POSIBILITĂŢI DE ANALIZĂ FINANCIARĂ MULTIDIMENSIONALĂ OFERITE DE CĂTRE O NOUĂ GENERAŢIE DE SISTEME INFORMATICE DE ASISTARE A DECIZIEI (SIAD) DE TIP DATA MINING

CONF. UNIV. DR. IONESCU IULIANA
Academia de Studii Economice Bucuresti, Adresa de corespondenta: Str 9 Mai nr. 5, bl. 36, sc.B, et. 2, ap. 25, sector 6, Bucureşti, tel/fax 0212213446, mobil 0723 320468, e-mail ionescu.iuliana@gmail.com

PROF. UNIV. DR. IONESCU BOGDAN
Academia de Studii Economice Bucuresti, Adresa de corespondenta: Str 9 Mai nr. 5, bl. 36, sc.B, et. 2, ap. 25, sector 6, Bucureşti, tel/fax 0212213446, mobil 0722 593695, e-mail ionescub@gmail.com

Possibilités d'analyse financière multidimensionnelle parmi une nouvelle génération des SIAD de type Data Mining
Les plus répandus outils informatiques pour mettre en évidence des liaisons dissimules entre les données sont les logiciels statistiques d'analyse de données. Ces logiciels peuvent conférer une dimension supplémentaire à l'analyse financière effectuée avec les modules classiques des Systèmes Interactifs d' Aide à la Décision, pour interpréter des grosses masses des données financiers. Le contexte multidimensionnel des données et l'analyse des données (Data Mining) du SIAD sont appliques dans un exemple d'analyse des plusieurs entreprises qui appartiennent au même secteur d'activité, qui transpose le modèle en trois dimensions: les variables d'analyse, les sociétés et les dimensions temporelles (les ans). L'analyse multidimensionnelle peut donnée un possible interprétation de la situation financière de l'entreprise, en partant d'un modèle qui se fonde sur un system de rapports entre plusieurs indicateurs.

Cele mai adecvate instrumente informatice pentru punerea în evidenţă a legăturilor ascunse dintre date sunt *aplicaţiile statistice specializate*. Majoritatea acestor instrumente, datorită aspectului lor prea specializat sunt complicate şi deci dificil de utilizat. Motivele acestei calificări rezidă în cel puţin trei motive: necesită cunoştinţe de specialitate, necesită o configuraţie informatică performantă şi costă destul de mult.

Un instrument mai puţin sofisticat din punct de vedere al utilizării, dar foarte apropiat de performanţele unor aplicaţii statistice de analiză a datelor este MULTIWIN. Acest produs informatic a fost derivat din pachetul SPSS, a fost simplificat şi a fost adaptat cerinţelor de utilizare din mediul universitar. Aplicaţia MultiWin va aduce o dimensiune suplimentară analizei financiare efectuate cu ajutorul modelelor clasice de SIAD, aceea de a interpreta multidimensional mari volume de date. Cu alte cuvinte, SIAD-ul pentru analiză financiară evoluează, trecând într-un nou stadiu de dezvoltare şi reprezentare, acela conferit de contextul Data Mining.

Aplicaţia MultiWin[24] conţine următoarele module principale: regresia multiplă pas cu pas; ajustarea matematică; analiza în componente principale; analiza factorială a corespondenţelor; analiza discriminantă.

Regresia multiplă pas cu pas constă în căutarea, plecând de la un ansamblu de observaţii, a ecuaţiei matematice care permite legarea a două variabile. Cunoscând valoarea luată de o variabilă, se poate calcula rapid valoarea celeilalte. Regresia multiplă extinde această noţiune la nivelul mai multor variabile, căutând o relaţie între o variabilă de explicat şi mai multe variabile explicative, clasându-le într-o ordine descrescătoare a capacităţilor lor explicative.

Ajustarea matematică este, în mod particular, interesantă pentru a genera modele explicative sau previzionale, mai ales în zona de influenţă a sistemelor de asistare a deciziei. Ajustarea matematică presupune, ca şi regresia, căutarea unei ecuaţii matematice care leagă două variabile, plecând de la un ansamblu de observaţii.

Analiza în componente principale (ACP) rezidă în faptul de a căuta dacă variaţia observată asupra variabilelor $X_1,X_2,X_3,\ldots\ldots X_n$, poate fi explicată printr-un număr redus de variabile necorelate, numite factori. Reprezentarea sub formă grafică reprezintă elementul principal ce permite reprezentarea de o manieră sintetică a maximum de informaţii conţinute într-un tablou de date unde: liniile corespund "indivizilor" (în sens statistic) studiaţi, iar coloanele corespund măsurătorilor efectuate asupra variabilelor cantitative, caracteristice fenomenelor studiate.

Analiza factorială a corespondenţelor (AFC) este, în principal, destinată să prelucreze şi să interpreteze tablouri de valori reieşite din încrucişarea a două variabile repartizate în clase. Aceste tablouri sunt numite "tablouri de contingenţă". Obiectivul analizei factoriale este de a căuta dacă există o dependenţă sau o corespondenţă între ansambluri de proprietăţi.

Potrivit unui demers orientat către "Data Mining", vom furniza în continuare o metodologie generală de urmat pentru analiza financiară, metodologie ce foloseşte tehnici statistice de interpretare. După cum s-a putut observa, nu există o ordine strictă de urmat pentru a începe, a continua sau a termina o analiză financiară. La realizarea analizei se poate utiliza o tehnică de investigaţie sau mai multe tehnici, acestea fiind complementare sau chiar concurente.

Pentru a lucra cu o aplicaţie de analiză a datelor, informaţiile trebuie transpuse de către aplicaţia furnizoare de date, într-o formă şi un format adecvat prelucrărilor statistice. De exemplu, dacă s-ar face o analiză financiară multidimensională a mai multor întreprinderi care aparţin aceluiaşi sector de activitate, datele transpuse vor avea trei dimensiuni : variabilele, societăţile comerciale şi anii. Orice aplicaţie (SIAD clasic) poate fi generată cu un procesor de tabele care într-un format adecvat exportului către aplicaţiile de analiză a datelor.

[24] Marcel PISTRE : MultiWin – Modules d'Analyse de Donnee sous Windows. Specimen p.13

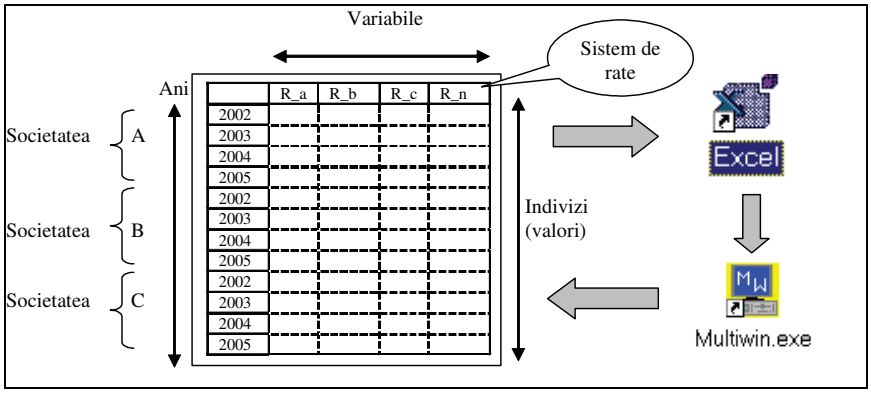


Figura 1 Structură tridimensională a unei matrici de analiză a datelor

Datele se pot calcula în Excel (rate de rentabilitate sau diverşi indici) iar prin intermediul Clipboard-ului pot fi exportate către un o aplicaţie de analiză a datelor.
Prelucrarea statistică a datelor, realizată cu o aplicaţie de analiză a datelor, reprezintă un demers structurat în patru etape:

Prima etapă vizează generarea rezultatelor statistice. Aceste rezultate sunt compuse din trei elemente pentru fiecare variabilă studiată (pentru sistemul de rate): media, ecartul tip şi coeficienţii de variaţie (ecartul tip/medie). Elementele menţionate sunt calculate automat de aplicaţia de analiză a datelor.

A doua etapă este aceea a calculării matricei de corelaţie. Respectiva matrice pune în evidenţă importanţa legăturii între două variabile. Astfel, toate elementele diagonalei sale sunt egale cu 1 (o variabilă este corelată maximal cu ea însăşi).

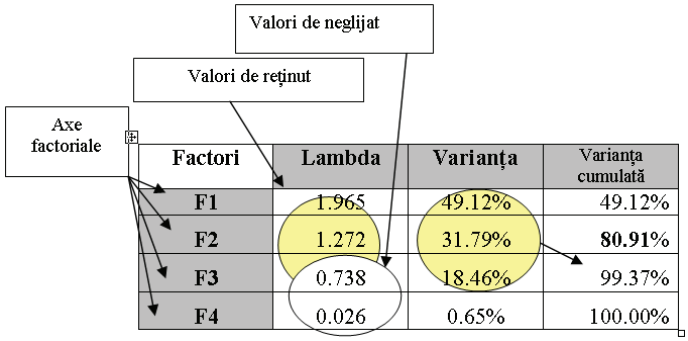
	Rata_"a"	**Rata_"b"**	**Rata_"c"**	**Rata_"n"**
Rata_"a"	**+1**			
Rata_"b"	-0.25	**+1**		
Rata_"c"	-0.01	+0.10	**+1**	
Rata_"n"	+0.41	-0.05	+0.88	**+1**

Tabelul 1 Matricea de corelaţie

Din matricea de corelaţie, exemplificată pe un caz ipotetic, se pot deduce următoarele corelaţii între variabile (între rate), plecând de la postulatul "cu cât valoarea indicelui de corelaţie este mai mare şi mai aproape de 1, cu atât variabilele sunt mai corelate, iar cu cât valoarea indicelui este mai mică în sens negativ, cu atât variabilele au un grad mai mare de libertate (sunt puţin legate între ele, sau deloc)".

- variabila "Rata_a" *nu este corelată deloc* (-0.25) de "Rata_b" (de exemplu "rata autonomiei financiare" faţă de "indicele de creştere economică");
- variabila "Rata_a" *este puternic corelată* (+0.41) cu "Rata_n" (de exemplu "indicele de creştere economică" cu "capacitatea de autofinanţare";
- variabila "Rata_c" *este foarte puternic corelată* (+0.88) cu "Rata_n" (de exemplu "valoarea adăugată" cu "capacitatea de autofinanţare", ambele rate având aceeaşi sorginte).

A treia etapă de parcurs este dată de generarea tabloului valorilor proprii (sau individuale). Acest tablou calculează "varianța valorilor" (sau a indivizilor-în sens statistic) denumită **λ**. Această valoare cuantifică importanța factorilor prin care se explică fenomenul analizat. Variațele **λ** apar în matricea valorilor proprii, în care toate elementele sunt egale cu zero, în afara celor situate pe diagonală a căror sumă este numită "amprentă sau urmă". Fiecare valoare **λ** raportată la valoarea "amprentei" furnizează procentajul varianței totale exprimate.

În general, primele axe factoriale fac să apară un procent mai mare a factorilor explicativi ai varianței, în timp ce ultimele axe factoriale sunt neglijabile, iar neluarea lor în considerare conduce la o pierdere mică de informație.	

Factori	Lambda	Varianța	Varianța cumulată
F1	1.965	49.12%	49.12%
F2	1.272	31.79%	**80.91%**
F3	0.738	18.46%	99.37%
F4	0.026	0.65%	100.00%

Figura 2 Valorile proprii sau individuale

Soluția cea mai simplă constă în a reține numai factorii la care coeficientul **λ** este mai mare ca 1. În exemplul ipotetic considerat, ilustrat în tabelul anterior, varianța cumulată a primilor doi factori explică aproximativ 81% din varianța totală.

A patra etapă o constituie generarea matricei factoriale reduse. Aceasta indică gradul de corelație care există între fiecare variabilă şi factorul corespondent.

Coloana "Comun" explicitează calitatea reprezentării fiecărei variabile prin factorii reținuți (F1+F2). Valoarea "Comun" reprezintă suma pătratelor coeficienților de corelație aferenți factorilor.

Factori	**F1**	**F2**	Comun
Rata_"a"	0.426	-0.711	0.686
Rata_"b"	-0.068	0.788	0.626
Rata_"c"	0.890	0.381	0.937
Rata_"n"	0.993	0.017	0.987

Tabelul 2 Matricea factorială redusă

Reprezentarea grafică este a 5-a etapă a demersului statistic de analiză aratelor. Graficul face să apară simultan, pentru combinația factorilor selectați (automat): **poziția**

variabilelor şi **poziţia indivizilor**. **Poziţia variabilelor** în plan factorial este reprezentată într-un sistem de coordonate polare ce pleacă din centrul (intersecţia) axelor (0,0). Astfel, ***poziţia faţă de cercul de corelaţie unitară, relevă calitatea reprezentării variabilei respective în plan factorial***, *iar poziţia unghiurilor formate de direcţia variabilelor, indică influenţa lor reciprocă*.

Poziţia indivizilor în plan factorial permite *vizualizarea indivizilor (valorilor) care participă cel mai mult la formarea axelor* (indivizii ai căror coordonate sunt de valoarea cea mai mare). În egală măsură, poziţia indivizilor permite *regruparea pe zone a indivizilor ce posedă aceleaşi caracteristici sau caracteristici asemănătoare*.

O astfel de reprezentare grafică, potrivit regulilor analizei în componente principale, este disponibilă în figura următoare, pentru valorile prezentate ca exemplu.

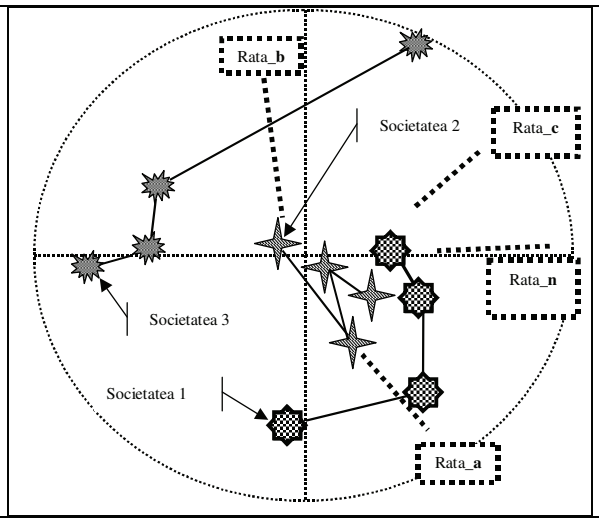


Figura 3 Reprezentarea în componente principale a modelului ratelor

Sistemul de coordonate al reprezentării polare permite interpretarea datelor pentru furnizarea unei stări de fapt a societăţilor analizate prin ansamblul ratelor.
În utilizarea metodei analizei în componente principale, variabilele sunt cu atât mai bine reprezentate cu cât ele se află poziţionate mai aproape de cercul de corelaţie.

Într-o primă abordare, se poate observa comportamentul diferit al societăţilor în raport cu diferitele variabile studiate. Aceasta înseamnă că nu există o uniformitate în ceea ce priveşte importanţa şi semnificaţia ratelor alese pentru a caracteriza mai multe societăţi (acestea au particularităţile lor, din care rezultă aspectele diferite judecate prin rate).
Reprezentarea de mai sus ilustrează o corelaţie foarte importantă între două elemente în raport cu axele factoriale.
Prima axă factorială reprezintă puţin în legătură cu "Rata_n" şi de o manieră mai puţin precisă "Rata_c". A doua axă defineşte "Rata_b" şi, în sens opus, "Rata_a".

Rezultatele se pot interpreta pentru fiecare societate în parte, astfel:
"Societatea C" prezintă o evoluţie semnificativă a elementului definit prin "Rata_c", indicator care a stagnat în primele trei perioade. "Rata_b" este globalmente în creştere, dar pentru "Societatea C" se observă o diminuare constantă a elementului definit prin "Rata_b".

“Societatea B” se identifică prin evoluția ei în “zig-zag”, evoluție inversă cu progresia grafică aferentă celorlalte două societăți. Evoluția respectivei societăți este considerată ca negativă, fiind vizibil ecartul tot mai mare față de elementele semnificative judecate prin cele patru rate.

Poziția “Societății A” în raport cu cei patru indicatori (rate) demonstrează o vizibilă diminuare a elementului exprimat prin “Rata_a”, începând de la a doua perioadă (2003), ca de altfel o stagnare a elementului exprimat prin “Rata_n” şi o creştere a indicatorului “Rata_c”.

Dacă cele trei societăți alese sunt reprezentative pentru sectorul de activitate, analiza în componente principale poate oferi judecăți generale în care se poate circumscrie evoluția oricărei întreprinderi din sector. Acest aspect constituie un pol important al asistării deciziei.

Analiza multidimensională poate fi extinsă pentru judecarea întregului sistem de indicatori ai întreprinderii, operându-se astfel cu *modele de rentabilitate* (rentabilitatea capitalurilor angajate, rata valorii adăugate, rata CAF, rentabilitatea economică, financiară), *de lichiditate* (gradul de lichiditate, solvabilitatea generală, lichiditatea generală, etc), *de structură financiară* (autonomie financiară, levier financiar, acoperirea cheltuielilor financiare, etc).

Analiza multidimensională prezentată a prezentat câteva puncte de vedere în ceea ce priveşte o posibilă interpretare a situației economico-financiare a unei întreprinderi, plecând de la un model ce se bazează pe un sistem de rate. Acest model a plecat de la date agregate (ratele), care prin tehnici specifice de analiză au putut furniza o expertiză independentă de modelele clasice de interpretare ale analizei financiare. Demersul prezentat se integrează perfect în operațiuni de “verificare” a modelelor clasice de interpretare şi face trecerea către operațiuni de “descoperire” care reprezintă o a doua abordare a tehnologiei Data Mining.

Segmentul decizional financiar – contabil, ca parte a deciziei strategice, poate fi influențat pozitiv în sens calitativ, de tehnicile de descoperire a informațiilor ascunse în date.

Datele financiar – contabile, prin caracterul lor exhaustiv, pot forma o imagine de ansamblu asupra funcționării unei întreprinderi (latură Data Mining orientată pe verificare), dar sunt mai sărace în conținut dacă se analizează latura prospectivă a acestora. Plecând de la abundența datelor operativ-contabile şi contabile, se pot descoperi relații noi între date, relații invizibile prin intermediul instrumentelor şi operațiilor clasice cu care operează tehnicile financiar - contabile. Astfel, noile legături “ascunse” dintre date pot conduce la informații noi, care să completeze sistemul informațional – decizional financiar – contabil.

Un exemplu ar putea proba cele afirmate. Dacă s-ar investiga relațiile datelor ce reflectă în perioade diferite, clienți sau furnizori, cu cifra de afaceri sau cu rezultatele financiare, s-ar putea trage concluzii care nu ar reieşi la o analiză clasică. Astfel, dacă anumiți furnizori sunt selectați, iar aceştia oferă componente calitative într-un ciclu de producție, calitatea produselor va fi mai bună, iar produsele vor fi mai vandabile, în consecință vor creşte vânzările. Această analiză poate conduce şi la alte explicații ce țin de ritmicitatea

aprovizionărilor, de pierderile cauzate de transport, de întreruperile ciclului de fabricaţie datorită calităţii necorespunzătoare a reperelor furnizate.

Creşterea cifrei de afaceri poate fi afectată în egală măsură şi de calitatea clienţilor (aceştia sunt sau nu cunoscuţi pe piaţă), de a căror influenţă pot depinde şi rezultatele financiare.

Datele ce reprezintă clienţii pot conduce la explicaţii referitoare la rezultatele financiare înregistrate. În urma colaborării cu anumiţi clienţi, o cifră de afaceri mare poate fi coroborată cu un rezultat financiar semnificativ, dacă marja de profit este mare, iar clientul este cunoscut în sectorul distribuţiei. De asemenea, rezultatele financiare por fi minime, chiar în condiţiile unei cifre de afaceri mari, dacă marja profitului este mică (negociere defectuoasă sau existenţa unor interese particulare), in condiţiile în care clientul este sau nu este cunoscut pe piaţă.

Tehnicile Data Mining orientate pe "descoperire" pot surprinde astfel de legături ascunse, care influenţează uneori decisiv sistemul decizional al unei firme, plecându-se de la date financiar – contabile.

Analiza financiară şi analiza datelor trebuie să fie la temelia noilor instrumente de asistare a deciziei în finanţe şi să fie integrate cu instrumentele de birotică sau cu bazele de date, care într-un viitor apropiat, vor permite obţinerea unui raport de expertiză, fără ca decidentul să posede cunoştinţe aprofundate asupra domeniilor implicate în cunoaştere.

BIBLIOGRAFIE

1. B. IONESCU, Sisteme de asistare a deciziei financiar contabile, Ed. InfoMega, Bucuresti, 2002
2. JOLIVET Felix
3. REBOUL Gerard, Informatique appliquée a la gestion, Bertrand Lacoste, Paris, 1996
4. MATHE Jean Charles, Diagnostic et dynamique de l'entreprise, Editions Comptables Malesherbes, Paris, 1990
5. NICULESCU Maria, Diagnostic global strategic, Ed. Economică, Bucureşti, 1997
6. PISTRE Marcel, Multi-Win – modules d'analyse de données sous Windows, Specimen ESUG Toulouse, 1996