



UNIUNEA EUROPEANĂ



GUVERNUL ROMÂNIEI



Instrumente Structurale  
2007 - 2013

## **Studiul de caz ca instrument metodologic in cadrul evaluarii interventiilor publice**

***Metode cantitative si calitative  
utilizate in selectarea cazurilor de  
studiat***

Ziua II, 9 februarie 2012

---

PROJECT COFINANCED BY EFRD THROUGH TAOP 2007-2013

AAM Management Information Consulting Ltd.

[www.aam.hu](http://www.aam.hu)

- Cate cazuri de studiat trebuie/putem sa selectam pentru a raspunde in mod adecvat intrebarilor de evaluare si a atinge obiectivele evaluarii?
- trebuie: generalizare (daca este urmarita)
- putem: timp si buget
- 1/4/10/15/16/ 27\* locatia nu caracteristica primordiala (desi DG Regio spune ca SdC pot fi tematice si regionale)
- REPREZENTATIV
- Exemple+cunostinte suficiente pentru a putea analiza OT/RI/selectarea de cazuri prin metode statistice
- Nu sa utilizati metodele statistice - expert



- Perspective (bases) prin care se selecteaza cazurile de studiat
- Reprezentativitatea – esantionarea



3 perspective (bases) prin care se selecteaza cazurile de studiat (GAO, WB, Standford)

Convenabil (convenience) – nu riguros

SCOP (purpose) - bracketing, best cases, worst cases, cluster, representative, typical and special interest

Probabilistice (probability/random - S) – nu fezabil

Perspectiva depinde de intrebarea la care trebuie sa raspundem

Daca se selecteaza cazurile prin perspectiva gresita – EROARE FATALA, in ciuda unui design si a unei implementari “perfecte”

COMBINATII



Convenabil: locatie usor de atins – date usor de colectat, resurse putine

- as adauga: date de calitate disponibile, limba, program cunoscut deja, manager de program cunoscut deja

Scop - extreme: program/proiect cu absorbtie sau eficacitate ridicata/scazuta

- mai comprehensiva – 2 dimensiuni/unitatile de analiza studiate ofera si alte variatii

Scop - exemple de succes

Scop - exemple de insucces

- pot avea si o alta trasatura comuna – se suplimenteaza numarul pentru a se testa care trasatura comuna(e) este/sunt explicative



## Scop - clustere

- Se realizeaza grupuri pe baza anumitor criterii (depind, evident, de intrebari si ipoteze), si din aceste grupuri se selecteaza o parte din cazuri
- Putine unitati – se poate si calitativ (vizual/consultativ/filtre excel) – mai eficient, poate nu neaparat mai riguros, dar rezultatele nu pot sa difere f. mult
- Unitati numeroase – peste 50/80 – se pot aplica metode statistice

## Scop - reprezentativitate

- Selectia depinde de intrebari si ipoteze (vezi mai jos)

## Scop - cazuri tipice – greu de definit si gasit

## Scop - cazuri speciale – coordonatele cazului, contestatii, outlier deosebit\* etc.

- “Cat de mult” sau “cat de extinsa este o problema”
- Daca este realizata corect – generabilizare si reprezentativitate
- Totusi, cel putin 30 de cazuri selectate (GAO)
- = toate unitatile vizate au o sansa cunoscuta, egala, de a fi selectate
- Random – la intamplare/lista completa a unitatilor

- **Ilustrativ:** tipice\*, representative (“anecdotal information”); numar mic; narative; visuals; colectie
- **Explorativ:** cel putin un caz pe variatie
- **Caz critic:** numar limitat/un caz – investigarea unei situatii aparte/unice&testarea unei teorii considerate valida
- **Implementare de program:** SCOP, nr. de cazuri depinde de diversitatea populatiei/daca se urmareste generalizarea – documentare si observatie
- **Efectele programului:** SdC foarte utilizate, in combinatie cu alte metode (analiza datelor si sondajul). Selectia depinde de diversitatea programului – daca e prea ridicata, nu e recomandat SdC – recomandat best, worst, typical.
- **Cumulative:** studiile la dispozitie (de ex. 15 – DG Regio, evaluare in inovare)



What is the “instance that we want to take as a whole”? !UA!- design

- A site (fishery at Songkla Lake in Thailand)
- A function (public sector management)
- A project (Daxinganling, China Forest Fire Rehabilitation Project; Indonesia, University Development Project; or Equatorial Guinea, Technical Assistance Project)
- A policy (promoting gender equality)
- An office or department (Department of Education)
- An event (Rwanda emergency relief effort)
- A region, nation, or organization (Northwest Brazil, South Asia and Pacific Region, India, UNESCO)
- “Nested” units in a large or complex case study (Tunisia, first through the seventh water supply project, or Malawi, Country Assistance Review).

- Selectarea studiilor de caz NU ESTE ESANTIONARE statistica
- Perspectiva probabilistica utilizata si in esantionarea statistica
- Pana si in esantionare statistica un % nu spune neaparat totul despre reprezentativitate (de ex. omnibus survey in cadrul evaluarii intermediare POSDRU)

## ESANTIONAREA STATISTICA

- Investigarea unei populatii mari intr-un timp redus si cu rezultate (cunostinte/informatii) imediate
- =setul de operatii cu ajutorul carora, din ansamblul populatiei se alege o parte pt analiza
- selectia – sa se obtina concluzii cu valabilitate generala=exprime caracteristicile intregului univers de ind/pop

- Esantionul trebuie sa fie reprezentativ=sa aiba capacitatea de a reproduce cat mai fidel **structurile si caracteristicile** populatiei din care este extras
- R poate fi exprimata cantitativ (eroare maxima, nivel de probabilitate/incredere)
- R nu in general, ci in raport cu o caracteristica data
- R o notiune relativa=un esantion este mai mult sau mai putin reprezentativ, nu reprezentativ/nereprezentativ; o anumita eroare este acceptata (0.95 nivel minim de probabilitate=5% eroare maxima)
- R depinde de: caracteristicile populatiei (nr. indirect proportional cu omogenitatea)/marimea esantionului (pana la un punct direct proportional cu R)/procedura de esantionare
- R NU DEPINDE DE MARIMEA POPULATIEI (uneori factori de corectie)
- 500-1000 – evident ca nu cazuri de studiat!

- Probabilistice (aleatoare) – elimina interventia subiectiva a omului
- Neprobabilistice (nealeatoare)
- **REPREZENTATIVITATEA** se poate calcula numai pentru esantioane selectate aleatoriu/probabilistic!

1. **Esantionarea simpla aleatoare (ESA):** loterie+teoria probabilitatii

2. **Esantionarea prin stratificare:** pt. fiecare strat se selecteaza un sub-esantion (%cu stratul) – suma acestora=esantionul pt toata populatia, reprezentativ in functie de criteriul de stratificare - stratificarea multipla – reprezentativitate superioara fata de 1

3. **Esantionarea multistadiala:** investigarea indivizilor grupati geografic – stadii: judet/oras/strada – grupuri “omogene” ex. o grupa ca esantion pt “populatia de studenti” - solutie practica dar reprezentativitate mai scazuta ca 1

- 2 scenarii:

(1) 20-30 (up to 50) de potentiali candidati

- a se evita o procedura exhaustiva de selectare care poate deveni un mini-studiu in sine
- criterii operationale
- interogarea persoanelor care au cunostinte despre candidati

(2) Un numar mai mare de potentiali candidati

- procedura structurata pe doua etape

2.1 colectarea de date cantitative relevante pentru toti+definirea de criterii relevante pt stratificarea si reducerea numarului de candidati

2.2 procedura simpla de la punctul (1)

- DG Regio evaluari ex-post -2 cazuri de analiza cluster/ la 1 s-a renuntat din cauza datelor
- Analiza univariata, bivariata, multivariata
- = toate tehnicile statistice care analizeaza simultan valorile mai multor variabile pentru fiecare din cazurile sau obiectele din esantion / variabile interdependente
- = explica variatia variabilelor
- AMV investigheaza modul specific in care variatia unui set de variabile influenteaza variatia altor variabile
  
- Tehnici de dependenta
- Tehnici de interdependenta

## Tehnici de dependenta

- Variabile dependente explicate/determinate de variabile independente: pentru a prezice valorile variabilei dependente si pt a intelege forma si masura determinarii acesteia de catre variabilele independente

## Tehnici de interdependenta

- Nu presupune o diferenta de statut ontologic al variabilelor (D/I)
- Folosite pentru a identifica structura datelor, fie prin reducerea variabilelor, fie prin gruparea obiectelor sau a cazurilor fie prin reprezentarea relatiilor dintre obiecte si attributele lor pe o harta perceptuala



Tehnici de dependenta [cate variabile sunt explicate/prezise]	Tehnici de interdependenta [structura relatiilor intre]
<b>1. Relatii multiple intre VD si VI</b> Modele cu ecuatii structurale/LISREL Analiza PATH	<b>1. Variabile</b> Analiza factoriala/componentelor principale
<b>2. Cateva VD in relatii individuale</b> <b>2.1. metrica</b> Analiza de corelatie canonica (VI metrice) MANOVA (VI non-metrice) <b>2.2. non-metrica</b> Analiza de corelatie canonica (V dummy)	<b>2. Cazuri</b> Analiza cluster
<b>3. O singura VD intr-o singura ecuatie</b> <b>3.1. metrica</b> Regresie multipla liniara si non-liniara Regresie cu date categoriale Analiza conjoint <b>3.2. non-metrica</b> Analiza discriminativa multipla Analiza loglineara	<b>3. Obiecte</b> <b>3.1. metric</b> Scalare multidimensionala <b>3.2. non-metric</b> Analiza de corespondenta/de omogenitate



- Analiza factoriala
- Analiza cluster
- Scalare multidimensională
- Analiza de corespondență/de omogenitate
  
- Abordare multicriterială în loc de cluster (Structural Change and Globalisation)
  
- Folosire complementar (factor+cluster+scalare(?))
  
- SPSS

= o colectie de metode statistice multivariate interdependente al carui scop principal este identificarea a unui set de date care descriu un concept ori un fenomen sau care caracterizeaza o populatie

= ajuta sa intelegem cum se structureaza o serie de variabile intr-un concept sau un fenomen si ne indica dimensiunile sale

Variabilele “transformate” in factori=variabile neobservabile direct, mai putine – sumarizare si reducere de date (ex. 20 variabile – 4/5 factori)

Variabilele care compun acelasi factor sunt mai puternic corelate intre ele si mai putin cu variabilele care compun alti factori

AF ne indica felul in care se structureaza  $V$  in sub-seturi de  $V$  puternic corelate, fiecare din sub-seturi corespunzand unui factor care poate fi interpretat, denumit si inteles

- Demers exploratoriu – cel mai des: AF ne clarifica, structureaza si simplifica intelegerea felului in care datele covariaza
- Demers confirmativ – ipoteza de lucru este ca un fenomen este explicat de un anumit nr. de factori – AF poate confirma/infirma acest lucru
- !!! O strategie este impartirea setului de date in 2 – un prim set pentru explorare – rezulta un model factorial, al doilea set este utilizat pentru confirmare/validarea modelului!!!

- DATE care caracterizeaza – POZITIA PARTIDELOR fata de anumite – ASPECTE
- Reducerea variabilelor la un număr mai mic de factori care să justifice/descrie diferentele dintre partide și să descrie dimensiunile fundamentale după care se structurează sistemul politic
- K. Janda 1980, 121 partide, ipoteza ca acestea se “clasifica” in 2 “dimensiuni”: stanga-dreapta
- 13 teme: proprietatea asupra mijloacelor de productie și produsului; rolul guvernului în planificarea economică; susținerea pentru armată, redistribuirea veniturilor, asistenta sociala, secularizarea societatii, alinierea cu blocurile Est- Vest, anticolonialismul, integrarea supranațională, integrarea națională, participare electorală, protectia drepturilor civile, interferența cu libertățile civile
- 4 factori: dimensiunea stanga economica, dimensiunea razboiul rece, dimensiunea liberala, dimensiunea 4 - drepturile civile
- a respins ipoteza ca o singura dimensiune stanga/dreapta ar putea explica diferentele dintre partide

- avem: obiecte (cazuri, ex. proiecte); variabile observabile VO (attribute, categorii, criterii); factori F (variabile latente) ( $F < VO$ ) (VO=variabile dependente/F=variabile independente – in tehnici de dependenta)
  
- Pasi:
  1. Matrice factoriala
  2. Model factorial
  3. Realizarea AF

- = matricea alcatuita din saturatiile factoriale pentru fiecare VO

**Saturatie factoriala**= saturatii ale F pentru VO (comparativi cu coeficienti de regresie standardizati)

**Pas principal** – SF ne spune in ce masura un factor “determina” o VO, ce factori determina cu preponderenta o VO si ce variabile “satureaza” cu preponderenta un factor

Cu alte cuvinte, saturatia factoriala ne indica *relatiile* dintre F si VO, precum si *intensitatea* lor

**Covariatie=corelatie**= daca doua VO (variatii ale VO) sunt determinate de acelasi (1) factor, cand F variaza, si cele 2 VO variaza, dar “corelat”.

**Comunalitate**= unei VO cu F comun este acea parte din varianța sa care se datoreaza F comun (vs. unicitate)

**Varianța**=este media pătratului abaterii variabilei de la media sa

- Dupa matricea factoriala avem o serii de relatii care pun in legatura corelatiile dintre VO (pe de o parte) si saturatiile factoriale (pe de alta parte)
- =Punctul de pornire pt modelul factorial
  
- Complexitatea factoriala a unei variabile=nr de factori care au saturatii nenule pt VO= nr de factori care determina VO respectiva
- Grad de determinare a VO=masura in care VO sunt determinate de F comun
  
- Aceste 2 aspecte determina alegerea celei mai bune solutii factoriale

- Model factorial ortogonal = F independenti intre ei
- Model factorial oblic = F corelati
- In realitate procesul este invers – nu cunoastem nr de factori, ci incercam sa ii calculam pornind de la VO ale obiectelor prinse in analiza
- Dorim sa obtinem structura factoriala (matricea saturatiilor, corelatiile F-VO si F-F, daca exista, dimensiunea lor) pornind de la corelatiile (covariatiile) cunoscute dintre variabilele observate
- Procesul contine “nedeterminari” (N) – cunoasterea COV nu duce imediat la cunoasterea F
- Reducerea N – verificare prin “teorie”!!!, SIMPLITATE (teorema rang), rotatie



1. Definirea problemei conceptuale
2. Matricea de corelatie
3. Extragerea factorilor
4. Rotatia factorilor
5. Interpretarea factorilor
6. Validarea analizei factoriale

## 1. Definirea problemei conceptuale

- Obiective
- Pregatirea setului de date
- Problema de cercetare - ipoteze

## 2. Matricea de corelatie

- Testam daca exista destula corelatie intre variabile:

(1) prezenta corelatiei (testul de sfericitate Bartlet),

(2) examinarea coeficientilor de corelatie partiali,

(3) masuri de adecvare ale esantionarii\* (Kaiser-Meyer-Olkin KMO)

### 3. Extragerea factorilor

Simplu cand stim F si VO – AF – procesul invers, mai complicat

Pe baza de criterii practice si statistice, estimam F

Stabilirea unui numar minim de F comuni care sa produca in mod satisfacator corelatiile dintre variabile = rangul matricei de corelatie ajustata

Se extrag atatia factori pana cand discrepanta dintre COR produse si cele observate sunt destul de mici pt a fi acceptate

#### METODE:

- (a) celor mai mici patrate
- (b) probabilitatii maxime
- (c) de extragere factoriala Alpha
- (d) analiza imaginii
- (e) factorilor principali
- (f) componentelor principale

## 4. Rotatia factorilor

Cand relatiile dintre factori si variabile nu este clara sau usor interpretabila

Sistemul de axe ortogonale/oblice reprezentat de  $F$  este rotit in jurul originii intr-o alta pozitie – varianta distribuita mai clar  
varimax; quartimax, equamax/direct oblimin (q), promax (v)

## 5. Interpretarea factorilor

## 6. Validarea analizei factoriale

- Impartirea in 2 esantioane: explorator si confirmator
- Adaugarea unui esantion

- Esentiale calitatea datelor si criteriile de definit pt stratificarea si reducerea variabilelor
- Analiza de “jos in sus” – exploratoriu/confirmativ
- Variabile corelate - factori ( $V > F$ ) – nu stim factorii – acestia se identifica pe baza corelatiilor (covariatiilor) cunoscute/identificate intre variabilele observate – matricea corelatiilor
- Matrice factoriala – calcularea saturatiilor factoriale – ce si in ce masura variabilele “satureaza” (influenta) un factor – dupa care verificare pe baza de “cunoastere” si extragerea factorilor – rotim pt a clarifica - modelul factorial
- Interpretam factorii – validam modelul
- SPSS

= varietate de proceduri folosite pentru obtinerea unei clasificari

= multimea metodelor de grupare a obiectelor pe baza similaritatii caracteristicilor acestora

Obiectiv:

Identificarea structurii unui set de date alcatuit dintr-o multime de obiecte si caracteristicile lor, prin definirea unor grupuri de obiecte, cat mai omogene intern si eterogene extern, in functie de similaritatea lor in raport cu anumite caracteristici

Numita si analiza Q (obiectele si relatiile dintre ele), tipologie, analiza de clasificare, taxonomie

Metoda de reducere a datelor – clase de omogenitate (nu factori latenti ca AF)

- Furnizeaza clasificari/tipologii
- Formularea unor scheme conceptuale
- Generare de ipoteze
- Testare de ipoteze

- Mai simplu decat AF, nu implica conditii statistice substantiale
- Componenta principala – algoritmul de grupare, dat de o regula simpla de alaturare, in pasi succesivi, a obiectelor (cazurilor) asemanatoare
- Procedurile difera in functie de: caracteristicile obiectelor, intensitatea asemanarilor dintre acestea, masura de similaritate folosita, forma algoritmului de grupare, numarul de grupuri obtinute



1. Formularea problemei de cercetare si stabilirea scopului analizei cluster
2. Selectia variabilelor care servesc drept criterii de grupare
3. Masuri de similaritate a obiectelor (explorare a datelor)
4. Metode (algoritmi de grupare)
5. Stabilirea numarului de grupuri
6. Interpretarea grupurilor

## 1. Formularea problemei de cercetare si stabilirea scopului analizei cluster

### Forma AC depinde de formularea PC (ipoteze) si a obiectivelor

Clasificarea obtinuta determinata de *forma datelor*, alegerea *criteriilor de grupare (variabile)*, *metoda de grupare* si de *definitia masurii de similaritate intre obiecte* (INTERREG ex-post – clasificari anterioare)

## 2. Selectia variabilelor care servesc drept criterii de grupare

### Dupa considerente teoretice/conceptuale si practice

In acord cu ipotezele cercetarii (teorie, cercetare anterioara, common sense\*)

AC nu distinge variabilele relevante de cele irelevante, ci doar grupeaza cazurile in functie de toate variabilele incluse in analiza – daca variabile irelevante – rezultatul AC alterat – AF sau selectie calitativa

Ponderea variabilelor – ex. PIB/capita vs. life exp - standardizate

- 2 obiecte sunt similare daca au masuri apropiate pentru caracteristicile in functie de care sunt comparate
- Simplu pt 1 caracteristica – mai multe? – tipul de date
- Estimările cantitative ale similarității dominate de conceptul de metrica (“cantitativ”): majoritatea masurilor de similaritate (MS) folosite in AC sunt metrice: cazurile reprezentate ca puncte in spatiu de coordonate=S sau D corespund distantelor metrice dintre ele (grafic flipchart)
- MS pt variabile cantitative: (1) coeficienti de corelatie si (2) distante
- MS pt variabile categoriale: (3) coeficienti de asociere



## Coeficienti de corelatie

= o masura intuitiva a similaritatii dintre doua obiecte (0-1);

Corelatia este o masura a similaritatii dintre profilurile celor doua obiecte dupa un numar de variabile

Ignora magnitudinea valorilor luate de obiecte pt variabile/fara inteles statistic

## Distante (cele mai comune)

=reprezinta similaritatea ca proximitate a obiectelor intr-un sistem de coordonate definit de variabile (cu cat mai apropiate, cu atat mai similare)

Distanta euclidiană (linie dreapta)

Patratal distantei euclidiene (evita extragerea radicalului DE)

Distanta Manhattan/city block (Cebisev), Distanta Minkowski, Dputere, Mahalanobis



## Coeficienti de asociere

Compararea obiectelor ale caror caracteristici sunt masurate pe scale non-metrice (“calitative” - nominale sau ordinale)

Stabilesc gradul de corespondenta intre obiecte in functie de prezenta sau absenta unor insusiri (V dihotomice)

Peste 30 de metode

CA simpla

Coeficientul lui Jaccard

Coeficientul lui Gower

- Pentru a contracara efectele I, tarile trebuie sa poata anticipa fenomenul – cine imigreaza de unde – prin raspunsuri la intrebari .: Care sunt regiunile de origine? Care sunt caracteristicile culturale și sociale ale populatiei de imigranți? La ce volum de populație se pot aștepta? Care va fi statutul legal al potențialilor imigranti?
- Identificare de grupuri de regiuni cu potential de emigratie catre EV si care ar ridica probleme economice, sociale și culturale similare statelor-gazdă (problema de cercetare)
- AC pentru 109 tari, in functie de o serie de variabile demografice, ec, soc, pol.,
- Volumul populatiei, speranta de viata, PIB, PIB/locuitor, rata mortalitatii, natalitatii etc.
- Explorare a datelor – similaritati, corelatii, asocieri (imaginile de mai jos)
- (va continua dupa (4) modele de grupare)

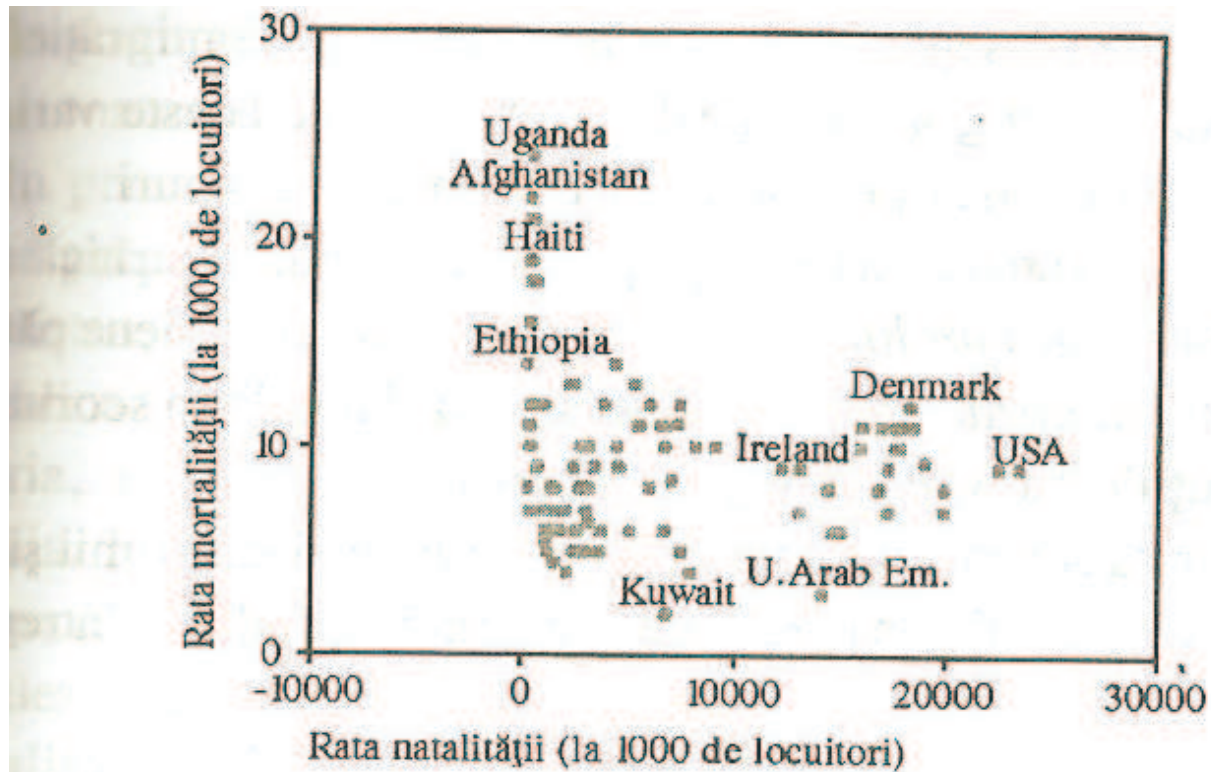


Figura 5. *Reprezentarea a 109 țări în funcție de PIB per capita și rata mortalității*

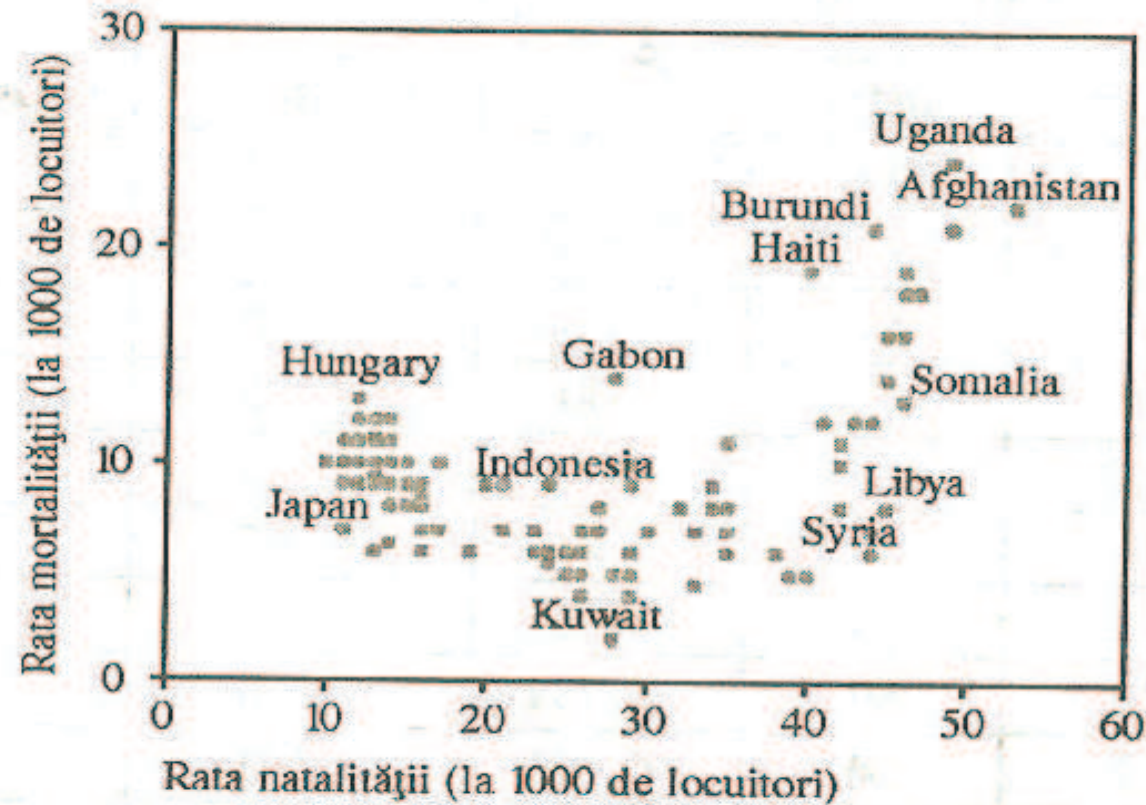


Figura 6. *Reprezentarea a 109 țări în funcție de rata natalității și rata mortalității*



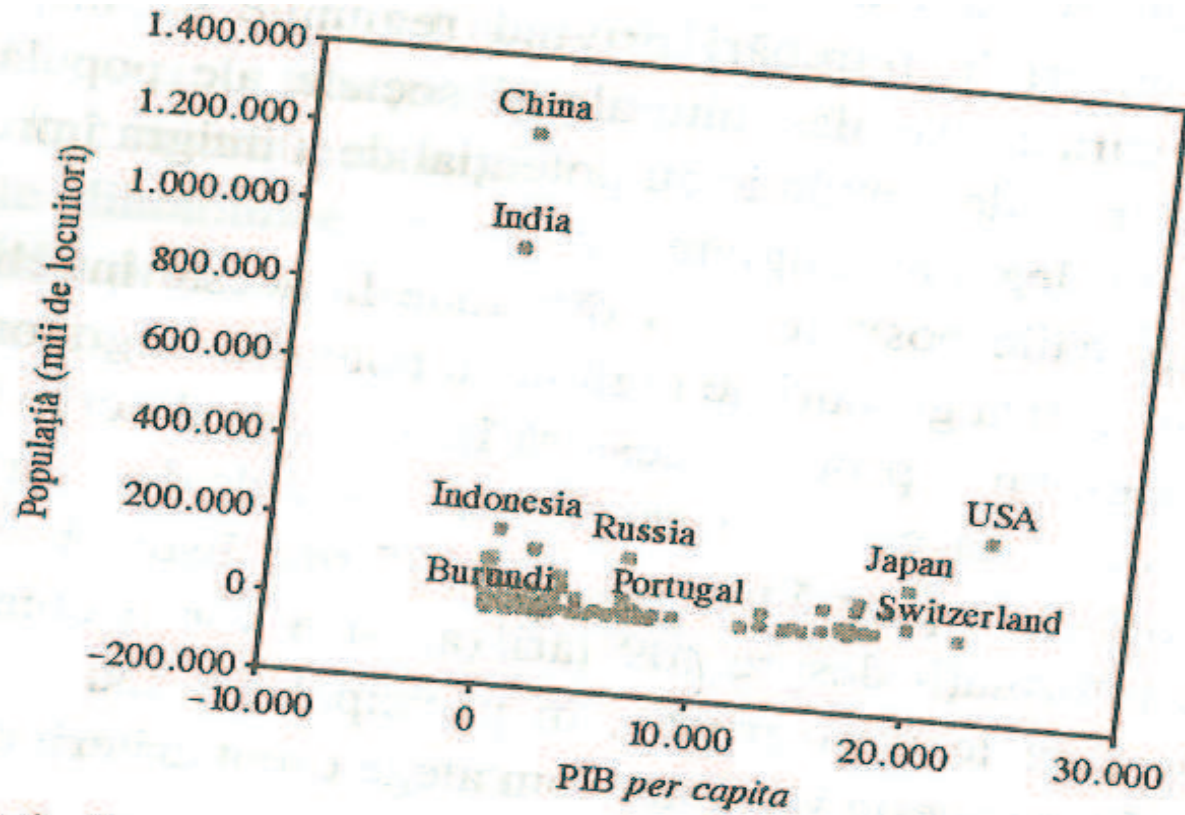


Figura 4. Reprezentarea a 109 țări în funcție de PIB per capita și volumul populației

- La inceput doar se stie ca grupurile de identificat trebuie sa fie omogene
- Nimic despre: nr, densitate, variant(z)a, forma sau gradul de separare intre ele
- Fiecare dintre metode va produce clusteri diferiti pentru acelasi set de date: 7 familii de MdG

(1) Ierarhice aglomerative

(2) Ierarhice divizive

(3) De partitionare iterativa

(4) De densitate

(5) Factoriale

(6) Non-exclusive (clumping)

(7) Bazate pe teoria grafurilor

- Cele mai populare (de jos in sus)
- Se porneste de la obiect-grup; la fiecare pas sunt alaturate doua dintre grupuri intr-unul singur, conform unui criteriu particular, pe baza unei matrici simetrice de similaritati
- Se unesc grupurile intre care distanta este cea mai mica
- Distanța: single linkage/complet linkage/average linkage between and within groups /centroid/Ward

- Pornesc de la o impartire initiala a obiectelor intr-un nr. de grupuri
- Partitionare a obiectului intr-un nr de grupuri; se calculeaza centroidul (centrul de cluster) pt fiecare grup; se re-organizeaza grupurile prin re-distribuirea obiectelor centroidului celui mai apropiat
- Procedura se repeta pana nu se mai produc schimbari in componenta grupurilor
  
- Avantaje fata de IA
- lucreaza direct cu datele brute, nu mai trebuie calculate matricea distantelor – se pot manevara seturi de date mai voluminoase
- Parcurgerea de mai multe ori a setului de date
- Se obtin grupuri “independente”
- PRB: alegerea centroizilor initiali (aleatoriu)

- Analiza factoriala inversa/Q-type factoring/factor analysis variants – psihologie
- Folosite matrice de similaritati intre obiecte (nu de corelatii intre variabile) pentru obtinerea unui nr. de factori, dupa regulile AF clasice, iar obiectele sunt alocate in grupuri conform scorurilor lor factoriale
- Nu in soft-uri statistice

Pt a selecta MG cea mai potrivita - examinarea datelor + identificarea outliers



## 5. Stabilirea numarului de grupuri

Examinarea distantelor intre clusteri la pasi succesivi : cand distanta creste brusc =la pasul respectiv sunt unite 2 grupuri sensibil diferite (revers si incheiere)

Daca distanta dintre grupuri e mica, putem considera unirea

Mai multe solutii produse (pasi/grupari)

## 6. Interpretarea grupurilor

Examinarea valorilor pe care le iau variabilele de grupare pt obiectele din fiecare grup rezultat, pt a intelege natura grupurilor

Analiza centroizilor

Profiluri/etichete\*

- Metoda ierarhica aglomerativa – distante euclidiene patrute (cea mai simpla!)
- Rezulta 4 grupuri: 20+63+23+2
- I: Afganistan, Bangladesh, Burkina Faso, Burundi, Cambodgia, Camerun, Republica Central-Africană, Etiopia, Gambia, Haiti, Kenya, Liberia, Nigeria, Pakistan, Rwanda, Senegal, Somalia, Tanzania, Uganda, Zambia
- II: tarile din Europa de Est și fosta Uniune Sovietică, America Latină, Orientul Mijlociu și câteva țări mai dezvoltate din Africa – factorul economic principal pt emigratie/educatie – cel mai numeros
- III: țările capitaliste din Vestul Europei și America de Nord, precum și tigrii asiatici
- IV: China și India – volum + situatie economica si politica!
- Analiza se poate relua prin metoda de partitionare iterativa/2 parti!
- K-means= SPSS metode de partitionare iterativa
- Hierarchical clusters= SPSS metode ierarhice aglomerative

- Tipologii, concepte ipoteze – formulare si validare
- Componenta principala – algoritmul de grupare, dat de o regula simpla de alaturare, in pasi succesivi, a obiectelor (cazurilor) asemanatoare
- Selectia variabilelor care servesc drept criterii de grupare fi importanta – standardizarea!
- Masura de similaritate: coeficienti de corelatie, distanta, asociere
- Metode (algoritmi de grupare): (1) Ierarhice aglomerative (mai simpla); (3) De partitionare iterative (mai multe date si seturi de analize); (5) Factoriale
- Stabilirea si analizarea grupurilor



Criteriile mentionate in caietul de sarcini:

- be successful in terms of contributing to the Lisbon and Gothenburg agendas
- facilitate the uptake of new ideas and thus contribute to the improvement of regional policies
- trigger larger joint implementation actions or investment
- be improving people's daily lives as well as
- bring people together in a "European experience", bringing Europe closer to the citizen and making regions look beyond their own horizon.

In implementare:

“ Programmes and projects were selected on an intuitive basis to reflect as far as possible the diversity of situations, while pinpointing some ‘success stories’ at the project level” – **CONVENABIL** – anexa 1 a studiului

Desk-based research only

- Nici un criteriu mentionat in CdS

Am propus/agreat 10 proiecte - IR

- One case study for each of the 10 sub-themes listed under the two programme priorities (i.e. 4 case studies for Priority 1 and 6 case studies for Priority 2) – “cazul tipic”/”reprezentativ” dpdv AP/”cluster”
- Per priority, at least one case study is realised on a capitalisation/fast track project (2 in total) – cazul “critic”
- The remaining 8 case studies will focus on regional initiative projects. Under each priority, a balanced representation of **regional initiative** projects showing different intensities of co-operation will be made (braket)

Realizat:

- General principles for selection (main criteria): Priority & sub-theme, Type of project, Level of intensity (only RIP)
- Additional criteria: Funding volume (total budget), Lead Partner (country), Approved under 1<sup>st</sup> or 2<sup>nd</sup> call

- Instrument introdus ceva mai tarziu si oarecum aleatoriu
  - 4 studii de caz – scop ilustrativ/detalii/bune practici
  - Selectia - impreuna cu AMPOSDRU si expertii non-cheie
  
  - Nu sunt mentionate in Raportul Final – nu au fost folosite?
  - VALOROASE – expertul coordonator
1. Project history & description
  2. Local and regional social, economic and political context
  3. Project Implementation
  4. Results
  5. Sustainability and Transferability
  6. Key success factors and lessons learned

- “Evaluarea tematică a vizat identificarea a opt proiecte de succes și realizarea unui raport privind aceste intervenții sprijinite prin programul Phare care, atât individual, cât și colectiv, ilustrează realizările pozitive ale programului Phare în România.”

PROCES riguros (?)

Criteriile utilizate pentru a defini o poveste de succes au fost:

- intervenția să influențeze pozitiv un grup extins de persoane.
- intervenția să fi influențat pozitiv o politică guvernamentală sau o strategie.
- ca intervenția să contribuie la dezvoltarea economică și socială a unei regiuni.
  
- SCOP – povesti de succes (intervenție, nu proiect) – descriptive – ilustrative
- Colectie de bune practici – sinteza de doua pagini



- CdS nu a cerut analiza cluster – a nominalizat 22 programe si a solicitat o selectie justificata a 12, cu conditia ca pt fiecare SM sa fie ales cel putin 1 program
- Raportul initial (p 26) – 4 indicatori, nu AC, propunere (populatia totala si dimensiunea asistentei primite (% al populatiei care locuieste in zona asistata), natura urbana/rurala a zonei asistate, FEDER in zonele asistate)
- Criterii calitative: importanta/dimensiunea unei anumite industrii in zona, potential competitiv
- Raportul intermediar – nici o mentiune despre analiza cluster
- Final report – date pt indicatorii relevanti nu existau la nivel NUTS 3, unde erau, in multe cazuri, localizate programele – granitele acestora nu coincideau cu cele ale regiunii NUTS 2
- Exemplu interesant de analiza a impactului, dar nu poveste de succes in ceea ce priveste selectarea studiilor de caz prin AMI

Task-ul 4 – Dezvoltarea unui grid analitic si a unei tipologii

- Indicator sintetic –intensitatea&profundimea cooperarii/gradul de inter-regionalitate – pana la 6 nivele - o serie de aspecte indicate in CdS
- CdS cere analiza factoriala/cluster - pt 6 A, 3 B, 2 C
- Indicele sintetic: 6 criterii, 16 sub-criterii, 19 sub-indicatori
- Datele au fost colectate din documente si prin sondaj (raspuns 73 din 81 de programe – programele care nu au raspuns au fost nominalizate in raport)
- 95% din date disponibile – s-au verificat in masura posibila

**Overview Table 3. 4.1:**

Step 1 Assessment Framework		Step 2 Component Results		Step 3 Synthetic Result	
Main Criteria (Components)	Related sub-criteria	Related sub-indicators (SI) & intensity level scoring values (ILSV)		Summing up of the ILSV for each component	For all programmes: Calculation of the "Real Rate", comparison of "Real Rate" & "Expected Rate" and scale placement according to weighted / non-weighted values
		SI	ILSV		
Historical Criterion	1.1.	SI 1 (quantitative)	0-100	$\Sigma$ ILSV (SI 1 - SI 3)	<p>Calculation of the "Real Rate" (i.e. the gross non-weighted value) according to:</p> $RR = \frac{1}{16} \cdot \sum_{i=4}^{19} SI_i$ <p>Comparative analysis between "Real Rate" and "Expected Rate" according to:</p> <ul style="list-style-type: none"> <li>Option 1: Estimating the coefficients of the historical variables on the real rate within a linear single equation regression (e.g. <math>Y=b_0+b_1X_1+b_2X_2+b_3X_3+u</math>).</li> <li>Option 2: Running separate linear single regressions for each of the historical variables (in case of multicollinearity under Option 1, which leads to high standard deviations of the regression coefficients).</li> </ul> <p>Scale placement of all programmes according to their weighted values and non-weighted values.</p>
	1.2.	SI 2 (qualitative)	0-100		
	1.3.	SI 3 (qualitative)	0-100		
INTERREG Criterion 1	2.1.	SI 4 (qualitative)	0-100	$\Sigma$ ILSV (SI 4 - SI 5)	
	2.2.	SI 5 (qualitative)	0-100		
INTERREG Criterion 2	3.1.	SI 6 (qualitative)	0-100	$\Sigma$ ILSV (SI 6 - SI 8)	
	3.2.	SI 7 (qualitative)	0-100		
	3.3.	SI 8 (qualitative)	0-100		
INTERREG Criterion 3	4.1.	SI 9 (quantitative)	0-100	$\Sigma$ ILSV (SI 9 - SI 11)	
	4.2.	SI 10 (quantitative)	0-100		
	4.3.	SI 11 (quantitative)	0-100		
INTERREG Criterion 4	5.1	SI 12 (quantitative)	0-100	$\Sigma$ ILSV (SI 12 - SI 17)	
		SI 13 (quantitative)	0-100		
	5.2	SI 14 (quantitative)	0-100		
		SI 15 (quantitative)	0-100		
	5.3	SI 16 (quantitative)	0-100		
INTERREG Criterion 5	6.1	SI 18: 18a, 18b, 18c (quantitative)	0-100	$\Sigma$ ILSV (SI 18 - SI 19)	
	6.2	SI 19 (quantitative)	0-100		



- Indicatorul sintetic – una din cele 4 componente ale metodologiei pentru dezvoltarea unei clasificari generale a programelor (overall tipology concept)
- (1) Factori (4) principali care sunt foarte relevanti pentru aspect-cheie ale INTERREG – per strand
  - (2) Pentru fiecare factor au fost definiti indicatori – pentru care existau date, colectate in cadrul task-ului 2, si indicatorul sintetic (A – 13 indicatori, B-11 indicatori, C-8)
  - (3) Datele nu au fost complete, dar stabilitatea bazei de date a fost ridicata
  - (4) Analiza cluster prin K-means a rezultat in 6 grupuri pt Strandul A, 3 pentru Strandul B, 1 pentru Strand-ul C



- Strand A – in etape din cauza datelor care variau de la program la program si a lipsei de date pt unele programe
- K-means - partitionare iterativa – valorile indicatorilor – in ordinale, din nominale (ex. A 1.1 si A 1.2. )
- Datele au fost “standardizate” – pt ca unele sa nu influenteze mai mult rezultatul, din cauza valorii lor (ex. PIB vs. life expectancy)
- S-a plecat de la 6 clusteri – nr. mare de programe – s-a masurat distanta pe baza datelor standardizate - **SPSS a ales centrozii!**
- Prima analiza K-m=baseline, rezultate acceptabile, nici un cluster cu nr mic de cazuri + 3 alte analize, fiecare fara o serie de indicatori/programe (pt care datele erau mai putin disponibile)
- Rezultatele acceptabile in toate cazurile, s-au pastrat centrozii alesi de SPSS – (K-means sensibil la outliers)
- Test ANOVA – identifica variabilele care determina mai mult o grupare pt. fiecare analiza
- in total 4 analize, fiecare in 2 pasi: MS – distantele + 2 analize pt 5/4 cluster

- (Î) **Cum** influențează capacitatea de implementare a beneficiarilor obținerea rezultatelor proiectelor finanțate din IS?
- (SÎ1) **Cum** afectează politicile, legislația, relațiile de putere și normele sociale relevante capacitatea de implementare a beneficiarilor?
- (SÎ2) **Cum** este afectată capacitatea de implementare a beneficiarilor de IS de politicile interne, regulamentele, procedurile și cadrul organizațional (dacă acestea există)?
- (SÎ3) Au beneficiarii experiența, cunoștințele și abilitățile tehnice relevante pentru implementarea proiectelor finanțate din IS?
- **3 SdC** – pt. a rafina analiza și ilustra mai bine rezultatele și recomandările evaluării
- **Teme !:** (i) dificultățile financiare ale beneficiarilor din sectorul privat (IMM-uri), (ii) motivația resurselor umane și beneficiarii din sectorul public și (iii) factorii externi și capacitatea de implementare a beneficiarilor
- “Practica evaluării a evidențiat faptul că pentru studiile pe tema capacității beneficiarilor, studiile calitative în profunzime bazate pe clustere de studii de caz sunt mai potrivite, din moment ce analiza cantitativă aduce doar o perspectivă parțială asupra acestui subiect”

- Studii de caz regionale: Regiunea Centru, Regiunea Nord Vest, Regiunea Sud Vest – sa reflecte experiența regiunilor în ceea ce privește coordonarea și elaborarea strategiilor și a planurilor/programelor regionale de dezvoltare.

**Criteriul de selectie** - gradul de identificare a anumitor practici regionale care pot constitui surse de “lecții” pentru viitoarea perioadă de programare.

- Studii dpdv al structuri instituționale din domeniul dezvoltării regionale ale altor state membre: Italia, Spania, Polonia, Republica Cehă, Ungaria

**Criteriul de selectie** - exemple de structuri instituționale în domeniul dezvoltării regionale

- Anexa II GAO ( pagina 108)
- Work Package 6a: enterprise and innovation – folosit analiza cluster, dar nu pentru a selecta studii de caz, exploratic si ilustrativ
- Work Package 9: Rural Development – tipologie a zonelor rurale, dar calitativ
- Work Package 9: Efficiency – unit costs of major projects – probleme in colectarea datelor – nu s-a putut realiza analiza statistica
- Ex-post Analysis of URBAN – s-a folosit analiza cluster - datele trebuie sa fie create/culese pt a putea aplica aceasta metoda

Avem: caiet de sarcini din AL 1

Sarcina 1: completarea CdS cu cerintele vis-à-vis de metodologie referitoare la SdC – cate, instance/unitatea de analiza, metoda de selectie – cat de concret?

Sarcina 2: selectam studii de caz

2.a. selectam/dezvoltam criterii “calitative” si de ce

2.b. dezvoltam factori/indicatori pe baza carora se pot colecta datele si realiza o analiza cluster

2.c. Selectarea unui caz de studiat in AL viitoare

- 3 perspective (bases) prin care se selecteaza cazurile de studiat (convenabil, scop, probabilitate)
- Scop - extreme; exemple de succes; exemple de insucces; clustere; reprezentativitate; cazuri tipice; cazuri speciale
- Tipul de SdC si selectia cazurilor
- Reprezentativitate/esantionare statistica
- Analiza multivariata – tehnici de dependenta; interdependenta
- Analiza factoriala, cluster, scalare multidimensionala, analiza de corespondenta/de omogenitate
- Exemple si atelier de lucru 2
  
- Este necesar sa folosim metode statistice pentru selectarea cazurilor de studiat?

- Ce v-a placut?
- Ce nu v-a placut?
- Ce ati pastra?
- Ce nu ati pastra?
- Ce ati aprofunda?