

# Analyse de données par la méthode ACP

## INSEA Rabat Maroc

OURAHOU MOHAMED  
LACHGAR MOHAMED

Master: Systèmes d'Informations et Systèmes intelligents  
Institut Nationale de Statistique et Économie Appliqué

February 27, 2022

- ① Introduction
- ② ACP
- ③ Mise en pratique de l'ACP
- ④ Interprétations et Conclusions

# Introduction

Lors de toute étude statistique, il est nécessaire de décrire et explorer les données avant d'en tirer de quelconques lois ou modèles prédictifs. Mais dans beaucoup de situations, les données sont trop nombreuses pour pouvoir être visualisables (nombre de caractéristiques trop élevées).

Il est alors nécessaire d'extraire l'information pertinente qu'elles contiennent ; Les techniques d'Analyse de Données répondent à ce besoin.

# Définition de l'ACP

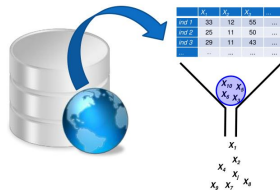
L'ACP a été inventée en 1901 par Karl Pearson. Actuellement, l'ACP est utilisée comme outil d'exploration et d'analyse de données ainsi que pour la conception de modèles.

L'ACP est basée sur une

transformation de l'espace de représentation des données, le nouvel espace étant de dimension inférieure à celle de l'espace initial.

Selon les domaines d'application, l'ACP est aussi connue sous le nom de transformation de Karhunen-Loève, de transformation de Hotelling ou encore de décomposition orthogonale

L'ACP met en oeuvre une décomposition en valeurs et vecteurs propres de la matrice de covariance des données ou encore une décomposition en valeurs singulières de la matrice des données.



# Les étapes de l'ACP

- ① Centrer et réduire la matrice des données;
- ② Déterminer les valeurs propres de la matrice de corrélation;
- ③ Classer les valeurs propres selon l'ordre décroissant;
- ④ Calculer les composantes principales;
- ⑤ Calculer les qualités et les contributions des individus dans les composantes principales;
- ⑥ Interpréter les résultats obtenus.

# Forme générale de données

Les données collectées sur un système ou après une statistique faite sur  $n$  individus et  $p$  variables peut être représentées sous forme d'une matrice de  $n$  lignes (individus) et  $p$  colonnes (variables).

$$X = \begin{matrix} & \begin{matrix} X_1 & X_2 & \dots & X_p \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{matrix} & \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix} \end{matrix}$$

où  $x_{ij}$  représente la valeur de  $X_j$  prise par l'individu  $x_i$ .

# Vecteurs variable et Vecteurs individu

- La variable  $X_j$ : Une colonne du tableau

$$X_j = \begin{pmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{pmatrix} \in \mathbb{R}^n$$

- L'individu  $x_i$ : Une ligne du tableau
- Les  $n$  individus sont décrits par un nuage de  $p$  variables.  
L'information représentée par un nuage correspond à la dispersion des  $n$  points.

# Centre d'inertie et tableau centre

- **Centre d'inertie:** c'est le vecteur  $g$  des moyennes arithmétiques de chaque variable:

$$g = (\bar{x}_1 \bar{x}_2 \dots \bar{x}_p)^T \in \mathbb{R}^p$$

où :

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}$$

- **Tableau centré:** il est obtenu en centrant les variables autour de leur moyenne:

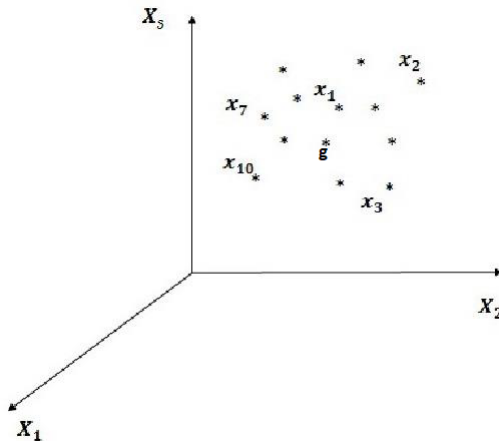
$$\bar{x}_{ij} = x_{ij} - \bar{x}_j$$

- **Inertie total du nuage des individus:** permet de mesurer la dispersion du nuage des individus par rapport à son centre de gravité.

$$I_g = \frac{1}{n} \sum_{i=1}^n d^2(x_i, g) = \sum_{j=1}^p \text{Var}(X_j) = \text{Tr}(\Sigma)$$



# Centre d'inertie et tableau centre



# Matrice de variance covariance

$$\Sigma = \begin{pmatrix} V(X_1) & cov(X_1, X_2) & \dots & cov(X_1, X_p) \\ cov(X_2, X_1) & V(X_2) & \dots & cov(X_2, X_p) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_p, X_1) & cov(X_n, X_2) & \dots & V(X_p) \end{pmatrix}$$

avec:

$$cov(X_i, X_j) = \frac{1}{n} \sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)$$

# Matrice de corrélation

C'est la matrice dont les éléments sont des coefficients de corrélations:

$$\sum_{corr} = \begin{pmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{n2} & \dots & 1 \end{pmatrix}$$

avec

$$r_{ij} = \frac{\text{cov}(X_i, X_j)}{\sigma_{X_i} \sigma_{X_j}}$$

# Les valeurs propres et les vecteur propres

- $\lambda$  est une valeur propre de la matrice  $A$  si et seulement si  $Av = \lambda v$ ;
- Le vecteur  $v$  dans la relation ci-dessus est appelé vecteur associé à  $\lambda$ ;
- Les valeurs propres s'obtiennent en résolvant le système d'équations  $\det(A - \lambda I) = 0$ ;
- Le nombre de valeurs propres,  $\lambda_1 > \dots > \lambda_p$  est égal au nombre de lignes = nombre de colonnes de la matrice  $A$ ;
- **La somme** des valeurs propres de  $A$  est égale à la variance contenue dans l'ensemble des données.

# Calcul des composantes principales

Les composantes principales s'écrivent:

$$Y_j(indiv_k) = \lambda_1 X_1(indiv_k) + \dots + \lambda_p X_p(indiv_k)$$

avec les  $x_k(indiv_k)$  est les valeurs centrées réduites de la matrice initiale des données X

$$\sum_{CP} = \begin{pmatrix} Y_1(indiv_1) & Y_2(indiv_1) & \dots & Y_p(indiv_1) \\ Y_1(indiv_2) & Y_2(indiv_2) & \dots & Y_p(indiv_2) \\ \vdots & \vdots & \ddots & \vdots \\ Y_1(indiv_n) & Y_2(indiv_n) & \dots & Y_p(indiv_n) \end{pmatrix}$$

# Calcul des composantes principales

La variance d'une composante principale est égale à l'inertie du nuage des points projetés sur l'axe correspondant. Par exemple, la variance de la première composante principale est égale à la première valeur propre ( $Var(Y_1) = \lambda_1$ ). Sachant que  $\lambda_1$  est par construction la plus grande des valeurs propres, on voit donc que c'est bien la première composante principale qui rend le mieux compte de la dispersion des individus.

# Choix des composantes principales

Pour le bon choix des composantes principales il faut tenir compte les deux règles:

- Une première règle proposée en 1960 par Kaiser stipule qu'on devrait retenir comme CP seulement celles pour lesquelles la valeur propre correspondante de  $\sum_{corr}$  est supérieur à l'unité  $\lambda_i > 1$ ;
- Une deuxième règle recommande d'extraire des facteurs de façon à expliquer au moins 80% de la variance totale.

# Choix des composantes principales

Nous désirons toutefois être certains de bien choisir le bon nombre de facteurs à extraire. Nous regardons donc le graphique des valeurs propres et examinons où se situe la rupture du coude de Cattell. Nous voyons dans la figure par exemple un changement après le sixième facteur. Nous ne retenons donc que six facteurs pour l'analyse, puisque ce critère est plus rigoureux que celui des valeurs propres.





# Qualité de la représentation des individus par les composantes principales

La qualité de la représentation des individus exprime la part de la variance des variables initiales qui est restituée par les composantes principales retenues.

Elle représente la répartition des variables sur les différents facteurs. La somme horizontale des éléments de la matrice sera égale à 100%.

$$QL_j(indiv_k) = \frac{Y_j(indiv_k)^2}{\sum_{i=1}^p Y_i(indiv_k)^2}$$

$$\sum_{QL} = \begin{pmatrix} QL_1(indiv_1) & QL_2(indiv_1) & \dots & QL_p(indiv_1) \\ QL_1(indiv_2) & QL_2(indiv_2) & \dots & QL_p(indiv_2) \\ \vdots & \vdots & \ddots & \vdots \\ QL_1(indiv_n) & QL_2(indiv_n) & \dots & QL_p(indiv_n) \end{pmatrix}$$

# Contribution des individus à la formation des composantes principales

La contribution relative d'un individu  $i$  à la formation de la composante principale  $k$  est l'inertie relative de cet individu sur l'axe principale  $j$ . Elle est définie par :

$$CTR_j(indiv_k) = \frac{Y_j(indiv_k)^2}{\sum_{k=1}^n Y_j(indiv_k)^2}$$

$$\sum_{CTR} = \begin{pmatrix} CTR_1(indiv_1) & CTR_2(indiv_1) & \dots & CTR_p(indiv_1) \\ CTR_1(indiv_2) & CTR_2(indiv_2) & \dots & CTR_p(indiv_2) \\ \vdots & \vdots & \ddots & \vdots \\ CTR_1(indiv_n) & CTR_2(indiv_n) & \dots & CTR_p(indiv_n) \end{pmatrix}$$

La somme verticale est de 100%. Les individus ont contribué à 100% à la construction des axes.

# Mise en pratique de l'ACP

Les données sur lesquelles on a appliqué l'ACP dans ce projet ont comme source la nouvelle plateforme *data.gov.ma*

Elles ont comme titre : **“Indicateurs sociaux : Emploi, activité et chômage”** et elles sont produites par le ministère de l'économie et du finance marocain. Ces données sont structurées comme un tableau excel ( extension.xlsx ) et elles contiennent plusieurs indicateurs, statistiques et taux concernant l'emploi, l'activité et le chômage au sein de la population active ayant 15 ans et elles sont associées aux années qu'on appelle Les Individus : Moyenne 2002-08, 2008, 2009, 2010, 2011, 2012 et 2013. Les indicateurs, statistiques et taux sur lesquelles on va se concentrer sont regroupés dans le tableau suivant avec une description générale puis les abréviations qu'on a choisies - pour pouvoir bien lire les graphiques par la suite- des sous groupes qui vont être nos Variables (au total 22 variables) de la matrice initiale de l'ACP:

# Mise en pratique de l'ACP

Indicateurs, Statistiques et Taux	Signification	Abréviation dans ce qui suit
Population active 15 ans et plus	"La population active se définit comme l'ensemble des personnes en âge de travailler qui sont disponibles sur le marché du travail, qu'elles aient un emploi (population active occupée) ou qu'elles soient au chômage" - wikipedia	<b>PA15+:Total</b> <b>PA15+Urb:Urbain</b> <b>PA15+Rur:Rural</b>
Population active occupée	"Comprend toutes les personnes, âgées de 15 ans et plus, participant à la production de biens et services pendant une brève période de référence spécifiée et toutes les personnes pourvues normalement d'un emploi, mais absentes de leur travail pour un empêchement temporaire. " - hcp	<b>PAO:Total</b> <b>PAOUrb</b> <b>PAORur</b>
Structure de l'emploi urbain selon les branches d'activité	Combien en % constitue une branche d'activité dans le total des emplois urbains	- <b>AgrForPêche:</b> Agriculture, forêt et pêche <b>-Indus:</b> industrie <b>-BatTrav:</b> Bâtiments et travaux publics <b>-comm:</b> Commerce de gros et de détail <b>-TraEntCommu:</b> Transports, entrepôts et communications <b>-Serv :</b> services <b>-MalDes:</b> Activité mal désignée

# Mise en pratique de l'ACP

Indicateurs, Statistiques et Taux	Signification	Abréviation dans ce qui suit
Population active en chômage	Ceux qui ne sont pas occupés parmi la population active	<b>PAC</b>
Taux de féminité de la population active en chômage	Pourcentage des femmes dans la population active en chômage	<b>TFPAC</b>
Taux de chômage urbain selon le diplôme	Pourcentages des diplômés en chômage, ceux sans diplôme en chômage etc...	<b>- ChSansDipUrb: Sans diplôme</b> <b>- ChNivMoyUrb: Niveau moyen</b> <b>- ChNivSupUrb: Niveau supérieur</b> <b>- ChDipUrb: Ayant un diplôme</b>
Taux de chômage rural selon le diplôme	Même chose que ci-dessus mais pour le milieu rural	<b>- ChSansDipRur: Sans diplôme</b> <b>- ChDipRur: Ayant un diplôme</b>

Figure: Description des données

# Mise en pratique de l'ACP

Démarche suivie pour obtenir un tableau de données pour appliquer l'ACP: Le tableau de données initialement était sous ce format:

	PA15+	PA15+Urb	PA15+Rur	PAO	PAOUrb	PAORur	AgrForPêche	Indus	BatTrav
2002-08	10 849	5 532	5 318	9 710	4 586	5 124	5,3	21,9	10,0
2008	11 267	5 674	5 393	10 189	5 013	5 176	5,5	20,9	11,2
2009	11 314	5 916	5 398	10 284	5 101	5 184	5,0	20,2	11,8
2010	11 415	5 966	5 449	10 405	5 169	5 235	4,8	20,2	12,4
2011	11 538	5 553	5 237	10 509	5 272	5 237	4,9	20,2	12,4
2012	11 549	6 655	5 404	10 511	5 320	5 190	5,1	18,7	11,9
2013	11 705	6 217	5 488	10 625	5 346	5 278	4,9	18,4	10,9

Figure: Extrait des données initiales

Les valeurs sont soit en milliers comme PA15+ de 2013 : 11 705 000, soit en %.

# Implémentation de l'ACP en R

voici le script tout entier de la méthode ACP et puis on va décortiquer chaque ligne:

---

```
library("FactoMineR")
library("xlsx")
library("factoextra")

setwd("C:\\Users\\hp\\Desktop\\M1S1\\analyse_des_données\\projet")

dataset2 <- read.xlsx("dataset2.xlsx",sheetIndex = 1, header=TRUE,
col_names=TRUE)

#Pour utiliser les années en tant que noms des individus au lieu de
#1,2,...,7
rownames(dataset2) <- dataset2[,1]

resultACP <- PCA(dataset2[1:7, 2:22], graph = TRUE)
```

---

# Implémentation de l'ACP dans R

---

```
resultACP$var
resultACP$ind
fviz_contrib(resultACP,choice = "ind", axes=1)
fviz_contrib(resultACP,choice = "ind", axes=2)
fviz_contrib(resultACP,choice = "ind", axes=3)

fviz_cos2(resultACP,choice = "ind", axe=1)
fviz_cos2(resultACP,choice = "ind", axe=2)
fviz_cos2(resultACP,choice = "ind", axe=3)

fviz_contrib(resultACP,choice = "var", axes=1)
fviz_contrib(resultACP,choice = "var", axes=2)
fviz_contrib(resultACP,choice = "var", axes=3)

fviz_cos2(resultACP,choice = "var", axe=1)
fviz_cos2(resultACP,choice = "var", axe=2)
fviz_cos2(resultACP,choice = "var", axe=3)
```

---



# Implémentation de l'ACP dans R

On commence par importer les bibliothèques qu'on va utiliser qui sont:

- **FactoMineR**: Nous fournit la fonction `PCA()` qui réalise l'ACP;
- **xlsx**: Puisque on travail sur un fichier Excel alors on aura besoin des fonctions de cette bibliothèque pour le lire;
- **FactoExtra**: Nous fournit la fonction `fviz()` qui permet de visualiser les résultats de l'ACP dans des graphes.

Puis on se place sur le répertoire de travail par le biais de la fonction `setwd()`, et on lit le fichier excel qui contient les données en utilisant la fonction `read.xlsx`, on a passer à cette fonction les paramètres suivants:

- **"dataset2.xlsx"** : le nom du fichier;
- **sheetIndex = 1** : indique sur quelle feuille du classeur excel on s'intéresse;

# Implémentation de l'ACP dans R

- **header=TRUE** : True car la première ligne est une entête, c'est à dire elle contient juste les noms des variables;
- **col\_names=TRUE**: True car les noms des colonnes existes.

Pour utiliser les années en tant que noms des individus au lieu des nombres 1,2,...,7 on doit ajouter cette ligne de code : **rownames(dataset2) <- dataset2[,1]**

Après il suffit juste d'appeler la fonction PCA() en passant les données et elle va faire toutes les calculs nécessaires: les valeurs propres, les axes, les contributions, les qualités de représentation. . . , cette fonction prend en paramètres:

# Implémentation de l'ACP dans R

- **dataset2[1:7, 2:22]** : Les données en spécifiant qu'elles commencent de la première ligne après l'entête jusqu'à la 7<sup>ème</sup> ligne (on a 7 individus) et de la deuxième colonne jusqu'à 22<sup>ième</sup> (on a 22 variables);
- **scale.unit = TRUE**: pour centrer et réduire avant d'appliquer l'ACP;
- **graph = TRUE**: pour générer les graphiques.

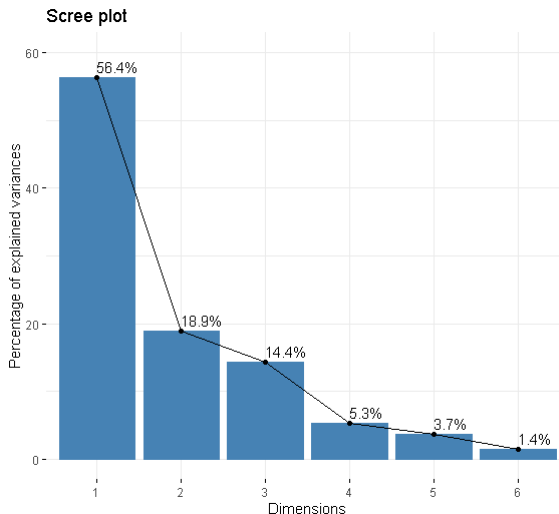
Puis, on affiche les résultats des variables et des individus et on génère leurs graphiques en utilisant la fonction **fviz\_contrib** pour les contributions et **fviz\_cos2** pour les qualités de représentation.

# Analyse des résultats

En suivant la règle de Cattell (1966) appelée test du talus (scree test) qui a dit que les dimensions qu'on accepte sont celles dont les pourcentages sont au-dessus de la courbe reliant les dernière valeurs propres, on ne choisira par la suite que les trois premières dimensions. On aura donc  $56.4\% + 18.9\% + 14.4\% = 89.7\%$  de l'information initiale présentée par ces trois dimensions.

Cherchons les axes qui résument le maximum de l'information initiale, pour cela examinons le graphique des pourcentage de variances des valeurs propres:

# Analyse des résultats



# Analyse des résultats

Maintenant étudions les résultats correspondant aux variables:

\$contrib

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
PA15.	7.83784315	0.98137813	0.000219132	0.8472038	2.193573957
PA15.Urb....	2.77720971	4.93170988	5.745382991	4.8752634	23.660554819
PA15.Rur	0.73690081	7.85656149	15.264186491	6.7090183	1.090025039
PAO	8.22412819	0.29952372	0.001767641	0.1777182	1.135957053
PAOurb	8.28373865	0.16477341	0.035902492	0.6449651	0.060529954
PAORur	5.82763261	1.45036615	0.530943078	2.6128998	26.727807368
AgrForPêche	4.11040061	0.41623703	4.982648989	30.7103591	0.116495189
Indus	6.12363974	5.89828480	0.334013937	0.9811454	2.368961585
BatTrav	5.18253615	8.25690613	0.163779992	2.4031427	1.599485181
comm	0.08084223	12.65181185	15.303598449	1.4457330	1.217146813
TraEntCommu	5.18900213	8.04906202	0.046192820	2.2209771	3.743360938
Serv	0.93198119	19.91779458	1.474391649	1.5831893	0.867667161
MalDes	3.52445650	2.64034535	2.581869698	22.7404427	15.618454491
PAC	5.77863595	5.53233504	0.096488160	2.9953952	7.115439258
TFPAC	2.14947054	3.26777035	17.269332069	7.5189683	1.184870002
ChsansDipurb	6.62725165	1.91941436	0.089292850	3.5786664	9.985913154
ChnivMoyurb	8.07304096	0.21615738	0.627936022	1.3225939	0.000301024
ChnivSupurb	8.38447354	0.03366986	0.042694436	0.1316238	0.162732661
Chdipurb	6.35701854	0.89102383	5.194655892	4.5382654	0.356938746
ChsansDipRur	1.01279004	0.56783464	26.804597680	1.1671936	0.775370462
ChdipRur	2.78700712	14.05703997	3.410105532	0.7952356	0.018415144

# Analyse des résultats

\$coord

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
PA15.	0.96327556	0.19725180	-0.002572355	0.09700129	0.130065330
PA15.Urb.....	0.57339855	0.44218250	0.416521550	0.23269263	-0.427167172
PA15.Rur	0.29536346	0.55810903	0.678914080	-0.27296884	0.091686131
PAO	0.98672741	0.10897287	0.007305923	0.04442726	0.093597956
PAOUrb	0.99029697	0.08082512	0.032926099	0.08463533	0.021605811
PAORur	0.83061226	0.23979583	-0.126619900	-0.17035110	0.454011730
AgrForPêche	-0.69758050	-0.12846155	0.387889596	0.58401772	0.029973634
Indus	-0.85144593	-0.48357703	-0.100429263	-0.10438792	0.135165052
BatTrav	0.78329152	-0.57215206	-0.070324784	-0.16337040	-0.111064564
comm	-0.09782985	0.70823788	-0.679789989	0.12671488	0.096885077
TraEntCommu	0.78378001	-0.56490501	-0.037347817	0.15705639	0.169908951
Serv	0.33216650	0.88863500	0.211000952	0.13260198	-0.081801636
MalDes	0.64594915	-0.32354386	0.279219114	0.50255435	0.347059899
PAC	-0.82711315	0.46833545	-0.053977777	0.18239404	0.234253595
TFPAC	0.50444986	-0.35993871	-0.722130609	0.28897662	-0.095591826
ChsansDipurb	-0.88576596	0.27585900	0.051926174	-0.19936283	0.277510423
ChnivMoyurb	-0.97762170	0.09257374	-0.137700552	0.12119839	-0.001523653
ChnivSupurb	-0.99630006	0.03653621	0.035905719	0.03823410	0.035426053
ChDipurb	-0.86751900	0.18795222	-0.396055798	0.22450629	-0.052466519
ChsansDiprur	-0.34626769	-0.15004231	0.899668739	0.11385577	0.077328580
ChDiprur	-0.57440908	-0.74653424	0.320894219	-0.09397914	-0.011917165

# Analyse des résultats

\$cor

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
PA15.	0.96327556	0.19725180	-0.002572355	0.09700129	0.130065330
PA15.Urb....	0.57339855	0.44218250	0.416521550	0.23269263	-0.427167172
PA15.Rur	0.29536346	0.55810903	0.678914080	-0.27296884	0.091686131
PAO	0.98672741	0.10897287	0.007305923	0.04442726	0.093597956
PAOurb	0.99029697	0.08082512	0.032926099	0.08463533	0.021605811
PAORur	0.83061226	0.23979583	-0.126619900	-0.17035110	0.454011730
AgrForPêche	-0.69758050	-0.12846155	0.387889596	0.58401772	0.029973634
Indus	-0.85144593	-0.48357703	-0.100429263	-0.10438792	0.135165052
BatTrav	0.78329152	-0.57215206	-0.070324784	-0.16337040	-0.111064564
comm	-0.09782985	0.70823788	-0.679789989	0.12671488	0.096885077
TraEntCommu	0.78378001	-0.56490501	-0.037347817	0.15705639	0.169908951
Serv	0.33216650	0.88863500	0.211000952	0.13260198	-0.081801636
MalDes	0.64594915	-0.32354386	0.279219114	0.50255435	0.347059899
PAC	-0.82711315	0.46833545	-0.053977777	0.18239404	0.234253595
TFPAC	0.50444986	-0.35993871	-0.722130609	0.28897662	-0.095591826
ChsansDipurb	-0.88576596	0.27585900	0.051926174	-0.19936283	0.277510423
ChnivMoyurb	-0.97762170	0.09257374	-0.137700552	0.12119839	-0.001523653
ChnivSupurb	-0.99630006	0.03653621	0.035905719	0.03823410	0.035426053
Chdipurb	-0.86751900	0.18795222	-0.396055798	0.22450629	-0.052466519
ChsansDipRur	-0.34626769	-0.15004231	0.899668739	0.11385577	0.077328580
ChdipRur	-0.57440908	-0.74653424	0.320894219	-0.09397914	-0.011917165

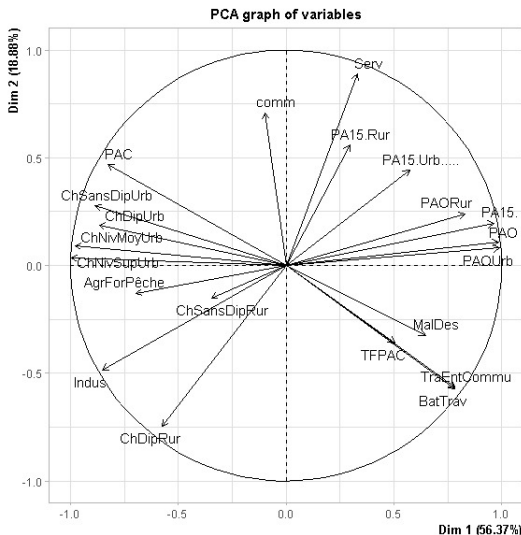


# Analyse des résultats

\$cos2

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
PA15.	0.92789980	0.038908273	6.617011e-06	0.009409251	1.691699e-02
PA15.Urb....	0.32878590	0.195525364	1.734902e-01	0.054145858	1.824718e-01
PA15.Rur	0.08723958	0.311485688	4.609243e-01	0.074511985	8.406347e-03
PAO	0.97363098	0.011875087	5.337651e-05	0.001973782	8.760577e-03
PAOurb	0.98068809	0.006532700	1.084128e-03	0.007163139	4.668111e-04
PAORur	0.68991673	0.057502038	1.603260e-02	0.029019499	2.061267e-01
AgrForPêche	0.48661855	0.016502369	1.504583e-01	0.341076700	8.984187e-04
Indus	0.72496017	0.233846741	1.008604e-02	0.010896839	1.826959e-02
BatTrav	0.61354561	0.327357978	4.945575e-03	0.026689886	1.233534e-02
comm	0.00957068	0.501600900	4.621144e-01	0.016056662	9.386718e-03
TraEntCommu	0.61431110	0.319117673	1.394859e-03	0.024666710	2.886905e-02
Serv	0.11033458	0.789672168	4.452140e-02	0.017583285	6.691508e-03
MalDes	0.41725031	0.104680627	7.796331e-02	0.252560874	1.204506e-01
PAC	0.68411616	0.219338089	2.913600e-03	0.033267586	5.487475e-02
TFPAC	0.25446966	0.129555875	5.214726e-01	0.083507487	9.137797e-03
chsansdipurb	0.78458134	0.076098189	2.696328e-03	0.039745538	7.701203e-02
chNivMoyurb	0.95574419	0.008569898	1.896144e-02	0.014689049	2.321518e-06
chNivsupurb	0.99261380	0.001334894	1.289221e-03	0.001461846	1.255005e-03
chdipurb	0.75258921	0.035326036	1.568602e-01	0.050403076	2.752736e-03
chsansdipRur	0.11990131	0.022512694	8.094038e-01	0.012963136	5.979709e-03
chdipRur	0.32994579	0.557313370	1.029731e-01	0.008832079	1.420188e-04

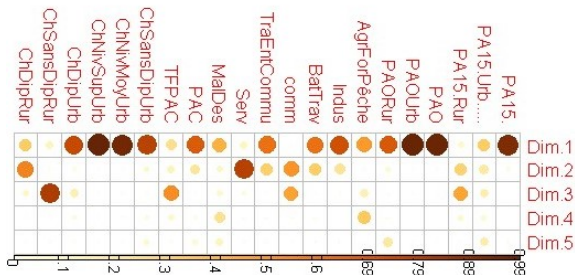
# Cercle de corrélation



# Cercle de corrélation

La première remarque est que presque toutes les variables sont bien représentées dans ce plan constitué par les deux premiers axes, car la distance entre la majorité des variables et l'origine du cercle de corrélation est grande. Puis on remarque que les variables s'arrangent en deux groupes opposés, ceux à droite de la deuxième dimension et ceux à gauche. Maintenant étudions la qualité de représentation des variables:

# Qualité de représentation des variables



On remarque que presque toutes les variables sont bien représentées par la première dimension, et que la qualité de représentation sur la deuxième dimension est soit faible soit moyenne.



# Étude sur les individus

\$contrib

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
2002-08	68.157890799	4.370334	10.69682873	0.05459475	1.037328
2008	4.230267068	10.242695	33.29221173	22.37623293	15.302490
2009	0.001366212	2.934589	8.64091321	25.51011871	16.444646
2010	3.337913997	3.756466	0.47509427	28.56760876	4.261650
2011	7.236680531	18.836402	46.58404658	3.02596469	3.356057
2012	9.201129662	2.291783	0.28392802	20.34237883	44.250203
2013	7.834751733	57.567731	0.02697745	0.12310132	15.347627

\$coord

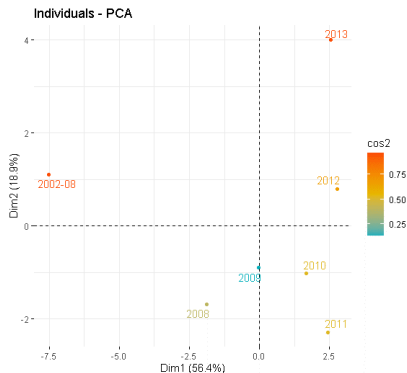
	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
2002-08	-7.51552541	1.1013089	-1.5036769	-0.06514904	-0.2366423
2008	-1.87234196	-1.6860053	2.6527622	1.31894318	0.9088987
2009	-0.03364808	-0.9024548	1.3514713	-1.40827944	-0.9422078
2010	1.66317844	-1.0210371	0.3168961	-1.49028573	0.4796487
2011	2.44889958	-2.2863925	-3.1379470	0.48502545	0.4256465
2012	2.76135251	0.7975144	0.2449803	1.25757376	-1.5455820
2013	2.54808491	3.9970664	0.0755140	-0.09782818	0.9102382

# Liaison entre les variables

\$cos2

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
2002-08	0.9406616713	0.02019915	0.0376550999	7.068558e-05	0.0009326088
2008	0.2196983063	0.17814525	0.4410146596	1.090206e-01	0.0517710912
2009	0.0001832634	0.13182754	0.2956438911	3.210206e-01	0.1436973146
2010	0.3791149327	0.14288158	0.0137634402	3.043913e-01	0.0315311502
2011	0.2773133586	0.24172992	0.4553234907	1.087822e-02	0.0083777434
2012	0.6107627872	0.05094551	0.0048071882	1.266764e-01	0.1913431778
2013	0.2773186962	0.68239226	0.0002435602	4.087703e-04	0.0353884421

# Graphique du nuage des individus



Les individus, 2002-08 et 2013 sont représentés avec une très bonne qualité. L'individu 2009 n'est pas bien représenté sur ce plan. Pour les autres, ils ont une qualité de représentation moyenne.



# Interprétations et Conclusions

D'après le graphique du nuage des individus, on peut construire des groupes des individus qui se ressemblent:

- **Groupe 1:** 2013 et 2012;
- **Groupe2:** 2010 et 2011;
- **Groupe3:** 2008 et 2009;
- **Groupe4:** 2002-08.

On peut conclure qu'à chaque deux ans la situation de l'emploi, activité et chômage change considérablement. Cela peut être dû aux transformations démographiques, sociales et économiques qui prennent place au cours de cette période.

# Interprétations et Conclusions

Aussi, on remarque que les deux groupes 1 et 2 s'opposent aux autres groupes 3 et 4 par rapport à la seconde dimension. Si on ajoute à cela quelques informations historique et politiques à savoir :

- **Le mouvement appelé “printemps arabe” qui a commencé dans le début de 2010 et qui a affecté plusieurs pays arabes et nords africains;**
- **Le nouveau constitution marocain de 2011;**
- **L'arrivée du gouvernement de la partie de justice et développement sous présidence de Benkirane en 2011.**

On peut donc dire qu'après ces événements critiques, la situation d'emploi, activité et chômage a subi un changement critique en la comparant à celle avant 2010.