

POLYTECH'LILLE: Département GIS

## Statistique Exploratoire Multidimensionnelle: Analyse en Composantes Principales Normée

#### Exercice 1 : Les données de dépenses publicitaires de certains pays européens.

Le tableau ci-dessous donne la répartition (en pourcentages) des dépenses publicitaires par média dans seize pays d'Europe en 1990. On se propose d'analyse les grandes tendances des dépenses publicitaires par média et par pays en utilisant l'analyse en composantes principales normée.

Table 1: Distribution of avertising expenditure by media in 1990

(as % of total per country)

			(43 /0	or total per coul	11u y <i>)</i>		
		Newspapers	Magazines	TV	Radio	Cinema	Outdoor/Transport
BEL	Belgium	35.3	20.5	27.5	1.2	1.4	14.2
DK	Denmark	76.1	11.8	9.6	0.8	0.4	1.3
FR	France	28.6	27.5	24.8	6.6	0.8	11.7
AL	Germany	46.0	27.9	15.8	5.1	1.2	3.9
GR	Greece	26.0	28.2	35.3	5.7	0.0	4.8
IR	Ireland	63.7	5.4	19.6	6.2	0.0	5.1
IT	Italy	27.8	24.0	43.02	1.4	0.0	3.6
NE	Neth	50.4	27.8	9.0	2.2	0.3	10.4
POR	Portugal	25.1	21.7	37.1	6.7	0.0	9.5
ESP	Spain	37.6	15.4	31.3	10.3	0.8	4.7
UK	U.K.	44.1	19.3	30.5	2.2	0.5	3.4
AUT	Austria	51.6	18.0	17.7	8.0	0.4	4.4
FIN	Finland	67.2	13.5	12.9	4.0	0.1	2.3
NOR	Norway	77.2	15.8	2.5	1.0	1.2	2.2
SUE	Sweden	78.6	14.3	2.2	0.0	0.6	4.3
SUI	Switz	61.1	16.8	6.7	1.7	0.9	12.7

(La somme de chaque ligne vaut 1 aux arrondis près). Source: Press Landscape update & Eur. Direct Marketing NTC, 1991

- 1. Importer les données dans R ou dans Rstudio.
- 2. Effectuer la statistique descriptive des variables du tableau : moyenne, écart-type, médian, minimum et maximum.
- 3. Calculer la matrice de corrélations entre les variables (colonnes) du tableau n°1 et tester leur significativité en prenant une erreur de 1ère espèce de 5%.
- 4. Mise en œuvre de l'analyse en composantes principales : ACP.
  - a. Faire l'ACP à l'aide de R ou Rstudio, ou autre
  - b. En utilisant le tableau des valeurs propres combien de composantes principales peut-on retenir ?
  - c. Afficher les cordonnées, les corrélations, les cosinus carrés (qualité de la représentation) et les contributions de chaque variable (média publicitaire) par rapport aux cinq premières composantes principales.
  - d. Interpréter les composantes retenues (voir question 4.b) par rapport aux médias publicitaires.
  - e. Afficher les cordonnées, les cosinus carrés (qualité de la représentation) et les contributions de chaque pays par rapport aux cinq premières composantes principales.
  - f. Interpréter les composantes retenues (voir question 4.b) par rapport aux pays européens concernés.
  - g. Représentation de la répartition des dépenses publicitaires (cercle des corrélations):
    - i. par rapport aux composantes principales 1 et 2 et analyser le graphique
    - ii. par rapport aux composantes principales 1 et 3 et analyser le graphique
  - h. Représentation de la répartition des pays par rapport aux composantes principales 1 et 2 et analyser le graphique.
  - Conclusion : Résumez en quelques lignes les grandes tendances issues de l'ACP Normée. Il s'agit de relier par des phrases les analyses effectuées par rapport aux médias et aux pays.
- 5. Reprendre la question 4 en mettant les dépenses « newspapers » en variable supplémentaire : cette variable ne participera pas aux calculs des valeurs propres et des vecteurs propres.

Exercice 2 : Le niveau scolaire d'une classe de 6ème d'un collège.

Analyse des notes de 27 élèves sur 14 disciplines d'une classe de 6ème : Tableau 2 :

eleves	ORTH	GRAM	EXPR	RECI	MATH	ANGL	HIST	BIOL	EDMU	ARTS	TECH	EPS	GEO	EXPO
EL01	13	10	2	4	9	9	8	7	7,5	1,5	14	10	10,5	13
EL02	6,5	8	8,5	14	13	7	11	8,5	16	4	18	18	16	15
EL03	14	6,5	8	5	11	8	9,5	8	18,5	9,5	14	16,5	14	13
EL04	13	7,5	9	5	10	10,5	10	16	16	11,5	0	11,5	15	18
EL05	15	7,5	10	14	12	11	9	11	16,5	13,5	16	13	15,51	17
EL06	5	8	5,5	6,5	16	12	9	7	13,5	5	16	12,5	13	17,5
EL07	12	6,5	9	16	18	13,5	9	10	15	11	16	13,5	14	18
EL08	8,5	2,5	9	13	12	9,5	12	13,5	16,5	8	13	12	14	17
EL09	15,5	7,5	12,5	16	15	13	12	13,5	17	14	15	16	15	17
EL10	20	14,5	16,5	10	18	16,5	15	10,5	18	13,5	12	14,5	13	15
EL11	6	4	11	9,5	13	12	11	7,5	17,5	12	18	13,5	13	17
EL12	0	5,5	6	6	9	3	7	2	12	1,5	15	12,5	14	0
EL13	15,5	12	12	8	17	17	11	4,5	16	9	13	5	12,5	16,5
EL14	15	6,5	11,5	10	13	12,5	7	11	14	7	14	11	15	13
EL15	11	9	7,5	11	12,5	11	9	11,5	18	8,5	17	15	15	15
EL16	6,5	9	10	10	10	6,5	7	12,5	18	14,5	16	13,5	13	18
EL17	7	4,5	6	6	15	13,5	11	7	7	6,5	11	10,5	13,5	16
EL18	11,5	8,5	7,5	12	10	7,5	8	10	17,5	7	11	16,5	15,5	14
EL19	4,5	7,5	9	12	12	9	7	11,5	17	12,5	12	16	13	18
EL20	5,5	8	8	11	10	12	8	7	12	5	18	10,5	13	15
EL21	16	13,5	14,5	9,5	16	17	11,5	12,5	19	13,5	6	11	16	16
EL22	5,5	2	7,5	14,5	11	10	7	8	15,5	8	16	13,5	13	15
EL23	14	8	6,5	6,5	10	14,5	9	5	9,5	3,5	0	18,5	13	0
EL24	5,5	8	4,5	15	8	7,5	5	5	13,5	8	6	9	0	13
EL25	11	4	10,5	6	11	12	9	17	14	12,5	8	10,5	17,5	15
EL26	4,5	7	10,5	11,5	14	8,5	7,5	7	18	14	13	15,5	18	13
EL27	9	8,5	7,5	10	14	11,5	9	15	17,5	9	13	13,5	17,5	16

Cette étude concerne les notes obtenues par 27 élèves d'une classe de 6<sup>ème</sup> d'un collège relativement aux 14 matières suivantes : ORTH (orthographe), GRAM (grammaire), EXPR (expression écrite), RECI (récitation), MATH (mathématique), ANGL (anglais), HIST (histoire), BIOL (biologie), EDMU (éducation musicale), ARTS (art plastique), TECH (technologie), EPS (éducation physique et sportive), GEO (géographie), EXPO (exposé).

- 1) Réaliser l'ACP à l'aide de R.
- 2) Quels renseignements peut-on déduire du tableau des statistiques sommaires ?
- 3) Relevez les coefficients de corrélation les plus importants (les plus significatifs par rapport à 5% d'erreur) et donnez une interprétation rigoureuse de ces coefficients.
- 4) Combien de composantes principales peut-on retenir et pourquoi ?
- 5) Interprétez les trois premiers axes par rapport aux disciplines enseignées.
- 6) Quel phénomène peut-on déduire du cercle des corrélations formé par les deux premiers facteurs

Quelles sont les disciplines les mieux représentées dans ce plan ?

- 7) Que peut-on déduire de ce phénomène pour les 27 élèves de cette classe de 6ème ?
- 8) Interprétez les contributions des élèves EL10 et EL12 par rapports aux axes factoriels 1 et 2.
- 9) Quels sont les élèves parmi les 27 qui expliquent les axes 1, 2 et 3 ?
- 10) Faites un résumé des résultats de l'ACP. Peut-on dégager des tendances quant aux matières enseignées dans cette classe.

# **ACP NORMEE – R version > 2.12.0**

#### **INTRODUCTION TO R**



R is a language and environment for statistical computing and graphics. It is a GNU project which is similar to the S language and environment which was developed at Bell Laboratories (formerly AT&T, now Lucent Technologies) by John Chambers and colleagues. R can be considered as a different implementation of S. There are some important differences, but much code written for S runs unaltered under R.

R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, ...) and graphical techniques, and is highly extensible. The S language is often the vehicle of choice for research in statistical methodology, and R provides an Open Source route to participation in that activity.

One of R's strengths is the ease with which well-designed publication-quality plots can be produced, including mathematical symbols and formulae where needed. Great care has been taken over the defaults for the minor design choices in graphics, but the user retains full control.

R is available as Free Software under the terms of the Free Software Foundation's GNU General Public License in source code form. It compiles and runs on a wide variety of UNIX platforms and similar systems (including FreeBSD and Linux), Windows and MacOS.

• • •

Source: https://www.r-project.org/about.html

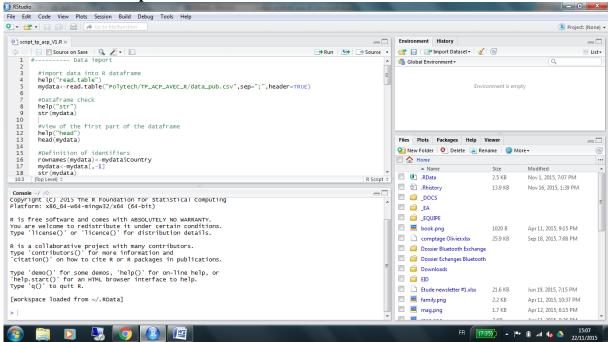
#### **INTRODUCTION TO RSTUDIO**



RStudio is an integrated development environment (IDE) for R. It includes a console, syntax-highlighting editor that supports direct code execution, as well as tools for plotting, history, debugging and workspace management.

Source: https://www.rstudio.com/products/rstudio/

> Lancer Rstudio à partir de la console : /usr/local/rstudio/bin/rstudio



- 1 : Editeur de script (ouvrir un script : File > New File > R script)
- 2 : Console : Exécution des lignes du scripts, affichage des résultats et des erreurs de compilation.
- 3 : Environnement de travail : Gestion des dataframes, variables et autres éléments de l'environnement de travail.
- 4 : Explorateur : Arborescence fichier, affichages des graphiques, gestion des packages (obtenir le chemin de l'arborescence via la commande 'getwd')
- > getwd()
- [1] "D:/Users/Gregory/Google Drive/MyFolder/Mes Documents"

#### **LES DONNEES DU CAS « PUB »**

Analyse des dépenses publicitaires par média dans les 16 principaux pays d'Europe en 1990.

#### Importer les données

```
> mydata<-read.table("Polytech/TP_ACP_AVEC_R/data_pub.csv",sep=";",
dec=".",header=TRUE)</pre>
```

Les données sont stockées dans le dataframe « mydata » grace à la fonction « read.table ».

```
> str(mydata)
'data.frame':
               16 obs. of 7 variables:
 $ Country
            : Factor w/ 16 levels "Austria", "Belgium", ...: 2 3 5 6 7 8 9 10
12 13 ...
                    35.3 76.1 28.6 46 26 63.7 27.8 50.4 25.1 37.6 ...
 $ Newspapers: num
                    20.5 11.8 27.5 27.9 28.2 5.4 24 27.8 21.7 15.4 ...
 $ Magazines : num
                    27.5 9.6 24.8 15.8 35.3 .
 $ TV
             : num
                    1.2 0.8 6.6 5.1 5.7 6.2 1.4 2.2 6.7 10.3 ...
 $ Radio
             : num
                    1.4 0.4 0.8 1.2 0 0 0 0.3 0 0.8
 $ Cinema
             : num
                    14.2 1.3 11.7 3.9 4.8 5.1 3.6 10.4 9.5 4.7 ...
 $ Outdoor
             : num
```

La fonction « str » permet d'obtenir le type d'un objet. « mydata » est bien un dataframe de 16 lignes et 7 variables.

```
> head(mydata)
                              rines TV Radio Cinema Outdoor
20.5 27.5 1.2 1.4 14.2
  Country Newspapers Magazines
1 Belgium
                   35.3
                   76.1
                              11.8
                                     9.6
                                            0.8
                                                              1.3
                                                     0.4
2 Denmark
                  28.6
                              27.5 24.8
                                            6.6
                                                     0.8
                                                             11.7
   France
                              27.9 15.8
                                                              3.9
4 Germany
                  46.0
                                            5.1
                                                     1.2
                                            5.7
                              28.2 35.3
   Greece
                  26.0
                                                     0.0
                                                              4.8
                               5.4 19.6
6 Ireland
                                             6.2
                                                     0.0
                                                              5.1
                  63.7
```

La fonction « head » permet d'afficher les premières lignes d'un dataframe. Taper simplement le nom du dataframe pour en afficher l'exhaustivité.

NB : Si la première colonne avait un nom dans le fichier importé, alors R ajoute un identifiant de ligne (1,2,3,...). Il est important de reformater le dataframe pour que le pays soit traité en tant qu'identifiant des lignes. Le script ci-dessous permet de faire la transformation.

```
> rownames(mydata)<-mydata$Country</pre>
> mydata<-mydata[,-1]</pre>
> head(mydata)
                                  TV Radio Cinema Outdoor
         Newspapers Magazines
               35.3
                          20.5 27.5
Belgium
                                        1.2
                                               1.4
                                                       14.2
Denmark
               76.1
                          11.8 9.6
                                        0.8
                                               0.4
                                                        1.3
                          27.5 24.8
France
               28.6
                                        6.6
                                               0.8
                                                       11.7
               46.0
                          27.9 15.8
                                        5.1
                                               1.2
                                                        3.9
Germany
                          28.2 35.3
Greece
               26.0
                                        5.7
                                               0.0
                                                        4.8
                            5.4 19.6
Ireland
               63.7
                                        6.2
                                               0.0
                                                        5.1
```

#### 1. Statistiques descriptives

#### > summary(mydata)

```
        Newspapers
        Magazines
        TV
        Radio
        Cinema
        Outdoor

        Min. : 25.10
        Min. : 5.40
        Min. : 2.20
        Min. : 0.000
        Min. : 0.0000
        Min. : 1.300

        1st Qu.: 33.62
        1st Qu.: 15.12
        1st Qu.: 9.45
        1st Qu.: 1.350
        1st Qu.: 0.0750
        1st Qu.: 3.550

        Median : 48.20
        Median : 18.65
        Median : 18.65
        Median : 3.100
        Median : 0.4500
        Median : 4.550

        Mean : 49.77
        Mean : 19.24
        Mean : 20.34
        Mean : 3.944
        Mean : 0.5375
        Mean : 6.156

        3rd Qu.: 64.58
        3rd Qu.: 24.88
        3rd Qu.: 30.70
        3rd Qu.: 6.300
        3rd Qu.: 0.8250
        3rd Qu.: 9.725

        Max. : 78.60
        Max. : 28.20
        Max. : 43.02
        Max. : 10.300
        Max. : 1.4000
        Max. : 14.200
```

De base, R permet d'accéder aux statistiques de bases d'un dataframe via la fonction « summary ». Il est possible d'aller plus loin en installant d'autres packages. « Psych », « Hmisc », « pastecs » sont des packages qui contiennent des fonctions de statistiques descriptives plus avancées.

```
L'ensemble des packages R sont disponibles via <a href="https://cran.r-project.org/">https://cran.r-project.org/</a>
<a href="https://cran.rstudio.com/bin/windows/contrib/3.2/psych_1.5.8.zip">https://cran.rstudio.com/bin/windows/contrib/3.2/psych_1.5.8.zip</a>
<a href="https://cran.rstudio.com/bin/windows/contrib/3.2/psych_1.5.8.zip</a>
<a href="https://cran.rstudio.com/bin/windows/contrib/3.2/psych_1.5.8.zip</a>
<a href="https://cran.rstudio.com/bin/windows/contrib/3.2/psych_1.5.8.zip</a>
<a href="https://cran.rstudio.com/bin/windows/contrib/3.2/psych_1.5.8.zip</a>
<a href="https://cran.rstudio.com
```

> library(psych)
> psych::describe(mydata)

```
sd median trimmed
0.00 48.20 49.48
                                                                 mmed mad min max range skew
49.48 25.57 25.1 78.60 53.50 0.17
19.59 7.04 5.4 28.20 22.80 -0.17
20.02 15.94 2.2 43.02 40.82 0.15
                 vars n mean sd r
1 16 49.77 19.00
                                                                                                                skew kurtosis se 0.17 -1.53 4.75
Newspapers
                      2 16 19.24 6.66
3 16 20.34 12.84
                                                                                                                              -0.95 1.67
-1.39 3.21
                                                    18.65
18.65
Magazines
Radio
                      4 16 3.94
                                        3.06
                                                     3.10
                                                                  3.77
                                                                          3.26
                                                                                     0.0 10.30 10.30
                                                                                                                 0.45
                                                                                                                               -1.14 0.76
                                                     0.45
4.55
                                                                            0.59
                                                                                     0.0 1.40 1.40
1.3 14.20 12.90
                                                                  0.51
                      5 16
                               0.54
                                          0.47
Cinema
                                                                                                                 0.35
                                                                                                                               -1.310.12
                                                                                                                              -1.10 \ 1.03
                      6 16
                                         4.10
                                                                  5.93
Outdoor
                               6.16
```

Il n'est pas obligatoire de préciser le nom de la librairie où se trouve la fonction, mais c'est préférable pour éviter la confusion si plusieurs fonction porte le même nom.

#### 2. Matrices des corrélations

> install.packages("Hmisc")
trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.2/Hmisc\_3.17-

The downloaded binary packages are in D:\Users\Gregory\AppData\Local\Temp\RtmpCq0FwY\downloaded\_packages > library(Hmisc)

Le chargement a nécessité le package : grid

> rcorr(as.matrix(mydata))

> rcorr(as.matrix(mydata))										
	Newspapers	Magazine	es TV	Radio	Cinema	Outdoor				
Newspapers	1.00	-0.69	-0.89	-0.47	0.14	-0.41				
Magazines	-0.69	1.00	0.37	0.07	0.10	0.35				
TV	-0.89	0.37	1.00	0.42	-0.35	0.10				
Radio	-0.47	0.07	0.42	1.00	-0.19	0.01				
Cinema	0.14	0.10	-0.35	-0.19	1.00	0.30				
Outdoor	-0.41	0.35	0.10	0.01	0.30	1.00				

n=16

#### P\_values

	Newspapers	Magazines	TV	Radio	Cinema	Outdoor
Newspapers	5	0.0032	0.0000	0.0688	0.5963	0.1107
Magazines	0.0032		0.1614	0.7969	0.7002	0.1823
TV	0.0000	0.1614		0.1027	0.1793	0.7108
Radio	0.0688	0.7969	0.1027		0.4727	0.9788
Cinema	0.5963	0.7002	0.1793	0.4727		0.2584
Outdoor	0.1107	0.1823	0.7108	0.9788	0.2584	

La fonction « rcorr » du package « Hmisc » permet d'obtenir la matrice des corrélations ainsi que les p-values.

#### 3. **ACP**

Une méthode d'ACP est incluse dans le package "FactoMineR".

FactoMineR is an R package dedicated to multivariate Exploratory Data Analysis. It is developped and maintained by F. Husson\*, J. Josse\*, S. Lê\*, from Agrocampus Rennes, and J. Mazet.

L'acp normée (scale.unit=TRUE) est effectuée sur le jeu de donnée « mydata ». Les 6 premiers axes sont retenus (ncp=6), les graphiques ne sont pas générés à cette étape (graph=FALSE). Les résultats sont stockés dans "mydata.pca" : valeurs propres, coordonnées, contributions, etc.

```
> str(mydata.pca)
```

List of 5

```
$ eig :'data.frame': 6 obs. of 3 variables:
 ..$ eigenvalue
                                 : num [1:6] 2.699 1.543 0.778 0.58 0.401 ...
 ..$ percentage of variance
                                   : num [1:6] 44.98 25.72 12.96 9.66 6.68 ...
 ..$ cumulative percentage of variance: num [1:6] 45 70.7 83.7 93.3 100 ...
$ var :List of 4
 ..$ coord : num [1:6, 1:6] -0.992  0.681  0.867  0.542 -0.213 ...
 ....- attr(*, "dimnames")=List of 2
 .....$ : chr [1:6] "Newspapers" "Magazines" "TV" "Radio" ...
 .....$ : chr [1:6] "Dim.1" "Dim.2" "Dim.3" "Dim.4" ...
 ..$ cor : num [1:6, 1:6] -0.992 0.681 0.867 0.542 -0.213 ...
 ....- attr(*, "dimnames")=List of 2
 .....$ : chr [1:6] "Newspapers" "Magazines" "TV" "Radio" ...
 .....$ : chr [1:6] "Dim.1" "Dim.2" "Dim.3" "Dim.4" ...
 ..$ cos2 : num [1:6, 1:6] 0.9842 0.4644 0.7522 0.2938 0.0456 ...
 ....- attr(*, "dimnames")=List of 2
 .....$ : chr [1:6] "Newspapers" "Magazines" "TV" "Radio" ...
 .....$ : chr [1:6] "Dim.1" "Dim.2" "Dim.3" "Dim.4" ...
 ..$ contrib: num [1:6, 1:6] 36.47 17.2 27.87 10.88 1.69 ...
 ....- attr(*, "dimnames")=List of 2
 .....$ : chr [1:6] "Newspapers" "Magazines" "TV" "Radio" ...
 .....$ : chr [1:6] "Dim.1" "Dim.2" "Dim.3" "Dim.4" ...
$ ind :List of 4
 ..$ coord : num [1:16, 1:6] 0.8 -2.41 1.98 0.29 2.26 ...
 ....- attr(*, "dimnames")=List of 2
 .....$ : chr [1:16] "Belgium" "Denmark" "France" "Germany" ...
 .....$ : chr [1:6] "Dim.1" "Dim.2" "Dim.3" "Dim.4" ...
 ..$ cos2 : num [1:16, 1:6] 0.0673 0.847 0.6317 0.0185 0.7385 ...
 ....- attr(*, "dimnames")=List of 2
 .....$ : chr [1:16] "Belgium" "Denmark" "France" "Germany" ...
 .....$ : chr [1:6] "Dim.1" "Dim.2" "Dim.3" "Dim.4" ...
 ..$ contrib: num [1:16, 1:6] 1.482 13.415 9.035 0.195 11.797 ...
 ....- attr(*, "dimnames")=List of 2
 .. ... $: chr [1:16] "Belgium" "Denmark" "France" "Germany" ...
 .....$ : chr [1:6] "Dim.1" "Dim.2" "Dim.3" "Dim.4" ...
 ..$ dist : Named num [1:16] 3.08 2.62 2.49 2.13 2.63 ...
 ....- attr(*, "names")= chr [1:16] "Belgium" "Denmark" "France" "Germany" ...
$ svd :List of 3
 ..$ vs: num [1:6] 1.643 1.242 0.882 0.761 0.633 ...
 ..$ U : num [1:16, 1:6] 0.487 - 1.465 1.202 0.177 1.374 ...
```

```
..$ V : num [1:6, 1:6] -0.604 0.415 0.528 0.33 -0.13 ...
$ call:List of 9
              : num [1:16] 0.0625 0.0625 0.0625 0.0625 0.0625 0.0625 0.0625 0.0625 0.0625 0.0625 ...
 ..$ row.w
 ..$ col.w
              : num [1:6] 1 1 1 1 1 1
 ..$ scale.unit: logi TRUE
             : num 6
 ..$ centre
            : num [1:6] 49.775 19.244 20.345 3.944 0.537 ...
 ..$ ecart.type: num [1:6] 18.397 6.449 12.436 2.961 0.457 ...
        :'data.frame': 16 obs. of 6 variables:
 ....$ Newspapers: num [1:16] 35.3 76.1 28.6 46 26 63.7 27.8 50.4 25.1 37.6 ...
 .. ..$ Magazines : num [1:16] 20.5 11.8 27.5 27.9 28.2 5.4 24 27.8 21.7 15.4 ...
                : num [1:16] 27.5 9.6 24.8 15.8 35.3 ...
 .. ..$ Radio
                 : num [1:16] 1.2 0.8 6.6 5.1 5.7 6.2 1.4 2.2 6.7 10.3 ...
                 : num [1:16] 1.4 0.4 0.8 1.2 0 0 0 0.3 0 0.8 ...
 .. ..$ Outdoor : num [1:16] 14.2 1.3 11.7 3.9 4.8 5.1 3.6 10.4 9.5 4.7 ...
 ..$ row.w.init: num [1:16] 1 1 1 1 1 1 1 1 1 ...
..$ call : language PCA(X = mydata, scale.unit = TRUE, ncp = 6, graph = FALSE)
- attr(*, "class")= chr [1:2] "PCA" "list "
```

L'accès aux différents objets de mydata.pca se font par le caractère '\$' : mydata.pca\$eig pour les valeurs propres, mydata.pca\$var pour les résultats sur les variables, mydata.pca\$ind pour les résultats sur les individus

#### > mydata.pca\$eig

```
eigenvalue percentage of variance cumulative percentage of variance
comp 1 2.699047e+00
                             4.498411e+01
comp 2 1.543048e+00
                              2.571747e+01
                                                                    70.70158
comp 3 7.775508e-01
                              1.295918e+01
                                                                    83.66076
comp 4 5.795016e-01
                              9.658359e+00
                                                                    93.31912
comp 5 4.008445e-01
                              6.680742e+00
                                                                    99.99986
comp 6 8.196494e-06
                              1.366082e-04
                                                                    100,00000
```

# > mydata.pca\$var

```
$coord
                       Dim.2
                                 Dim.3
                                           Dim.4
                                                      Dim.5
             Dim.1
                                                                 Dim. 6
Newspapers -0.9920905 -0.06106564 0.03906945 -0.01788785 0.10087660 2.226887e-03
Magazines 0.6814347 0.43314328 -0.34658589 0.34464343 0.33035155 7.822721e-04
          0.8672855 -0.29429274 -0.10679773 -0.01519317 -0.38674131 1.506755e-03
Radio
          Cinema
         Outdoor
          0.3986410 0.70197410 0.16202942 -0.56226037 0.07698831 4.764187e-04
$cor
             Dim.1
                       Dim.2
                                 Dim.3
                                           Dim.4
                                                      Dim.5
Newspapers -0.9920905 -0.06106564 0.03906945 -0.01788785 0.10087660 2.226887e-03
Magazines 0.6814347 0.43314328 -0.34658589 0.34464343 0.33035155 7.822721e-04
          0.8672855 -0.29429274 -0.10679773 -0.01519317 -0.38674131 1.506755e-03
Radio
          0.5420114 -0.39190197 0.70001163 0.11434215 0.22258926 3.535579e-04
Cinema
         Outdoor
          0.3986410 0.70197410 0.16202942 -0.56226037 0.07698831 4.764187e-04
$cos2
             Dim.1
                       Dim.2
                                 Dim.3
                                            Dim.4
                                                      Dim.5
Newspapers 0.98424354 0.003729012 0.001526422 0.0003199752 0.01017609 4.959027e-06
Magazines 0.46435327 0.187613102 0.120121781 0.1187790923 0.10913215 6.119496e-07
         0.75218408 0.086608217 0.011405755 0.0002308325 0.14956884 2.270309e-06
Radio
         0.29377633 0.153587153 0.490016282 0.0130741275 0.04954598 1.250032e-07
Cinema
         0.04557491 0.618742985 0.128227043 0.1309607972 0.07649426 3.229819e-09
Outdoor
         0.15891467 0.492767642 0.026253535 0.3161367271 0.00592720 2.269748e-07
```

\$contrib

```
        Newspapers
        36.466338
        0.2416653
        0.1963115
        0.05521559
        2.538662
        60.50180679

        Magazines
        17.204343
        12.1586036
        15.4487370
        20.49676863
        27.225555
        7.46599267

        TV
        27.868508
        5.6128008
        1.4668823
        0.03983294
        37.313431
        27.69854485

        Radio
        10.884448
        9.9534909
        63.0204832
        2.25609879
        12.360398
        1.52508143

        Cinema
        1.688556
        40.0987488
        16.4911463
        22.59886912
        19.083275
        0.03940488

        Outdoor
        5.887807
        31.9346907
        3.3764397
        54.55321493
        1.478678
        2.76916938
```

#### > mydata.pca\$ind

\$coord

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5	Dim.6
Belgium	0.7999605	2.6030438	0.22221512	-0.64215853	-1.2796850	4.089553e-03
Denmark	-2.4069194	-0.8147392	-0.56817432	0.06154194	-0.2369873	2.315095e-03
France	1.9752304	1.2878230	0.60454838	-0.02346439	0.4998936	-2.439455e-03
Germany	0.2902451	1.0390177	0.30229846	1.78764624	0.3254998	-2.698690e-03
Greece	2.2570889	-0.8624721	-0.81852317	0.41699202	0.4652988	2.415656e-04
Ireland	-1.0392994	-1.9075174	0.96276331	-1.23703819	-0.2132120	-2.841754e-03
Italy	1.7030910	-0.9539679	-1.84192235	0.11245204	-0.7855195	-4.537277e-03
Neth	0.1807401	1.1374650	-0.89190180	-0.50679197	1.4049972	3.577543e-03
Portugal	2.3435560	-0.6836342	0.04353914	-0.86439797	0.1058213	1.599333e-03
Spain	1.1621252	-0.9046069	1.96866508	0.59464953	-0.4954016	1.530921e-03
U.K.	0.2689405	-0.4336052	-0.74443542	0.37997931	-0.7990084	2.358139e-03
Austria	0.1316536	-0.8945222	0.99010137	0.30388160	0.6045277	2.161258e-03
Finland	-1.3623025	-1.3765700	-0.08778381	-0.14891750	0.3589521	4.958207e-05
Norway	-2.6374899	0.7497973	0.06641563	1.02781000	-0.2689420	-1.350239e-03
Sweden	-2.6050841	0.2441422	-0.54034120	-0.14459109	0.1559853	1.788193e-03
Switz	-1.0615360	1.7703461	0.33253557	-1.11759303	0.1577801	-5.843767e-03

\$cos2

Dim.1 Dim.2 Dim.3 Dim.4 Dim. 5 Belgium 0.067254640 0.712111779 0.0051895826 4.333808e-02 0.172104157 1.757668e-06 Denmark 0.846987954 0.097049060 0.0471973190 5.537279e-04 0.008211156 7.835954e-07 0.631730139 0.268539661 0.0591777189 8.914861e-05 0.040462369 9.635667e-07 Germany 0.018487056 0.236910285 0.0200544112 7.012958e-01 0.023250885 1.598248e-06 0.738468050 0.107826246 0.0971172418 2.520518e-02 0.031383269 8.458718e-09 Greece Ireland 0.149575116 0.503865962 0.1283562598 2.119065e-01 0.006295080 1.118281e-06 0.370296738 0.116182763 0.4331285357 1.614391e-03 0.078774943 2.628239e-06 0.007504737 0.297236638 0.1827512347 5.900457e-02 0.453499876 2.940329e-06 Neth Portugal 0.817313112 0.069548106 0.0002820964 1.111899e-01 0.001666418 3.806410e-07 0.203285949 0.123174631 0.5833714383 5.322598e-02 0.036941652 3.527827e-07 Spain 0.045281169 0.117704773 0.3469438217 9.039096e-02 0.399675791 3.481319e-06 Austria 0.007684263 0.354747279 0.4346064069 4.093979e-02 0.162020192 2.070857e-06 Finland 0.474702156 0.484697431 0.0019710743 5.672384e-03 0.032956954 6.288158e-10 0.804046036 0.064981090 0.0005098471 1.221026e-01 0.008360198 2.107272e-07 Norway Sweden 0.944758481 0.008297799 0.0406455926 2.910451e-03 0.003387231 4.451496e-07 Switz 0.199602730 0.555153958 0.0195872417 2.212404e-01 0.004409622 6.048991e-06

#### \$contrib

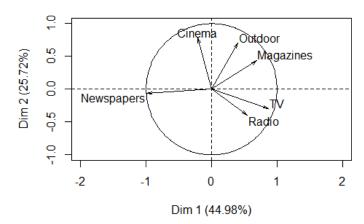
Dim.1 Dim.2 Dim.3 Dim.4 Dim.5 Dim.6 1.48185820 27.4450179 0.39691587 4.447438180 25.5334928 12.752743044 Belaium Denmark 13.41506264 2.6886715 2.59486301 0.040847710 0.8756978 4.086859246 9.03452044 6.7175807 2.93774004 0.005938051 3.8963619 4.537719167 France 0.19507396 4.3726670 0.73455292 34.465816534 1.6519827 5.553387649 Germany Greece 11.79687359 3.0129412 5.38534076 1.875343859 3.3757312 0.044496111 Ireland 2.50121463 14.7379671 7.45058375 16.504091792 0.7088060 6.157788596 6.71653568 3.6861083 27.27054839 0.136382947 9.6209511 15.697937409 Italv 0.07564476 5.2405474 6.39418675 2.770032449 30.7790353 9.759366396 Neth 3.12734269 3.3145176 31.15264376 3.813709190 3.8266513 1.787136471 Spain  $0.16748740 \quad 0.7615344 \quad 4.45456487 \quad 1.557203325 \quad 9.9542102 \quad 4.240242384$ U.K. Austria 0.04013608 3.2410283 7.87971593 0.995942378 5.6981833 3.561763914 Finland 4.29750789 7.6753316 0.06194127 0.239175608 2.0089865 0.001874568

```
16.10835586 2.2771322 0.03545617 11.393340811 1.1277721 1.390186947
Sweden
        15.71495385 0.2414272 2.34686120 0.225480228
                                                        0.3793775 2.438263341
Switz
         2.60939013 12.6945389 0.88884791 13.470781342
                                                        0.3881579 26.039807760
$dist
Belgium Denmark
                   France Germany
                                     Greece Ireland
                                                       Italy
                                                                 Neth Portugal
3.084661 2.615309 2.485146 2.134671 2.626533 2.687268 2.798742 2.086349 2.592275 2.577503 1.263856
Austria Finland Norway
                            Sweden
                                      Switz
1.501867 1.977256 2.941375 2.680164 2.376028
```

L'accès aux graphiques de l'ACP se fait via la fonction « plot.PCA », elle aussi incluse dans le package « FactoMineR ».

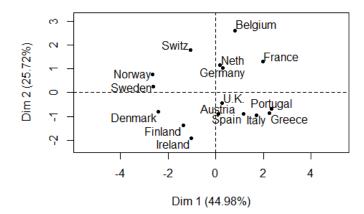
> plot.PCA(mydata.pca, axes=c(1,2), choix="var")

### Variables factor map (PCA)



> plot.PCA(mydata.pca, axes=c(1,2), choix="ind")

#### Individuals factor map (PCA)

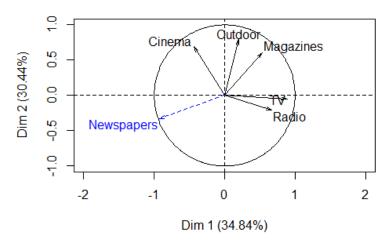


#### 4. Pour aller plus loin ...

Effectuer l'ACP avec les dépenses newspapers en variables supplémentaire. Indiquer le numéro de la variable dans le paramètre « quanti.sup ».

- > mydata.pca = PCA(mydata, scale.unit=TRUE, ncp=6,graph=FALSE,
  quanti.sup=1)
- > plot.PCA(mydata.pca, axes=c(1,2), choix="var")

# Variables factor map (PCA)



Effectuer l'ACP avec les dépenses de la Belgique et du Danemark en individus supplémentaires. Indiquer le numéro des individus dans le paramètre « ind.sup ». > mydata.pca = PCA(mydata, scale.unit=TRUE, ncp=6,graph=FALSE, ind.sup=c(1,2)) > plot.PCA(mydata.pca, axes=c(1,2), choix="ind")

# Individuals factor map (PCA)

