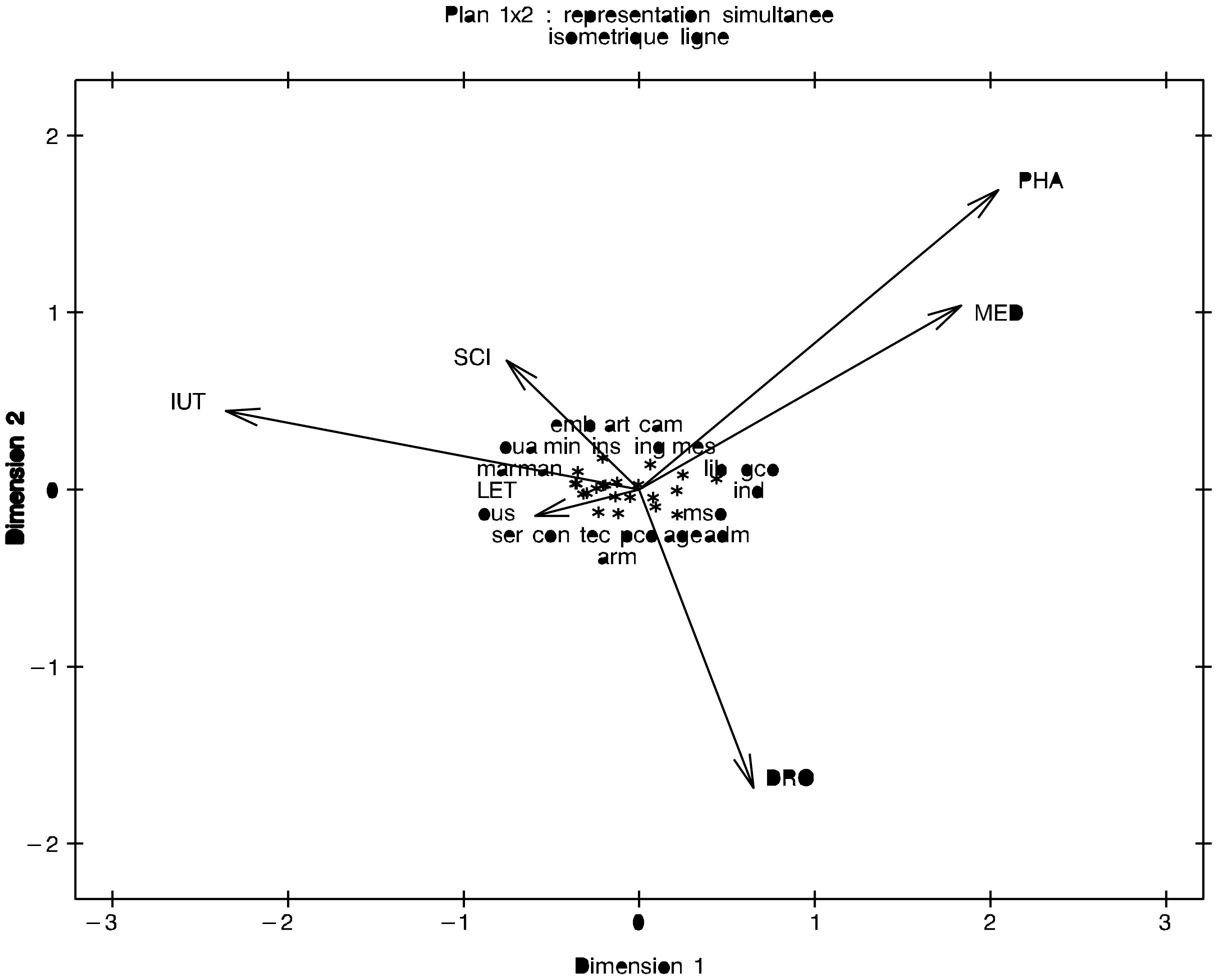
	Dim1	Dim2	Dim3
adm	0.437071	0.483784	0.072746
age	0.730793	0.010163	0.160963
arm	0.415048	0.525590	0.002826
art	0.870434	0.097384	0.006665

```
Column Coordinates

               Dim1           Dim2           Dim3

DRO        0.138733       -.140311       0.030680
IUT        -.500889       0.037051       0.629807
LET        -.126509       -.012292       -.049852
MED        0.390065       0.086375       -.017204

Summary Statistics for the Column Points

            Quality           Mass        Inertia

DRO        0.990697       0.209699       0.140204
IUT        0.972333       0.010857       0.120362
LET        0.937254       0.336505       0.111173
MED        0.979480       0.120000       0.325426

Partial Contributions to Inertia for the Column Points

               Dim1           Dim2           Dim3

DRO        0.089043       0.596155       0.034619
IUT        0.060095       0.002152       0.755312
LET        0.118817       0.007341       0.146671
MED        0.402806       0.129280       0.006229

Squared Cosines for the Column Points

               Dim1           Dim2           Dim3

DRO        0.478186       0.489126       0.023386
IUT        0.375929       0.002057       0.594346
LET        0.804704       0.007596       0.124953
MED        0.931969       0.045698       0.001813
```

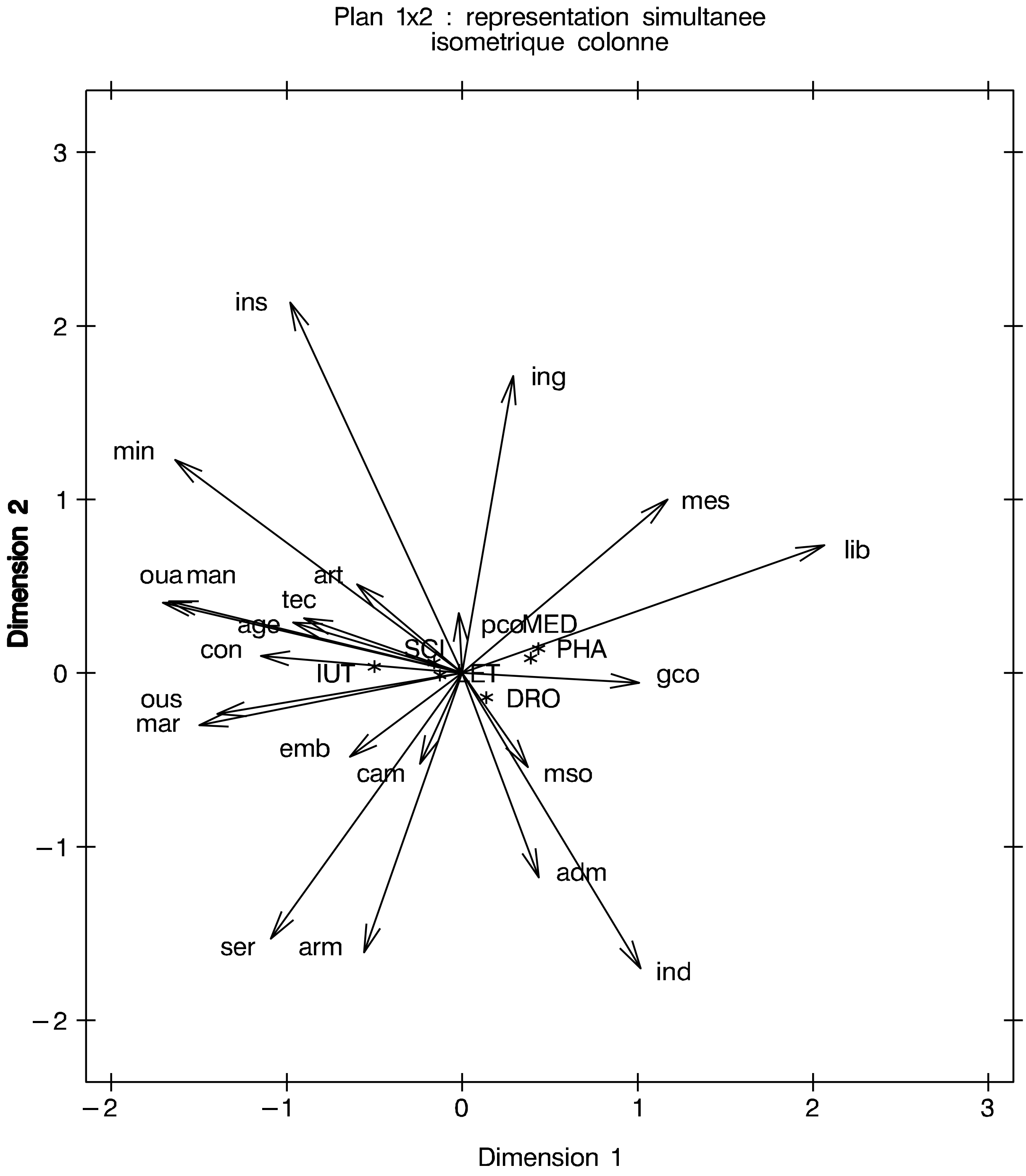
```
options nodate center;
title 'Plan 1x2 : representation simultanee';

%plotit(data=graphe,plotvars=dim2 dim1,labelvar=_name_,
        datatype=corresp,
        method=print,post=plan12.eps,
        gopprint=dev=psepsf gsfmode=replace gaccess=gsasfile,
        monochro=black);
```



```
* -------- Représentation simultanée isométrique ligne;

title 'Plan 1x2 : representation simultanee';
title2 ' isometrique ligne';


Proc corresp data=CSP outc=graphe dim=5 profile=row noprint;
 table origine,etude;
 weight poids;
run;

%plotit(data=graphe,plotvars=dim2 dim1,labelvar=_name_,
        datatype=row,method=print,post=plan12l.eps,
        gopprint=dev=psepsf gsfmode=replace gaccess=gsasfile,
        monochro=black);
```


Plan 1x2 : representation simultanee
isometrique ligne
PHA
MED
SCI
IUT
DRO
emb art cam
oua min ins ing mes
marman
lib gco
LET
ind
ous
mse
ser con tec pco ageadm
arm
Dimension 1
Dimension 2
2
1
0
−1
−2
−3
3

```
* -------- Représentation simultanée isométrique colonne;

title 'Plan 1x2 : representation simultanee';
title2 ' isometrique colonne';

Proc corresp data=CSP outc=graphe dim=5 profile=column noprint;
 table origine,etude;
 weight poids;
run;


%plotit(data=graphe,plotvars=dim2 dim1,labelvar=_name_,
        datatype=column,method=print,post=plan12c.eps,
        gopprint=dev=psepsf gsfmode=replace gaccess=gsasfile,
        monochro=black);
```


Plan 1x2 : representation simultanee
isometrique colonne
Dimension 1
Dimension 2
ins
ing
min
mes
lib
oua man
art
tec
age
pcoMED
SCI
PHA
con
IUT
LET
gco
DRO
ous
mar
emb
cam
mso
adm
ser
arm
ind
3
2
1
0
−1
−2
−2
−1
0
1
2
3