

Comprendre les résultats de l'ACP en réalisant une étude sur les indicateurs de développements économiques sur les pays

AUTHOR

Axel-Cleris GAILLOTY

1 Généralités sur l'ACP

L'analyse en composantes principales (ACP) est une méthode statistique utilisée pour simplifier des données complexes en les transformant en un ensemble de variables moins nombreux, appelées composantes principales. Son objectif principal est de réduire la dimensionnalité des données tout en préservant autant que possible l'information originale.

1.1 Une composante

Une composante, dans le contexte de l'analyse en composantes principales (ACP), est une nouvelle variable qui est une combinaison linéaire des variables originales de manière à capturer la variation maximale des données.

Les composantes sont ordonnées en fonction de la quantité de variation qu'elles capturent, de sorte que la première composante représente la plus grande variation des données, la deuxième composante représente la deuxième plus grande variation, et ainsi de suite.

Les composantes permettent de réduire la dimensionnalité des données tout en conservant l'essentiel de l'information, ce qui facilite l'interprétation et l'analyse des données.

2 Le jeu de données tabulaire

Un jeu de données tabulaire est une structure de données organisée sous forme de tableau où les données sont disposées en lignes et en colonnes. Chaque ligne correspond à une observation ou un enregistrement distinct dans le jeu de données, tandis que chaque colonne représente une variable spécifique ou un attribut de ces observations.

Nous allons réaliser une ACP sur un jeu de données portant sur les indicateurs de développement dans les pays du monde. Le jeu de données contient 167 pays et 23 variables. Le jeu de données est consultable sur le [lien suivant](#)

Voici un aperçu du jeu de données.

	GDP_CAPITA	BALANCE	MANUFACTURE_IMPORTS	SERVICE_EMPLOYMENT	FOREIGN_INVESTMENT
Afghanistan	2082	5.7e+09	45	38	0.66
Albania	13498	3.4e+09	48	43	7.95
Algeria	11909	7.2e+09	66	59	0.84
American Samoa	22613	1.0e+09	66	55	-5.73
Andorra	22613	1.0e+09	68	55	-5.73
Angola	7149	-2.5e+10	59	42	-8.30
Argentina	23291	7.4e+08	79	78	2.23
Armenia	13019	1.7e+09	65	51	2.14
Australia	50169	-2.1e+10	73	78	4.29
Austria	56960	-1.4e+09	76	71	-6.28
Azerbaijan	14549	-9.8e+09	68	49	2.98
Bahrain	47486	1.1e+09	49	64	0.29
Bangladesh	5322	1.7e+10	66	40	0.75
Belarus	19428	2.5e+09	54	58	2.38
Belgium	52535	2.1e+09	70	78	-7.66

2.1 Les variables actives

Dans le contexte de l'ACP, une variable active est une variable qui est incluse dans le processus de calcul des composantes principales. Ces variables sont utilisées pour déterminer les directions (ou axes) dans lesquelles les données présentent le plus de variation.

Notre jeu de données est composée de 23 variables dont voici les noms : GDP_CAPITA, BALANCE, MANUFACTURE_IMPORTS, SERVICE_EMPLOYMENT, FOREIGN_INVESTMENT, FUEL_EXPORTS, GROSS_SAVINGS, INFLATION, EXPORT_TELECOM, LIFE_EXPECTANCY, ELECTRICITY_ACCESS, FERTILITY_RATE, HUMAN_CAPITAL_INDEX, FEMALE_LABOR_FORCE, POPULATION_65_ABOVE, EDUCATION_EXPENDITURE, URBAN_POPULATION, CO2_EMISSIONS, ARMED_FORCE, SURFACE_AREA, EXPORTS_GOOD_SERVICES, REGION, INCOME_GROUP. La liste des variables est décrite dans l'annexe.

2.2 Les variables illustratives

Les variables illustratives, également appelées variables supplémentaires ou variables de projection, sont des variables qui ne sont pas incluses dans le calcul des composantes principales lors d'une Analyse en Composantes Principales (ACP), mais qui sont projetées sur les graphiques de l'ACP pour fournir une perspective supplémentaire ou faciliter l'interprétation des résultats.

Ces variables peuvent être quantitatives ou qualitatives et représentent souvent des informations supplémentaires sur les individus ou les variables dans l'ensemble de données.

3 Tableau des valeurs propres

	eigenvalue	percentage of variance	cumulative percentage of variance
comp 1	5.57	27.85	28
comp 2	2.00	9.99	38
comp 3	1.46	7.30	45
comp 4	1.30	6.48	52
comp 5	1.19	5.95	58
comp 6	1.10	5.48	63
comp 7	1.06	5.31	68
comp 8	0.95	4.73	73
comp 9	0.91	4.56	78

	eigenvalue	percentage of variance	cumulative percentage of variance
comp 10	0.76	3.81	81
comp 11	0.75	3.73	85
comp 12	0.71	3.54	89
comp 13	0.63	3.13	92
comp 14	0.53	2.64	95
comp 15	0.30	1.50	96
comp 16	0.24	1.20	97
comp 17	0.19	0.94	98
comp 18	0.17	0.84	99
comp 19	0.11	0.55	100
comp 20	0.09	0.45	100

Dans ce tableau, chaque ligne représente une composante principale extraite de l'Analyse en Composantes Principales (ACP), avec les informations suivantes :

eigenvalue : C'est la valeur propre de chaque composante. Elle mesure la quantité de variance expliquée par cette composante. Dans ce cas, pour la première composante (comp 1), l'eigenvalue est de 5.57, ce qui signifie qu'elle explique 5.57 unités de variance dans les données.

percentage of variance : C'est le pourcentage de variance expliquée par chaque composante par rapport à la variance totale des données. Par exemple, pour la première composante (comp 1), elle explique 27.85% de la variance totale des données.

cumulative percentage of variance : C'est le pourcentage cumulatif de variance expliquée jusqu'à cette composante. Il indique la proportion totale de variance expliquée par les composantes jusqu'à cette composante particulière. Par exemple, pour la première composante (comp 1), le pourcentage cumulatif de variance est de 28%, ce qui signifie que les deux premières composantes (comp 1 et comp 2) expliquent ensemble 38% de la variance totale des données.

3.1 La règle de Kaiser

En ce qui concerne la règle de Kaiser, elle stipule que dans une ACP, seules les composantes dont les valeurs propres (eigenvalues) sont supérieures à 1 doivent être retenues. Cette règle est basée sur l'idée que les composantes avec des valeurs propres élevées capturent plus de variance que ce que l'on pourrait obtenir avec une seule variable, donc elles méritent d'être considérées comme significatives. Cependant, la règle de Kaiser est souvent utilisée comme une règle de pouce plutôt que comme une règle stricte, et d'autres critères peuvent être pris en compte pour décider du nombre de composantes à retenir.

4 Le tableau descriptif des dimensions

4.1 Dimension 1

\$Dim.1

Link between the variable and the continuous variables (R-square)

=====

	correlation	p.value
LIFE_EXPECTANCY	0.92	3.4e-70
HUMAN_CAPITAL_INDEX	0.87	1.3e-53
SERVICE_EMPLOYMENT	0.86	4.4e-51
ELECTRICITY_ACCESS	0.83	4.3e-44
GDP_CAPITA	0.81	3.3e-40
URBAN_POPULATION	0.79	3.5e-36
POPULATION_65_ABOVE	0.77	2.2e-34
EXPORTS_GOOD_SERVICES	0.50	8.9e-12
MANUFACTURE_IMPORTS	0.37	1.2e-06
EDUCATION_EXPENDITURE	0.23	2.9e-03
GROSS_SAVINGS	0.21	7.2e-03
EXPORT_TELECOM	0.19	1.4e-02
INFLATION	-0.34	9.4e-06
FERTILITY_RATE	-0.88	2.2e-54

Link between the variable and the categorical variable (1-way anova)

=====

	R2	p.value
INCOME_GROUP	0.79	6.8e-55
REGION	0.63	2.5e-32

Link between variable and the categories of the categorical variables

=====

	Estimate	p.value
INCOME_GROUP=High income	3.0	3.3e-29
REGION=Europe & Central Asia	1.4	6.2e-11
REGION=North America	2.5	3.8e-02
INCOME_GROUP=Lower middle income	-0.8	2.7e-06
INCOME_GROUP=Low income	-3.1	3.7e-20
REGION=Sub-Saharan Africa	-3.3	1.3e-29

Dans ce résultat d'une Analyse en Composantes Principales (ACP), la dimension 1 (Dim.1) est présentée avec les informations suivantes :

Link between the variable and the continuous variables (R-square) : Cela indique la corrélation entre chaque variable et la dimension 1 de l'ACP, ainsi que la significativité statistique de cette corrélation (p-valeur).

correlation : Il s'agit du coefficient de corrélation entre chaque variable et la dimension 1 de l'ACP. Par exemple, pour LIFE_EXPECTANCY, la corrélation est de 0.92, ce qui indique une forte corrélation positive entre LIFE_EXPECTANCY et la première dimension de l'ACP.

p.value : Il s'agit de la p-valeur associée à la corrélation entre chaque variable et la dimension 1. Une p-valeur faible (inférieure à un seuil spécifié, souvent 0.05) indique une corrélation statistiquement significative. Par exemple, pour LIFE_EXPECTANCY, la p-valeur est de 3.4e-70, ce qui est extrêmement faible, indiquant une corrélation très significative entre LIFE_EXPECTANCY et la première dimension de l'ACP.

Link between variable and the categories of the categorical variables : Cela indique la relation entre les variables continues et les différentes catégories des variables catégorielles. Dans ce cas, les catégories des variables catégorielles sont INCOME_GROUP et REGION.

Estimate : Il s'agit de l'estimation de l'effet moyen de chaque catégorie de variable catégorielle sur la variable continue. Par exemple, pour la catégorie "High income" de la variable INCOME_GROUP, l'estimation est de 3.0. Cela signifie que cette catégorie est associée à une augmentation moyenne de 3.0 unités dans la composante.

p.value : Il s'agit de la p-valeur associée à l'estimation de l'effet de chaque catégorie de variable catégorielle sur la variable continue. Une p-valeur faible (généralement inférieure à 0.05) indique une association statistiquement significative. Par exemple, pour la catégorie "High income" de la variable INCOME_GROUP, la p-valeur est de 3.3e-29, ce qui est extrêmement faible, indiquant une association très significative entre cette catégorie et la variable continue.

Dans cette première dimension de l'Analyse en Composantes Principales (ACP), une opposition notable émerge entre plusieurs groupes de variables :

1. **Variables liées au développement économique et social** (LIFE_EXPECTANCY, HUMAN_CAPITAL_INDEX, SERVICE_EMPLOYMENT, ELECTRICITY_ACCESS, GDP_CAPITA, URBAN_POPULATION, POPULATION_65_ABOVE) :
 - Ces variables sont fortement corrélées positivement dans cette dimension, ce qui suggère qu'elles partagent des caractéristiques communes associées à un développement économique et social avancé. Par exemple, une espérance de vie plus élevée, un meilleur capital humain, un accès à l'électricité et une urbanisation croissante sont souvent associés à des économies plus développées.
2. **Variables économiques plus spécifiques** (EXPORTS_GOOD_SERVICES, MANUFACTURE_IMPORTS, EDUCATION_EXPENDITURE, GROSS_SAVINGS, EXPORT_TELECOM, INFLATION) :
 - Ces variables ont une corrélation moins forte avec la première dimension mais montrent tout de même une association significative. Elles peuvent être considérées comme reflétant des aspects plus spécifiques du développement économique, tels que les échanges commerciaux, les investissements en éducation et les politiques monétaires.
3. **Variables démographiques et comportementales** (FERTILITY_RATE, INCOME_GROUP, REGION) :
 - Ces variables présentent une opposition claire par rapport aux variables économiques et sociales. Par exemple, le taux de fécondité (FERTILITY_RATE) est fortement corrélé négativement, indiquant une relation inverse avec le développement économique. De même, les différences régionales et de revenu sont marquées, avec des valeurs contrastées entre les catégories, telles que les pays à revenu élevé par rapport à ceux à revenu faible ou moyen, ou les régions géographiques comme l'Europe et l'Afrique subsaharienne.

Ainsi, cette première dimension met en évidence les tensions et les disparités qui existent entre les différents aspects du développement économique, social et démographique dans les données analysées.

Dans la section **Link between the variable and the continuous variables (R-square)**, les corrélations entre les variables initiales et la composante étudiée sont affichées dans l'ordre décroissant.

4.2 Dimension 2

\$Dim.2

Link between the variable and the continuous variables (R-square)

=====		
	correlation	p.value
FUEL_EXPORTS	0.66	4.7e-22
CO2_EMISSIONS	0.58	1.1e-16
GROSS_SAVINGS	0.39	1.7e-07
ARMED_FORCE	0.34	8.0e-06
ELECTRICITY_ACCESS	0.21	5.3e-03
URBAN_POPULATION	0.18	2.2e-02
MANUFACTURE_IMPORTS	-0.18	2.1e-02
BALANCE	-0.20	8.5e-03
EXPORT_TELECOM	-0.31	4.0e-05
POPULATION_65_ABOVE	-0.41	3.6e-08
FEMALE_LABOR_FORCE	-0.73	7.5e-29

Link between the variable and the categorical variable (1-way anova)

=====		
	R2	p.value
REGION	0.43	2.5e-17
INCOME_GROUP	0.09	1.5e-03

Link between variable and the categories of the categorical variables

=====		
	Estimate	p.value
REGION=Middle East & North Africa	2.44	5.1e-19
INCOME_GROUP=Lower middle income	0.47	1.9e-02
INCOME_GROUP=High income	-0.27	3.8e-02
REGION=Europe & Central Asia	-0.56	1.3e-02
INCOME_GROUP=Low income	-0.62	1.3e-02
REGION=Sub-Saharan Africa	-0.62	1.2e-02

Dans cette deuxième dimension de l'Analyse en Composantes Principales (ACP), plusieurs oppositions et relations sont observées :

1. Variables liées aux activités économiques et environnementales (FUEL_EXPORTS, CO2_EMISSIONS, GROSS_SAVINGS, ARMED_FORCE) :

- Les exportations de carburants (FUEL_EXPORTS) et les émissions de dioxyde de carbone (CO2_EMISSIONS) présentent une corrélation positive significative avec cette dimension. Cela suggère une relation entre les activités économiques basées sur les combustibles fossiles et les émissions de CO2, mettant en évidence une facette du développement économique liée à l'impact environnemental.
- Les épargnes brutes (GROSS_SAVINGS) et les effectifs militaires (ARMED_FORCE) montrent également une corrélation positive, mais moins forte. Ces variables peuvent être associées à des aspects de la politique économique et de sécurité nationale.

2. Variables démographiques et sociales (URBAN_POPULATION, POPULATION_65_ABOVE, FEMALE_LABOR_FORCE) :

- La proportion de population urbaine (URBAN_POPULATION) et la proportion de population âgée de 65 ans et plus (POPULATION_65_ABOVE) ont une corrélation négative significative avec cette dimension. Cela suggère une opposition entre les dynamiques urbaines et démographiques spécifiques analysées.
- Le pourcentage de main-d'œuvre féminine (FEMALE_LABOR_FORCE) est fortement corrélé négativement, indiquant une relation inverse avec cette dimension. Cette opposition peut refléter des différences dans les tendances de l'emploi entre les genres.

3. Variables liées aux échanges internationaux et à la catégorie économique (MANUFACTURE_IMPORTS, EXPORT_TELECOM, BALANCE, REGION, INCOME_GROUP) :

- Les importations de produits manufacturés (MANUFACTURE_IMPORTS) et les exportations de services de télécommunications (EXPORT_TELECOM) présentent une corrélation négative significative, suggérant une relation inverse entre ces activités et cette dimension.
- Les déséquilibres commerciaux (BALANCE) sont également négativement corrélés, ce qui peut indiquer une relation avec les flux commerciaux et les relations économiques internationales.
- L'analyse des variables catégorielles montre des différences marquées entre les régions géographiques et les groupes de revenu, mettant en évidence des oppositions et des disparités dans les contextes économiques et sociaux analysés. Par exemple, les régions du Moyen-Orient et d'Afrique du Nord sont fortement positivement associées à cette dimension, tandis que les régions d'Europe et d'Asie centrale sont négativement associées. De même, les groupes de revenu inférieurs sont négativement associés, tandis que le groupe de revenu élevé est légèrement négatif.

4.3 Dimension 3

\$Dim.3

Link between the variable and the continuous variables (R-square)

=====		
	correlation	p.value
SURFACE_AREA	0.71	4.3e-27
BALANCE	0.49	1.2e-11
EDUCATION_EXPENDITURE	0.37	6.3e-07
CO2_EMISSIONS	0.24	1.7e-03
ELECTRICITY_ACCESS	0.16	3.8e-02
GDP_CAPITA	-0.18	2.0e-02
ARMED_FORCE	-0.21	5.3e-03
EXPORTS_GOOD_SERVICES	-0.60	7.2e-18

Link between the variable and the categorical variable (1-way anova)

=====		
	R2	p.value
REGION	0.234	1.1e-07
INCOME_GROUP	0.047	4.8e-02

Link between variable and the categories of the categorical variables

=====		
	Estimate	p.value
REGION=North America	3.2	7e-09

Dans cette troisième dimension de l'Analyse en Composantes Principales (ACP), plusieurs relations et contrastes sont observés :

1. **Variables liées à la géographie et à l'environnement** (SURFACE_AREA, CO2_EMISSIONS, ELECTRICITY_ACCESS) :
 - La superficie terrestre (SURFACE_AREA) et les émissions de dioxyde de carbone (CO2_EMISSIONS) présentent une corrélation positive significative avec cette dimension. Cela suggère que des pays ou des régions avec une plus grande superficie et des émissions de CO2 plus élevées peuvent être associés à certaines caractéristiques géographiques ou environnementales spécifiques.

- L'accès à l'électricité (ELECTRICITY_ACCESS) montre également une corrélation positive, bien que moins forte, indiquant une relation avec les infrastructures énergétiques et le développement économique.
2. **Variables économiques et sociales** (BALANCE, EDUCATION_EXPENDITURE, GDP_CAPITA, ARMED_FORCE, EXPORTS_GOOD_SERVICES) :
- Le solde commercial (BALANCE) et les dépenses en éducation (EDUCATION_EXPENDITURE) présentent des corrélations positives significatives, suggérant des relations avec l'activité économique et sociale.
 - Le PIB par habitant (GDP_CAPITA) montre une corrélation négative avec cette dimension, indiquant une relation inverse entre le niveau de richesse et cette dimension spécifique de l'ACP.
 - Les effectifs militaires (ARMED_FORCE) présentent également une corrélation négative, indiquant une relation inverse entre la force militaire et cette dimension.
 - Les exportations de biens et services (EXPORTS_GOOD_SERVICES) présentent une forte corrélation négative, ce qui peut refléter des relations commerciales ou économiques spécifiques.
3. **Variables catégorielles** (REGION, INCOME_GROUP) :
- Les régions géographiques ont une influence significative sur cette dimension, avec des variations marquées entre les régions. Par exemple, l'Amérique du Nord est fortement associée à cette dimension.
 - Les groupes de revenu ont également une influence, bien que moins prononcée que les régions géographiques. Par exemple, les groupes de revenu les plus élevés ont tendance à être légèrement négativement associés à cette dimension, tandis que les groupes de revenu plus bas ont des associations plus mitigées.

4.4 Dimension 4

\$Dim.4

Link between the variable and the continuous variables (R-square)

=====		
	correlation	p.value
GROSS_SAVINGS	0.44	2.2e-09
MANUFACTURE_IMPORTS	0.41	5.0e-08
FUEL_EXPORTS	0.40	8.4e-08
INFLATION	0.34	6.6e-06
SURFACE_AREA	0.34	8.6e-06
GDP_CAPITA	0.32	1.9e-05
FERTILITY_RATE	0.21	7.2e-03

FOREIGN_INVESTMENT	0.18	2.3e-02
EXPORTS_GOOD_SERVICES	0.16	3.4e-02
ELECTRICITY_ACCESS	-0.27	5.3e-04
CO2_EMISSIONS	-0.27	4.9e-04
ARMED_FORCE	-0.37	9.7e-07

Link between the variable and the categorical variable (1-way anova)

=====

	R2	p.value
INCOME_GROUP	0.11	0.00029
REGION	0.10	0.00762

Link between variable and the categories of the categorical variables

=====

	Estimate	p.value
INCOME_GROUP=Low income	0.51	0.0115
INCOME_GROUP=High income	0.27	0.0124
REGION=Sub-Saharan Africa	0.28	0.0177
REGION=North America	1.24	0.0432
INCOME_GROUP=Lower middle income	-0.33	0.0457
REGION=South Asia	-0.95	0.0385
INCOME_GROUP=Upper middle income	-0.46	0.0055

Dans cette quatrième dimension de l'Analyse en Composantes Principales (ACP), nous observons plusieurs relations contrastées :

1. **Variables économiques et environnementales** (GROSS_SAVINGS, MANUFACTURE_IMPORTS, FUEL_EXPORTS, INFLATION, SURFACE_AREA, GDP_CAPITA, CO2_EMISSIONS, FOREIGN_INVESTMENT) :

- Les économies avec un fort taux d'épargne brute (GROSS_SAVINGS), des importations importantes de produits manufacturés (MANUFACTURE_IMPORTS) et des exportations de carburants (FUEL_EXPORTS) sont positivement corrélées avec cette dimension. Cela peut suggérer une relation entre les économies plus tournées vers l'exportation et celles axées sur l'industrie manufacturière et l'épargne.
- L'inflation et la superficie terrestre (SURFACE_AREA) montrent également une corrélation positive, bien que moins forte. Cela peut refléter des relations économiques et spatiales complexes.
- En revanche, des taux de fécondité élevés (FERTILITY_RATE) sont associés négativement, indiquant une relation inverse avec cette dimension.

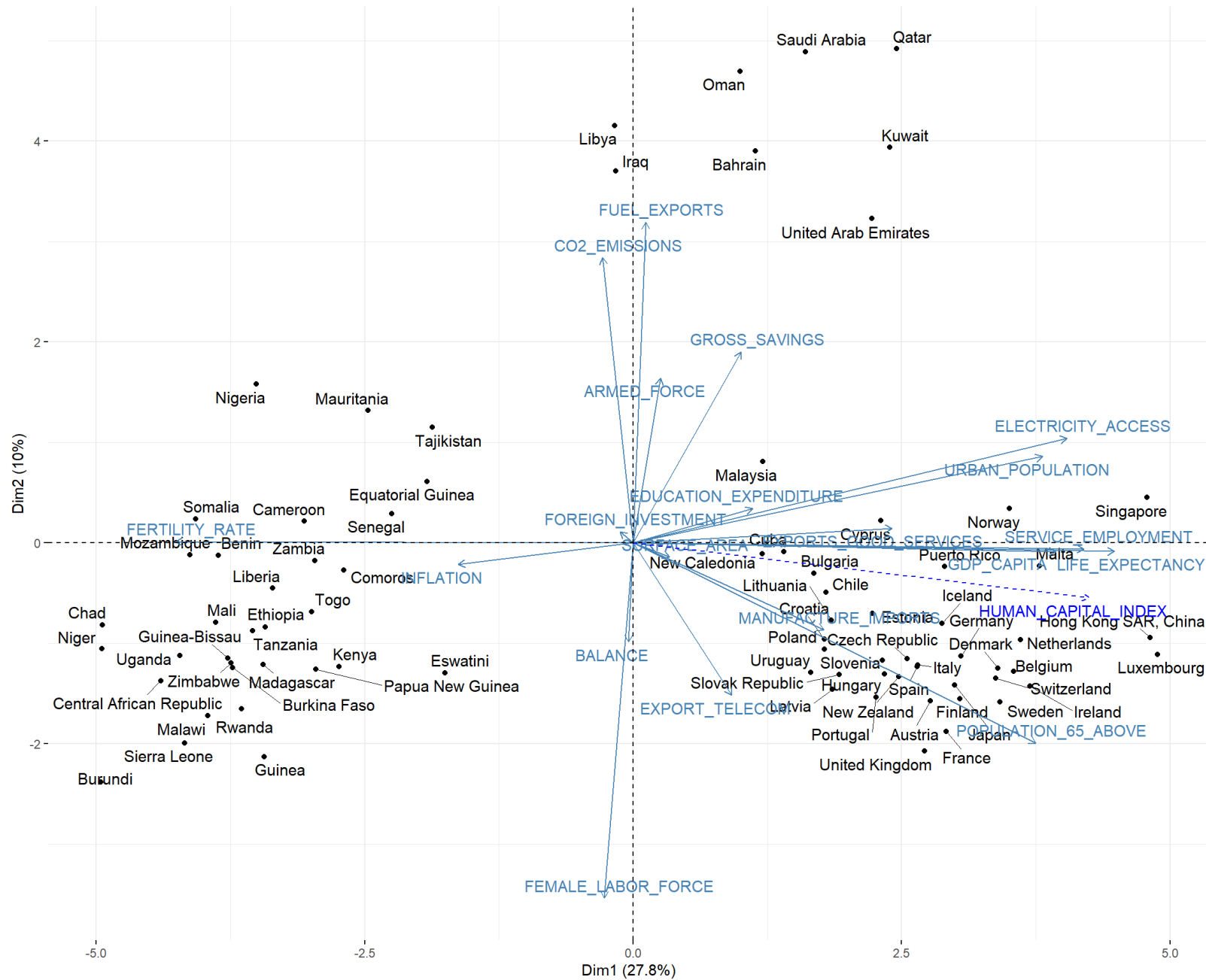
- Les émissions de CO2 (CO2_EMISSIONS) et les dépenses en électricité (ELECTRICITY_ACCESS) montrent une corrélation négative, suggérant une relation inverse avec cette dimension, ce qui peut être lié à des politiques environnementales ou à des schémas de consommation spécifiques.
2. **Variables sociodémographiques et militaires** (ARMED_FORCE, REGION, INCOME_GROUP) :
- La force militaire (ARMED_FORCE) présente une corrélation négative significative, ce qui peut indiquer une relation entre la répartition des ressources militaires et cette dimension spécifique de l'ACP.
 - Les variables catégorielles montrent des contrastes régionaux et économiques significatifs. Par exemple, les régions d'Amérique du Nord sont fortement associées positivement à cette dimension, tandis que les régions d'Asie du Sud et d'Afrique subsaharienne ont des associations négatives. De même, les groupes de revenu les plus élevés ont des associations positives, tandis que les groupes de revenu les plus bas ont des associations négatives avec cette dimension. Ces relations mettent en évidence les disparités économiques et géographiques dans le contexte analysé.

5 Le plan factoriel

5.1 Le plan factoriel des variables et individus

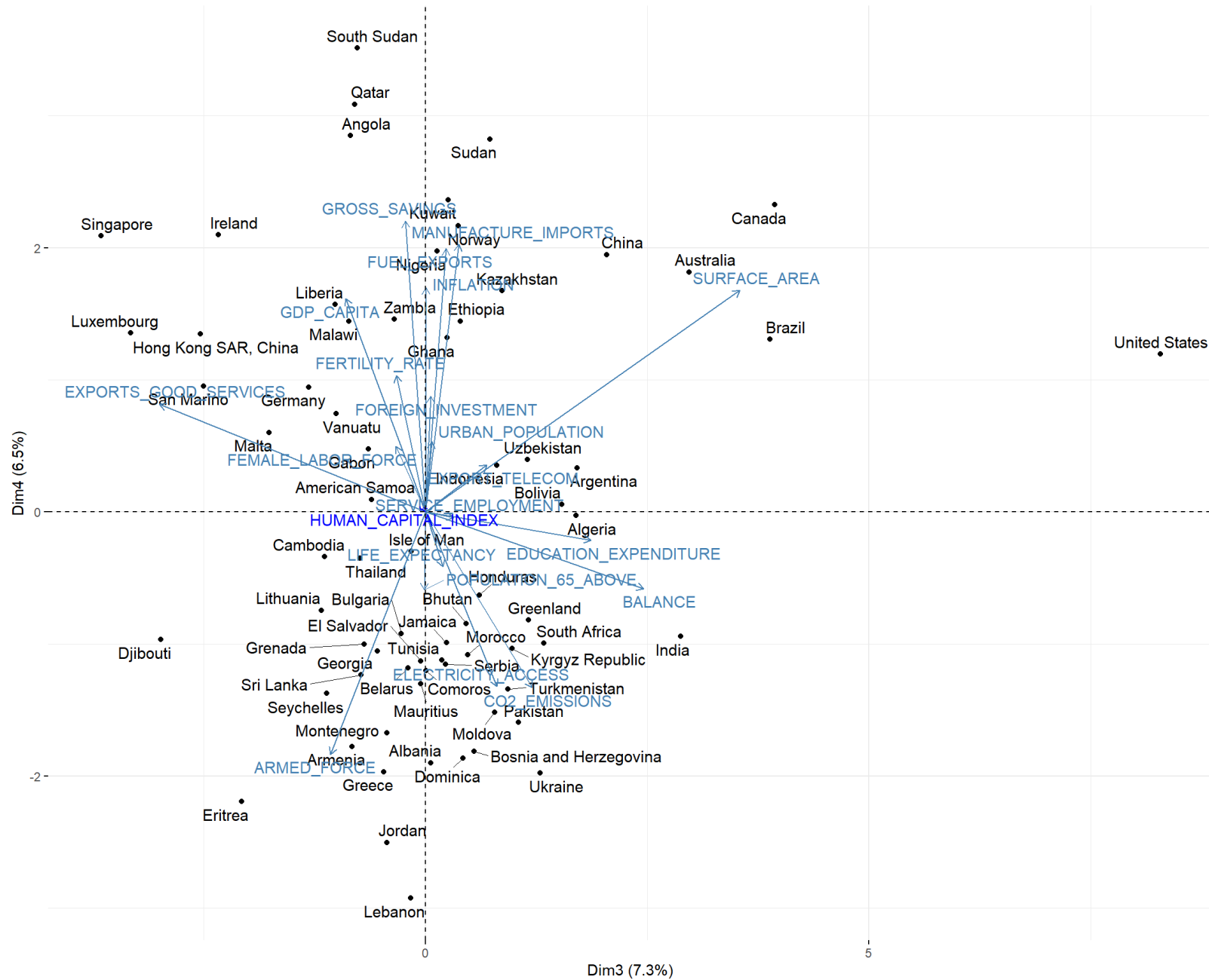
5.1.1 Premier plan factoriel (composante 1 et composante 2)

PCA - Biplot



5.1.2 Deuxième plan factoriel (composante 2 et composante 3)

PCA - Biplot



6 Les coordonnées

6.1 Les coordonnées des variables

	Dim.1	Dim.2	Dim.3	Dim.4
GDP_CAPITA	0.81	-0.02	-0.18	0.32
BALANCE	-0.01	-0.20	0.49	-0.12
MANUFACTURE_IMPORTS	0.37	-0.18	0.08	0.41
SERVICE_EMPLOYMENT	0.86	-0.01	0.06	-0.01
FOREIGN_INVESTMENT	-0.02	0.02	0.01	0.18
FUEL_EXPORTS	0.02	0.66	0.05	0.40
GROSS_SAVINGS	0.21	0.39	-0.05	0.44
INFLATION	-0.34	-0.04	0.00	0.34
EXPORT_TELECOM	0.19	-0.31	0.14	0.07
LIFE_EXPECTANCY	0.92	-0.02	0.04	-0.08
ELECTRICITY_ACCESS	0.83	0.21	0.16	-0.27
FERTILITY_RATE	-0.88	0.00	-0.07	0.21
FEMALE_LABOR_FORCE	-0.06	-0.73	-0.07	0.10
POPULATION_65_ABOVE	0.77	-0.41	0.00	-0.12
EDUCATION_EXPENDITURE	0.23	0.07	0.37	-0.04
URBAN_POPULATION	0.79	0.18	0.01	0.11
CO2_EMISSIONS	-0.06	0.58	0.24	-0.27
ARMED_FORCE	0.05	0.34	-0.21	-0.37
SURFACE_AREA	0.07	-0.03	0.71	0.34
EXPORTS_GOOD_SERVICES	0.50	0.03	-0.60	0.16

6.2 Les coordonnées des individus

	Dim.1	Dim.2	Dim.3	Dim.4
Afghanistan	-2.49	1.38	-0.33	-1.52
Albania	0.17	-0.45	0.06	-1.90
Algeria	0.58	2.53	1.70	-0.03
American Samoa	0.52	0.16	-0.61	0.09
Andorra	0.57	-0.53	-0.59	-0.06
Angola	-3.39	1.39	-0.85	2.85
Argentina	1.61	-1.20	1.71	0.33
Armenia	0.28	-0.07	-0.83	-1.78
Australia	2.72	-0.63	2.97	1.81
Austria	2.77	-1.57	-0.31	0.55
Azerbaijan	0.06	2.13	-0.65	1.12
Bahrain	1.14	3.90	-1.09	-0.37
Bangladesh	-0.85	0.48	0.74	0.16
Belarus	1.28	1.03	-0.19	-1.18
Belgium	3.55	-1.28	-0.35	0.50

Pour interpréter les positions des pays sur chaque dimension, nous examinons leurs coordonnées sur les différentes dimensions de l'ACP en relation avec les corrélations des variables avec chaque dimension.

- Afghanistan se positionne négativement sur la première dimension, ce qui suggère un niveau de vie plus faible, avec une espérance de vie inférieure, un capital humain limité et une fécondité élevée. Sur la deuxième dimension, il se situe positivement, ce qui peut indiquer une implication dans les exportations de carburant et un taux de fécondité élevé.
- L'Algérie se positionne positivement sur la première et la deuxième dimension, ce qui peut indiquer un niveau relativement élevé de développement économique, mais peut-être avec des défis en matière de fécondité.

- L'Argentine se positionne positivement sur la première et la troisième dimension, ce qui suggère un PIB par habitant élevé et peut-être une forte participation des femmes dans la population active, ainsi qu'une préoccupation environnementale, comme indiqué par les émissions de CO2.
- L'Arménie se positionne positivement sur la première dimension, suggérant un niveau de vie modéré, et négativement sur la troisième dimension, indiquant une faible préoccupation environnementale.
- L'Australie se positionne positivement sur toutes les dimensions, ce qui suggère un niveau de développement élevé, une urbanisation importante et une préoccupation environnementale, mais aussi un faible taux de fécondité.
- L'Autriche se positionne positivement sur la première dimension, indiquant un niveau de vie élevé, mais négativement sur la deuxième et la troisième dimension, suggérant une participation relativement faible aux exportations de carburant et une préoccupation environnementale moindre.
- L'Azerbaïdjan se positionne positivement sur la deuxième dimension, suggérant une implication dans les exportations de carburant, et positivement sur la quatrième dimension, indiquant peut-être des investissements étrangers élevés.

7 Les indicateurs de qualité de représentation et de contribution

7.1 Le tableau des cosinus carré des variables

	Dim.1	Dim.2	Dim.3	Dim.4
GDP_CAPITA	0.66	0.00	0.03	0.11
BALANCE	0.00	0.04	0.24	0.01
MANUFACTURE_IMPORTS	0.13	0.03	0.01	0.17
SERVICE_EMPLOYMENT	0.75	0.00	0.00	0.00
FOREIGN_INVESTMENT	0.00	0.00	0.00	0.03
FUEL_EXPORTS	0.00	0.43	0.00	0.16

	Dim.1	Dim.2	Dim.3	Dim.4
GROSS_SAVINGS	0.04	0.15	0.00	0.20
INFLATION	0.11	0.00	0.00	0.12
EXPORT_TELECOM	0.04	0.10	0.02	0.01
LIFE_EXPECTANCY	0.85	0.00	0.00	0.01
ELECTRICITY_ACCESS	0.69	0.05	0.03	0.07
FERTILITY_RATE	0.77	0.00	0.00	0.04
FEMALE_LABOR_FORCE	0.00	0.53	0.00	0.01
POPULATION_65_ABOVE	0.60	0.17	0.00	0.01
EDUCATION_EXPENDITURE	0.05	0.00	0.14	0.00
URBAN_POPULATION	0.62	0.03	0.00	0.01
CO2_EMISSIONS	0.00	0.34	0.06	0.07
ARMED_FORCE	0.00	0.11	0.05	0.14
SURFACE_AREA	0.00	0.00	0.51	0.11
EXPORTS_GOOD_SERVICES	0.25	0.00	0.36	0.03

Vous avez raison, le cosinus carré ne peut pas être négatif. Reprenons donc l'interprétation en corrigeant cette erreur :

- Pour la première dimension (Dim.1), les variables telles que l'espérance de vie, le PIB par habitant, l'indice du capital humain et l'emploi dans les services ont des valeurs de cosinus carré proches de 1. Cela signifie qu'elles sont fortement corrélées avec cette dimension, indiquant que la première dimension capture principalement des aspects liés au niveau de développement économique et social des pays.
- Pour la deuxième dimension (Dim.2), les variables telles que les exportations de carburant, les émissions de CO2 et les forces armées ont des valeurs de cosinus carré élevées, indiquant une forte corrélation avec cette dimension. Cela suggère que la deuxième dimension peut être associée à des aspects tels que l'impact environnemental et la puissance militaire des pays.

- Pour la troisième dimension (Dim.3), les variables telles que la superficie, la balance commerciale et les dépenses d'éducation ont des valeurs de cosinus carré élevées, indiquant une forte corrélation avec cette dimension. Cela suggère que la troisième dimension peut représenter des aspects liés à la taille du pays, à son équilibre commercial et à son investissement dans l'éducation.
- Pour la quatrième dimension (Dim.4), les variables telles que l'épargne brute, les importations de produits manufacturés et les émissions de CO2 ont des valeurs de cosinus carré élevées, suggérant une forte corrélation avec cette dimension. Cela peut indiquer que la quatrième dimension capture des aspects liés aux politiques économiques et environnementales des pays.

7.2 Le tableau des cosinus carré des individus

	Dim.1	Dim.2	Dim.3	Dim.4
Afghanistan	0.25	0.08	0.00	0.09
Albania	0.00	0.02	0.00	0.40
Algeria	0.02	0.35	0.16	0.00
American Samoa	0.14	0.01	0.18	0.00
Andorra	0.08	0.07	0.09	0.00
Angola	0.33	0.06	0.02	0.23
Argentina	0.21	0.12	0.24	0.01
Armenia	0.01	0.00	0.08	0.35
Australia	0.24	0.01	0.29	0.11
Austria	0.61	0.20	0.01	0.02

7.3 Le tableau des contributions des variables

	Dim.1	Dim.2	Dim.3	Dim.4
GDP_CAPITA	11.80	0.02	2.23	8.11
BALANCE	0.00	2.06	16.71	1.05
MANUFACTURE_IMPORTS	2.40	1.59	0.39	12.74
SERVICE_EMPLOYMENT	13.41	0.01	0.28	0.01
FOREIGN_INVESTMENT	0.01	0.03	0.01	2.39
FUEL_EXPORTS	0.01	21.64	0.15	12.35
GROSS_SAVINGS	0.77	7.67	0.14	15.07
INFLATION	2.02	0.10	0.00	8.96
EXPORT_TELECOM	0.64	4.88	1.32	0.40
LIFE_EXPECTANCY	15.28	0.02	0.10	0.54
ELECTRICITY_ACCESS	12.43	2.31	1.78	5.43
FERTILITY_RATE	13.81	0.00	0.30	3.31
FEMALE_LABOR_FORCE	0.06	26.52	0.30	0.76
POPULATION_65_ABOVE	10.72	8.43	0.00	1.08
EDUCATION_EXPENDITURE	0.94	0.25	9.59	0.14
URBAN_POPULATION	11.07	1.57	0.01	0.88
CO2_EMISSIONS	0.06	17.11	3.98	5.50
ARMED_FORCE	0.05	5.71	3.16	10.46
SURFACE_AREA	0.09	0.04	34.69	8.75
EXPORTS_GOOD_SERVICES	4.43	0.05	24.85	2.07

Dimension 1 (Dim.1) : - Le PIB par habitant (GDP_CAPITA) contribue de manière significative à la première dimension de l'ACP, indiquant son importance dans la variation des données dans cette dimension. - Les variables telles que l'espérance de vie (LIFE_EXPECTANCY), l'accès à l'électricité (ELECTRICITY_ACCESS), et le pourcentage de la population urbaine (URBAN_POPULATION) ont également une contribution notable à cette dimension. - Les variables comme le taux de

fécondité (FERTILITY_RATE), l'inflation (INFLATION), et le pourcentage de la population de plus de 65 ans (POPULATION_65_ABOVE) ont des contributions plus faibles mais non négligeables à cette dimension.

Dimension 2 (Dim.2) : - La force de travail féminine (FEMALE_LABOR_FORCE) est la variable qui contribue le plus à la deuxième dimension de l'ACP, suggérant son rôle important dans la variation des données le long de cette dimension. - Les exportations de carburant (FUEL_EXPORTS), les émissions de CO2 (CO2_EMISSIONS), et les importations de produits manufacturés (MANUFACTURE_IMPORTS) ont également des contributions significatives à cette dimension. - D'autres variables telles que le solde commercial (BALANCE), les économies brutes (GROSS_SAVINGS), et l'accès à l'électricité (ELECTRICITY_ACCESS) contribuent également à cette dimension, bien que dans une moindre mesure.

Dimension 3 (Dim.3) : - La surface terrestre (SURFACE_AREA) est la variable qui contribue le plus à la troisième dimension de l'ACP, indiquant son rôle prépondérant dans la variation des données dans cette dimension. - Les variables telles que l'éducation (EDUCATION_EXPENDITURE), les exportations de services (EXPORTS_GOOD_SERVICES), et les forces armées (ARMED_FORCE) ont également des contributions significatives à cette dimension. - D'autres variables comme le taux de fécondité (FERTILITY_RATE), les émissions de CO2 (CO2_EMISSIONS), et l'accès à l'électricité (ELECTRICITY_ACCESS) ont des contributions moins importantes mais néanmoins discernables à cette dimension.

Dimension 4 (Dim.4) : - Les importations de produits manufacturés (MANUFACTURE_IMPORTS) ont la contribution la plus importante à la quatrième dimension de l'ACP, indiquant leur rôle prédominant dans la variation des données dans cette dimension. - Les variables telles que les émissions de CO2 (CO2_EMISSIONS), les forces armées (ARMED_FORCE), et les économies brutes (GROSS_SAVINGS) ont également des contributions significatives à cette dimension. - D'autres variables comme le PIB par habitant (GDP_CAPITA) et l'accès à l'électricité (ELECTRICITY_ACCESS) ont des contributions moins importantes mais non négligeables à cette dimension.

7.4 Le tableau des contributions des individus

	Dim.1	Dim.2	Dim.3	Dim.4
Afghanistan	0.66	0.57	0.04	1.06
Albania	0.00	0.06	0.00	1.67
Algeria	0.04	1.91	1.18	0.00
American Samoa	0.03	0.01	0.15	0.00

	Dim.1	Dim.2	Dim.3	Dim.4
Andorra	0.04	0.08	0.14	0.00
Angola	1.23	0.58	0.30	3.75
Argentina	0.28	0.43	1.20	0.05
Armenia	0.01	0.00	0.28	1.46
Australia	0.80	0.12	3.63	1.52
Austria	0.82	0.74	0.04	0.14

Les contributions des pays aux différentes dimensions de l'Analyse en Composantes Principales (ACP) indiquent dans quelle mesure chaque pays participe à la variation des données le long de chaque dimension. Voici une explication pour chaque pays :

- **Afghanistan :**
 - Contribue le plus à la première dimension (Dim.1), indiquant une forte variation des données dans cette dimension, suivie par une contribution notable à la deuxième dimension (Dim.2).
 - Sa contribution à la quatrième dimension (Dim.4) est également notable mais moins prononcée que dans les deux premières dimensions.
- **Albanie :**
 - Ne contribue pratiquement pas à la première dimension (Dim.1), mais a une contribution significative à la quatrième dimension (Dim.4), indiquant une variation plus importante le long de cette dimension.
 - La contribution à la deuxième dimension (Dim.2) est également présente mais relativement faible.
- **Algérie :**
 - Contribue principalement à la deuxième dimension (Dim.2), suggérant une forte variation des données le long de cette dimension.
 - Sa contribution à la troisième dimension (Dim.3) est également significative, indiquant une variation notable dans cette dimension.
- **American Samoa :**
 - A une contribution relativement faible à la troisième dimension (Dim.3), indiquant une variation moindre le long de cette dimension.
- **Andorre :**

- Sa contribution est relativement faible dans toutes les dimensions, avec une légère prépondérance dans la troisième dimension (Dim.3).
- **Angola :**
 - Contribue de manière significative à la première dimension (Dim.1), suivie par une contribution notable à la quatrième dimension (Dim.4), indiquant des variations importantes le long de ces dimensions.
- **Argentine :**
 - A une contribution significative à la troisième dimension (Dim.3), suggérant une variation notable le long de cette dimension.
- **Arménie :**
 - A une contribution principalement à la quatrième dimension (Dim.4), indiquant une variation plus importante le long de cette dimension.
- **Australie :**
 - Contribue de manière significative à la troisième dimension (Dim.3), suivie par une contribution notable à la première dimension (Dim.1), indiquant des variations importantes le long de ces dimensions.
- **Autriche :**
 - A une contribution significative à la première dimension (Dim.1), suivie par une contribution notable à la deuxième dimension (Dim.2), suggérant des variations importantes le long de ces dimensions.

8 Annexes

8.1 Description des variables

1. **GDP_CAPITA:** Le Produit Intérieur Brut (PIB) par habitant, qui mesure la valeur totale des biens et services produits par un pays divisée par sa population.
2. **BALANCE:** La balance commerciale, qui représente la différence entre les exportations et les importations d'un pays sur une période donnée.
3. **MANUFACTURE_IMPORTS:** Les importations de produits manufacturés, qui sont des biens produits par l'industrie manufacturière et achetés par un pays auprès d'autres pays.

4. **SERVICE_EMPLOYMENT:** Le pourcentage de la population active employée dans le secteur des services, qui comprend des activités telles que les services financiers, le commerce de détail et les soins de santé.
5. **FOREIGN_INVESTMENT:** Les investissements étrangers directs (IDE), qui représentent les investissements réalisés par des entreprises étrangères dans des entreprises ou des actifs situés dans un pays donné.
6. **FUEL_EXPORTS:** Les exportations de carburants, qui désignent les produits pétroliers ou autres combustibles exportés par un pays.
7. **GROSS_SAVINGS:** L'épargne brute, qui représente le revenu disponible d'un pays moins la consommation finale.
8. **INFLATION:** Le taux d'inflation, qui mesure la variation du niveau général des prix des biens et services dans une économie sur une période donnée.
9. **EXPORT_TELECOM:** Les exportations de services de télécommunications, qui incluent les services de télécommunications fournis par un pays à d'autres pays.
10. **LIFE_EXPECTANCY:** L'espérance de vie à la naissance, qui est la durée moyenne de vie prévue d'une population donnée.
11. **ELECTRICITY_ACCESS:** Le taux d'accès à l'électricité, qui mesure la proportion de la population ayant accès à l'électricité dans un pays.
12. **FERTILITY_RATE:** Le taux de fécondité, qui représente le nombre moyen d'enfants qu'une femme aurait au cours de sa vie.
13. **HUMAN_CAPITAL_INDEX:** L'indice du capital humain, qui mesure les compétences et les connaissances des individus dans une population donnée.
14. **FEMALE_LABOR_FORCE:** Le pourcentage de la population active qui est féminine, mesurant la participation des femmes sur le marché du travail.
15. **POPULATION_65_ABOVE:** Le pourcentage de la population âgée de 65 ans et plus, qui représente la proportion de personnes âgées dans la population totale d'un pays.
16. **EDUCATION_EXPENDITURE:** Les dépenses d'éducation, qui incluent les fonds dépensés par un gouvernement ou d'autres entités pour l'éducation.

17. **URBAN_POPULATION:** Le pourcentage de la population vivant dans des zones urbaines, qui mesure le degré d'urbanisation d'un pays.
18. **CO2_EMISSIONS:** Les émissions de dioxyde de carbone, qui sont la quantité de CO2 libérée dans l'atmosphère par les activités humaines telles que la combustion de combustibles fossiles.
19. **ARMED_FORCE:** Les effectifs militaires, qui représentent le nombre total de membres des forces armées d'un pays.
20. **SURFACE_AREA:** La superficie du territoire d'un pays, qui mesure la taille physique du pays en kilomètres carrés.
21. **EXPORTS_GOOD_SERVICES:** Les exportations de biens et de services, qui représentent la valeur totale des biens et des services vendus par un pays à d'autres pays.
22. **REGION:** La région géographique à laquelle appartient un pays.
23. **INCOME_GROUP:** Le groupe de revenu auquel appartient un pays, basé généralement sur le revenu national brut (RNB) par habitant.