**ACADEMIA DE STUDII ECONOMICE DIN BUCUREȘTI, FACULTATEA DE CIBERNETICĂ, STATISTICĂ ȘI INFORMATICĂ ECONOMICĂ**

A logo with wings and a letter

Description automatically generated

**PROIECT**

**Dezvoltare Software**

**pentru analiza datelor**

**PROF. COORDONATOR: STUDENTI:**

**Vinte Claudiu Homos Cristian-Cosmin**

**Puscasu Tiberiu**

**Grupa: 1101E**

Cuprins

[Introducere 3](#_Toc188226768)

[Data source 3](#_Toc188226769)

[Descrierea variabilelor 3](#_Toc188226770)

[Abordarea analizei datelor 5](#_Toc188226771)

[Principal component analysis (PCA) 6](#_Toc188226772)

[Analiza Cluster Ierarhică (HCA) 7](#_Toc188226773)

[Corelograme 12](#_Toc188226774)

[Concluzii 15](#_Toc188226775)

[Bibliografie 16](#_Toc188226776)

**Figuri**

[Figure 1. Correlograma of comunalitatilor 11](#_Toc188228724)

[Figure 2. Correlograma incarcaturilor factorilor 12](#_Toc188228725)

[Figure 3. Correlograma contributiei observatiilor la varianta axelor 12](#_Toc188228726)

[Figure 4. Correlograma calitatii repprezentarii punctelor 13](#_Toc188228727)

[Figure 5. Correlograma scorurilor 14](#_Toc188228728)

[Figure 6. Explained Variance by the Principal Components 14](#_Toc188228729)

# Introducere

Acest proiect explorează dinamica socio-economică complexă din țările Uniunii Europene, utilizând un set de date preluat din sursa EUROSTAT pentru anul 2022. Setul de date include o gamă variată de indicatori esențiali, precum sănătatea, demografia, educația, sărăcia, stilul de viață, infrastructura și sustenabilitatea mediului.

Prin aplicarea **Analizei Componentelor Principale (PCA)** și **Analizei Cluster Ierarhice (HCA)**, proiectul își propune să descopere tipare ascunse, corelații și relații în cadrul setului de date. PCA este utilizată pentru a reduce dimensiunea datelor, păstrând variația semnificativă, și pentru a evidenția factorii socio-economici dominanți. În același timp, HCA organizează țările în clustere pe baza similitudinilor, oferind o perspectivă structurată asupra peisajului socio-economic din Europa.

Rezultatele acestei analize contribuie la o mai bună înțelegere a structurilor sociale și a condițiilor de mediu în 2022. Aceste concluzii pot servi ca bază pentru elaborarea unor politici informate, intervenții sociale direcționate și cercetări ulterioare privind tendințele socio-economice din Uniunea Europeană.

# Data source

Mortalitatea infantilă, un indicator esențial al bunăstării unei societăți, reflectă pierderea tragică a copiilor înainte de a împlini vârsta de un an. Acest indicator reprezintă un barometru pentru evaluarea calității sistemului de sănătate al unei țări, a condițiilor socio-economice și a stării generale de sănătate publică. În această analiză, sunt utilizate **Analiza Componentelor Principale (PCA)** și **Analiza Cluster Ierarhică (HCA)** pentru a explora factorii diverși care influențează rata mortalității infantile în țările europene în anul 2022.

Aceste metodologii analitice au ca scop descoperirea tiparelor ascunse, corelațiilor și informațiilor critice din setul de date, oferind o înțelegere aprofundată a dinamicii socio-economice și a sistemelor de sănătate care influențează mortalitatea infantilă. Rezultatele obținute urmăresc să sprijine formularea de politici și intervenții eficiente pentru reducerea ratelor mortalității infantile.

Setul de date a fost preluat din multiple baze de date EUROSTAT (vezi bibliografia).

# Descrierea variabilelor

În cadrul proiectului nostru de analiză a mortalității infantile, fiecare variabilă selectată joacă un rol esențial în evidențierea diverselor factori care influențează sănătatea și supraviețuirea copiilor. Aceste variabile oferă o perspectivă cuprinzătoare asupra determinanților socio-economici, ai infrastructurii de sănătate și ai mediului înconjurător care afectează mortalitatea infantilă.

**BabiesBornWithDiseases** - această variabilă reflectă prezența afecțiunilor genetice sau congenitale la naștere, oferind informații despre riscurile inerente cu care se confruntă nou-născuții. Înțelegerea acestor afecțiuni ajută la identificarea factorilor genetici și de mediu care pot influența supraviețuirea copiilor. Monitorizarea acestei variabile este esențială pentru dezvoltarea unor strategii și intervenții medicale direcționate.

**GynoAndObs -** această variabilă reprezintă disponibilitatea și accesibilitatea la servicii ginecologice în cadrul unei populații. Serviciile ginecologice sunt cruciale pentru sănătatea maternă și infantilă, asigurând îngrijiri esențiale de reproducere și îndrumare prenatală. Accesul adecvat la ginecologi reduce riscurile unor complicații precum greutatea mică la naștere și nașterile premature, contribuind la îmbunătățirea sănătății neonatale.

**Babies<2kg - a**ceastă variabilă urmărește numărul nou-născuților cu o greutate mai mică de 2 kilograme, un factor critic asociat cu riscuri crescute de mortalitate infantilă. Identificarea populațiilor în care acest fenomen este frecvent permite implementarea de intervenții țintite pentru îmbunătățirea sănătății materne și fetale.

**Births<37 Weeks -** această variabilă măsoară numărul copiilor născuți înainte de 37 de săptămâni de gestație. Nașterile premature sunt asociate cu diverse riscuri de sănătate pentru nou-născuți, evidențiind necesitatea unor îngrijiri specializate. Împreună cu datele despre greutatea mică la naștere, această variabilă oferă informații esențiale pentru strategiile de reducere a riscurilor asociate.

**TotalFertilityRate -** această metrică demografică indică numărul de nașteri vii la 100 de femei aflate la vârsta fertilă într-un an dat. Tendințele privind rata fertilității oferă informații despre demografia maternă și sănătatea reproductivă. Variațiile în rata fertilității pot influența bunăstarea mamelor și a copiilor, oferind un context valoros pentru înțelegerea dinamicii sănătății populației.

**MortalityByMomEducation -** nivelul educațional este un determinant socio-economic important al rezultatelor în sănătate. Această variabilă evidențiază provocările cu care se confruntă mamele cu un nivel scăzut de educație, cum ar fi accesul redus la îngrijire medicală și dificultățile în adoptarea practicilor sănătoase de creștere a copiilor. Aceste aspecte pot afecta semnificativ ratele de supraviețuire infantilă.

**AtRiskOfPoverty - a**ceastă variabilă monitorizează nașterile în populațiile dezavantajate din punct de vedere economic, evidențiind disparitățile în accesul la resurse și îngrijire medicală. Înțelegerea impactului sărăciei asupra sănătății mamelor și copiilor ajută la orientarea intervențiilor care vizează reducerea inegalităților.

**Smoking și AlcoholConsumption - a**ceste variabile surprind factori esențiali de stil de viață care influențează sănătatea copiilor. Fumatul și consumul de alcool în timpul sarcinii sunt factori de risc bine documentați pentru rezultate nefavorabile, inclusiv probleme de dezvoltare la copii. Împreună, aceste variabile explorează intersecția dintre comportamentele de sănătate, statutul socio-economic și mortalitatea infantilă.

**HospitalBeds -** această variabilă reflectă numărul de paturi disponibile în unitățile de terapie intensivă neonatală, un indicator esențial al infrastructurii de sănătate. Disponibilitatea paturilor este vitală pentru gestionarea cazurilor critice care implică nou-născuți prematuri sau bolnavi. Această variabilă influențează direct ratele de supraviețuire infantilă.

**GreenhouseGasEmission - d**eși neconvențională în studiile tradiționale despre mortalitatea infantilă, această variabilă ia în considerare contextul mai larg al mediului. Emisiile de gaze cu efect de seră, un indicator al stresorilor de mediu, pot afecta indirect sănătatea populației, inclusiv a nou-născuților. Includerea acestei variabile adaugă o dimensiune ecologică analizei, evidențiind legătura dintre sustenabilitatea mediului și bunăstarea copiilor.

Această selecție de variabile oferă un cadru nuanțat și multidimensional pentru înțelegerea factorilor care influențează mortalitatea infantilă, permițând dezvoltarea unor strategii informate pentru îmbunătățirea rezultatelor.

**Descrierea observatiilor**

Observațiile din setul nostru de date includ informații colectate din 27 de țări membre ale Uniunii Europene, fiecare reprezentând un studiu de caz unic în contextul mortalității infantile. Alegerea acestor țări oferă o imagine diversă a regiunilor europene, evidențiind variații culturale, economice, în sistemele de sănătate și în condițiile de mediu. Setul de date surprinde o perspectivă asupra stării de sănătate maternă și infantilă din fiecare țară, permițând o analiză comparativă a factorilor care influențează ratele mortalității infantile.

Țările selectate reflectă o gamă variată de profiluri demografice, sisteme de sănătate și condiții socio-economice, contribuind la bogăția și diversitatea datasetului. Prin includerea unor state cu niveluri diferite de dezvoltare și infrastructură medicală, ne propunem să descoperim tipare și tendințe care depășesc granițele naționale, oferind perspective valoroase pentru strategii regionale și pan-europene destinate îmbunătățirii sănătății copiilor.

Aceste observații constituie componente esențiale în analiza noastră, oferind context pentru aplicarea **Analizei Componentelor Principale (PCA)** și **Analizei Cluster Ierarhice (HCA)**. Prin aceste metode, explorăm în detaliu natura complexă și multifactorială a mortalității infantile pe continentul european, identificând corelații și grupări relevante care pot ghida decizii informate și intervenții eficiente.

# Abordarea analizei datelor

În selecția atentă a variabilelor din setul nostru de date, obiectivul a fost să construim un cadru cuprinzător care să includă factori genetici, demografici, socio-economici, de sănătate și de mediu. Acest set divers de variabile permite o înțelegere detaliată a interacțiunilor complexe care contribuie la peisajul multifactorial al mortalității infantile.

Analiza noastră utilizează **Analiza Componentelor Principale (PCA)** și **Analiza Cluster Ierarhică (HCA)** pentru a explora relațiile dintre variabile și a identifica tipare relevante. PCA reduce dimensiunea datasetului, evidențiind structuri latente și interdependențe care ar putea fi dificil de observat direct. HCA organizează țările în clustere, oferind perspective asupra grupurilor cu caracteristici comune și contribuind la o analiză mai clară a profilurilor regionale.

## Principal component analysis (PCA)

Conform criteriului Kaiser, afirmăm că doar 4 componente principale (C1, C2, C3, C4) care au valoarea proprie mai mare decât 1 sunt statistic importante. Vizualizarea variației explicate de componentele principale arată că înclinarea curbei indică importanța fiecărei componente în explicarea variației, C1 explicând cea mai mare parte, în timp ce importanța scade până la C11.

Din punct de vedere statistic, matricea încărcăturilor factoriale **în Analiza Componentelor Principale (PCA)** oferă informații despre modul în care fiecare variabilă originală contribuie la fiecare componentă principală. Încărcăturile factoriale reprezintă coeficienții dintr-o combinație liniară a variabilelor originale care formează fiecare componentă principală. Matricea încărcăturilor factoriale în PCA este un element fundamental pentru înțelegerea relațiilor dintre variabilele originale și componentele principale.

Toate aceste variabile, luând în considerare locația (țările europene), reflectă cum arată mortalitatea infantilă (copii născuți cu greutate mai mică de 2 kilograme, părinți aflați în risc de sărăcie care nu își permit îngrijire prenatală, numărul de paturi disponibile pentru îngrijirea neonatală, copii născuți cu malformații genetice și boli și numărul de ginecologi.

**C1 (Caracteristici ale nașterilor, factori socio-economici și acces la sănătate)** este determinat de caracteristici precum greutatea la naștere, factorii socio-economici (de exemplu, persoane aflate în risc de sărăcie), accesul la îngrijire medicală, anomaliile congenitale și numărul de medici (ginecologi). Această componentă are o contribuție mai mare în **Germania, Franța, Italia și Spania**, ceea ce este sugerat de corelograma calității punctelor. Aceste țări ar putea avea în comun:

* **Tipare socio-economice similare:** distribuția veniturilor, niveluri de educație sau structuri ocupaționale;
* **Sisteme de sănătate comparabile:** accesibilitate, calitatea serviciilor și infrastructura medicală;
* **Norme culturale sau sociale comune:** diversitate etnică și culturală;
* **Similitudini demografice:** distribuția pe vârste sau densitatea populației;
* **Politici similare:** cadrul legislativ sau intervențiile guvernamentale în domenii precum sănătatea, educația sau asistența socială.

Contribuțiile pozitive ale variabilelor legate de greutatea la naștere, factori socio-economici și accesul la sănătate sugerează un risc mai scăzut de mortalitate infantilă.

**C2 (Educație și impact de mediu)** reflectă factori educaționali (de exemplu, România - cel mai mare număr de mame minore din Europa care au abandonat școala. Valorile pozitive mari în C2 sugerează o corelație cu nașterile din familii cu nivel scăzut de educație și expunerea la factori de mediu, cum ar fi poluarea.

**C3 (Durata gestației și stilul de viață)** este mai ridicat în **Austria, Croația și Malta**. Această componentă evidențiază importanța intervențiilor în sănătatea publică care vizează reducerea fumatului în timpul sarcinii. Programele educaționale, accesul la servicii de sprijin și măsurile politice pentru reducerea fumatului la femeile însărcinate sunt esențiale. Fumatul în timpul sarcinii este asociat cu rezultate negative precum nașteri premature, greutate mică la naștere și probleme de dezvoltare.

**C4 (Rata fertilității și malformații congenitale)** este mai prezentă î**n Austria, Croația și Franța**. În Austria, provocările legate de practicile medicale, testările genetice și alegerile de stil de viață pot explica această corelație. De exemplu, testările neonatale nu sunt gratuite, iar screeningul genetic nu este acoperit de asigurarea de sănătate obligatorie. În Franța, lipsa cerinței pentru testări genetice prenatale explică această corelație.

## Analiza Cluster Ierarhică (HCA)

Utilizând metoda Ward și metrica euclidiană, determinăm și analizăm inițial partiția optimă. Ulterior, selectăm o partiție pe baza examinării graficului dendrogramă. După cum se observă, partiția optimă constă în două clustere.

Analiza clusterelor implică examinarea distribuției fiecărui indicator în cadrul fiecărui cluster. Acest proces ne permite să identificăm particularitățile și diferențele dintre clustere. Vom prezenta mai multe distribuții care evidențiază diferențe clare între cele două clustere din partiția optimă, detaliate mai jos:

|  |  |  |  |
| --- | --- | --- | --- |
| Austria | C1 | Ireland | C1 |
| Belgium | C1 | Italy | C0 |
| Bulgaria | C1 | Latvia | C1 |
| Croatia | C1 | Lithuania | C1 |
| Cyprus | C1 | Luxembourg | C1 |
| Czechia | C1 | Malta | C1 |
| Denmark | C1 | Netherlands | C1 |
| Estonia | C1 | Poland | C0 |
| Finland | C1 | Portugal | C1 |
| France | C0 | Romania | C1 |
| Germany | C0 | Slovakia | C1 |
| Greece | C1 | Slovenia | C1 |
| Hungary | C1 | Spain | C0 |
|  |  | Sweden | C1 |

Conform informațiilor furnizate de Eurostat și partiției cu 2 clustere tarile din clusterul C0 prezintă valori mai scăzute ale indicatorilor de sănătate (‘BabiesBornWithDiseases’, ‘MortalityByMomEducation’, ‘Births<37 Weeks) și valori mai mici ale factorilor socio-economici (‘AtRiskOfPoverty’).

În schimb, în Germania, Spania, Franța si Italia, imigrația are un impact negativ semnificativ, datorită fluxului de persoane din țări mai sărace, cu standarde de viață și educație mai scăzute.

Pentru resursele de sănătate tarile din C1 au o infrastructură de sănătate mai dezvoltată, însă impactul asupra mediului este mai ridicat.

Pentru factorii de stil de viață consumul de alcool și fumatul variază, dar sunt mai ridicate în țările din C1.

Clusterul C1, care include principalele economii europene, tinde să aibă indicatori de sănătate mai ridicați, provocări socio-economice mai mari, resurse de sănătate mai bune și un impact mai mare asupra mediului comparativ cu clusterul C0, care constă din țări cu valori generale mai scăzute ale acestor factori. Diferențele pot fi atribuite variațiilor în dezvoltarea economică, sistemele de sănătate și alegerile legate de stilul de viață.

|  |  |  |  |
| --- | --- | --- | --- |
| Austria | C2 | Ireland | C2 |
| Belgium | C2 | Italy | C1 |
| Bulgaria | C2 | Latvia | C2 |
| Croatia | C2 | Lithuania | C2 |
| Cyprus | C2 | Luxembourg | C2 |
| Czechia | C2 | Malta | C2 |
| Denmark | C2 | Netherlands | C2 |
| Estonia | C2 | Poland | C1 |
| Finland | C2 | Portugal | C0 |
| France | C1 | Romania | C0 |
| Germany | C1 | Slovakia | C2 |
| Greece | C2 | Slovenia | C2 |
| Hungary | C2 | Spain | C1 |
|  |  | Sweden | C0 |

Analizând partiția cu 3 clustere:

* **Clusterul C1**, care include principalele economii europene, are indicatori de sănătate mai ridicați, provocări socio-economice mai mari și resurse de sănătate superioare.
* **Clusterul C2** include țări cu valori mai scăzute pentru majoritatea factorilor analizați.

|  |  |  |  |
| --- | --- | --- | --- |
| Austria | C0 | Ireland | C4 |
| Belgium | C4 | Italy | C1 |
| Bulgaria | C4 | Latvia | C4 |
| Croatia | C4 | Lithuania | C4 |
| Cyprus | C4 | Luxembourg | C4 |
| Czechia | C4 | Malta | C4 |
| Denmark | C4 | Netherlands | C4 |
| Estonia | C4 | Poland | C1 |
| Finland | C4 | Portugal | C3 |
| France | C2 | Romania | C3 |
| Germany | C2 | Slovakia | C4 |
| Greece | C4 | Slovenia | C4 |
| Hungary | C4 | Spain | C1 |
|  |  | Sweden | C3 |

Analizând partiția cu 5 clustere putem observa ca Suedia formează un cluster distinct, caracterizat de:

* **Număr redus de ginecologi**;
* **Număr scăzut de paturi de spital**;
* **Număr mai mare de nou-născuți cu boli congenitale**.

|  |  |  |  |
| --- | --- | --- | --- |
| Austria | C0 | Ireland | C3 |
| Belgium | C3 | Italy | C1 |
| Bulgaria | C2 | Latvia | C3 |
| Croatia | C2 | Lithuania | C3 |
| Cyprus | C3 | Luxembourg | C3 |
| Czechia | C3 | Malta | C2 |
| Denmark | C3 | Netherlands | C3 |
| Estonia | C3 | Poland | C1 |
| Finland | C3 | Portugal | C5 |
| France | C4 | Romania | C5 |
| Germany | C4 | Slovakia | C2 |
| Greece | C2 | Slovenia | C2 |
| Hungary | C2 | Spain | C1 |
|  |  | Sweden | C5 |

Analizând partiția cu 6 clustere putem observa ca:

* Spania și Italia formează clustere noi, având date similare pentru majoritatea factorilor. Se poate concluziona că demografia, cultura și rădăcinile comune joacă un rol crucial.
* Franța și Germania, fiind cele mai puternice țări implicate politic în Uniunea Europeană, formează propriul cluster, având politici și legi similare.

Analiza clusterelor aplicată variabilelor arată că:

* **Clusterul 0** evidențiază relația dintre sănătate, condiții socio-economice și greutatea scăzută la naștere;
* **Clusterul 1** pune accent pe interacțiunea dintre factorii legați de populație și considerațiile de mediu;
* **Clusterul 2** sugerează o conexiune între factorii educaționali și anumite rezultate ale sarcinii;
* **Clusterul 3** evidențiază legătura dintre alegerile de stil de viață (fumat și consum de alcool) și implicațiile socio-economice și asupra sănătății.

# Corelograme

Acest grafic prezintă comunalitățile fiecărei variabile cu componentele principale. Valoarea comunalităților indică proporția variației unei variabile explicate de toate componentele principale.

A red and white chart

Description automatically generated

Figure 1. Correlograma of comunalitatilor

Reprezintă încărcăturile factoriale, adică corelația dintre variabilele originale și componentele principale. Este util pentru a identifica ce variabile contribuie cel mai mult la fiecare componentă.

A screenshot of a diagram

Description automatically generated

Figure 2. Correlograma incarcaturilor factorilor

Afișează contribuția fiecărei observații (țară) la variația fiecărei componente principale. Permite evidențierea țărilor care influențează cel mai mult anumite axe.

A screenshot of a math test

Description automatically generated

Figure 3. Correlograma contributiei observatiilor la varianta axelor

Arată calitatea reprezentării fiecărei observații în spațiul componentelor principale. Valorile mai mari indică o mai bună reprezentare.

A screenshot of a computer

Description automatically generated

Figure 4. Correlograma calitatii repprezentarii punctelor

Prezintă poziționarea fiecărei observații (țară) în spațiul definit de componentele principale. Este util pentru identificarea grupurilor de observații similare.

A red blue and white squares

Description automatically generated

Figure 5. Correlograma scorurilor

Graficul arată proporția din variabilitatea totală explicată de fiecare componentă principală. Linia roșie indică pragul de semnificație conform criteriului Kaiser.

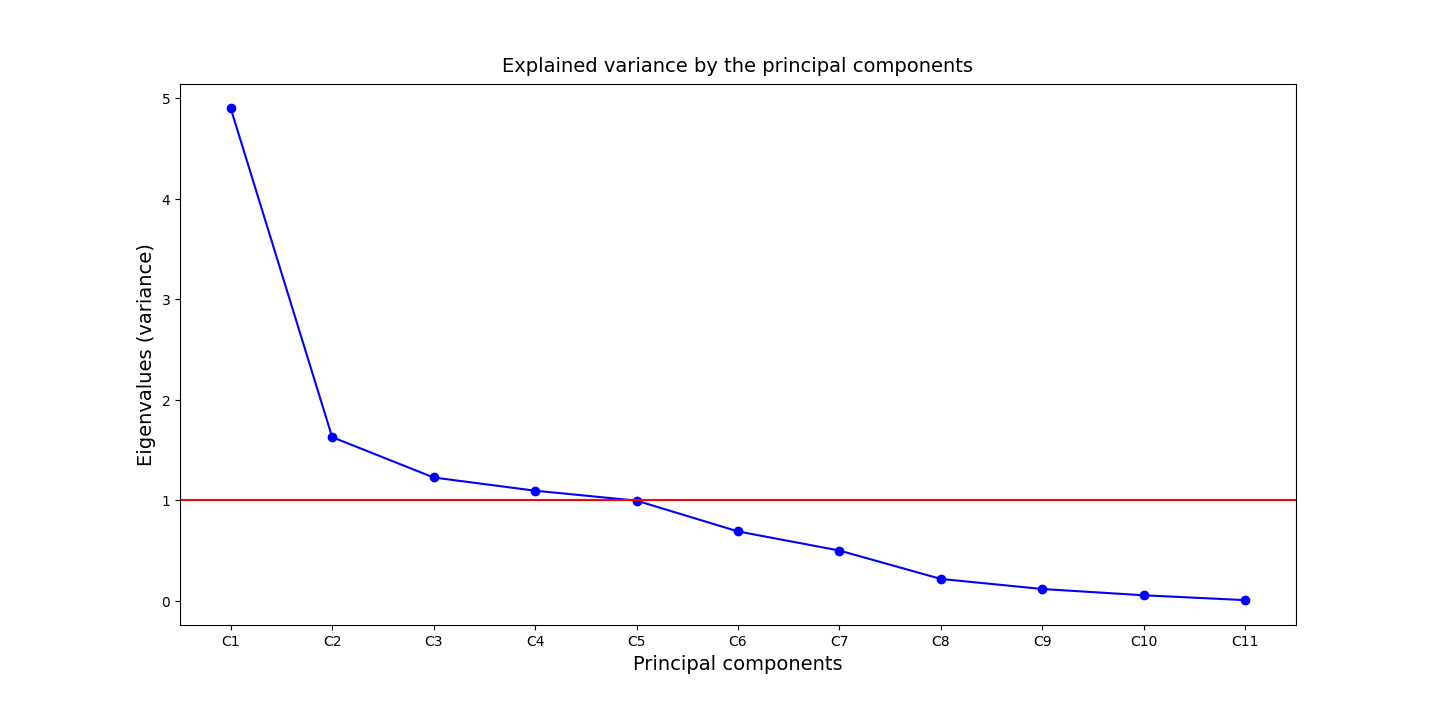


Figure 6. Explained Variance by the Principal Components

# Concluzii

În acest proiect de dezvoltare software pentru analiza datelor privind mortalitatea infantilă în țările membre ale Uniunii Europene, s-a realizat o explorare cuprinzătoare pentru a descifra dinamica complexă a factorilor socio-economici. Setul de date, obținut de la EUROSTAT pentru anul 2022, a inclus o gamă largă de indicatori ce acoperă sănătatea, demografia, educația, sărăcia, stilul de viață, infrastructura și sustenabilitatea mediului. Metode analitice robuste, precum **Analiza Componentelor Principale (PCA)** și **Analiza Clusterelor (HCA)**, au fost utilizate pentru a descoperi tipare și corelații ascunse în setul de date, având ca scop furnizarea de perspective valoroase pentru formularea de politici și intervenții sociale.

Variabilele selectate au fost alese cu atenție pentru a reprezenta factori genetici, demografici, socio-economici, de sănătate și de mediu. Fiecare variabilă, precum *‘BabiesBornWithDiseases’*, *‘GynoAmdObs’*, *‘*Babies<2kg, *‘*Births<37 Weeks’, *‘TotalFertilityRate’*, *‘*MortalityByMomEducation’, *‘AtRiskOfPoverty’*, *‘Smoking’*, *‘AlcoholConsumption’*, *‘HospitalBeds* și *‘GreenhouseGasEmission’*, a avut un rol unic în înțelegerea peisajului complex al mortalității infantile. Observațiile, colectate din cele 27 de țări membre ale Uniunii Europene, au oferit o perspectivă diversă asupra contextelor culturale, economice, de sănătate și de mediu, contribuind la un set de date bogat și cuprinzător.

Aplicarea **Analizei Componentelor Principale (PCA)** și **Analizei Clusterelor (HCA)** a scos la iveală tipare ascunse, corelații și grupări distincte în setul de date, contribuind la o înțelegere detaliată a dinamicii mortalității infantile.

**Analiza Componentelor Principale (PCA)** a evidențiat factori compoziți ce reprezintă caracteristicile nașterilor, condițiile socio-economice, accesul la sănătate, alegerile de stil de viață și impactul asupra mediului. Interpretând rezultatele:

* Țări precum **Germania, Franța, Italia, Spania și România** au prezentat valori mai ridicate în factorii compoziți, ceea ce sugerează un acces mai bun la infrastructura de sănătate și riscuri mai scăzute de mortalitate infantilă.
* Factorii analizați includ corelații între caracteristici precum greutatea la naștere, riscul de sărăcie, accesul la unități NICU, anomaliile congenitale și stilul de viață (fumat, consum de alcool).

**Analiza Clusterelor (HCA)** a clasificat cele 27 de țări în grupuri distincte pe baza similarităților indicatorilor de sănătate, provocărilor socio-economice, resurselor de sănătate, impactului asupra mediului și factorilor de stil de viață:

* **Clusterul C1** include țări precum Germania, Franța, Italia și Spania, care se caracterizează prin infrastructuri de sănătate superioare, provocări socio-economice mai mari și un impact mai ridicat asupra mediului.
* **Clusterul C0** include majoritatea celorlalte țări, caracterizate prin indicatori de sănătate mai scăzuți, provocări socio-economice mai mici și un impact mai redus asupra mediului.
* Grupările au facilitat identificarea unor provocări comune, oferind un punct de plecare pentru intervenții adaptate la nevoile specifice fiecărui cluster.

Acest proiect a aplicat cu succes tehnici avansate de analiză a datelor, precum PCA și HCA, pentru a obține perspective detaliate asupra factorilor multifactoriali care influențează mortalitatea infantilă în cele 27 de țări membre ale Uniunii Europene în anul 2022. Rezultatele subliniază importanța accesului la îngrijire medicală, condițiilor socio-economice, stilului de viață și factorilor de mediu în modelarea rezultatelor de sănătate ale copiilor.

Clusterele identificate oferă o înțelegere detaliată a provocărilor specifice fiecărei țări și a oportunităților pentru intervenții. Aceste perspective pot sprijini formularea de politici și strategii adaptate pentru reducerea ratei mortalității infantile în Europa, promovând astfel sănătatea și bunăstarea generațiilor viitoare.

# Bibliografie

* <https://ec.europa.eu/eurostat/databrowser/view/hlth_cd_ainfo__custom_9196501/default/table?lang=en>
* <https://ec.europa.eu/eurostat/databrowser/view/tps00199/default/table?lang=en>
* <https://ec.europa.eu/eurostat/databrowser/view/demo_minfedu/default/table?lang=en>
* <https://ec.europa.eu/eurostat/databrowser/view/demo_fweight__custom_9197290/default/table?lang=en>
* <https://ec.europa.eu/eurostat/databrowser/view/tec00113/default/table?lang=en>
* <https://ec.europa.eu/eurostat/databrowser/view/hlth_ehis_sk3e/default/table?lang=en>
* <https://ec.europa.eu/eurostat/databrowser/view/hlth_ehis_al1e/default/table?lang=en>
* <https://ec.europa.eu/eurostat/databrowser/view/hlth_rs_bds1__custom_9198589/default/table?lang=en>
* <https://ec.europa.eu/eurostat/databrowser/view/sdg_13_10/default/table?lang=en>
* <https://ec.europa.eu/eurostat/databrowser/view/hlth_rs_physcat__custom_9332197/default/table?lang=en>