



Dimensiunile latente ale dezvoltării: O analiză multivariată a indicatorilor socio-economici globali

ILIESCU TUDOR

IVAN TIBERIU

AMMAR ISLAM

ILINCĂI LUCIAN

ACADEMIA DE STUDII ECONOMICE DIN BUCUREȘTI

Facultatea de Cibernetică, Statistică și Informatică Economică

Cuprins

1. Introducere	Error! Bookmark not defined.
2. Extracția setului de date	3
3. Descrierea datelor	4
4. Analiza componentelor principale	6
5. Analiza factorială	11
6. Concluzie	14
7. Referințe.....	15
8. Anexe	Error! Bookmark not defined.

Abstract

Lucrarea analizează dimensiunile latente ale dezvoltării socio-economice globale, utilizând metode multivariate precum analiza componentelor principale (ACP) și analiza factorială (AF), aplicate asupra unui set de date format din 13 indicatori statistici relevanți pentru 217 țări, furnizați de Banca Mondială pentru anul 2023. Indicatorii includ aspecte esențiale precum PIB-ul per capita, speranța de viață, rata șomajului, accesul la internet, fertilitatea, urbanizarea și comerțul internațional. Prin ACP au fost identificate trei dimensiuni principale: „Dezvoltare economică și calitatea vieții”, „Piața muncii și stabilitate economică” și „Comerț și urbanizare”, ce sintetizează structura internă a datelor și reduc complexitatea acestora. Analiza factorială a confirmat existența a două factori majori, evidențiind diferențele dintre economiile dezvoltate și cele emergente din perspectiva infrastructurii moderne, respectiv a caracteristicilor demografice și a dependenței de agricultură. Studiul oferă o reprezentare clară a modului în care țările se grupează în funcție de aceste dimensiuni și propune o abordare integrată pentru înțelegerea dezvoltării globale. Rezultatele pot contribui la fundamentarea politicilor economice diferențiate, direcționate spre reducerea disparităților și optimizarea resurselor în funcție de nevoile fiecărei regiuni.

Cuvinte cheie: dezvoltare socio-economică, analiză multivariată, analiza componentelor principale, analiza factorială, indicatori socio-economici, urbanizare

Abstract

This paper analyzes the latent dimensions of global socio-economic development using multivariate methods such as Principal Component Analysis (PCA) and Factor Analysis (FA). The study is based on a dataset comprising 13 relevant statistical indicators for 217 countries, provided by the World Bank for the year 2023. The selected indicators cover key aspects such as GDP per capita, life expectancy, unemployment rate, internet access, urbanization, and international trade. The PCA identified three main dimensions: "Economic Development and Quality of Life," "Labor Market and Economic Stability," and "Trade and Urbanization," which synthesize the internal structure of the data and reduce its complexity. The factor analysis confirmed the presence of two dominant factors, highlighting the differences between developed and emerging economies from the perspective of modern infrastructure, demographic structure, and agricultural dependency. The study provides a clear overview of how countries cluster based on these dimensions and proposes an integrated approach to understanding global development. The findings may contribute to the design of differentiated economic policies aimed at reducing disparities and optimizing resource allocation according to regional needs.

Keywords: socio-economic development, multivariate analysis, principal component analysis, factor analysis, socio-economic indicators, urbanization

1. Introducere

În ultimele decenii, dezvoltarea socio-economică a devenit un subiect central al cercetărilor globale, pe fondul persistenței și chiar accentuării decalajelor dintre state. Abordările moderne nu se mai limitează la indicatori economici izolați, precum PIB-ul, ci vizează o perspectivă integrată asupra calității vieții și funcționării sustenabile a societăților. În acest context, analiza indicatorilor socio-economici la nivel internațional devine esențială pentru înțelegerea proceselor de dezvoltare.

Transformările sociale și economice recente, influențate de dinamica demografică, tehnologică și educațională, se reflectă în variabile statistice diverse. Indicatori precum speranța de viață, rata șomajului, fertilitatea sau accesul la internet reflectă realități profunde, care necesită o abordare analitică riguroasă pentru a fi corect interpretate.

Prezenta lucrare urmărește explorarea acestor dimensiuni pe baza unui set extins de date globale, permițând identificarea unor tipare comune, corelații relevante și diferențieri între state. Scopul principal este de a evidenția structura latentă a dezvoltării socio-economice, prin surprinderea relațiilor dintre indicatori care nu sunt vizibile la o analiză de suprafață.

Prin această abordare, cercetarea oferă o perspectivă sistemică asupra dezvoltării, depășind nivelul descriptiv și concentrându-se pe identificarea unor modele interpretative relevante. Rezultatele pot contribui la fundamentarea politicilor publice și la înțelegerea diferențelor structurale dintre țări într-o lume din ce în ce mai interdependentă.

Abordările moderne nu se mai limitează la indicatori economici izolați, precum PIB-ul, ci vizează o perspectivă integrată asupra calității vieții și funcționării sustenabile a societăților. În acest context, analiza indicatorilor socio-economici la nivel internațional devine esențială pentru înțelegerea proceselor de dezvoltare, iar tehnicile de analiză multivariată sunt esențiale pentru înțelegerea relațiilor complexe dintre variabile, în special atunci când aceste relații sunt interdependente (Hair et al., 2019, p. 215).¹














1. Extractia setului de date

Setul de date utilizat în această lucrare a fost extras din baza de date a World Bank și corespunde anului 2023. Acesta include 13 variabile economice și sociale esențiale, înregistrate pentru 217 țări, oferind o imagine globală cuprinzătoare asupra nivelului de dezvoltare.

Indicatorii selectați vizează aspecte cheie precum PIB-ul per capita, rata șomajului, speranța de viață, fertilitatea, accesul la internet sau urbanizarea. Fiecărui indicator i-a fost atribuit un simbol de lucru (X1–X13), utilizat pentru claritate și uniformitate în cadrul analizei.

Setul de date permite investigarea relațiilor dintre acești factori și conturarea unor perspective relevante asupra dezvoltării socio-economice la scară globală.

¹ **Note de subsol:** Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning, p. 240.

 X1 - PIB per capital	<i>GDP per capita (current US\$)</i>
 X2-Rata somajului	<i>Unemployment, total (% of total labor force)</i>
 X3-Rata inflatiei	<i>Inflation, consumer prices (annual %)</i>
 X4-Speranta de viata la nastere	<i>Life expectancy at birth, total (years)</i>
 X5-Acces la electricitate	<i>Access to electricity (% of population)</i>
 X6-Rata fertilitatii	<i>Fertility rate, total (births per woman)</i>
 X7-Import	<i>Imports of goods and services (current US\$)</i>
 X8-Export	<i>Exports of goods and services (current US\$)</i>
 X9-Acces la apa potabila	<i>People using at least basic drinking water services (% of population)</i>
 X10-Acces la internet	<i>Individuals using the Internet (% of population)</i>
 X11-Rata de participare la munca (+15ani)	<i>Labor force participation rate, total (% of total population ages 15+)</i>
 X12-Populatie urbana	<i>Urban population</i>
 X13-Contributia agriculturii la PIB	<i>Agriculture value added (% of GDP)</i>

Tabelul 1. Variabilele utilizate în setul de date

2. Descrierea datelor

```
> summary(date_finales)
```

X1		X2		X3		X4		X5		X6		X7		X8	
Min.	: 803.8	Min.	: 0.250	Min.	: 1.750	Min.	:58.92	Min.	: 65.4	Min.	:0.780	Min.	: 17.56	Min.	: 4.62
1st Qu.	: 4295.4	1st Qu.	: 3.690	1st Qu.	: 6.250	1st Qu.	:72.19	1st Qu.	: 99.9	1st Qu.	:1.470	1st Qu.	: 35.65	1st Qu.	:28.90
Median	: 9065.5	Median	: 5.590	Median	: 8.400	Median	:74.58	Median	:100.0	Median	:1.780	Median	: 52.01	Median	:43.27
Mean	:15038.1	Mean	: 6.221	Mean	: 9.159	Mean	:74.86	Mean	: 97.6	Mean	:1.995	Mean	: 53.15	Mean	:47.23
3rd Qu.	:21508.4	3rd Qu.	: 8.070	3rd Qu.	:11.900	3rd Qu.	:78.00	3rd Qu.	:100.0	3rd Qu.	:2.300	3rd Qu.	: 68.99	3rd Qu.	:63.05
Max.	:52084.7	Max.	:15.300	Max.	:19.870	Max.	:84.00	Max.	:100.0	Max.	:4.590	Max.	:104.65	Max.	:99.30

X9		X10		X11		X12		X13	
Min.	: 69.59	Min.	: 32.95	Min.	:43.77	Min.	:17.62	Min.	: 0.640
1st Qu.	: 95.09	1st Qu.	: 73.84	1st Qu.	:57.62	1st Qu.	:55.75	1st Qu.	: 2.620
Median	: 98.65	Median	: 83.90	Median	:61.30	Median	:66.91	Median	: 6.460
Mean	: 96.18	Mean	: 79.42	Mean	:60.90	Mean	:65.25	Mean	: 7.549
3rd Qu.	: 99.88	3rd Qu.	: 89.07	3rd Qu.	:64.38	3rd Qu.	:78.71	3rd Qu.	:10.900
Max.	:100.00	Max.	:100.00	Max.	:76.02	Max.	:98.15	Max.	:24.390

Figura 1. Statisticile descriptive pentru fiecare variabilă

Din analiza statisticilor descriptive se observă o variabilitate semnificativă între țări, ceea ce reflectă diferențele majore de dezvoltare la nivel global. Spre exemplu, PIB-ul per capita

variază între 803,8 și 52.084,7 unități monetare, evidențiind decalajul economic între economiile dezvoltate și cele emergente sau aflate în curs de dezvoltare. În mod similar, rata fertilității prezintă o amplitudine largă, cu valori cuprinse între 0,78 și 4,59 nașteri per femeie, ceea ce indică diferențe culturale, sociale și de acces la servicii medicale și educaționale. De asemenea, variabile precum accesul la electricitate și accesul la internet înregistrează valori ridicate în medie, dar cu prezența unor extreme notabile în rândul țărilor cu venituri mici sau infrastructură deficitară.

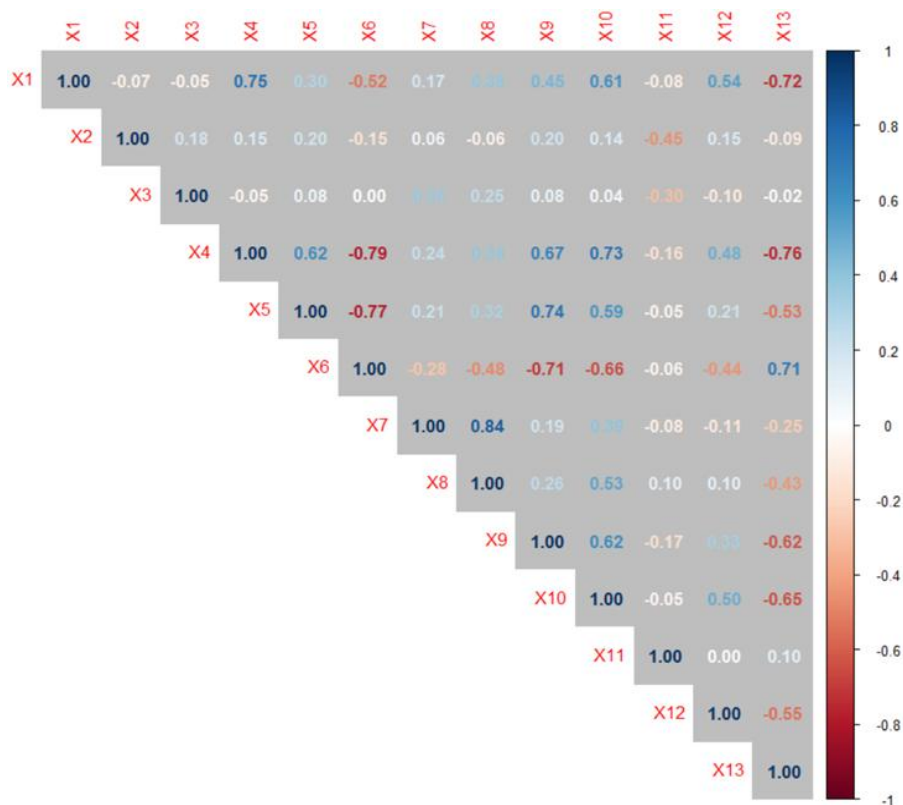


Figura 2. Reprezentare grafică a matricei de corelație

Pentru a înțelege modul în care aceste variabile interacționează, a fost analizată matricea de corelație, care oferă o perspectivă globală asupra relațiilor dintre indicatori. Se evidențiază corelații pozitive puternice, cum ar fi între PIB per capita și speranța de viață, ceea ce sugerează că un nivel economic mai ridicat este, în general, asociat cu condiții de trai mai bune și un sistem de sănătate mai performant. Pe de altă parte, sunt identificate și corelații negative relevante, cum este cazul dintre rata fertilității și accesul la internet, ceea ce ar putea reflecta diferențele în nivelul de educație și dezvoltare tehnologică între țări.

Un aspect important evidențiat de reprezentarea grafică a matricei de corelație este existența unei corelații directe și foarte puternice între variabilele X7 și X8, cu o valoare de 0.84. Acest lucru indică faptul că, în general, țările cu un volum ridicat de importuri sunt și mari exportatoare, ceea ce reflectă un nivel ridicat de integrare în economia globală. Totuși, o astfel de corelație înaltă poate genera redundanță informațională în analizele ulterioare, afectând acuratețea rezultatelor obținute prin tehnici multivariate, în special din cauza riscului de coliniaritate.

Prin urmare, pentru a simplifica modelul analitic și a asigura robustețea interpretărilor, se propune excluderea variabilei X7 din analiza principală. Această decizie permite reducerea influenței distorsiunilor statistice și contribuie la o înțelegere mai clară a structurilor latente din setul de date.

3. Analiza componentelor principale

Pentru a înțelege mai bine structura datelor noastre și pentru a evita interpretarea separată a celor treisprezece variabile, am ales să aplicăm analiza componentelor principale. Ne-am dat seama încă de la început că mulți dintre indicatorii socio-economici pe care îi folosim sunt

interdependenți și, prin urmare, pot fi grupați în dimensiuni comune care exprimă aspecte mai generale ale dezvoltării. În loc să analizăm fiecare variabilă în parte, ACP ne-a oferit o metodă prin care am redus dimensiunea setului de date, păstrând în același timp cea mai mare parte din informația relevantă. Pentru a înțelege mai bine structura datelor noastre și pentru a evita interpretarea separată a celor treisprezece variabile, am ales să aplicăm analiza componentelor principale (PCA). Această tehnică a fost esențială pentru reducerea dimensiunii setului de date și păstrarea informației relevante. (Abdi & Williams, 2010, p. 440)²

Am pornit de la ipoteza că dezvoltarea globală poate fi explicată prin câteva axe esențiale, construite pe baza combinației dintre indicatorii existenți. După standardizarea variabilelor, am evaluat relevanța fiecărei componente prin trei criterii uzuale: criteriul procentului de acoperire, criteriul valorilor proprii și criteriul pantei.

```
> procent_cumulat<-cumsum(procent_info)
> procent_cumulat
  Comp.1  Comp.2  Comp.3  Comp.4  Comp.5  Comp.6  Comp.7  Comp.8  Comp.9  Comp.10  Comp.11  Comp.12
45.69222 59.62618 69.73120 78.25265 83.86463 88.55492 91.53749 93.97613 96.15132 97.85083 99.20785 100.00000
> |
> valproprii
  Comp.1  Comp.2  Comp.3  Comp.4  Comp.5  Comp.6  Comp.7  Comp.8  Comp.9  Comp.10  Comp.11  Comp.12
5.48306690 1.67207476 1.21260267 1.02257392 0.67343786 0.56283410 0.35790865 0.29263704 0.26102287 0.20394073 0.16284193 0.09505857
> |
```

Figura 3. Valorile proprii și procentul cumulat

Observăm că prima componentă reține 45.69% din informație, primele două componente rețin 59.63% din informație, iar primele trei componente rețin 69.73% din informația totală. Primele patru componente rețin 78.25% din informație.

Aplicând criteriul lui Kaiser, conform căruia se păstrează doar componentele principale cu valori proprii mai mari decât 1, au fost selectate patru componente. Acestea rețin o cantitate semnificativă de informație, superioară celei oferite de variabilele standardizate individual, și sunt considerate suficiente pentru a explica structura de bază a datelor analizate.

Conform criteriului procentului de acoperire, un criteriu comun pentru selecția componentelor principale este ca acestea să rețină minim 75% din informația totală (procent cumulat). Astfel, în acest caz, păstrăm primele patru componente, deoarece acestea rețin peste 75% din informație.

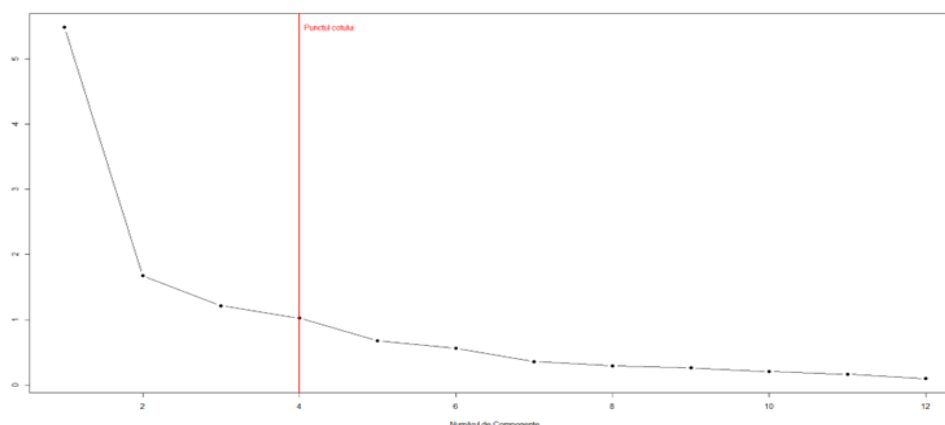


Figura 4. Graficul criteriului pantei

² **Note de subsol:** Kaiser, H.F. (1960). *The Application of Electronic Computers to Factor Analysis*. Educational and Psychological Measurement, 20(1), 141–151, p. 145.

În urma aplicării criteriului pantei, observăm o schimbare pronunțată a pantei la componenta 4, evidențiată pe grafic. Acest punct marchează începutul unei scăderi mai lente a variației explicate de componentele următoare, sugerând că primele 4 componente capturează cea mai mare parte a informației esențiale din date. Astfel, este recomandat să păstrăm aceste 4 componente pentru analiză.

Pentru a înțelege mai bine semnificația fiecărei componente extrase, am analizat matricea încărcăturilor factorilor, care reflectă corelațiile dintre variabilele inițiale și componentele principale. Această analiză ne-a permis să identificăm ce variabile contribuie cel mai mult la fiecare axă și, implicit, să atribuim fiecărei componente o interpretare conceptuală relevantă în contextul dezvoltării socio-economice.



Figura 5. Reprezentarea grafică a matricei factor

Prima componentă este marcată de valori ridicate ale coeficienților pentru variabilele X1 (PIB per capita), X4 (Speranța de viață), X5 (Acces la electricitate), X9 (Acces la apă potabilă) și X10 (Acces la internet). Aceste variabile indică infrastructura, condițiile de trai și accesul la servicii esențiale, toate asociate cu nivelul de dezvoltare economică și bunăstarea populației. Astfel, această axă a fost denumită „Dezvoltare economică și calitatea vieții”, reflectând dimensiunea fundamentală a progresului unei societăți moderne.

Componenta a doua este dominată de două variabile esențiale pentru piața muncii: X2 (Rata șomajului), cu o încărcătură negativă, și X11 (Rata de participare la muncă), cu o încărcătură pozitivă. Acestea reflectă gradul de implicare a populației în activitatea economică și capacitatea unei țări de a susține un nivel ridicat de ocupare. Componenta captează, așadar, o dimensiune legată de „Piața muncii și stabilitate economică”, relevând aspecte esențiale pentru funcționarea eficientă a economiei.

Componenta a treia include variabile precum X3 (Rata inflației), X8 (Exporturi) și X12 (Populația urbană). Deși la prima vedere par variate, acestea se leagă printr-o dimensiune comună ce ține de structura economică și demografică: stabilitatea prețurilor, integrarea în comerțul internațional și gradul de urbanizare. În acest sens, am interpretat această axă ca reflectând aspecte legate de „Comerț și urbanizare”, sugerând modul în care economiile se adaptează la dinamica globală și la procesele interne de modernizare.

În schimb, componenta a patra nu prezintă încărcături semnificative sau coerente pentru nicio variabilă, iar influențele sunt dispersate și slabe. Acest lucru indică faptul că această componentă nu contribuie cu valoare analitică suplimentară și nu conturează o direcție clară de interpretare. Prin urmare, am decis să excludem componenta a patra din analiza ulterioară, păstrând doar cele trei componente relevante, care explică în mod clar și coerent structura latentă a datelor.

Această etapă ne-a oferit o înțelegere detaliată a modului în care cele 13 variabile socio-economice se grupează în trei dimensiuni fundamentale, fiecare reflectând o latură esențială a dezvoltării: infrastructură și bunăstare, activitate economică și ocupare, comerț și urbanizare. Pe baza acestor componente sintetice, vom continua analiza prin reprezentări grafice și clasificări, pentru a evidenția diferențele dintre state și modelele distincte de dezvoltare la nivel global.

După identificarea componentelor relevante, am reprezentat grafic variabilele în planul format de prima și a doua componentă principală, pentru a înțelege mai bine distribuția acestora și relațiile dintre ele. În acest scop, am utilizat cercul de corelație, un instrument vizual esențial în interpretarea ACP.

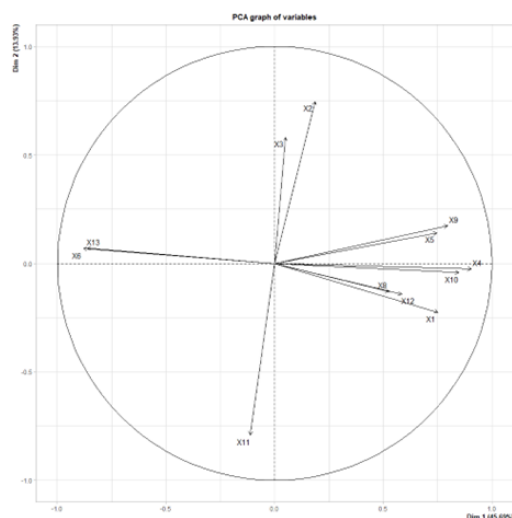


Figura 6. Reprezentarea grafică a cercului corelației

În graficul generat, unghiul dintre vectorii variabilelor arată natura corelației dintre acestea. Vectorii orientați în aceeași direcție, cum este cazul X1 (PIB per capita), X4 (Speranța de viață), X5 (Acces la electricitate), X9 (Apă potabilă) și X10 (Internet), indică o corelație pozitivă puternică, sugerând că aceste variabile cresc împreună și contribuie major la prima componentă. În contrast, vectorii opuși ca direcție – precum X2 (Șomaj) și X6 (Fertilitate) – au corelații negative cu celelalte, ceea ce înseamnă că nivelurile ridicate ale dezvoltării sunt asociate cu valori mai mici pentru aceste variabile.

Variabilele poziționate perpendicular, cum ar fi X11 (Participarea la muncă) față de X1 sau X4, indică lipsa unei corelații semnificative. Lungimea săgeților evidențiază contribuția fiecărei variabile: vectorii lungi, precum cei ai variabilelor X4 și X10, arată o implicare importantă în formarea componentelor, în timp ce vectorii scurți indică o influență redusă.

Această reprezentare vizuală susține și validează interpretările anterioare, oferind o imagine clară asupra modului în care variabilele se grupează și cum contribuie la formarea dimensiunilor principale identificate în cadrul ACP.

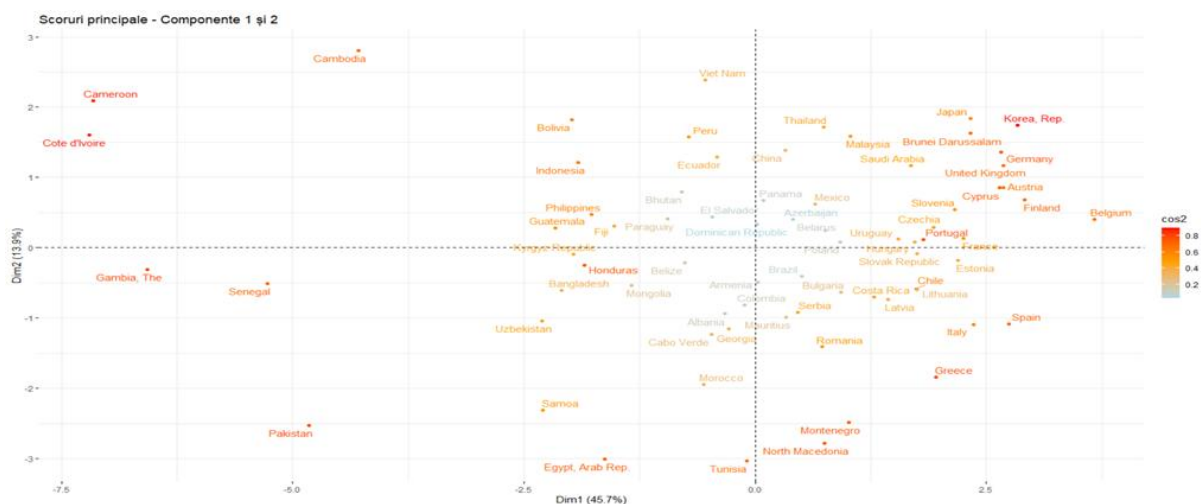


Figura 7. Reprezentarea scorurilor principale

Reprezentarea scorurilor principale pentru componentele 1 și 2 oferă o imagine de ansamblu asupra modului în care țările sunt poziționate în funcție de dimensiunile latente identificate anterior. Țările dezvoltate, precum Coreea de Sud, Japonia sau Germania, se află în partea dreaptă a graficului, confirmând asocierea lor cu infrastructura avansată, accesul extins la servicii și o calitate ridicată a vieții. La polul opus, țări precum Camerun, Coasta de Fildeș sau Pakistan au scoruri negative pe prima componentă, reflectând provocări semnificative în ceea ce privește dezvoltarea socio-economică.

În același plan, dimensiunea a doua aduce un nivel suplimentar de diferențiere, evidențiind aspecte regionale, culturale sau structurale care nu sunt explicate exclusiv de nivelul de dezvoltare economică. Astfel, se observă că unele state din aceeași zonă geografică se diferențiază clar între ele, în funcție de participarea la muncă, inflație sau dinamica demografică.

Biplotul ce suprapune scorurile țărilor cu vectorii variabilelor consolidează și mai mult aceste relații: statele dezvoltate sunt aliniate în direcția variabilelor precum speranța de viață (X4), accesul la apă (X9) sau internet (X10), în timp ce statele în curs de dezvoltare se află în proximitatea variabilelor precum fertilitatea (X6) sau agricultura în PIB (X13). Această distribuție evidențiază clar polarizarea dezvoltării globale și oferă o bază solidă pentru o clasificare ulterioară a țărilor, în funcție de dimensiunile latente.

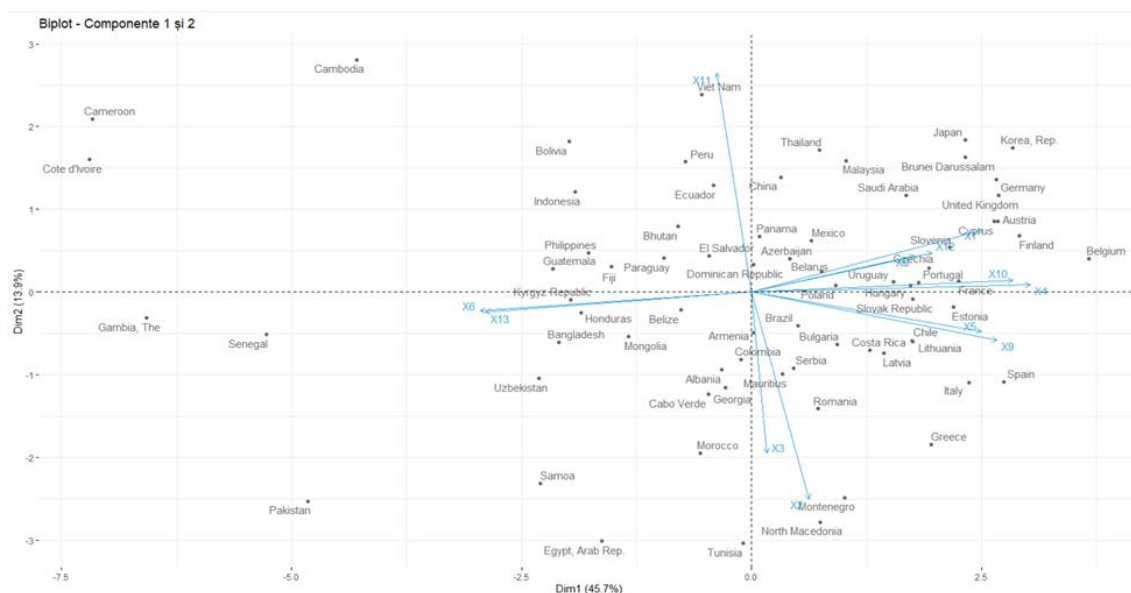


Figura 8. Reprezentare grafică a biplotului pentru componentele 1 și 2.

Prin urmare, analiza componentelor principale a permis o reducere eficientă a complexității celor 13 variabile inițiale, sintetizând informația în trei dimensiuni principale, interpretabile și relevante. Această abordare oferă nu doar o perspectivă integrată asupra relațiilor dintre indicatori, ci și o imagine clară a modului în care țările se diferențiază în peisajul global al dezvoltării.

În continuare, pentru a consolida structura identificată prin Analiza componetelor principale și a valida consistența acestor dimensiuni, vom apela la Analiza Factorială, o metodă complementară, care permite explorarea în profunzime a relațiilor dintre variabile și factorii comuni care le determină.

4. Analiza factorială (AF)

Pentru a înțelege mai bine dimensiunile latente care stau la baza dezvoltării economice și sociale a țărilor analizate, am aplicat o analiză factorială. Obiectivul nostru a fost să extragem factori comuni care să sintetizeze informația conținută în cele 13 variabile, oferind astfel o imagine mai clară asupra structurii ascunse a acestor indicatori. Am considerat că această metodă este potrivită pentru că variabilele sunt corelate între ele, iar gruparea lor în factori ne ajută să înțelegem mai bine diferențele dintre țări. Pentru a înțelege mai bine dimensiunile latente care stau la baza dezvoltării economice și sociale a țărilor analizate, am aplicat o analiză factorială. Această metodă ne-a permis să extragem factori comuni care sintetizează informația din variabilele socio-economice (Kaiser, 1960, p. 145).³

Înainte de a trece la selecția numărului de factori, am verificat dacă setul de date este adecvat pentru o analiză factorială. Testul KMO a arătat o valoare generală de 0.81, ceea ce indică o corelație suficient de mare între variabile pentru a putea fi grupate în factori comuni. În același timp, testul Bartlett a confirmat existența unei structuri interne relevante, respingând ipoteza că variabilele ar fi necorelate. Aceste rezultate ne-au oferit încrederea că putem continua cu extracția factorilor. Pentru a asigura robustețea rezultatelor, am aplicat analiza factorială

³ **Note de subsol:** Kaiser, H.F. (1960). *The Application of Electronic Computers to Factor Analysis*. Educational and Psychological Measurement, 20(1), 141–151, p. 145.

folosind atât metoda axelor principale, cât și metoda verosimilității maxime. Am observat că rezultatele obținute prin cele două metode sunt similare în ceea ce privește semnificația factorilor extrași, însă modelul bazat pe axele principale a oferit o mai bună stabilitate și o interpretare mai clară a structurii interne. După aplicarea metodelor de selecție a numărului de factori, atât Scree Plot-ul cât și analiza paralelă au indicat că extragerea a doi factori este optimă. Acești doi factori explică cea mai mare parte a variației totale și oferă o bază solidă pentru interpretare. Modelul rezultat surprinde două dimensiuni importante care separă clar statele în funcție de caracteristicile economice și sociale. Factorul întâi pare să reflecte dimensiunea dezvoltării economice și a infrastructurii, în timp ce factorul al doilea este asociat mai degrabă cu aspectele demografice și cu dependența de agricultură.

```
> cor<-cor(date_std)
> cortest.bartlett(cor,diag=TRUE,n=nrow(date_std))
$chisq
[1] 530.4097

$ p.value
[1] 1.030448e-73

$df
[1] 66
```

Figura 9. Testul Bartlett

```
> kmo(date_std)
Kaiser-Meyer-Olkin factor adequacy
Call: kmo(r = date_std)
Overall MSA = 0.81
MSA for each item =
  X1  X2  X3  X4  X5  X6  X8  X9  X10 X11 X12 X13
0.82 0.57 0.48 0.84 0.83 0.81 0.65 0.90 0.85 0.44 0.75 0.93
```

Figura 10. Testul KMO

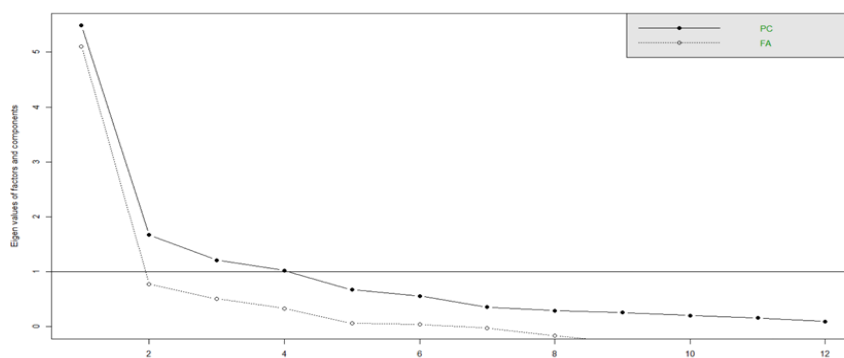


Figura 11. Reprezentare grafică a analizei paralele.

În cadrul modelului estimat prin metoda axelor principale, am observat că variabile precum PIB-ul per capital, speranța de viață, accesul la electricitate, apa potabilă și internet sunt cele mai puternic corelate cu primul factor. Acest grup de variabile descrie în mod clar o dimensiune a dezvoltării moderne și a infrastructurii sociale. Țările care au scoruri mari pe acest factor sunt, de regulă, cele dezvoltate, cu sisteme bine puse la punct și acces extins la servicii de bază. De cealaltă parte, variabilele cu încărcături negative sau mai slabe, cum ar fi contribuția agriculturii la PIB sau rata fertilității, nu sunt bine explicate de acest prim factor.

Cel de-al doilea factor a fost puternic influențat de rata fertilității și contribuția agriculturii la PIB, având o relație inversă cu gradul de urbanizare. Acest lucru ne-a dus la concluzia că al doilea factor reflectă o dimensiune demografică și agrară, prin care putem înțelege cât de mult depind anumite țări de agricultură și ce impact are acest lucru asupra structurii populației și nivelului de urbanizare. Țările care au scoruri mari pe acest factor tind să fie mai puțin

urbanizate, cu economii bazate pe agricultură și populații tinere, în timp ce țările dezvoltate înregistrează valori negative, sugerând o tranziție spre economie modernă și urbanizată.

Interpretarea modelului s-a bazat nu doar pe încărcăturile variabilelor, ci și pe comunalitate și unicitate. Variabilele precum rata fertilității, accesul la internet și contribuția agriculturii sunt foarte bine reprezentate de cei doi factori, ceea ce ne arată că explică bine structura latentă a datelor. În schimb, alte variabile precum inflația sau șomajul nu au fost la fel de relevante, având comunalități mai scăzute, ceea ce indică o contribuție mai slabă la modelul general.

Pentru a valida modelul și pentru a evalua calitatea estimării, am comparat cele două metode utilizate: axele principale și verosimilitatea maximă. Am observat că modelul obținut prin axele principale explică mai bine structura datelor, având o distribuție mai echilibrată a încărcăturilor și o interpretare mai clară. În plus, indicatorii de ajustare au arătat că acest model oferă o potrivire acceptabilă, reflectând relații semnificative între variabilele originale și factorii extrași.

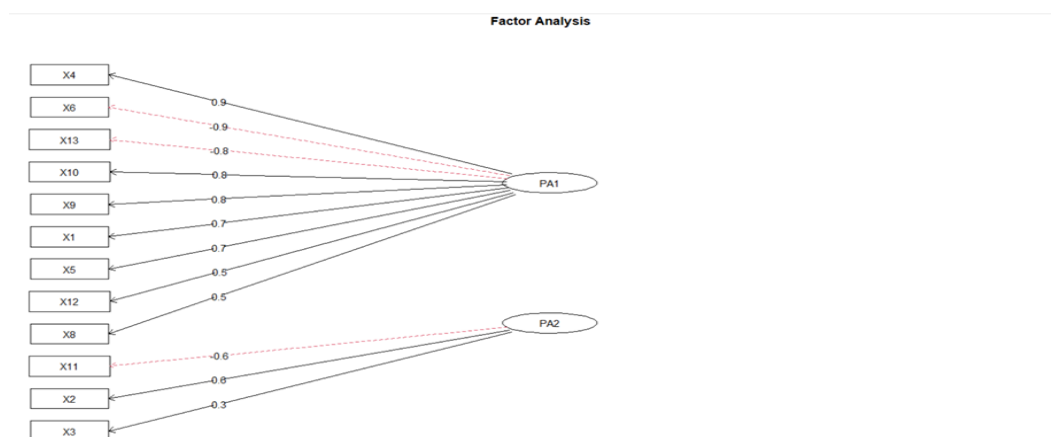


Figura 12. Diagrama modelului factorial prin metoda axelor principale.

Reprezentarea grafică a modelului ne-a oferit o perspectivă vizuală importantă. Proiecția țărilor în planul celor doi factori a arătat grupări clare. Țările din dreapta, cu scoruri mari pe factorul întâi, sunt cele dezvoltate, precum Italia, Spania sau România, caracterizate prin infrastructură puternică și indicatori socio-economici buni. În partea stângă regăsim țările cu infrastructură slabă și dependență economică ridicată de agricultură. Dimensiunea verticală, dată de factorul al doilea, separă țările în funcție de structura demografică și economică, cu țări precum Senegal și Gambia plasate în partea superioară, reflectând o economie agricolă activă și o populație tânără.

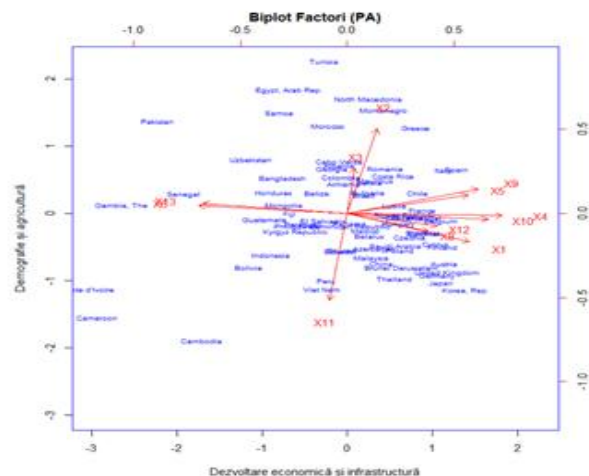


Figura 13. Biplot PA

Această analiză factorială ne-a ajutat să înțelegem că, dincolo de cifrele brute, există dimensiuni ascunse care influențează profund dezvoltarea unei țări. Factorii extrași ne arată că există două axe mari în jurul cărora gravitează indicatorii analizați: una legată de infrastructură și modernizare, cealaltă de tranziția demografică și economică. Țările care punctează bine pe ambele dimensiuni sunt cele care reușesc să combine dezvoltarea economică cu o structură socială echilibrată, în timp ce cele care înregistrează scoruri scăzute ar putea beneficia de strategii direcționate pe zonele lor vulnerabile. Analiza factorială poate fi utilizată pentru a identifica structura latentă care stă la baza unui set de variabile observate și poate ajuta la reducerea complexității datelor multivariate.” (Hair et al., 2019)⁴

5. Concluzie

Lucrarea de față a urmărit investigarea dimensiunilor latente ale dezvoltării socio-economice globale prin intermediul unor tehnici de analiză multivariată, aplicate pe un set de date extins, codificat și prelucrat din surse oficiale. Prin aplicarea Analizei Componentelor Principale și a Analizei Factoriale, am reușit să sintetizăm informațiile provenite din 13 indicatori relevanți, identificând structuri interne semnificative și relații ascunse între variabile.

Analiza componentelor principale a evidențiat existența a trei dimensiuni majore care descriu dezvoltarea globală: infrastructura și calitatea vieții, dinamica pieței muncii și stabilitatea economică, precum și gradul de urbanizare și integrare comercială. Ulterior, analiza factorială a confirmat existența a două factori dominanți: unul axat pe modernizarea economică și infrastructurală, iar celălalt pe structura agrară și caracteristicile demografice.

Rezultatele au arătat diferențieri clare între state, reflectând disparitățile majore în ceea ce privește accesul la resurse, nivelul de urbanizare, gradul de digitalizare și orientarea economică. Țările dezvoltate s-au distins prin scoruri ridicate la factorii legați de infrastructură și servicii moderne, în timp ce statele mai puțin dezvoltate au fost asociate cu o dependență mai mare de agricultură și o structură demografică specifică.

Prin integrarea acestor metode, lucrarea a oferit un cadru coerent pentru înțelegerea dezvoltării ca fenomen multidimensional. Concluziile obținute pot constitui un punct de














⁴**Note de subsol:** Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning.

plecare pentru politici publice diferențiate, adaptate nevoilor specifice ale fiecărei categorii de țări, și pot contribui la conturarea unor strategii mai eficiente de reducere a inegalităților globale.

6. Referințe

1. Abdi, H., & Williams, L.J. (2010). Principal Component Analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4), 433–459.
2. Hair, J.F., Black, W.C., Babin, B.J., & Anderson, R.E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning.
3. Kaiser, H.F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, 20(1), 141–151.
4. Maer, M. (2011). *Analiza datelor statistice*. Editura Universitară, București.
5. Stancu, S. (2011). *Econometrie. Teorie și aplicații economice utilizând EViews*. Editura ASE, București.
6. Țițan, E. (2011). *Statistica în analiza și modelarea fenomenelor economico-sociale*. Editura ASE, București.
7. World Bank. (2023). *World Development Indicators*. <https://data.worldbank.org/>

9. Anexe

 X1 - PIB per capital	<i>GDP per capita (current US\$)</i>
 X2-Rata somajului	<i>Unemployment, total (% of total labor force)</i>
 X3-Rata inflatiei	<i>Inflation, consumer prices (annual %)</i>
 X4-Speranta de viata la nastere	<i>Life expectancy at birth, total (years)</i>
 X5-Acces la electricitate	<i>Access to electricity (% of population)</i>
 X6-Rata fertilitatii	<i>Fertility rate, total (births per woman)</i>
 X7-Import	<i>Imports of goods and services (current US\$)</i>
 X8-Export	<i>Exports of goods and services (current US\$)</i>
 X9-Acces la apa potabila	<i>People using at least basic drinking water services (% of population)</i>
 X10-Acces la internet	<i>Individuals using the Internet (% of population)</i>
 X11-Rata de participare la munca (+15ani)	<i>Labor force participation rate, total (% of total population ages 15+)</i>
 X12-Populatie urbana	<i>Urban population</i>
 X13-Contributia agriculturii la PIB	<i>Agriculture value added (% of GDP)</i>

Tabelul 1. Variabilele utilizate în setul de date

```
> summary(date_finale)
```

X1		X2		X3		X4		X5		X6		X7		X8	
Min.	: 803.8	Min.	: 0.250	Min.	: 1.750	Min.	:58.92	Min.	: 65.4	Min.	:0.780	Min.	: 17.56	Min.	: 4.62
1st Qu.	: 4295.4	1st Qu.	: 3.690	1st Qu.	: 6.250	1st Qu.	:72.19	1st Qu.	: 99.9	1st Qu.	:1.470	1st Qu.	: 35.65	1st Qu.	:28.90
Median	: 9065.5	Median	: 5.590	Median	: 8.400	Median	:74.58	Median	:100.0	Median	:1.780	Median	: 52.01	Median	:43.27
Mean	:15038.1	Mean	: 6.221	Mean	: 9.159	Mean	:74.86	Mean	: 97.6	Mean	:1.995	Mean	: 53.15	Mean	:47.23
3rd Qu.	:21508.4	3rd Qu.	: 8.070	3rd Qu.	:11.900	3rd Qu.	:78.00	3rd Qu.	:100.0	3rd Qu.	:2.300	3rd Qu.	: 68.99	3rd Qu.	:63.05
Max.	:52084.7	Max.	:15.300	Max.	:19.870	Max.	:84.00	Max.	:100.0	Max.	:4.590	Max.	:104.65	Max.	:99.30

X9		X10		X11		X12		X13	
Min.	: 69.59	Min.	: 32.95	Min.	:43.77	Min.	:17.62	Min.	: 0.640
1st Qu.	: 95.09	1st Qu.	:73.84	1st Qu.	:57.62	1st Qu.	:55.75	1st Qu.	: 2.620
Median	: 98.65	Median	:83.90	Median	:61.30	Median	:66.91	Median	: 6.460
Mean	: 96.18	Mean	: 79.42	Mean	:60.90	Mean	:65.25	Mean	: 7.549
3rd Qu.	: 99.88	3rd Qu.	:89.07	3rd Qu.	:64.38	3rd Qu.	:78.71	3rd Qu.	:10.900
Max.	:100.00	Max.	:100.00	Max.	:76.02	Max.	:98.15	Max.	:24.390

Figura 1. Statisticile descriptive pentru fiecare variabilă

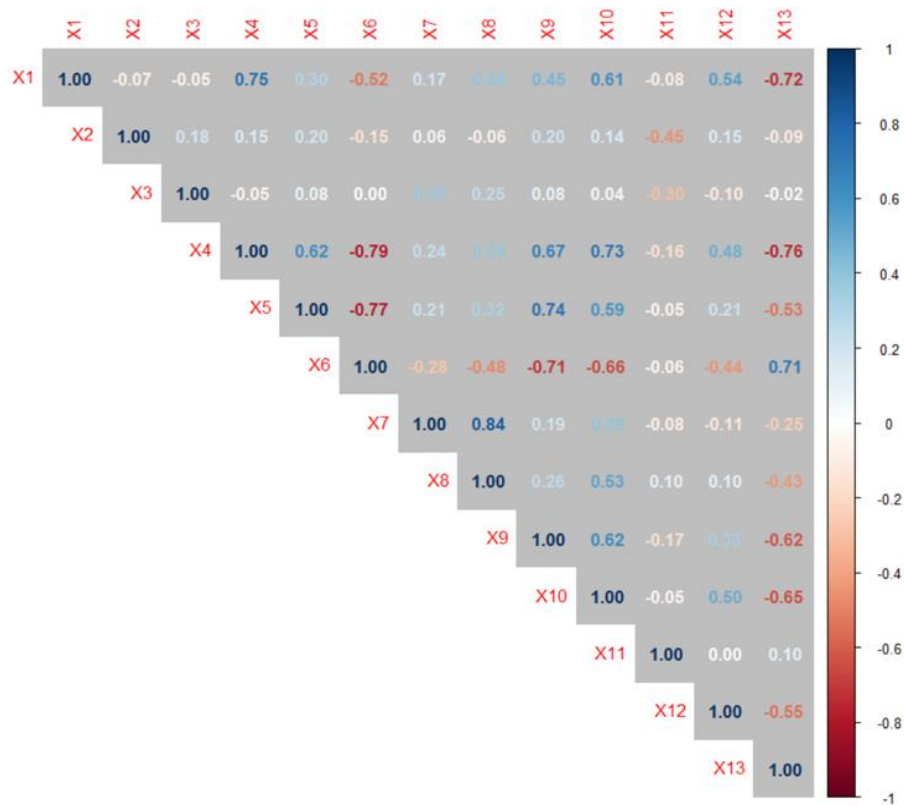


Figura 2. Reprezentare grafică a matricei de corelație

```
> procent_cumulat<-cumsum(procent_info)
> procent_cumulat
  Comp.1   Comp.2   Comp.3   Comp.4   Comp.5   Comp.6   Comp.7   Comp.8   Comp.9   Comp.10   Comp.11   Comp.12
45.69222 59.62618 69.73120 78.25265 83.86463 88.55492 91.53749 93.97613 96.15132 97.85083 99.20785 100.00000
>
> valproprii
  Comp.1   Comp.2   Comp.3   Comp.4   Comp.5   Comp.6   Comp.7   Comp.8   Comp.9   Comp.10   Comp.11   Comp.12
5.48306690 1.67207476 1.21260267 1.02257392 0.67343786 0.56283410 0.35790865 0.29263704 0.26102287 0.20394073 0.16284193 0.09505857
>
```

Figura 3. Valorile proprii și procentul cumulat

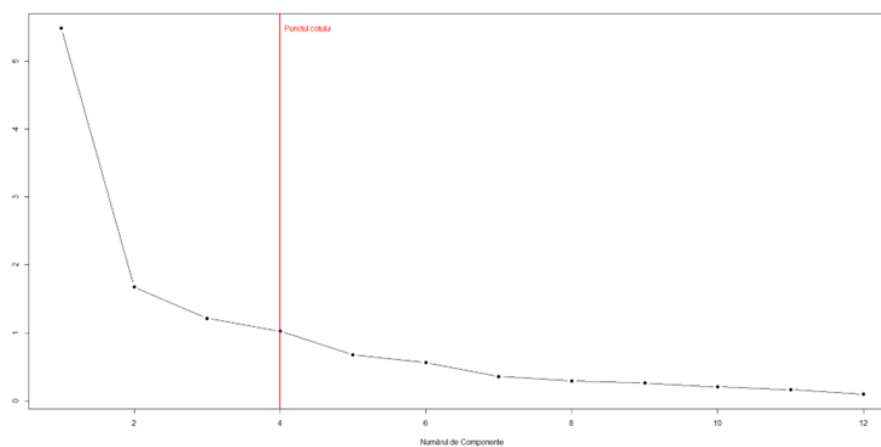


Figura 4. Graficul criteriului pantei



Figura 5. Reprezentarea grafică a matricei factor

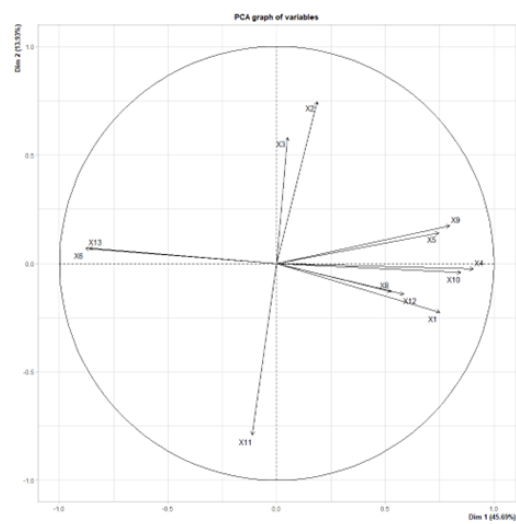


Figura 6. Reprezentarea grafică a cercului corelației

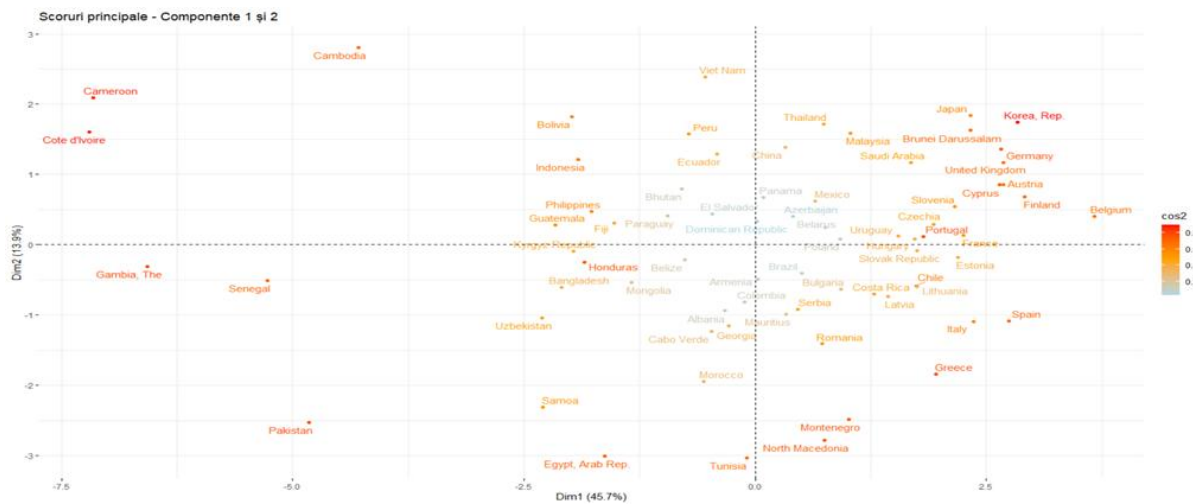


Figura 7. Reprezentarea scorurilor principale

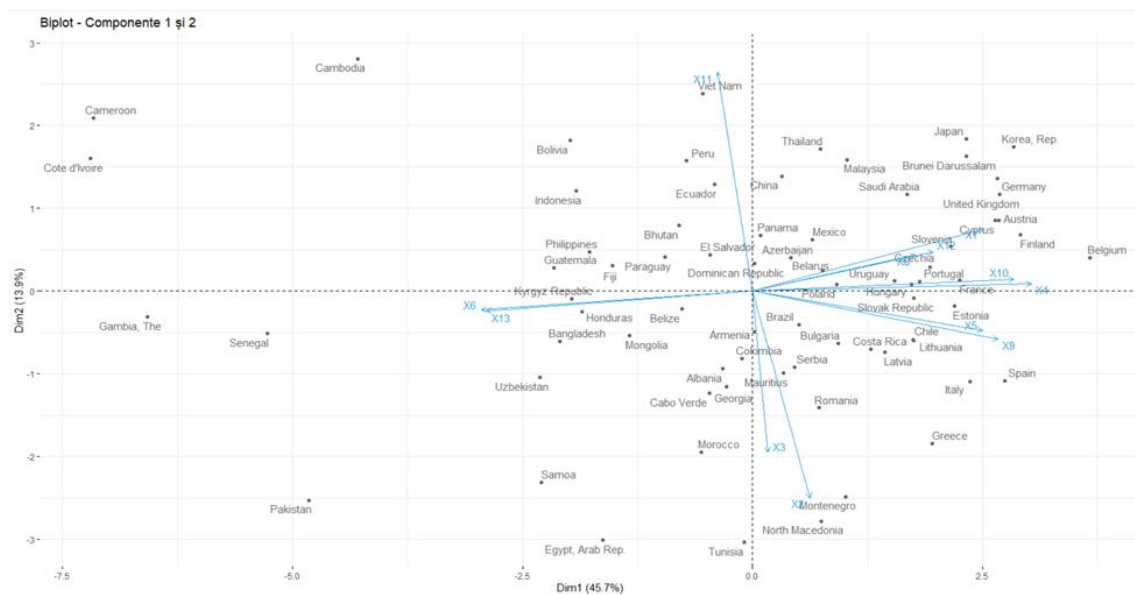


Figura 8. Reprezentare grafică a biplotului pentru componentele 1 și 2.

```
> cor<-cor(date_std)
> cortest.bartlett(cor,diag=TRUE,n=nrow(date_std))
$chisq
[1] 530.4097

$ p.value
[1] 1.030448e-73

$df
[1] 66
```

```
> kmo(date_std)
Kaiser-Meyer-Olkin factor adequacy
Call: kmo(r = date_std)
Overall MSA = 0.81
MSA for each item =
  x1  x2  x3  x4  x5  x6  x8  x9  x10 x11 x12 x13
0.82 0.57 0.48 0.84 0.83 0.81 0.65 0.90 0.85 0.44 0.75 0.93
```

Figura 9. Testul Bartlett

Figura 10. Testul KMO

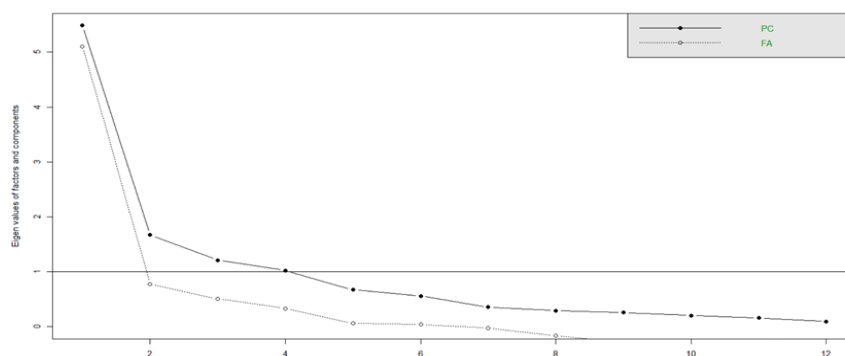


Figura 11. Reprezentare grafică a analizei paralele.

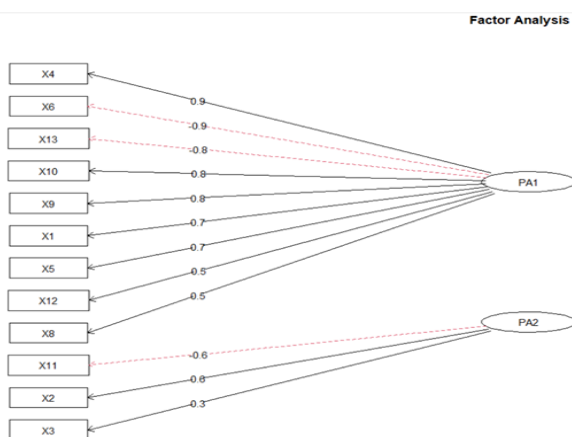


Figura 12. Diagrama modelului factorial prin metoda axelor principale.

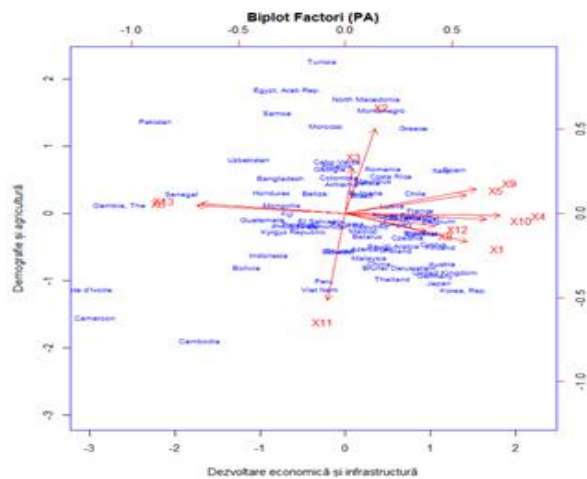


Figura 13. Biplot PA

