

Exemples d'analyses en composantes principales

Exemple 1

On cherche à étudier les différentes méthodes de production d'électricité dans différents pays du monde. On restreint notre étude à des pays qui produisent de l'énergie nucléaire. Cette étude a pour objectif de mettre en évidence les pays qui basent principalement leur production d'électricité sur des types particuliers d'énergie. On pourra par la suite envisager d'aider au développement de nouveaux modes de production moins polluants ou plus sécuritaires mais aussi mettre en avant les modèles de production basés sur de énergies renouvelables.

Pour chacun de ces pays, la production d'électricité est décrite au travers de 4 variables:

Hydraulique	proportion d'énergie hydraulique
Nucléaire	proportion d'énergie nucléaire
Fossile	proportion d'énergie fossile
Autres.Renouv	proportion d'autres types d'énergies renouvelables (géothermie, biomasse, marine, solaire, éolienne).

On dispose également d'une cinquième variable IDH correspondant à l'indice de développement humain.

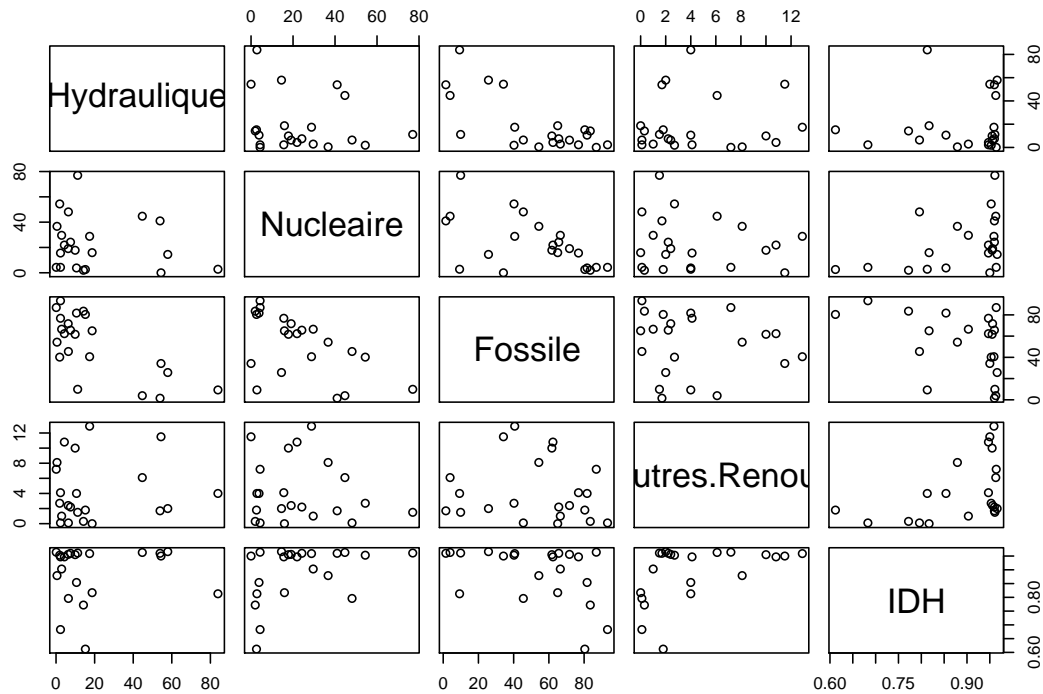
On a créé dans le logiciel R un tableau `tab3` contenant les données.

```
> tab3
```

	Hydraulique	Nucleaire	Fossile	Autres.Renouv	IDH
Canada	57.9	14.5	25.7	2.0	0.966
USA	6.3	19.1	71.7	2.4	0.956
Mexique	10.6	3.8	81.7	4.0	0.854
Bresil	83.9	2.8	9.3	4.0	0.813
Allemagne	4.3	21.9	62.3	10.8	0.947
Belgique	1.9	54.4	40.2	2.7	0.953
Espagne	9.9	17.8	61.7	10.0	0.955
Finlande	17.4	28.8	40.6	12.9	0.959
France	11.2	77.0	9.9	1.5	0.961
Pays-Bas	0.1	4.3	86.9	7.2	0.964
Royaume-Uni	2.3	15.6	76.8	4.1	0.947
Suede	44.7	44.7	3.9	6.1	0.963
Suisse	53.9	41.0	1.5	1.7	0.960
Tchequie	2.9	29.6	66.5	1.0	0.903
Hongrie	0.5	36.7	54.3	8.1	0.879
Russie	18.7	15.9	65.0	0.0	0.817
Ukraine	6.4	48.1	45.5	0.1	0.796
Afrique.du.Sud	2.3	4.3	93.3	0.1	0.683
Chine	14.2	2.0	83.5	0.3	0.772
Japon	7.5	24.2	65.6	2.2	0.960
Inde	15.2	2.6	80.4	1.8	0.612
Nouvelle-Zelande	54.4	0.0	34.2	11.5	0.950

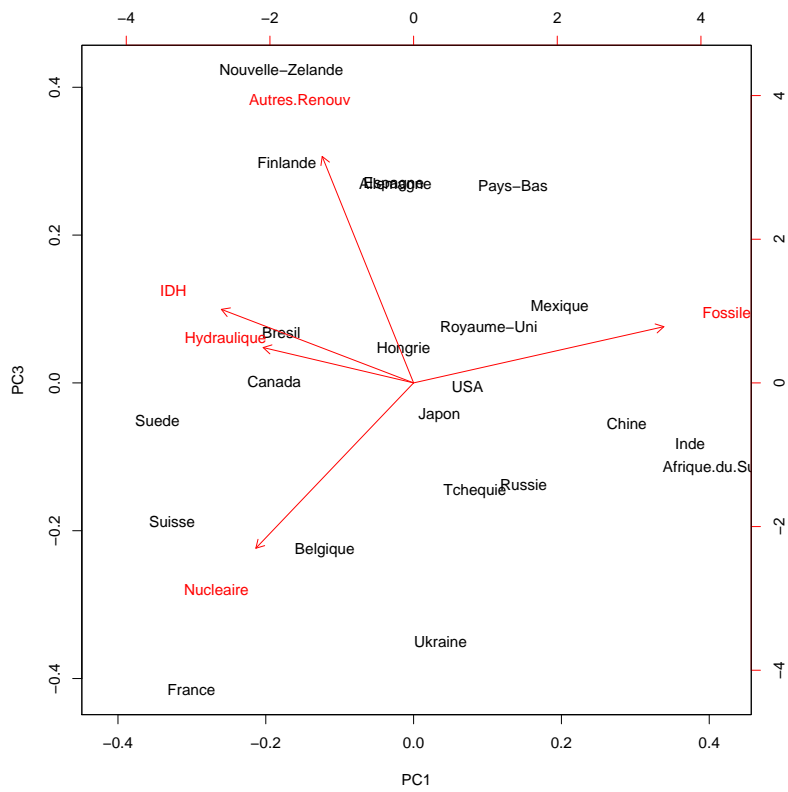
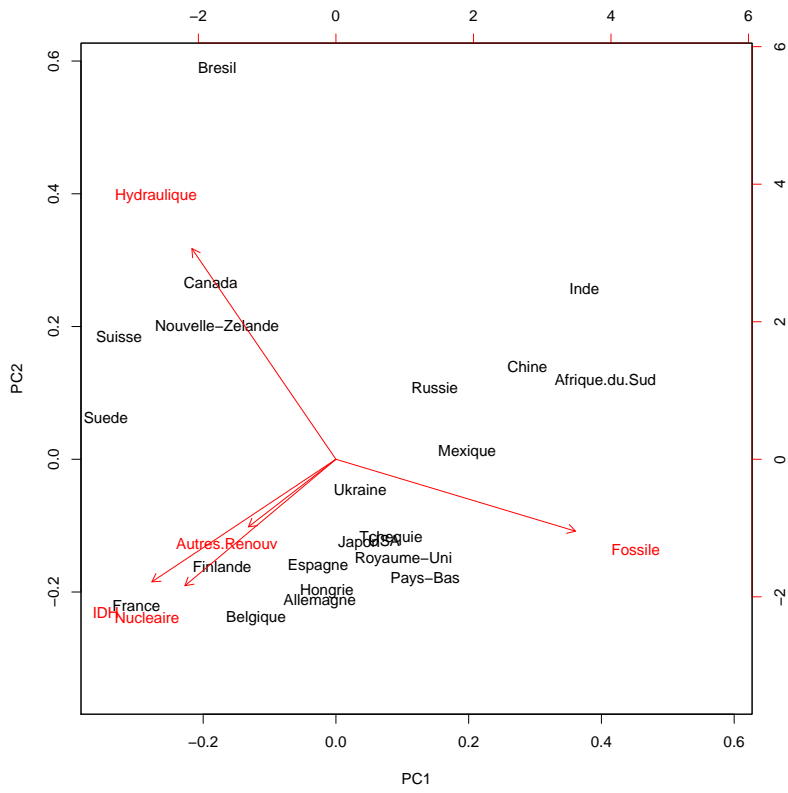
```

> cor(tab3)
                Hydraulique  Nucleaire  Fossile  Autres.Renov  IDH
Hydraulique      1.0000000 -0.1679988 -0.7093665   0.06245836  0.07670006
Nucleaire       -0.16799877  1.0000000 -0.5608144  -0.12107544  0.39431619
Fossile         -0.70936651 -0.5608144  1.0000000  -0.10946411 -0.41674349
Autres.Renov    0.06245836 -0.1210754 -0.1094641   1.00000000  0.44989479
IDH             0.07670006  0.3943162 -0.4167435   0.44989479  1.00000000
> plot(data.frame(tab3))
    
```



```

> acp<-prcomp(tab3,scale=T)
> summary(acp)
Importance of components:
                PC1  PC2  PC3  PC4  PC5
Standard deviation   1.463 1.131 1.103 0.6037 0.01097
Proportion of Variance 0.428 0.256 0.243 0.0729 0.00002
Cumulative Proportion 0.428 0.684 0.927 1.0000 1.00000
> biplot(acp)
> biplot(acp,choices=c(1,3))
    
```



Exemple 2

Le tableau suivant contient les données météorologiques de 46 villes françaises. Les colonnes correspondent à

- Nom de la ville
- Pluviométrie annuelle
- Nombre de jours de pluie
- Durée d'insolation (h/an)
- Nombre de jours avec faible ensoleillement
- Nombre de jours avec fort ensoleillement

> tab

	Pluvio	Jours pluie	Insolation	Faible ensoleil.	Fort ensoleil.
Biarritz	1483	143	1877	130	69
Aurillac	1232	135	2084	126	103
Besancon	1138	140	1797	152	69
Pau	1133	130	1852	134	72
Brest	1109	155	1492	164	38
Limoges	1049	136	1860	146	78
Grenoble	965	110	2020	126	84
Mont de Marsan	948	126	1852	128	61
Bordeaux	923	125	1992	115	75
Deauville	870	138	1600	100	70
Lyon	843	107	1932	136	75
Bastia	835	68	2533	70	115
Nice	803	63	2668	66	142
Dieppe	799	129	1600	100	70
Nantes	788	118	1690	142	49
Rouen	785	131	1518	168	39
Millau	779	101	2121	125	102
Metz	766	124	1605	167	49
Albi	761	106	2077	119	94
Nancy	759	124	1638	163	53
La Rochelle	755	115	1600	100	70
Dijon	744	115	1789	147	66
Bourges	739	119	1787	147	63
Saint-Auban	738	77	2758	65	160
Alencon	735	119	1615	157	48
Caen	711	123	1624	151	39
Le Havre	709	125	1600	100	70
Saint-Etienne	706	98	2007	124	83
Tours	694	113	1799	145	62
Carcassonne	688	92	2106	113	88
Cherbourg	688	119	1600	100	70
Lille	686	121	1617	157	42
Toulon	685	61	2793	64	160
Dunkerque	676	118	1600	100	70
Montpellier	655	60	2618	77	142
Toulouse	655	146	2010	121	81
Rennes	649	115	1626	150	41
Paris	641	111	1630	156	47

Ajaccio	640	72	2726	64	150
Orleans	637	113	1710	151	54
Strasbourg	610	111	1633	167	53
Avranches	607	119	1600	100	70
Clermont-Ferrand	592	90	1898	133	69
Marseille	555	56	2801	63	161
Perpignan	547	56	2392	86	114

```
> cor(tab)
```

```

          Pluvio Jours pluie Insolation Faible ensoleil.
Pluvio      1.000000  0.5252274 -0.1267316      0.2122035
Jours pluie 0.5252274  1.0000000 -0.8239493      0.6879070
Insolation  -0.1267316 -0.8239493  1.0000000     -0.7615970
Faible ensoleil. 0.2122035  0.6879070 -0.7615970      1.0000000
Fort ensoleil. -0.1367431 -0.7838400  0.9573736     -0.8631445

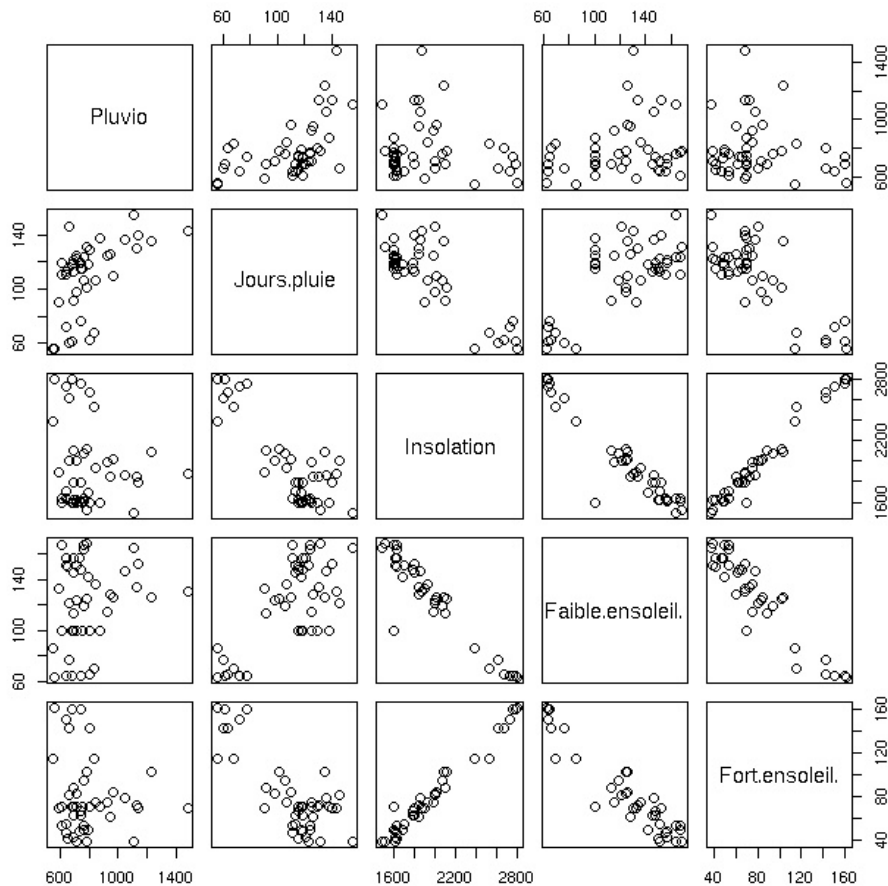
```

```

          Fort ensoleil.
Pluvio      -0.1367431
Jours pluie -0.7838400
Insolation   0.9573736
Faible ensoleil. -0.8631445
Fort ensoleil.  1.0000000

```

```
> plot(data.frame(tab))
```



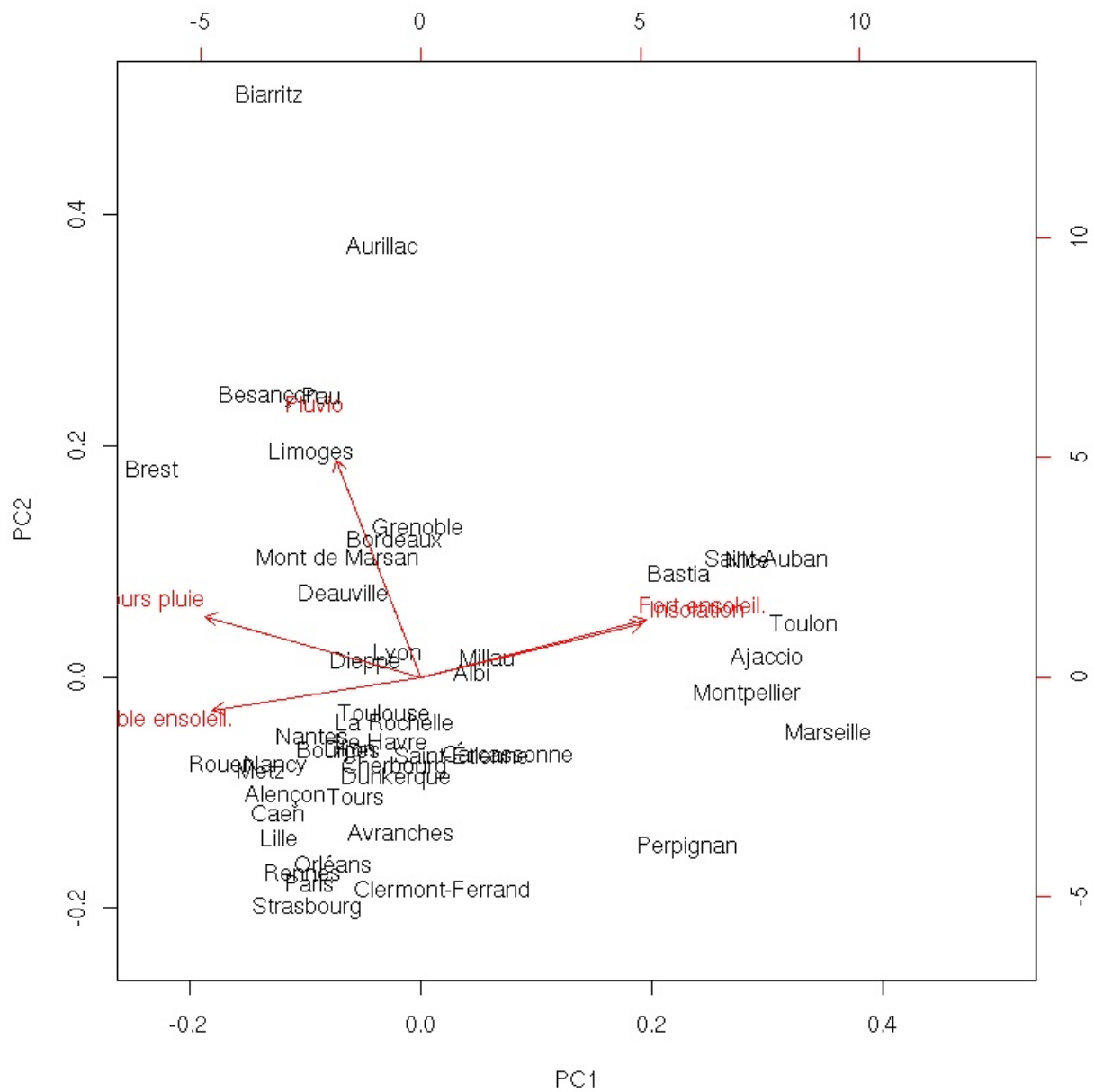
```
> acp<-prcomp(tab, scale=TRUE)
```

```
> summary(acp)
```

Importance of components:

	PC1	PC2	PC3	PC4	PC5
Standard deviation	1.882	1.023	0.5440	0.3103	0.14342
Proportion of Variance	0.708	0.209	0.0592	0.0192	0.00411
Cumulative Proportion	0.708	0.917	0.9766	0.9959	1.00000

```
> biplot(acp)
```



Exemple 3

Le crâne d'un animal préhistorique appartenant à la famille des canidés a été découvert il y a quelques années, dans la région de Jussac (Auvergne). L'une des questions que se posaient les paléontologues était de savoir si cet animal se rapprochait plus d'un chien ou d'un loup.

On a mesuré 6 grandeurs caractéristiques sur des crânes de chiens de même taille que celle de l'animal inconnu, et sur des crânes de loups. Ces variables sont

LCB longueur condylo-basale
 LMS longueur de la mâchoire supérieure
 LBM largeur bi-maxillaire
 LP longueur de la carnassière supérieure
 LM longueur de la première molaire supérieure
 LAM largeur de la première molaire supérieure

On a créé dans le logiciel R un tableau `Jussac` contenant les données. Ici la lettre C désigne les mesures effectuées sur des chiens, la lettre L celles effectuées sur des loups, et l'étoile celles effectuées sur l'animal inconnu.

```
> Jussac
  LCB LMS LBM  LP  LM  LAM
C 129  64  95 17.5 11.2 13.8
C 154  74  76 20.0 14.2 16.5
C 170  87  71 17.9 12.3 15.9
C 188  94  73 19.5 13.3 14.8
C 161  81  55 17.1 12.1 13.0
C 164  90  58 17.5 12.7 14.7
C 203 109  65 20.7 14.0 16.8
C 178  97  57 17.3 12.8 14.3
C 212 114  65 20.5 14.3 15.5
C 221 123  62 21.2 15.2 17.0
C 183  97  52 19.3 12.9 13.5
C 212 112  65 19.7 14.2 16.0
C 220 117  70 19.8 14.3 15.6
C 216 113  72 20.5 14.4 17.7
C 216 112  75 19.6 14.0 16.4
C 205 110  68 20.8 14.1 16.4
C 228 122  78 22.5 14.2 17.8
C 218 112  65 20.3 13.9 17.0
C 190  93  78 19.7 13.2 14.0
C 212 111  73 20.5 13.7 16.6
C 201 105  70 19.8 14.3 15.9
C 196 106  67 18.5 12.6 14.2
C 158  71  71 16.7 12.5 13.3
C 255 126  86 21.4 15.0 18.0
C 234 113  83 21.3 14.8 17.0
C 205 105  70 19.0 12.4 14.9
C 186  97  62 19.0 13.2 14.2
C 241 119  87 21.0 14.7 18.3
C 220 111  88 22.5 15.4 18.0
C 242 120  85 19.9 15.3 17.6
```

```
L 199 105 73 23.4 15.0 19.1
L 227 117 77 25.0 15.3 18.6
L 228 122 82 24.7 15.0 18.5
L 232 123 83 25.3 16.8 15.5
L 231 121 78 23.5 16.5 19.6
L 215 118 74 25.7 15.7 19.0
L 184 100 69 23.3 15.8 19.7
L 175 94 73 22.2 14.8 17.0
L 239 124 77 25.0 16.8 27.0
L 203 109 70 23.3 15.0 18.7
L 226 118 72 26.0 16.0 19.4
L 226 119 77 26.5 16.8 19.3
* 210 103 72 20.5 14.0 16.7
```

```
> cor(Jussac)
```

	LCB	LMS	LBM	LP	LM	LAM
LCB	1.0000000	0.9587410	0.3481835	0.6129486	0.7179356	0.5872510
LMS	0.9587410	1.0000000	0.2003331	0.6610016	0.7359561	0.5946533
LBM	0.3481835	0.2003331	1.0000000	0.3699619	0.3502798	0.3547771
LP	0.6129486	0.6610016	0.3699619	1.0000000	0.8935121	0.7626431
LM	0.7179356	0.7359561	0.3502798	0.8935121	1.0000000	0.7892164
LAM	0.5872510	0.5946533	0.3547771	0.7626431	0.7892164	1.0000000

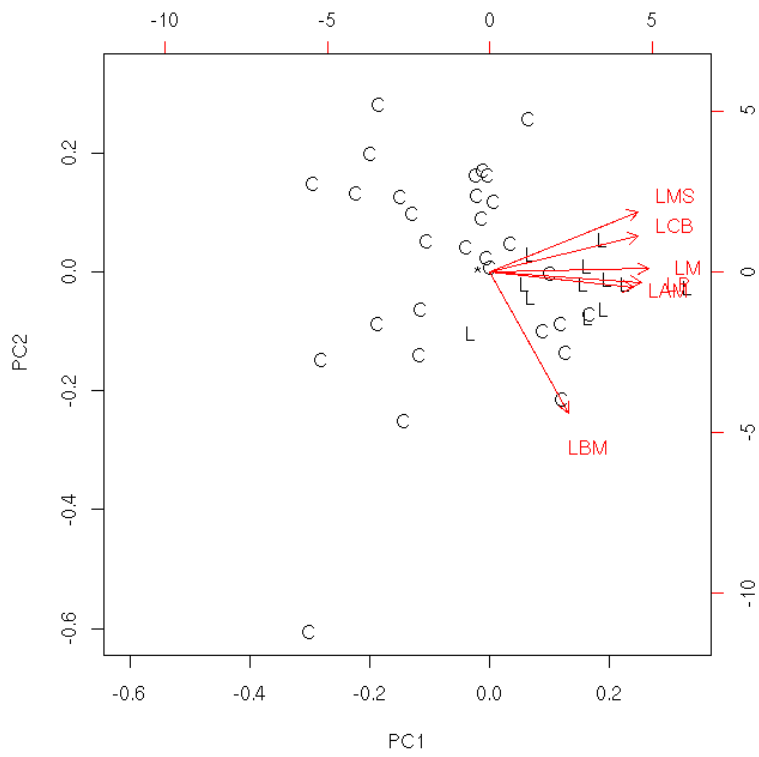
```
> acp<-prcomp(Jussac, scale=TRUE)
```

```
> summary(acp)
```

Importance of components:

	PC1	PC2	PC3	PC4	PC5	PC6
Standard deviation	2.025	0.939	0.799	0.5094	0.3119	0.14799
Proportion of Variance	0.683	0.147	0.106	0.0432	0.0162	0.00365
Cumulative Proportion	0.683	0.830	0.937	0.9801	0.9963	1.00000

```
> biplot(acp)
```

La commande suivante représente les composantes principales 1 et 3.

```
> biplot(ACP, choices=c(1,3))
```

