

TP3 - ADM

Guillaume Bernard-Reymond et Lorenzo Gaggini

Décembre 2023

Dans ce TP, nous avons à notre disposition un jeu données concernant 100 villes françaises dont on connaît 54 indicateurs répartis en quatre grands thèmes : Economie, Risques, Nature et Culture.

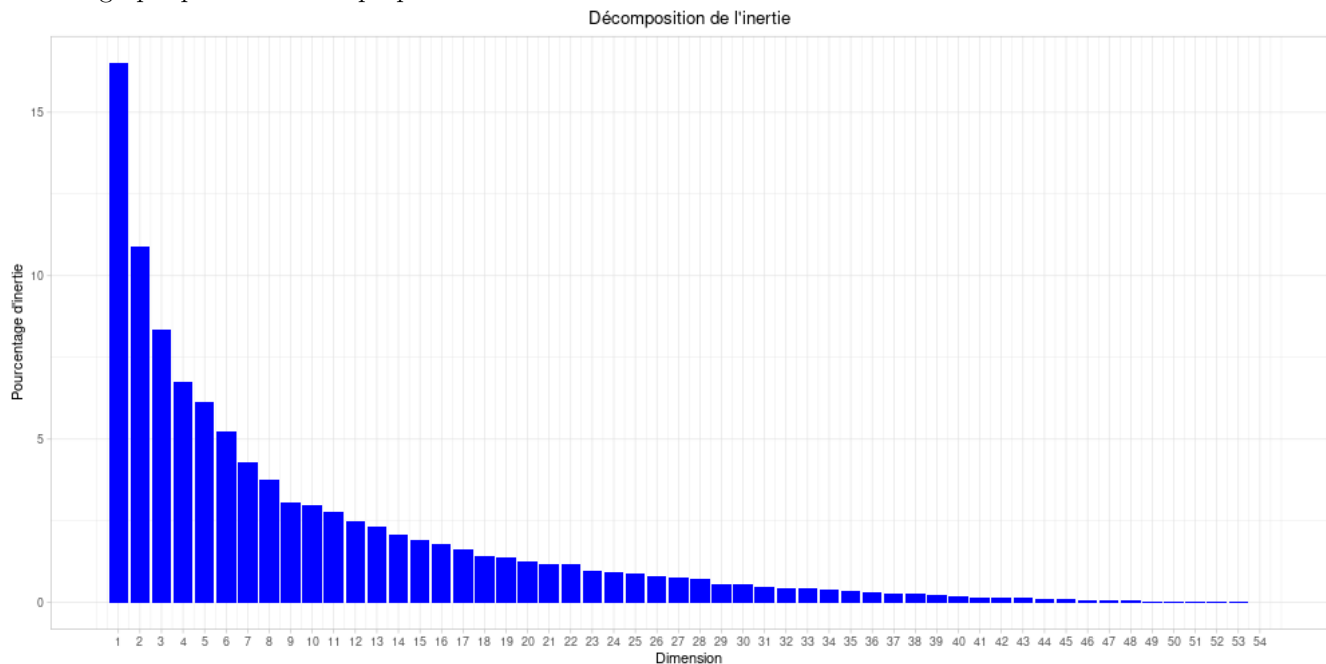
Nous avons décidé d'utiliser la bibliothèque Factoshiny pour son confort d'utilisation. On trouvera le code R utilisé en appendice.

1 ACP normée sur l'ensemble des variables

1.1 Etude de la variance cumulée

Dans un premier temps, nous nous intéressons à la variance cumulée et notamment aux grands décrochements. Nous allons devoir faire un choix qui ne sera malheureusement pas sans perte d'informations.

Voici le graphique des valeurs propres :



Voici le tableau des variances cumulées :

Valeurs propres	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5	Dim.6	Dim.7	Dim.8
Variance	8.914	5.889	4.511	3.646	3.306	2.820	2.326	2.041
% of var.	16.508	10.906	8.353	6.752	6.121	5.222	4.307	3.780
Cumulative % of var.	16.508	27.414	35.767	42.519	48.641	53.862	58.169	61.949

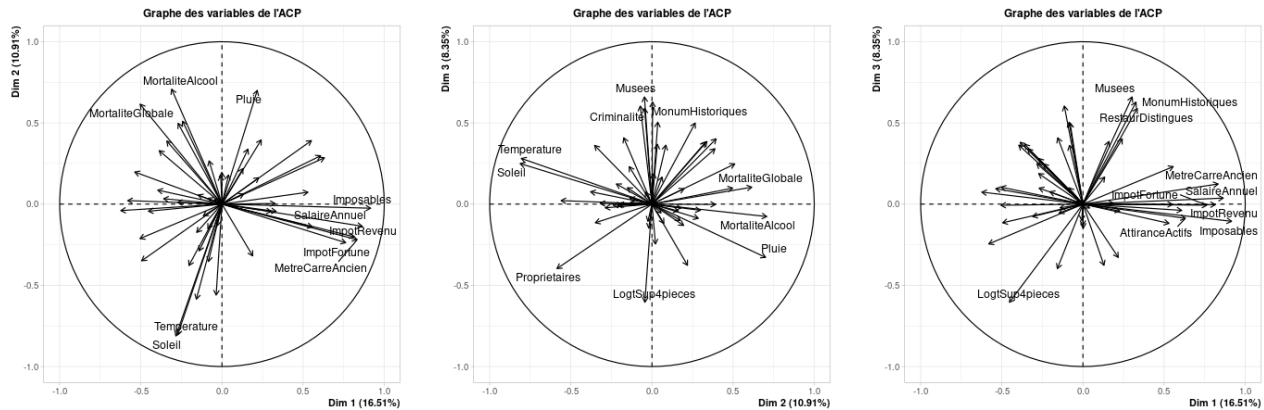
Le nombre d'axes choisis pour mener l'étude de l'ACP est le même pour les individus que les variables car les valeurs propres des opérateurs d'inertie directe et duale sont les mêmes à un certain nombre de valeurs propres nulles près. Ainsi dans toute la suite du rapport, nous ne distinguerons plus, sauf mention explicite du contraire, la variance

cumulée des individus et celle des variables.

On peut observer un décrochement après la troisième valeur propre. Nous nous contenterons donc des trois premiers axes avec une inertie cumulée de seulement 35.767%. Pour être plus exhaustif, on pourrait prendre davantage de valeurs propres. Toutefois, ne pouvant observer de véritables décrochements dans les valeurs propres, nous serions obligés d'en garder environ 8 car la différence entre deux valeurs successives est faible. Le temps de traitement en deviendrait trop long.

1.2 Etude des variables

Sur les graphiques suivants sont marqués les 10 variables les plus contributives à la fabrication des axes :



Axe 1

• Contribution

Tout d'abord, observons que $\frac{1}{54} \approx 2\%$ ce qui nous pousse à ne considérer que les variables dont la contribution est très supérieure à 2%. On trouve alors : *Imposables* (9,41%), *SalairesAnnuels* (8,38%), *MetreCarreAncien* (7,75%), *ImpotRevenu* (7,46%) et *ImpotFortune* (6,52%). En additionnant, leur contribution on obtient : 39,52% ce qui est très supérieur à $5 \times 2\% = 10\%$. Ce sont donc des variables économiques qui contribuent en une part importante à la fabrication de l'axe 1. On voit d'ailleurs sur le plan (1,2), ce groupe de variables allant dans la même direction et formant un faisceau.

• Originalité : COS2

Du fait de la proportionnalité entre la contribution et l'originalité, on retrouve ici les mêmes variables que précédemment et dont l'originalité est principalement décrite par l'axe 1 : *Imposables* (0,84), *SalairesAnnuels* (0,75), *ImpotFortune* (0,66), *MetreCarreAncien* (0,65).

On peut en conclure que cet axe 1 décrit des marqueurs de richesse.

Axe 2

• Contribution

Parmi les variables dont la contribution est très supérieure à 2%, il y a : *Soleil* (11,20%), *Température* (11,01%), *MortaliteAlcool* (8,50%), *Pluie* (8,33%) et *MortaliteGlobale* (6,43%). De manière cumulée, on obtient 45,47%. Sur la représentation du plan (1,2), on voit ces variables sont négativement corrélées : *Soleil* et *Température* d'un côté, les trois restantes de l'autre.

- **Originalité : COS2**

L'originalité des variables *Soleil* et *Temperature* est capturée à hauteur de 0,66 et 0,65 par l'axe 2. Si on leur ajoute leurs valeurs sur les axes 1 et 3, on trouve *Soleil* : 0,8 et *Temperature* : 0,8 aussi. Pour la *MortaliteAlcool*, on a 0,50 sur l'axe 2 et 0,61 si l'on considère les trois axes. De même, le COS2 de la mortalité globale est de 0,38 le long de l'axe 2 et 0,64 si l'on considère les 3 axes. Enfin pour la pluie, on trouve 0,49 sur l'axe 2 et 0,65 pour les 3 axes.

Ici on commence à avoir une idée de ce que semble représenter cet axe : une opposition entre le Nord et le Sud de la France. L'étude ultérieure des individus nous permettra d'en dire certainement plus.

Axe 3

- **Contribution**

Voici un tableau des variables les plus contributives à la fabrication de l'axe 3 :

Variables	Musées	MonumentHistorique	LogtSup4Pieces	Criminalite	RestauDistingues	Total
Contribution	9,65	8,15	8,10	8,08	7,70	41,68

- **Originalité : COS2**

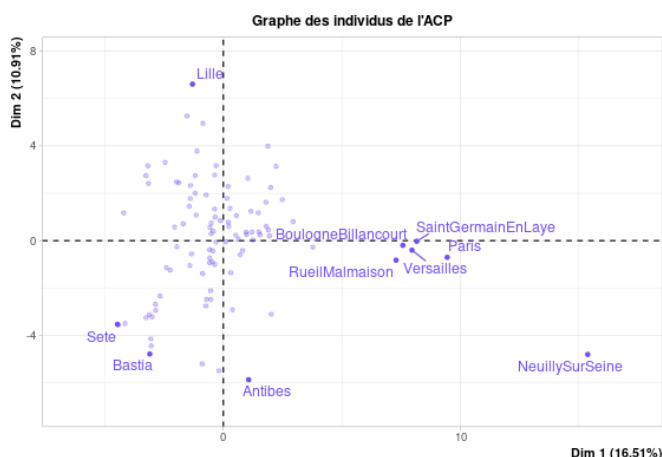
Pour ce qui est de l'originalité, on ne retrouve rien de très significatif. Celle-ci semble être diluée dans davantage d'axes que les 3 considérés.

Variables	Musées	MonumentHistorique	LogtSup4Pieces	Criminalite	RestauDistingues
Originalité sur l'axe 3	0,44	0,39	0,37	0,36	0,35
Originalité sur les 3 axes	0,53	0,5	0,58	0,38	0,46

Si l'on observe les plans (2,3) et (1,3), on peut voir un faisceau entre des variables culturelles *Musées*, *MonumentHistorique* et *RestauDistingues*, la variable *LogtSup4Pieces* étant elle négativement corrélée aux trois précédentes. L'originalité de la variable criminalité n'étant finalement pas suffisamment décrite par l'axe 3, on peut supposer que l'axe 3 reflète plutôt une tendance culturelle et qui va opposer grandes et petites villes.

1.3 Etude des individus

Voici le graphique du plan (1,2) avec les 10 individus les plus contributifs :



Pour ce qui concerne la contribution, nous comparerons les valeurs avec la contribution moyenne $\frac{1}{100} = 0,01$.

Axe 1

- **Contribution** Six villes ont une forte contribution pour l'axe 1 dont une plus particulièrement :

Individus	Neuilly-Sur-Seine	Paris	Saint-Germain-en-Layes	Versailles	Boulogne-Billancourt	Rueil-Malmaison	Total
Contribution	26,51	10,0	7,46	7,10	6,43	5,96	63,46

La contribution de ces six villes est très largement supérieure à la contribution moyenne de 6%.

- **Originalité : COS2**

Les villes dont l'originalité est largement supportée par l'axe 1 sont les mêmes que précédemment à l'exception notable de *Paris* :

Individus	Versailles	Saint-Germain-en-Layes	Neuilly-Sur-Seine	Boulogne-Billancourt	Rueil-Malmaison
Originalité	0,73	0,67	0,65	0,61	0,57

L'originalité de *Paris* le long de cet axe est seulement de 0,25. Ceci semble indiquer que Paris n'est pas seulement originale par sa dimension économique mais que son originalité est multiple.

Ce que l'on peut conclure de cette étude de l'axe 1, est que ce dernier représente effectivement des variables économiques et que quelques villes de l'Ouest parisien contribue fortement à son orientation. Si l'on met côte à côte le plan (1,2) du graphe de l'ACP des variables et celui des individus, on remarque évidemment une direction similaire entre les variables étudiées et les villes contributives.

Axe 2

- **Contribution**

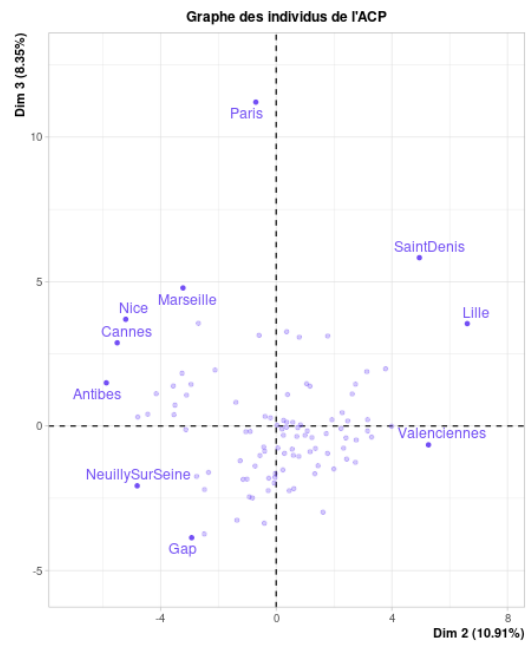
Les principales contributions à la construction de l'axe 2 sont données par 6 villes : *Lille* (7,42%), *Antibes* (5,85%), *Cannes* (5,13%), *Valenciennes* (4,71%), *Nice* (4,60%) et *Saint-Denis* (4,17%). En additionnant leur contribution, on obtient 31,88%.

- **Originalité : COS2**

Les cinq villes dont l'originalité est décrite partiellement par cet axe 2 sont : *Lille* (0,47), *Antibes* (0,44), *Troyes* (0,42), *Nice* (0,40) et *Valenciennes* (0,38).

De cet axe 2 se dégage un axe Nord-Sud. Au Nord, des villes pluvieuses avec une mortalité globale supérieure et au Sud des villes chaudes et ensoleillées.

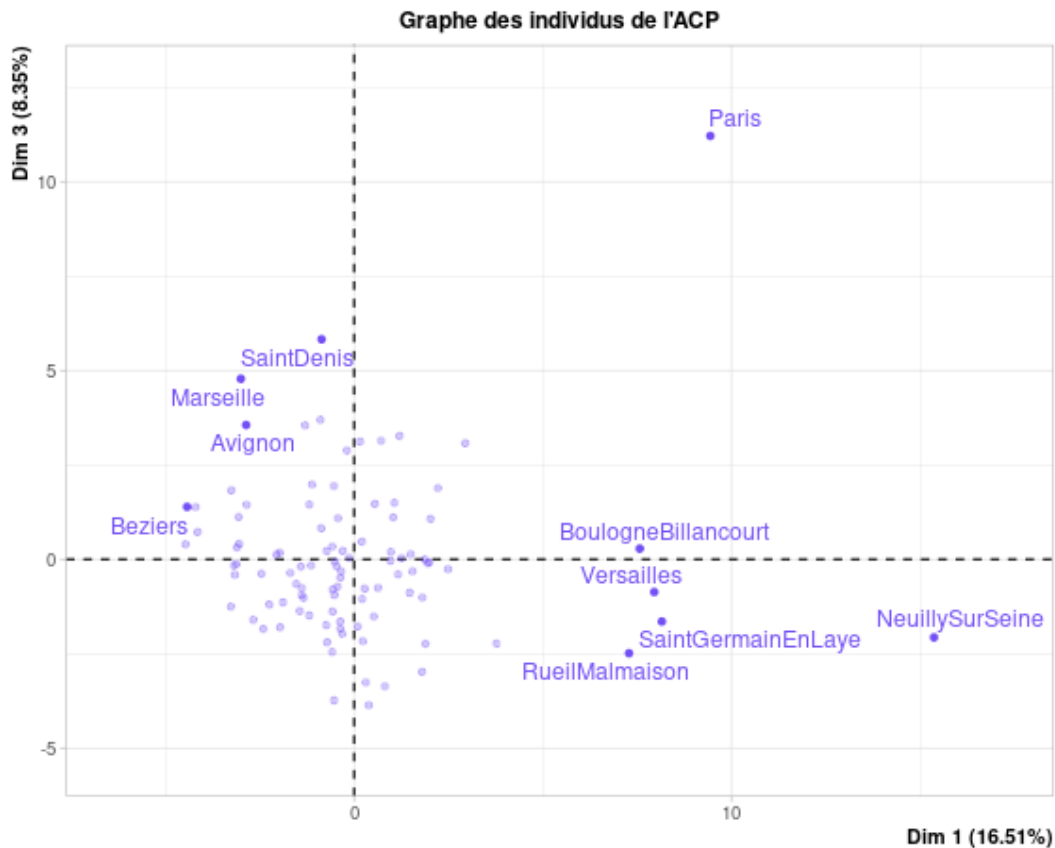
L'édition du plan (2,3) nous montre un regroupement de villes du Sud dans le quart supérieur gauche :



Ce regroupement se trouve être justement dans la même direction que celui des variables *Soleil* et *Temperatures*.

Axe 3

Voici le plan (1,3) des individus :



- **Contribution**

On trouve ici quelques contributeurs importants et notamment une ville singulière *Paris* contribuant à hauteur de 27,86%. Les deux autres villes sont *Saint-Denis* (7,53%) et *Marseille* (5,06%). Dans une moindre mesure, *Gap* (3,31%) et *Cholet* (3,09%) participent au positionnement de cet axe.

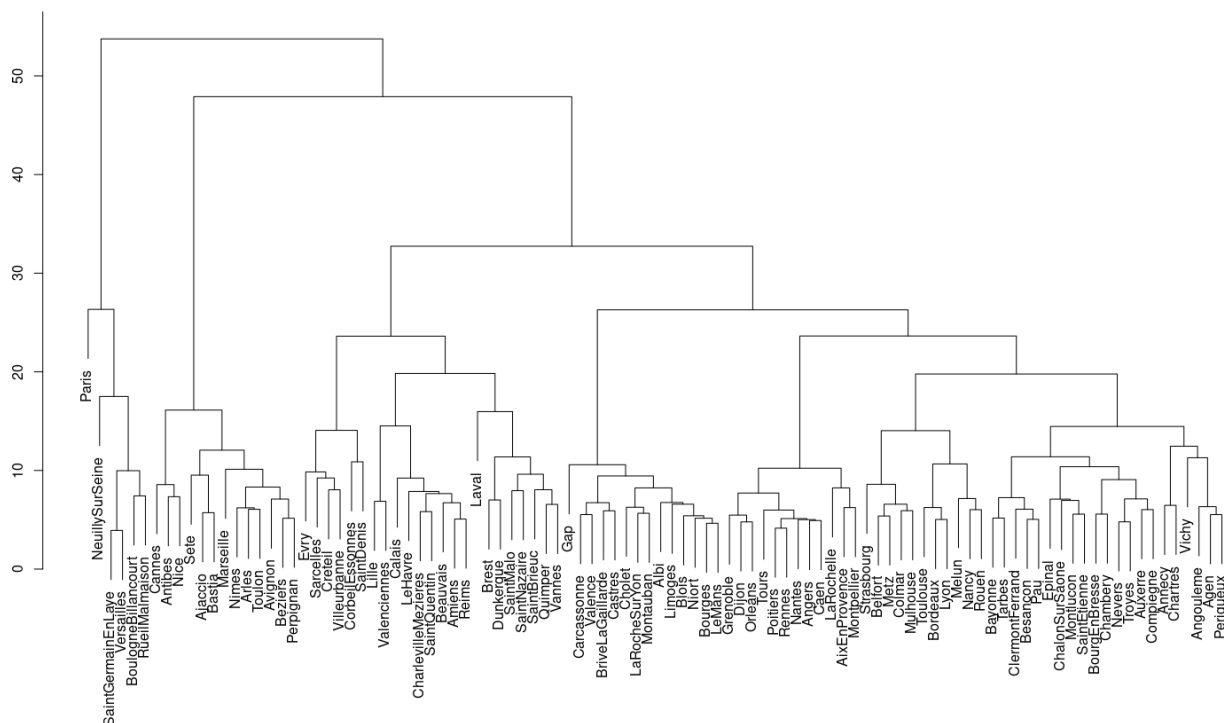
- **Originalité COS2**

Aucune ville ne semble être majoritairement décrite par cet axe 3. Toutefois on peut remarquer un fait notable. Nous avons dit que Paris, bien que contributeur de l'axe 1, ne voyait pas son originalité y être décrite. Paris se trouve être plus original selon cet axe 3 (0,35) auquel il participe grandement à la construction. Les villes contributrices tirent toutefois une part relativement faible de leur originalité sur cet axe : *Marseille* (0,29), *Cholet* (0,28) *Saint-Denis* (0,26) et *Gap* (0,22).

En superposant les plans où se trouvent l'axe 3, on remarque évidemment que les variables contributives à sa construction et la position de *Paris* ne sont pas étrangères. La direction est la même découlant du fait que Paris a une contribution majeure à cet axe. Elle en tire d'ailleurs son originalité principale. Cet axe semble être celui des marqueurs culturels.

1.4 Conclusion

On retrouve dans cette ACP les résultats obtenus dans le TP2 où le dendrogramme donnait une impression visuelle de former 3 classes en coupant au-dessus de 35 :



On retrouve dans cette ACP, 3 axes principaux :

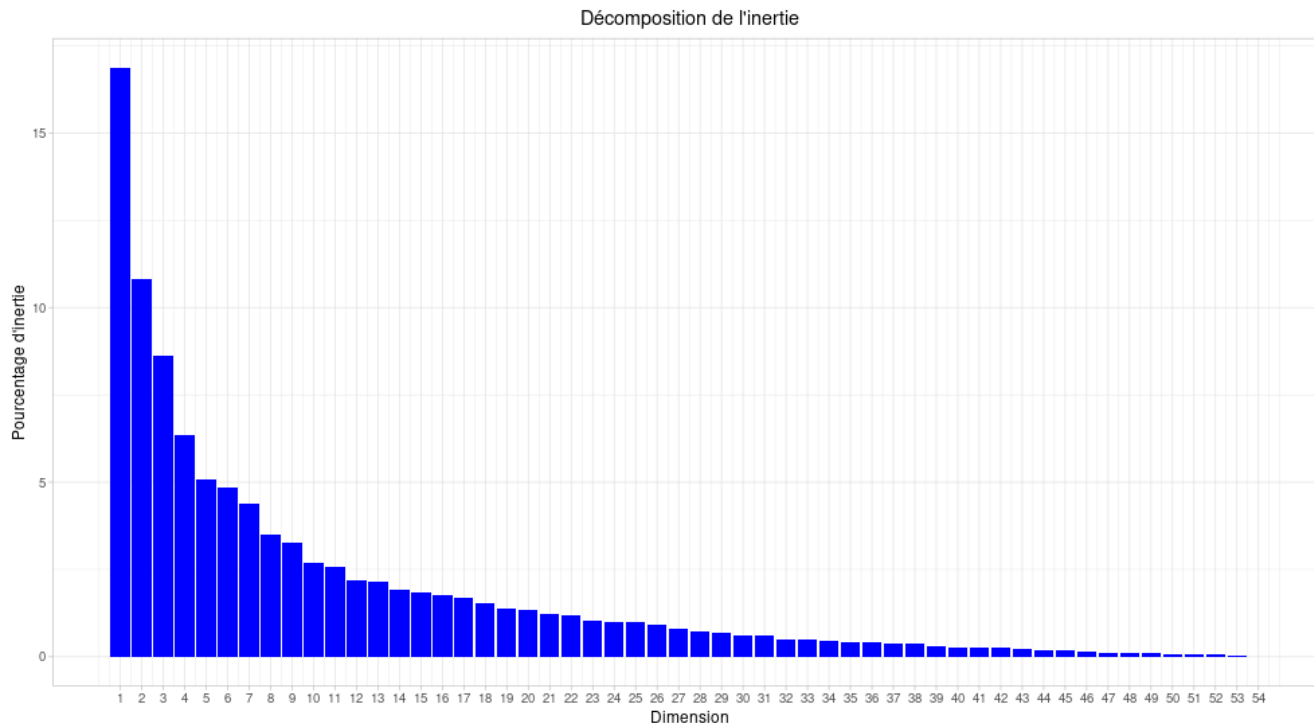
- **Axe 1 :** axe à dominante économique avec un individu singulier *Neuilly-Sur-Seine* et plusieurs gros contributeurs se caractérisant par une position géographique bien particulière : l'Ouest parisien.
- **Axe 2 :** celui-ci oppose le Sud ensoleillé et chaud au Nord pluvieux dont certaines variables au caractère plus morbide sont contributrices.
- **Axe 3 :** c'est la culture qui se distingue mais la culture parisienne. En effet, il s'agit d'un contributeur atypique à sa fabrication.

2 ACP de rang sur l'ensemble des variables

Après avoir obtenu la matrice des rangs, nous lançons l'ACP avec Factoshiny.

2.1 Etude de la variance cumulée

Nous nous intéressons, comme pour l'ACP normée, aux valeurs propres dont voici le graphique :



et le tableau de valeurs :

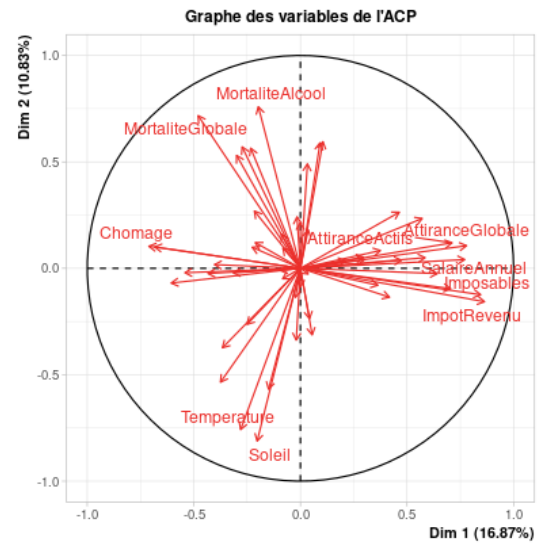
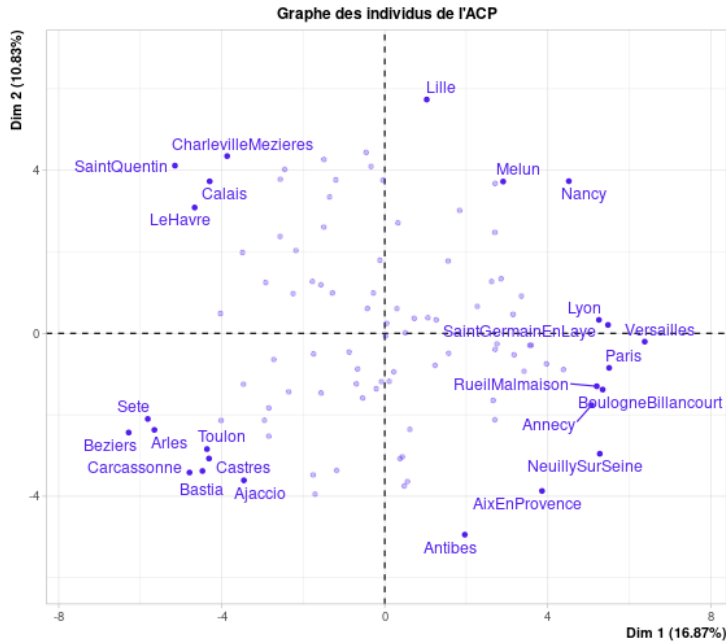
Valeurs propres	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5	Dim.6	Dim.7	Dim.8
Variance	9.110	4.511	4.662	3.422	2.752	2.616	2.358	1.879
% of var.	16.871	10.825	8.634	6.338	5.096	4.844	4.366	3.480
Cumulative % of var.	16.871	27.696	36.330	42.668	47.764	52.608	56.974	60.455

On a un décrochement entre la 4-ème et la 5-ème valeur ce qui nous pousse à nous intéresser aux quatre premiers axes. C'est déjà une différence avec l'ACP normée.

2.2 Etude des différents plans

Plan (1,2)

Côte à côte voici les graphiques des individus et celui des variables avec les 25 plus gros contributeurs pour les individus et les 10 plus importantes pour les variables :



On retrouve le même faisceau des variables de richesse que sur l'ACP normée, auxquelles viennent se rajouter *AttiranceActifs* et *AttiranceGlobale*. Ces variables sont à mettre en opposition à la variable *Chomage* qui décrit davantage la pauvreté. Dans la même direction se trouvent les villes trouvées lors de l'ACP normée : celle de l'Ouest parisien. Des villes attirantes comme *Lyon*, *Annecy* ou bien *Aix-en-Provence* viennent compléter la liste.

Tableau des variables les plus contributrices pour l'axe 1 :

Côté -				Côté +			
Variables	Coordonnées	CTR	COS2	COS2	CTR	Coordonnées	Variables
Chômage	-0.71	5.55	0.51	0.75	8.19	0.86	ImpotRevenu
				0.71	7.85	0.85	Imposables
				0.61	6.67	0.78	AttiranceGlobale

Tableau des individus les plus contributeurs pour l'axe 1 :

Côté -				Côté +			
Villes	Coordonnées	CTR	COS2	COS2	CTR	Coordonnées	Villes
Beziers	-6.28	4.33	0.57	0.58	4.47	6.38	Versailles
Sète	-5.81	3.71	0.41	0.44	3.33	5.51	Paris
				0.40	3.30	5.48	Saint-Germain-En-Laye

Sur cet axe 1, il y a plusieurs variables contributives, mais peu sont illustratives car seuls les COS2 de *ImpotRevenu* et *Imposables* sont importants. Enfin, il est important de noter que Neuilly-Sur-Seine a perdu sa singularité et contribue beaucoup plus faiblement au positionnement de l'axe 1.

L'axe 2 met en opposition des villes du Sud et du Nord avec deux groupes bien distincts. Ceci correspond à une opposition entre les variables *Température*, *Soleil* et les variables *MortaliteAlcool* et *MortaliteGlobale*. C'est clairement visible sur les deux tableaux suivants :

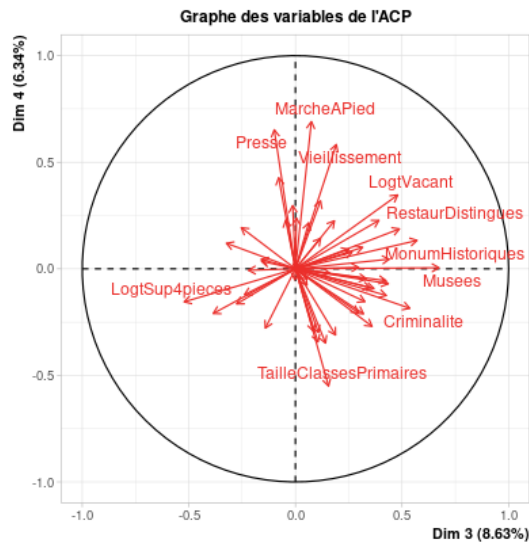
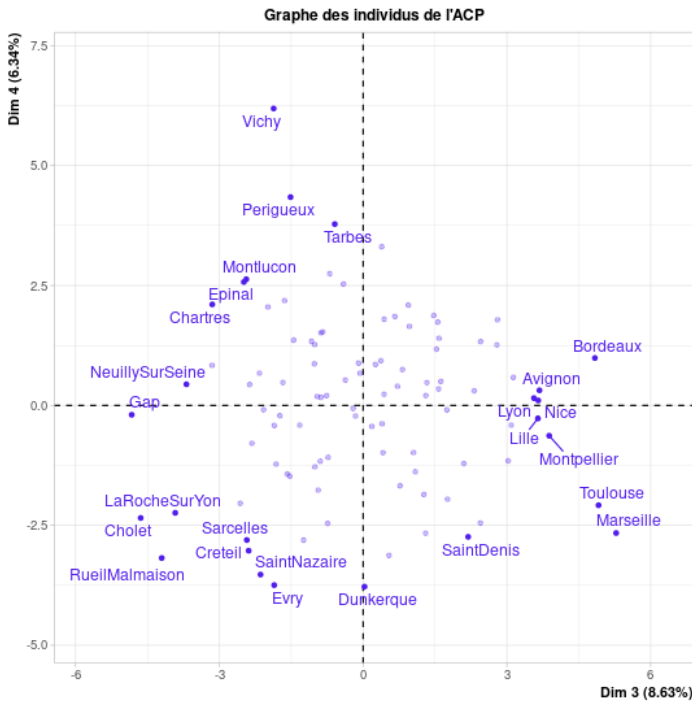
Tableau des variables les plus contributrices pour l'axe 2 :

Côté -				Côté +			
Variables	Coordonnées	CTR	COS2	COS2	CTR	Coordonnées	Variables
Soleil	-0.81	11.27	0.66	0.58	9.85	0.76	MortaliteAlcool
Temperature	-0.76	9.83	0.57	0.51	8.80	0.72	MortaliteGlobale
Vieillessement	-0.57	5.52	0.32				

Tableau des individus les plus contributeurs pour l'axe 2 :

Côté -				Côté +			
Villes	Coordonnées	CTR	COS2	COS2	CTR	Coordonnées	Villes
Antibes	-4.94	4.18	0.33	0.52	5.64	5.74	Lille
				0.33	3.38	4.44	Valenciennes
				0.22	3.12	4.27	Saint-Denis

Plan (3,4)



En observant ces deux graphiques, il ressort que les variables les mieux représentées pour l'axe 3 sur le plan sont *Musées*, *MonumHistoriques*, *Criminalité* et *RestauDistingue*. Ces variables sont corrélées négativement à *Logt-Sup4Pieces*. En regardant de plus près les individus contributeurs, on trouve des grandes villes françaises à droite qui font la part belle à la culture mais où la criminalité est importante. Par criminalité, il faut entendre le nombre d'infractions et délits commis. Les infractions relatives au code de la route y sont très nombreuses dans les grandes villes.

Tableau des variables les plus contributrices pour l'axe 3 :

Côté -				Côté +			
Variables	Coordonnées	CTR	COS2	COS2	CTR	Coordonnées	Variables
LogtSup4Pieces	-0.52	5.85	0.27	0.45	9.72	0.67	Musées
				0.33	6.99	0.57	MonumHistoriques
				0.29	6.18	0.54	Criminalité
				0.32	5.11	0.49	RestauDistingues

Tableau des individus les plus contributeurs pour l'axe 3 :

Côté -				Côté +			
Villes	Coordonnées	CTR	COS2	COS2	CTR	Coordonnées	Villes
Gap	-4.83	5	0.33	0.44	5.98	5.28	Marseille
Cholet	-4.64	4.62	0.33	0.43	5.17	4.91	Toulouse
				0.40	5.02	4.84	Bordeaux

La principale différence avec l'ACP normée est que l'on ne retrouve pas Paris sur cet axe alors qu'il en était le principal contributeur précédemment. Le fait de faire une ACP de rang 3 à éliminer une grande partie de la

singularité de Paris et les variables culturelles qui la décrivait. Ici les COS2 sont plutôt faibles et l'originalité des villes est partiellement mis en évidence ici.

Sur l'axe 4, trois variables ont une forte contribution : *MarcheAPied* (13.95), *Presse* (12.43) et *Vieillesse* (9.94). Pour ce qui est des villes, deux sortent du lot *Vichy* (11.20) et *Tarbes* (4.17). L'originalité de Vichy est tirée à 50% par ces variables et à 31% pour Tarbes. Concernant les variables, *MarcheAPied* tire 48% de son originalité le long de cet axe, 43% pour *Presse* et 34% pour *Vieillesse*.

Ces trois variables semblent décrire des variables liées à une population de retraités. Il faudrait regarder plus en détails les villes concernées pour en avoir le cœur net.

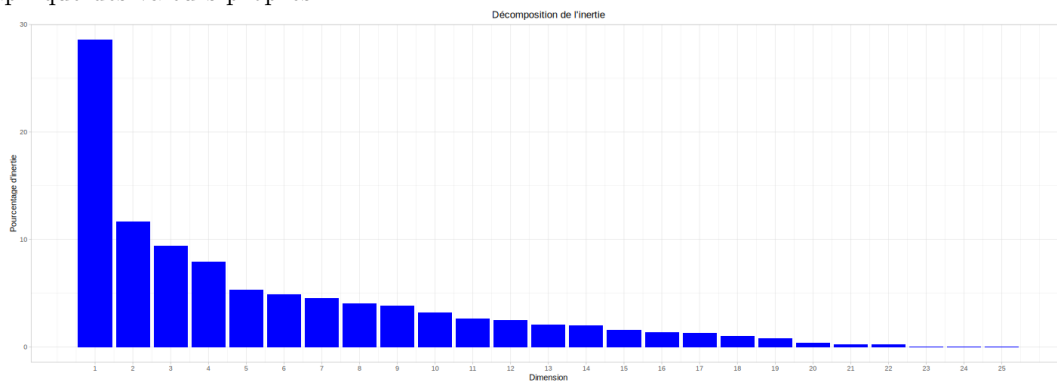
2.3 Conclusion

Cette ACP de rang a permis de dégager d'autres résultats. En effet sur l'axe 1 de l'ACP normée, *Neuilly-Sur-Seine* était un individu très singulier de part sa contribution disproportionnée. L'ACP de rang a permis de gommer cette singularité. De même, l'axe 3 s'est trouvé modifier car la contribution de Paris a été amoindrie permettant de faire émerger d'autres villes où la culture est bien présente.

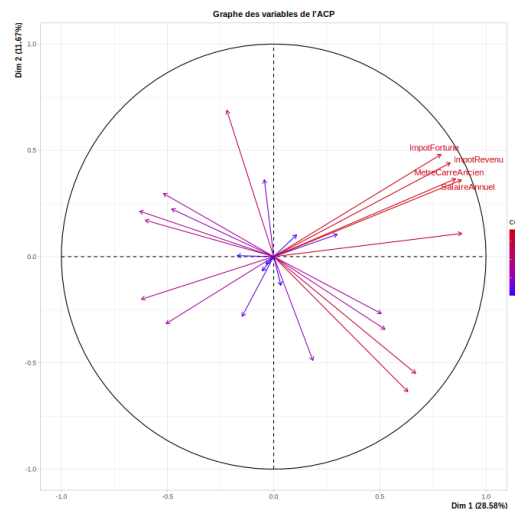
3 ACP normée pour les différents thèmes

3.1 Thème Economie

Voici le graphique des valeurs propres :

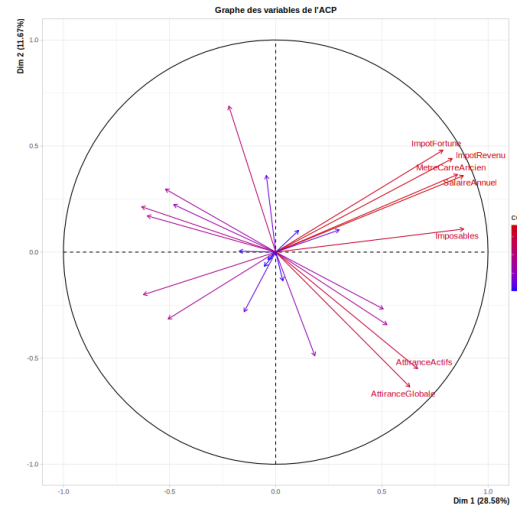


N'ayant de décrochement clair qu'après la 4-ème valeur propre, on se restreindra à 2 axes qui permettront d'explorer 40,25% de la variance totale pour le thème économie.



Se limiter aux individus et variables les mieux représentés dans un premier temps ($\text{COS2} > 0,85$), dégage un premier cluster *SalairesAnnuel*, *ImpotRevenu*, *ImpotFortune*, *MetreCarreAncien* porté par les villes de l'Ouest parisien.

Ce plan est alors fortement déterminé par les indicateurs du niveau de richesse de ces villes. Cette considération nous sert de levier d'interprétation pour la suite :



En étant plus tolérant ($\text{COS2} > 0,6$) : un second cluster *AttiranceGlobale*, *AttiranceActifs* se détache, et deux nouveaux groupes s'opposant apparaissent :

- *Lyon*, *Nancy*, *Melun* et *Rouen*, qui ont permis la détermination de ce cluster par leur point commun : une forte valeur sur les variables *AttiranceGloale* et *AttiranceAtif*. Leur différence commune est qu'elles sont toutes loin derrière l'Ouest parisien en niveau de richesse.

Notons que dans le tableau de données brutes, le top 10 dans ces 2 catégories est presque partagé entre ce groupe et l'Ouest parisien. Ceci renforce l'idée que c'est cette grande différence sur l'attirance de ces villes qui a permis l'émergence de ce second groupement.

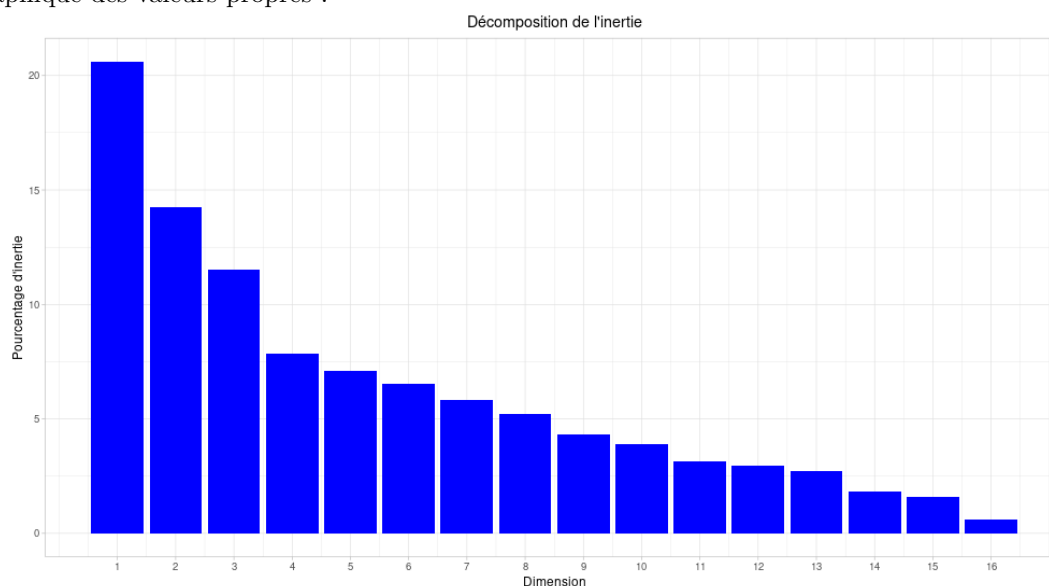
- *Le Havre*, *Arles*, *Béziers* forment un groupe antagoniste aux deux précédents. Ces trois villes sont dans le bas du classement dans la plupart des indicateurs en matière de richesse et d'attirance.

L'apparition de ce dernier groupe éclaire et améliore nos premiers travaux sur ce jeu de données. On retrouve l'intuition de notre première interprétation des 3 classes obtenues lors du TP2 par la CAH du thème *Economie*. La riche région parisienne formait déjà une classe à part entière (classe 3). *Lyon*, *Nancy*, *Melun* et *Rouen* étaient bien dans la même classe des "villes relativement attractives" (classe 1).

Finalement, nous obtenons une meilleure compréhension entre ce qui fait les points communs de la classe 1 et 3 (une certaine part de richesse et une haute attirance Global/Actif) et ce qui fait de la classe 2 (dont *Le Havre*, *Arles*, *Béziers* font partis) une classe si différentes des deux autres : son faible niveau de richesse et une faible attirance globale ou des actifs.

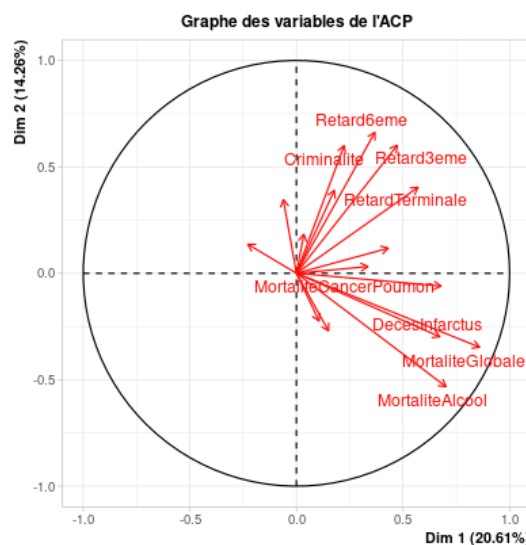
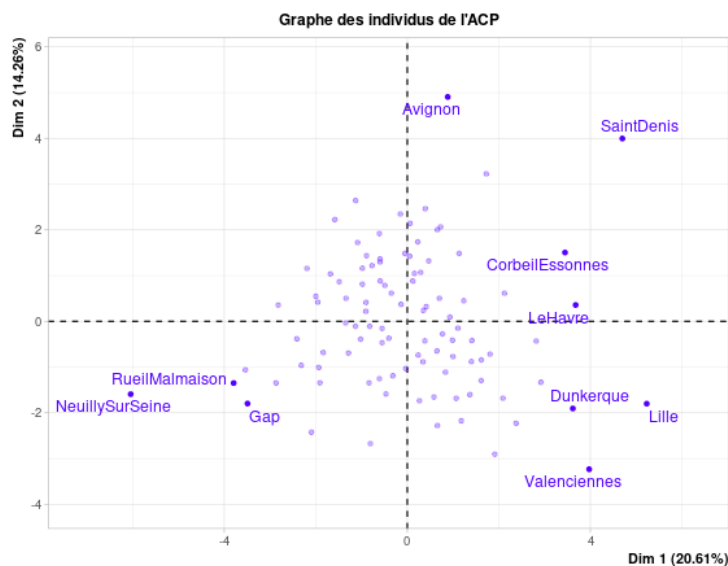
3.2 Thème Risques

Voici le graphique des valeurs propres :



On constate un décrochement entre la 3-ème et la 4-ème valeur propre. De plus, la variance cumulée sur les 3 premières valeurs propres est de 46,41%. Nous considérerons 3 axes.

Plan (1,2)



Voici les graphiques des individus avec les 10 plus gros contributeurs et celui des variables où sont indiquées les 8 variables les plus contributrices.

Du côté des variables, se dessinent deux faisceaux avec des thèmes propres. Un premier faisceau, qui nous parle de mortalité : *MortaliteGlobale* (CTR : 22,39%), *MortaliteAlcool* (14,92%), *MortaliteCancerPoumon* (13,95%) et *DecesInfarctus* (13,71%). *MortaliteGlobale* a d'ailleurs un COS2 assez élevé sur la première dimension : 0,74. Cette variable est donc assez bien corrélée avec la composante 1.

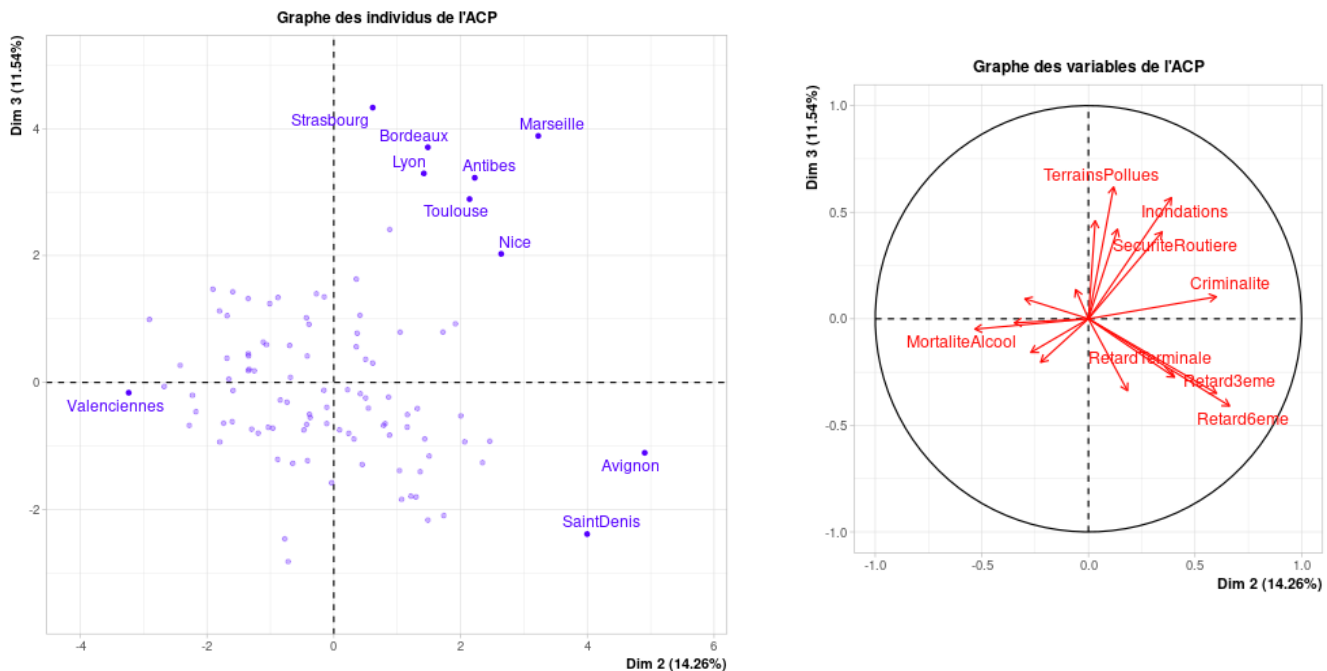
Trois villes de la même région pointent dans la même direction que les 4 variables précédemment citées : *Lille*

(CTR : 8,30% ; COS2 : 0,67), *Valenciennes* (CTR1 : 4,79% ; COS2 : 0,43) et *Dunkerque* (CTR1 : 3,97 ; COS2 : 0,41). Une autre ville est très contributrice au positionnement de l'axe 1, mais pointe dans la direction opposée : *Neuilly-Sur-Seine* (CTR1 : 11,08%).

Neuilly-Sur-Seine est aussi opposée au second faisceau de variables qui nous parle davantage de retard scolaire. Ces variables sont très contributrices au positionnement de l'axe 2. Il paraît malheureusement logique que ce retard scolaire dès la 6^{ème} (CTR2 : 19,21%), soit perpétué en 3^{ème} (CTR2 : 15,85%) puis en terminale (CTR2 : 7,15%). La variable *Criminalite* participe aussi très fortement au positionnement de l'axe 2 avec une contribution de 15,78%. Malgré de fortes contributions, l'étude des COS2 ne montre rien de très significatif.

Du côté des individus, Avignon tire une bonne partie son originalité de cet axe avec un COS2 égale à 0,67. Elle pointe malheureusement dans la même direction que la criminalité et le retard scolaire.

Plan (2,3)

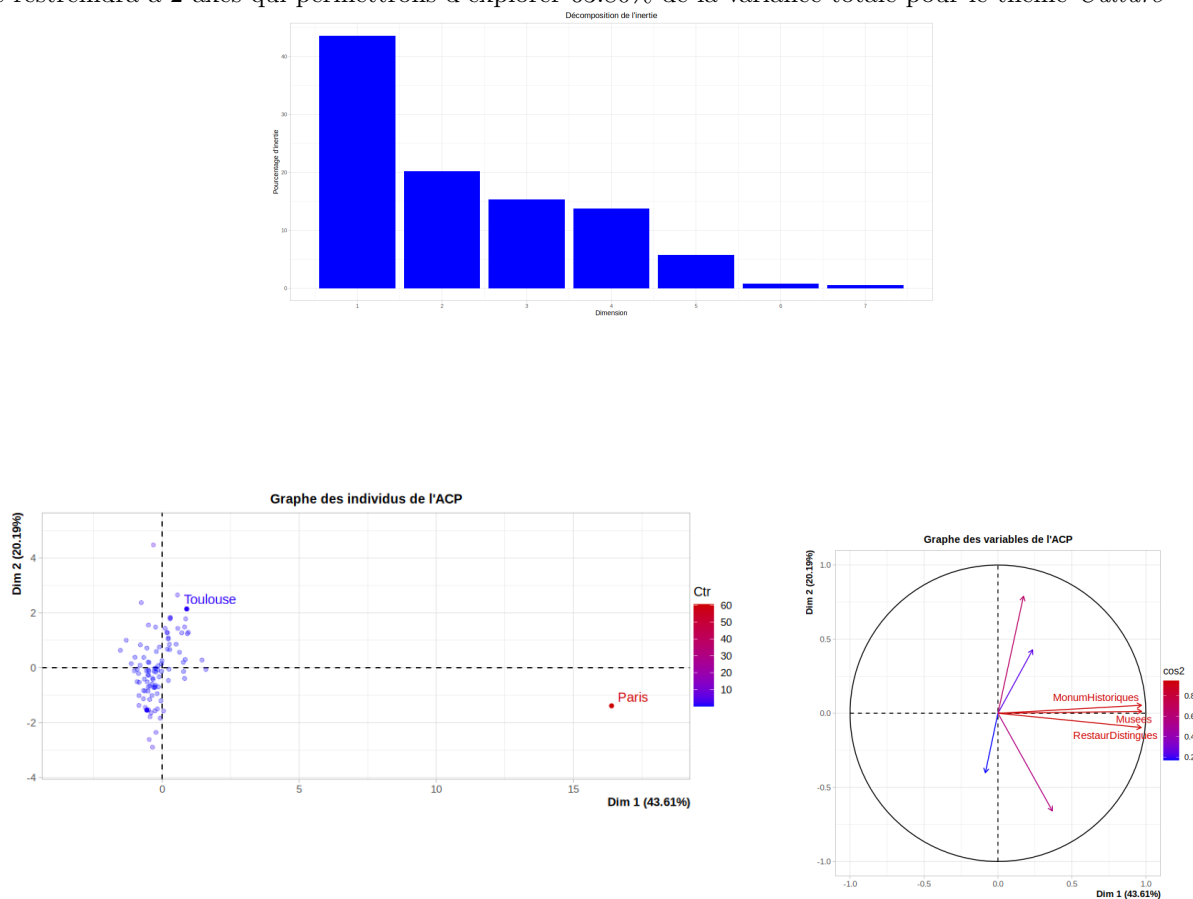


Sur ce plan, on distingue clairement un regroupement de plusieurs grandes villes de France qui contribuent au positionnement de l'axe 3 : *Strasbourg* (CTR3 : 10,16%) *Marseille* (CTR3 : 8,18%), *Bordeaux* (CTR3 : 7,44%), *Lyon* (CTR3 : 5,88%) et *Toulouse* (CTR3 : 4,52%). L'originalité de Bordeaux est d'ailleurs bien visible sur cet axe avec un COS2 de 0,69.

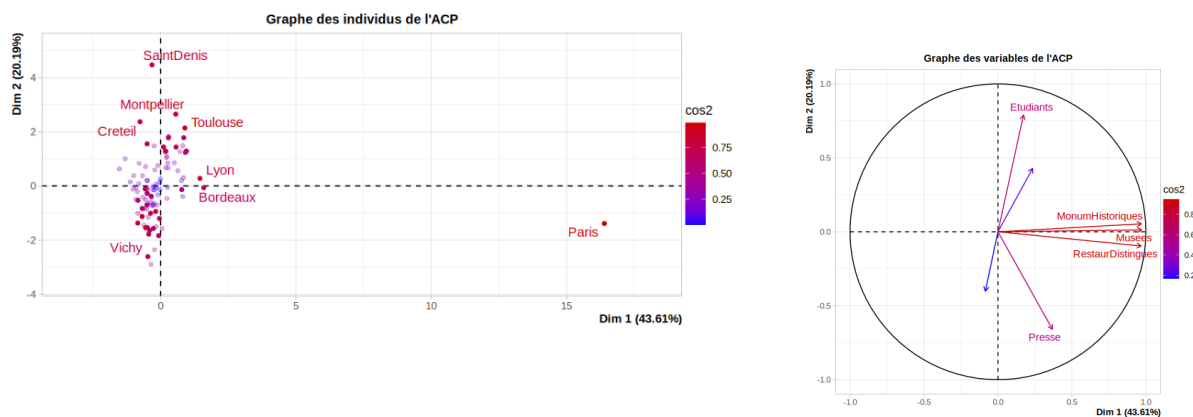
On retrouve parmi les variables les plus contributrices pointant dans la direction de ces villes : *TerrainPollue* (CTR3 : 20,71%), *Inondations* (CTR3 : 17,50%) et *UsinesRisques* : 11,45%. Toutefois il n'y pas de forte corrélation entre ces variables et l'axe 3, leur COS2 étant assez faible.

3.3 Thème Culture

On se restreindra à 2 axes qui permettront d'explorer 63.80% de la variance totale pour le thème *Culture*



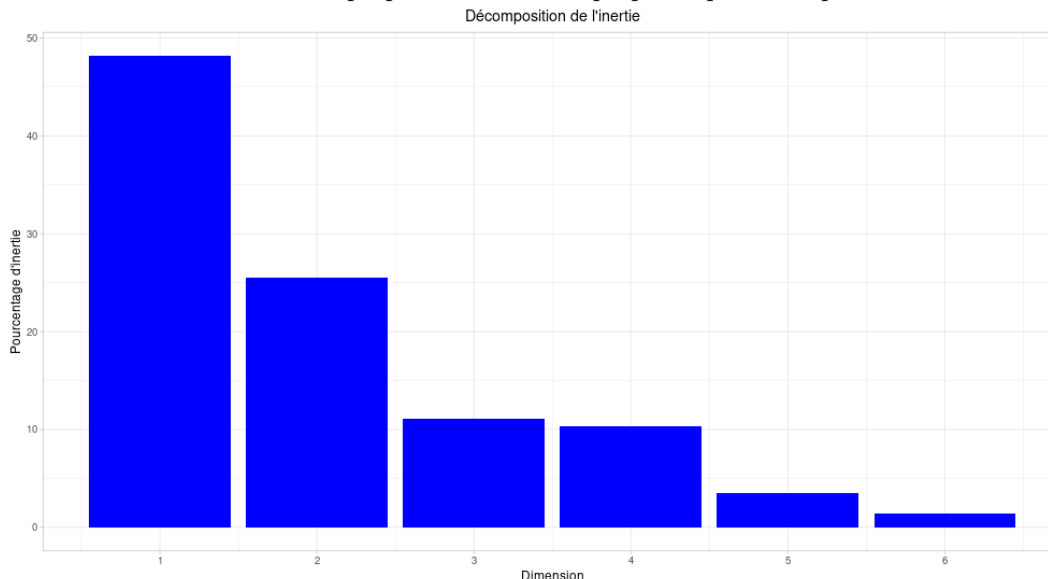
Dans cette première ACP du thème *Culture*, la place énorme qu'occupe Paris dans la détermination du plan permet difficilement d'étudier l'étalement vertical qui se dessine en élargissant notre tolérance à ($\text{COS2} > 0,5$) :



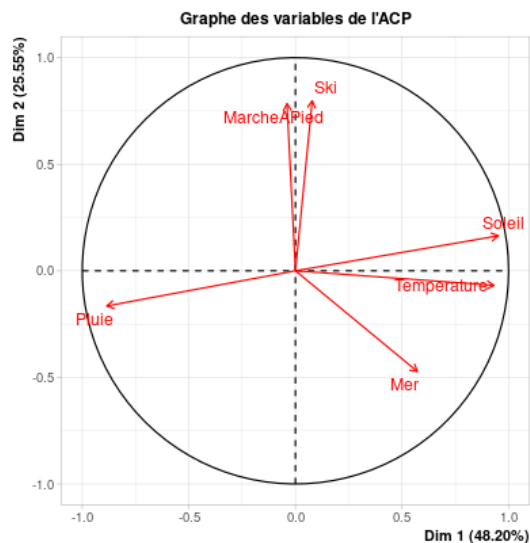
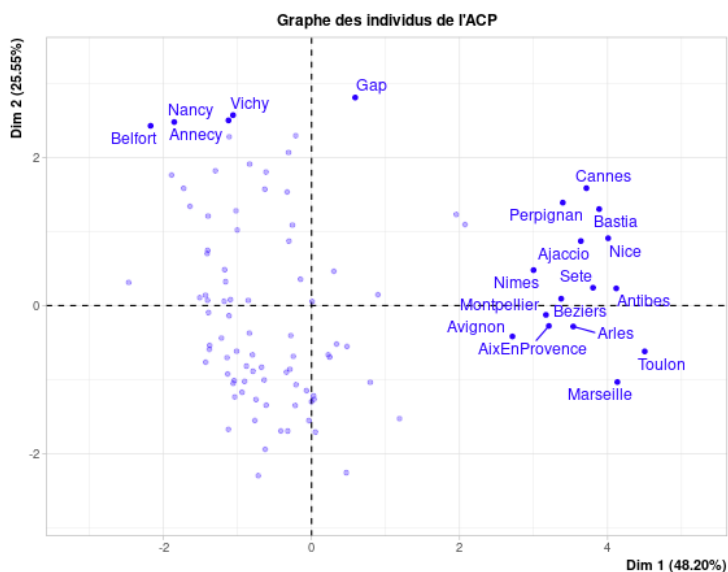
Cela pourra être mieux exploré dans l'ACP par rang qui gommara la singularité de Paris.

3.4 Thème Nature

Sur ce thème nous ne regarderons que les deux premiers axes. En effet ils capturent 73,75% de l'information. On observe un décrochement entre la valeur propre 2 et la valeur propre 3 qui est de plus inférieure à 1.



Plan (1,2)



Ici l'étude est assez simple. On remarque 3 variables qui ont positionnées l'axe 1 : *Soleil* (CTR : 31,33%), *Température* (CTR : 30,00%) et *Pluie* (CTR : 27,09%). La variable *Pluie* est logiquement négativement corrélée aux deux autres. Ces trois variables sont très corrélées avec cet axe. Leur COS2 est respectivement de 0,91 ; 0,87 et 0,78.

Sur le plan des individus, on voit un grand groupe de villes du Sud de la France dans la même direction que les variables *Soleil* et *Température*. On peut citer par exemple : *Toulon* (CTR1 : 7% ; COS2 : 0,96), *Marseille* (CTR1 : 5,90 ; COS2 : 0,92), *Antibes* (CTR1 : 5,86% ; COS2 : 0,88) et *Nice* (CTR1 : 5,55% ; COS2 : 0,89). Pour ces villes citées et celles du groupe des villes du Sud, l'originalité visible sur cet axe représente quasiment la totalité.

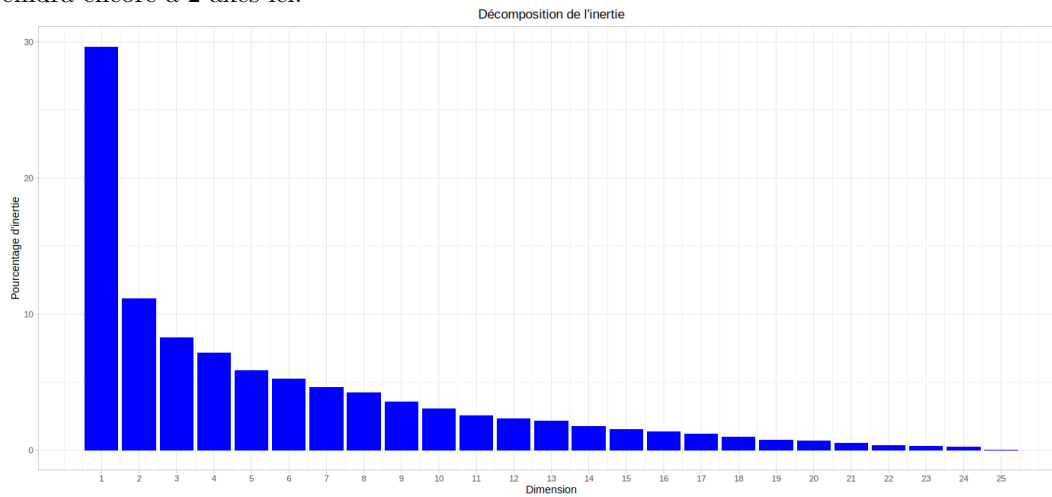
Pour l'axe 2, les variables *Ski* (CTR2 : 41,47% ; COS2 : 0,64) et *MarcheAPied* (CTR2 : 40,11% ; COS2 : 0,62) ont clairement positionné notre axe même si leur corrélation avec cet axe n'est pas si grande.

Gap (CTR2 : 5,14% ; COS2 : 0,76), *Vichy* (CTR2 : 4,31% , COS2 : 0,82), *Annecy* (CTR2 : 4,08% , COS2 : 0,74) et *Nancy* (CTR2 : 4,00%, COS2 : 0,60) ont contribué à positionner l'axe qui parle d'elles. Ces villes n'ont pas tout à fait le même profil, car si Gap et Annecy sont, de part leur positionnement géographique, des villes tournées vers le ski, il n'en est clairement pas question pour Vichy et Nancy où la variable *MarcheAPied* a du jouer un rôle important.

4 ACP de rang pour les différents thèmes

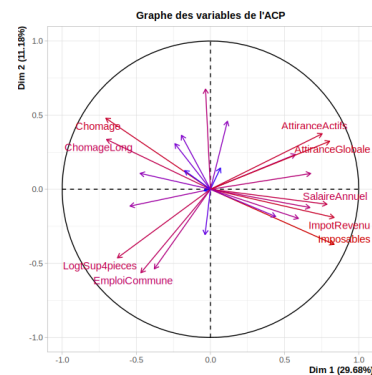
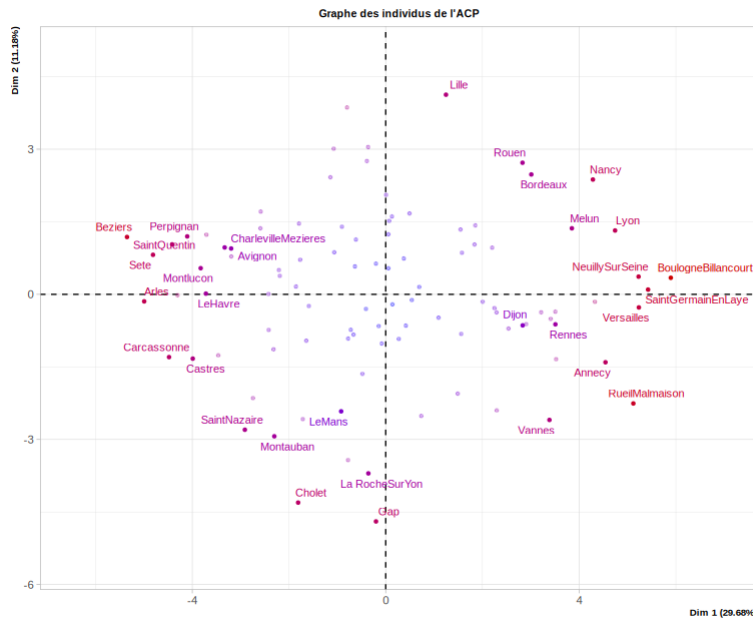
4.1 Thème Economie

On se restreindra encore a 2 axes ici:



Une ACP de rang pour ce thème permet une analyse plus fine, en réduisant drastiquement l'impact des individus très atypiques, on peut désormais observer d'autres disparités.

En se limitant à COS2 > 0,5, on peut désormais mieux explorer la structure des villes à faible richesse et attirance observées dans l'ACP normée :



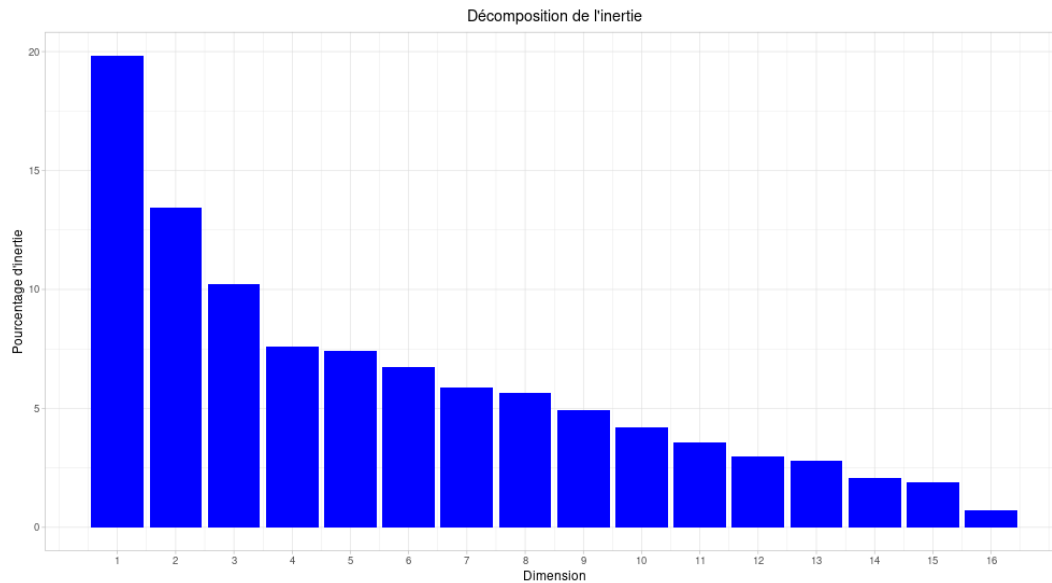
Dans le cadran inférieur gauche, on trouve des villes au haut taux d'emploi communal et avec un grand nombre de larges habitats (*LogtSup4Piece*). Parmi ces villes, on peut citer : *Montauban*, *Cholet* et *Gap*. L'opposition aux

grandes métropoles se fait moins par leur niveau de richesse que par leur faible capacité à attirer de nouveaux habitants. On y retrouve d'ailleurs les villes en bordure de zones rurales et d'anciens bassins ouvriers, nous confortant ainsi dans cette interprétation.

Dans le cadran supérieur gauche se trouvent par exemple : *Béziers, Perpignan, Avignon*. Les villes de ce cadran ont un haut taux de chômage et de chômage long. Leur opposition aux grandes métropoles se fait moins par leur capacité à attirer de nouveaux habitants que par leur faible niveau de richesse. On retrouve aussi des villes du littoral, touristique, avec une forte affluence de retraités, nous confortant dans cette interprétation.

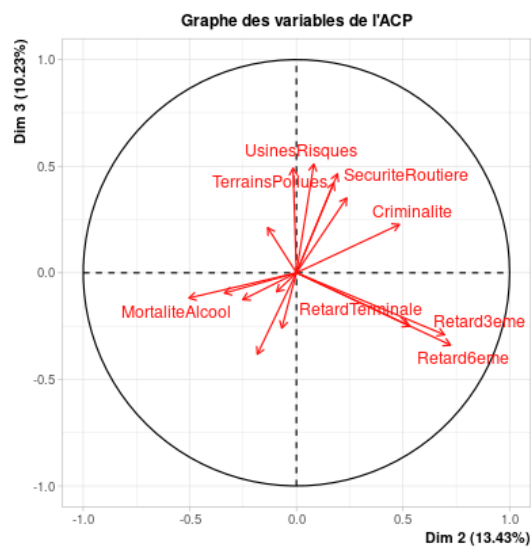
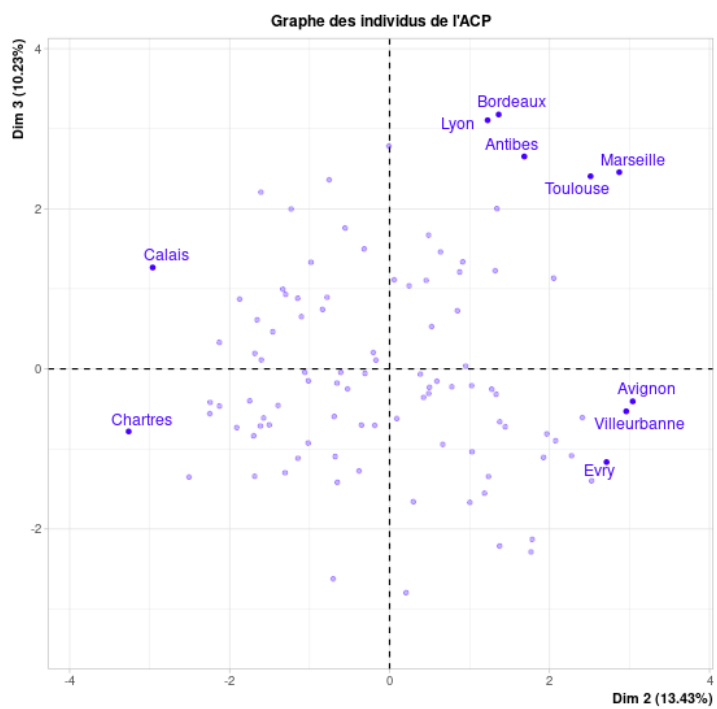
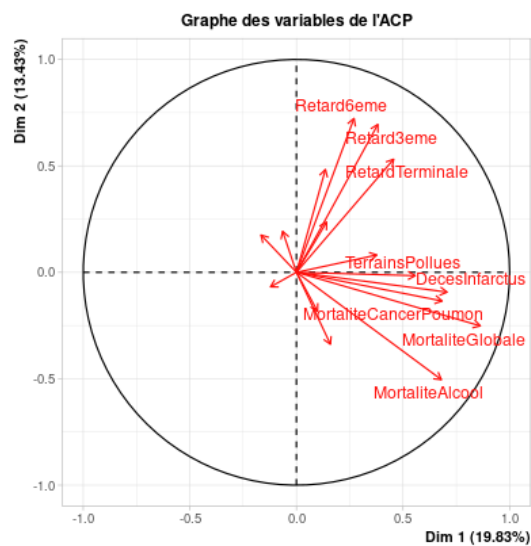
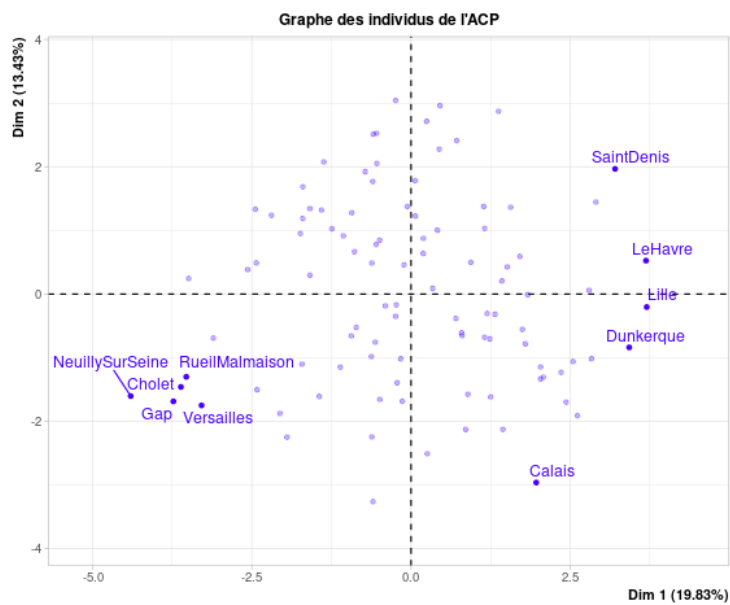
Bien que ces différences soit beaucoup moins tranchées que pour l'ACP normée, elle suggère de nouvelles dynamiques qui n'y étaient pas mises en évidence lorsque la riche région parisienne polarisait autant les axes.

4.2 Thème Risques



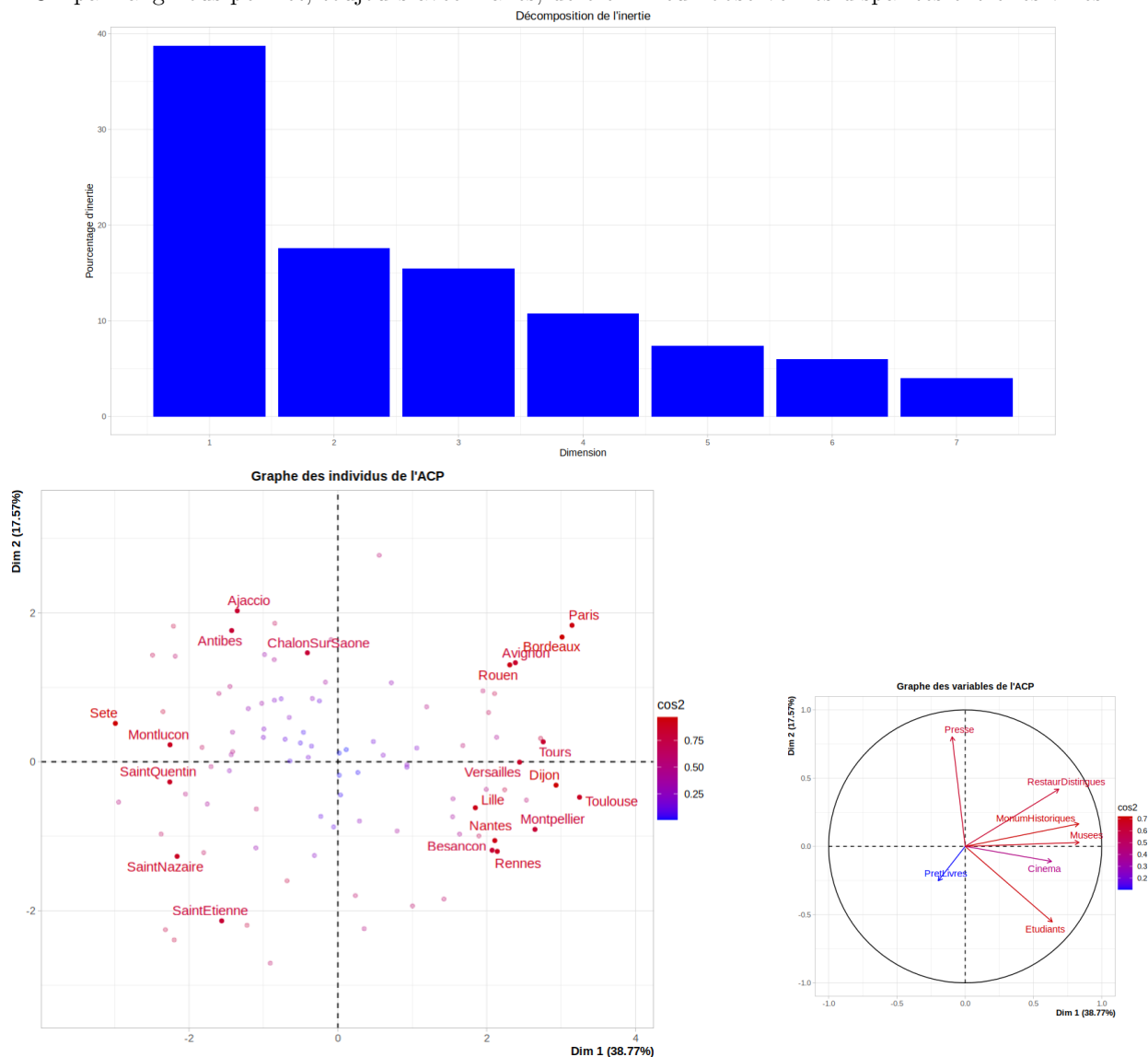
Il n'y a pas de changement notable dans l'étude des valeurs propres. On observe le même décrochement entre la 3-ème et la 4-ème valeur.

L'étude des plans (1,2) et (2,3) dont les graphiques se trouvent ci-dessous ne nous montrent rien de plus. En effet, on retrouve les mêmes faisceaux de variables avec peu ou prou les mêmes villes. Ceci s'explique par le fait qu'il n'y avait pas d'individu singulier par rapport aux variables du thème.



4.3 Thème Culture

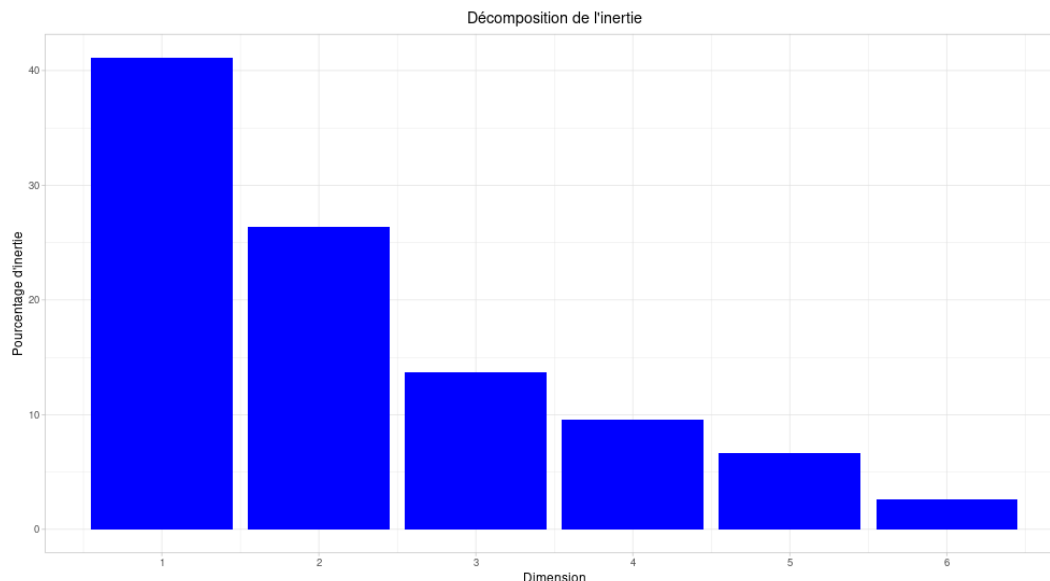
Ici, l'ACP par rang nous permet, toujours avec 2 axes, de bien mieux observer les disparités entre les villes :



En se limitant à $\text{COS2} > 0,9$, on peut désormais mieux comprendre l'étalement des groupes de villes déterminant la position des variables *Presse* et *Etudiants*.

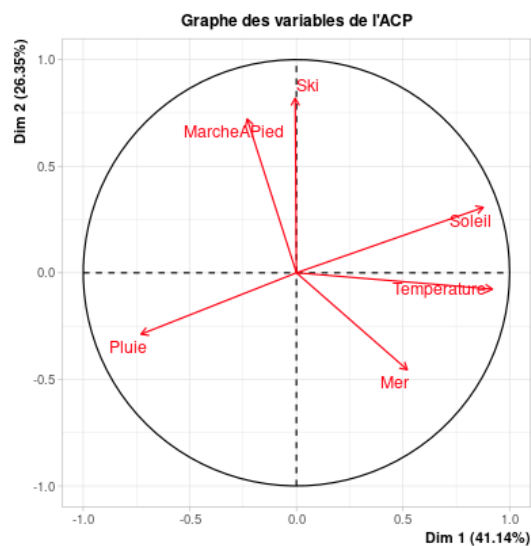
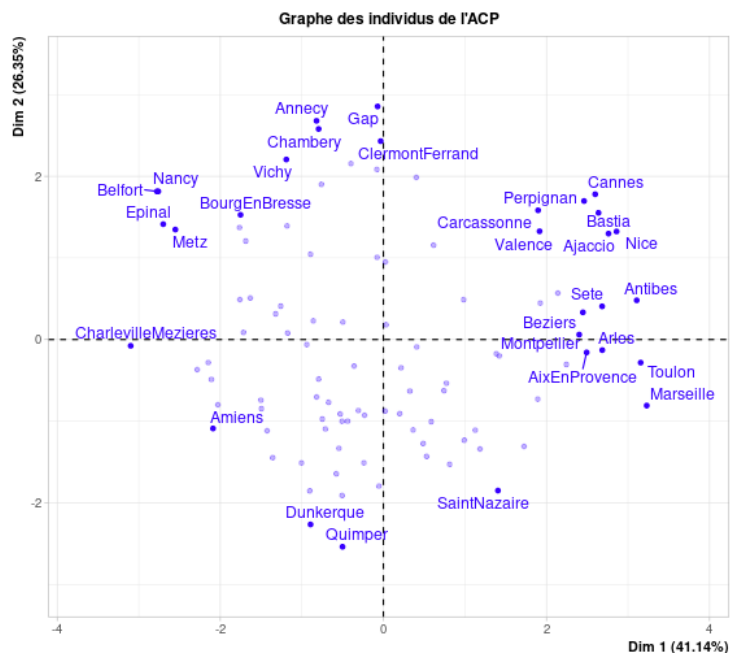
En effet, on découvre alors que les villes hors Paris ayant de hautes valeurs pour *Presse*, sont majoritairement les villes accueillant de grands événements culturels, (*Cannes*, *Angoulême*...), villes ne disposant souvent pas des grandes infrastructures pour attirer les milieux étudiants (grandes écoles, université). On remarque de plus que la majorité des variables et des grandes villes françaises sont globalement orientées dans la même direction.

4.4 Thème Nature



Bien que l'inertie capturée par les deux premiers axes soit inférieure à celles obtenue lors de l'ACP normée, nous ne garderons de nouveaux que deux axes. Ils capturent 67,49% de l'inertie.

Certes quelques différences de valeurs apparaissent mais on retrouve sensiblement les mêmes résultats que lors de l'ACP normée, notamment le groupe des villes du Sud et le positionnement opposé sur l'axe 2 pour d'un côté des villes typées montagne (Annecy, Gap, Chambéry, Clermont-Ferrand) et de l'autre trois villes de bord de mer (Dunkerque, Quimper, Saint-Nazaire).



A Code R

```
library(Factoshiny)
library(FactoMineR)

villesbrut = read.table("~/ADM/ADM-TP1/TP3/villes.csv", header = TRUE, row.names=1)

villes_rang = apply(villesbrut, MARGIN = 2 , FUN = rank)

villes_eco = villesbrut[,c('Chomage', "ChomageJeunes", "ChomageLong", "EvoluEmploiCree", "Activite
    "EmploiCommune", "DefaillEntreprise", "SalaireAnnuel", "ImpotRevenu", "Impo
    "MetreCarreAncien", "TaxeHabitation", "FoncierBati", "MetreCubeEau", "EvolD
    "AttiranceGlobale", "AttiranceActifs", "Proprietaires", "LogtSup4pieces", "
    "LogtConstruction")]

villes_ris = villesbrut[,c('Criminalite', 'EvolutionCrimes', 'SecuriteRoutiere', 'Inondations', 'T
    'MortaliteInfantile', 'MortaliteCancerPoumon', 'MortaliteAlcool', 'Dece
    'MortaliteGlobale', 'TailleClassesPrimaires', 'Retard6eme', 'Retard3eme

villes_nat = villesbrut[ , c('Mer', 'Ski', 'Soleil', 'Pluie', 'Temperature', 'MarcheAPied')]

villes_nat_rang = apply(villes_nat, MARGIN = 2 , FUN = rank)

acp_nat_rang = Factoshiny(villes_nat_rang)
```