ŞCOALA NAŢIONALĂ DE STUDII POLITICE ŞI ADMINISTRATIVE

FACULTATEA DE COMUNICARE ŞI RELAŢII PUBLICE

ANALIZA DATELOR ŞI GESTIONAREA INFORMATICĂ A RESURSELOR UMANE

Lector univ. dr. Mihai Gavotă
Lector univ. dr. Cornelia Maxim

Bucureşti - 2009

CUPRINS

Prezentare generală, obiectivele şi conţinutul cursului

În prezent orice manager modern se confruntă în activitatea curentă cu foarte multe probleme de decizie. El trebuie să fie în măsură să rezolve aceste probleme într-un mod rapid şi eficient. De cele mai multe ori se impune ca deciziile pe care le ia să fie fundamentate şi optimizate conform unor metode proprii cercetărilor operaţionale. Între problemele de decizie pe care un manager trebuie să le rezolve se numără problemele din domeniile resurselor umane, planificării şi prognozei, alocării resurselor, selecţiei, repartiţiei optime.

În prezent în majoritatea universităţilor americane şi europene există cursuri asemănătoare (din categoria *Management Science*) care asigură studenţilor o pregătire sistematică corespunzătoare ce îşi propune să facă posibile fundamentarea şi luarea viitoarelor decizii într-un orizont scientizat, pe baza unui instrumentar software adecvat.

Obiective

În acest sens actualul curs şi-a propus următoarele obiective:

- să ofere studenţilor un bagaj de cunoştinţe teoretice care să le permită abordarea, rezolvarea şi interpretarea unor probleme de modelare şi decizie cu ajutorul unui instrumentar software specializat;
- descrierea unor modele consacrate din domeniul cercetărilor operaţionale, a claselor şi tipurilor de probleme posibil de a fi rezolvate pe această cale;
- să ofere elemente ajutătoare pentru formularea şi formalizarea problemelor de decizie în vederea găsirii soluţiilor optime cu ajutorul algoritmilor clasici implementaţi în cadrul unui instrumentar software specializat;
- dezvoltarea unor abilităţi practice utile lucrului cu pachete de programe expert de calcul, modelare şi analiză statistică destinate rezolvării problemelor specifice de decizie şi optimizare;

Notă

Pentru majoritatea studiilor de caz şi tipurilor de probleme prezentate va exista un subcapitol care va trata rezolvarea acestora cu ajutorul unui produs software specializat. În acest sens vor fi date exemple de rezolvări realizate cu produse software cum sunt: Microsoft Excel, Microsoft Project, QSB, LINDO, SPSS, What's Best!. Felurile în care se introduc datele şi se obţin rezultatele sunt asemănătoare pentru toate produsele informatice din această categorie (care sunt peste 400). Operarea este asemănătoare indiferent de produs şi nu implică nici un fel de dificultate informatică.

Întotdeauna cel mai dificil este să se formuleze corect problema de decizie iar apoi să se găsească modelul adecvat acesteia. Identificarea obiectivelor, a resurselor, a variabilelor şi a restricţiilor ce trebuie respectate, scrierea modelului matematic sunt faze absolut obligatorii ce trebuie parcurse înainte de a se folosi un produs informatic. De asemenea oricine va utiliza un produs software specializat va trebui să fie în măsură să poată interpreta rezultatele furnizate de acesta (situaţiile finale).

Iată de ce prezentul curs şi-a propus în primul rând să ofere studenților o informație care să-i ajute să surprindă corect modelul adecvat unei anumite probleme de decizie, să formalizeze corect problema, să analizeze şi să interpreteze rezultatele obținute. Prezentul curs nu şi-a propus să insiste pe partea informatică decât pentru exemplificări şi nu doreşte „să-i lege” pe studenți de un anumit produs software. Parcurgerea cursului nu necesită cunoştințe avansate de matematică sau informatică.

Cursul are un profund caracter practic şi poate fi parcurs fără a necesita lucrul în paralel cu un anumit produs informatic. Informațiile şi abilitățile dobândite dau studenților posibilitatea ca să poată utiliza oricând un produs informatic specializat pentru a obține rezolvări numerice. În acest sens nu se vor solicita răspunsuri numerice nici la evaluările pe parcurs nici la evaluarea finală.

Conținutul cursului

În prima parte după o scurtă introducere şi un istoric al acestei ştiințe vor fi prezentate câteva elemente teoretice generale legate de lucrul cu modele.

Apoi un următor capitol se va ocupa de formularea şi rezolvarea unor probleme de planificare şi prognoză pe baza unor modele de regresie şi a metodelor de interpolare. Vor fi prezentate:

- Principalele modele de regresie. Tipuri, definiții, particularități şi parametrii, necesari acestor modele. Exemple de utilizare.
- Seriile cronologice, eroarea standard de estimare, interpretarea trend-ului.
- Analiza şi interpretarea rezultatelor. Rezolvarea acestor probleme cu ajutorul calculatorului prin intermediul unor produse software specializate.

O altă secțiune importantă va prezenta câtiva algoritmi clasici de optimizare utilizați în cercetările operaționale. Nu se va insista pe rezolvarea acestor probleme prin metode „creion – hârtie” pe baza algoritmilor prezentați ci pe formalizarea şi parametrizarea lor pentru a putea fi implementate, analizate şi rezolvate în modalități asistate de calculator cu ajutorul unor produse software special destinate pentru astfel de rezolvări moderne. Vor fi prezentate:

- principalele clase de probleme specifice din domeniile:
    - programării liniare,
    - programării scopurilor,
    - lanțurilor Markov,
    - problemelor de tip rețea,
    - problemelor de transport,
    - problemelor cozilor de aşteptare,
    - problemelor de alocare,
    - problemelor comisului voiajor.
- formalizarea şi parametrizarea problemelor de decizie;
- găsirea variantelor optime;
- analiză statistică;
- interpretarea rezultatelor;
- rezolvarea acestor probleme prin intermediul unor produse software;
- exemple de utilizare.

Introducere

În prezent decizia trebuie privită ca fiind una dintre cele mai importante componente ale managementului modern. Ştiinţa managementului este o certitudine în afara căreia orice activitate de conducere devine empirică şi poate avea efecte dezastruase. Managementul actual presupune fundamentarea ştiinţifică a deciziilor. În majoritatea cazurilor ritmul alert al evenimentelor impune operativitate. Implicit luarea rapidă şi corectă a deciziilor devine o activitate care necesită instrumente adecvate, moderne şi performante. Între acestea instrumentele matematice şi informatice au o importanţă majoră. Deciziile operative, fundamentate ştiinţific, prin efectele lor imediate sau propagate în timp fac posibilă gestionarea corectă a resurselor şi implicit asigură realizarea scopului propus. În acest sens instrumentele software actuale bazate pe modele matematice consacrate sunt de un real folos în luarea deciziei din orice domeniu de activitate.

În literatura americană ştiinţa managementului, *Management Science* (MS) este sinonim cu ceea ce şi la noi se studiază în cadrul unei discipline aparte numită Cercetări operaţionale - *Operations Research* (OR). Alte denumiri sub care mai poate fi întâlnită această ştiinţă sunt: OR/MS sau ORMS, *Industrial Engineering* (IE) şi *Decision Science* (DS). Ştiinţa managementului îşi propune să rezolve prin intermediul modelelor matematice cantitative, problemele calitative sau cantitative impuse de luarea deciziei. Câmpul problemelor rezolvate astfel acoperă aplicaţii din domeniul previziunii, planificării, managementului proiectelor.

Perioada de început a acestei ştiinţe este plasată în timpul celui de-al doilea război mondial când mai multe echipe formate din oameni de ştiinţă de diferite specialităţi, matematicieni, fizicieni, ingineri şi psihologi au fost puse să lucreze împreună pentru rezolvarea unor probleme militare de previziune şi planificare. Cercetările acestor echipe s-au numit *Operations Research* şi au fundamentat luarea unor decizii care s-au dovedit corecte şi care au dus la victorie. Ulterior, după război, cercetările lor şi-au găsit aplicarea în domeniul economic, al afacerilor, în planificarea producţiei, în agricultură, în domeniul monetar, în politică, în administraţie ş.a. Practic cercetările operaţionale au fost şi pot fi utilizate în orice domeniu care presupune cu necesitate optimizarea deciziei manageriale. Instrumentele software create special în acest scop, uşurează şi fac posibilă utilizarea ştiinţei managementului în orice domeniu.

Scurt istoric

MS este o disciplină relativ nouă care a apărut la începutul anului 1936 când ministerul britanic al războiului a creat pe coasta de est, lângă Suffolk, stația de cercetare Bawdsey. Această stație trebuia să fie centrul unde să se realizeze toate experimentele legate de radare. În acel timp echipamente radar experimentale deveniseră performante ele putând avea raze de acțiune de peste 100 de mile. Era foarte clar că descoperirea şi utilizarea radarelor ridica o serie de probleme noi ca de exemplu directivele ce trebuiau date luptătorilor pe baza informațiilor ce urmau a fi obținute cu ajutorul radarelor. O altă problemă ce trebuia rezolvată era aceea de a integra informațiile oferite de radar cu cele venite de la sol.

Primele experimente s-au derulat în vara anului 1937. Baza de cercetare Bawdsey a devenit operațională iar informațiile ce veneau de la ea erau introduse în sistemul general de avertizare şi control aerian. La acel moment exercițiile erau încurajatoare însă informațiile obținute de la radar după filtrarea transmisiei prin rețeaua de control nu erau satisfăcătoare.

În iulie 1938 s-a desfăşurat un mare exercițiu de apărare aviatică. Încă patru stații de radar au fost inaugurate pe coastă şi se spera că în acel moment Anglia va avea un sistem de detectare şi de apărare aeriană foarte performant. Fals! Exercițiul a arătat că apăruse o nouă mare problemă. Era nevoie să se coordoneze şi să se confrunte datele obținute de la toate stațiile radar deoarece de multe ori acestea erau diferite. Cum apropierea războiului era iminentă, devenise din ce în ce mai clar că ceva nou era necesar pentru ca problemele legate de radar să fie rezolvate. Era nevoie de o nouă manieră de abordare.

După ce exercițiile au luat sfârşit, de la stația de cercetare Bawdsey s-a comunicat faptul că s-a constatat fezabilitatea tehnică a radarului în detectarea obiectelor zburătoare însă că din punct de vedere operațional el nu se ridică la standardele necesare. S-a propus imediat inițierea unui program operațional - opus celui tehnic – pentru cercetarea aspectelor sistemului. Termenul "cercetare operațională" (*Operations Research* - OR) a fost considerat potrivit pentru această nouă direcție de cercetare. Echipa de specialişti implicați în cercetare a fost selectată în aceeaşi zi.

În vara anului 1939 Anglia a desfăşurat ceea ce a fost numit cel mai mare exercițiu de apărare de dinaintea războiului. El a implicat 33000 de militari, 1300 instrumente de zbor, 110 arme de foc, 700 de lumini de căutare şi 100 de baloane. Acest exercițiu a adus mari îmbunătățiri pentru proiect. Contribuția echipelor OR a fost atât de vizibilă încât şeful forțelor armate a ordonat ca atunci când războiul va începe, aceştia să-l însoțească la Stanmore.

Inițial a fost creată "Secția de cercetare Stanmore", dar în 1941 ea a fost redenumită în "Secția de cercetări operaționale" (ORS). Termenul *Operations Research* fusese acceptat oficial. Simultan au fost create astfel de secții şi în alte baze militare. Responsabilitatea acestor secții era legată de felul în care trebuia făcută de către avioane survolarea unor zone extinse, astfel încât să fie descoperit un număr cât mai mare de submarine germane pentru a fi distruse. Printre problemele ce trebuiau rezolvate se numărau:

- Organizarea zborurilor şi inspecțiilor.

Problema care se ridica aici era o aceea de a se găsi un nou plan de organizare a zborurilor şi de întreținere a avioanelor astfel încât escadrilele să fie utilizate la

maxim. În acest sens ORS a propus un nou sistem de planificare a zborurilor în care un echipaj să poate pilota mai multe aeronave. Avantajul acestui sistem era acela că astfel se putea mări numărul de ore de zbor. Dezavantajul era că în acest fel se rupeau legăturile dintre avion şi piloți care ar fi preferat desigur să aibă de fiecare dată "avionul lor". Într-o perioadă de încercare de 5 luni tipul de organizare ORS a adus o creştere a numărului de ore de zbor operațional cu 61%. Noul sistem propus a fost acceptat şi implementat.

- Îmbunătățirea probabilităților de atac şi distrugere a submarinelor germane.

Experiențele au arătat că erau necesare 170 de ore de muncă umană la sol pentru a asigura o oră de zbor operațional şi mai mult de 200 de ore de zbor pentru a ataca un submarin german. Deci peste 34000 de ore de munca umană erau necesare doar pentru a ataca un submarin. La începutul anului 1941 probabilitatea de atac şi distrugere era de 2-3%. În acest domeniu, ORS a avut un mare aport. Erau necesare îmbunătățiri. Arma principală în distrugerea submarinelor era aruncarea de încărcături explozive pe o anumită direcție. După ce loveau apa, încărcăturile se scufundau în timp ce erau deviate de inerție. Atunci când atingeau o anumită adâncime explodau, distrugând orice submarin aflat pe o rază 5 – 6 metri. Şase variabile erau considerate a influența probabilitatea de distrugere:

- stabilirea adâncimii la care va exploda încărcătura;
- raza de distrugere;
- erorile de țintă la lansarea încărcăturii;
- orientarea încărcăturii;
- intervalul între lansările succesive;
- indicatoarele de lansare a bombei.

În urma cercetărilor realizate, probabilitatea de distrugere a crescut de la 2-3% la peste 40%.

Primii specialişti OR veneau din diferite domenii de activitate precum fizică, psihologie, matematică, etc. Ceea ce au adus în plus a fost capacitatea lor de a face presupuneri, de a formula probleme, de a gândi logic, a exploata ipoteze, a analiza datele, a imagina experimente, a colecta date, a găsi modele adecvate de simulare etc. Mulți dintre aceşti specialişti erau foarte valoroşi. După război cel puțin patru din ei au fost laureați ai premiului Nobel.

În Anglia mulți dintre specialiştii care lucraseră în OR s-au întors la ocupațiile lor de dinainte de război. Aici în anii de după război OR nu s-a extins prea mult în industrie şi economie. În schimb, în USA OR a luat o foarte mare amploare şi s-a integrat practic în toate domeniile de activitate. În prezent în majoritatea universităților americane există cursuri (de tipul *Management Science* MS) care asigură o pregătire sistematică în această direcție.

Modele

În activitatea lor managerii se confruntă permanent cu probleme care solicită: analiza şi măsurarea performanţelor realizate, reducerea costurilor operaţiilor păstrând acelaşi nivel de calitate şi aceleaşi profituri, oferirea unor servicii mai bune fără a creşte costurile, etc.

Pentru a identifica metode de îmbunătăţire a sistemului condus ei (sau analiştii lor) trebuie să se realizeze o reprezentare sintetică, un model al sistemului fizic care să fie folosit pentru a descrie efectele diferitelor soluţii propuse, un model de simulare. Modelul poate fi gândit astfel încât să surprindă elementele esenţiale ale sistemului fără a-l reconstitui integral pe acesta. De exemplu o campanie de promoţie a unui produs poate fi utilizată ca model al răspunsului clienţilor. O ecuaţie matematică poate fi folosită pentru a modela energia conţinută de un anumit material. În fiecare din aceste exemple modelul surprinde un anumit aspect al realităţii pe care încearcă sa-l reprezinte.

Din moment ce un model surprinde doar anumite aspecte ale realităţii el nu poate fi utilizat în orice situaţie pentru că, în situaţia respectivă ar putea surprinde elemente greşite. De exemplu, temperatura este un model al condiţiilor climatice dar daca cineva este interesat de presiunea barometrică acest model ar fi greşit. O ecuaţie care prezice vânzările anuale ale unui produs este un model al acelui produs dar nu este de nici un folos dacă ne interesează costul de producere al acelui produs. Deci utilitatea modelului este dependentă de aspectul din realitate pe care îl reprezintă.

Un model se poate dovedi inadecvat chiar şi atunci când surprinde aspectele corecte ale realităţii dacă o face într-o manieră distorsionată. Un termometru care indică greşit temperatura nu este de nici un folos pentru diagnoza medicală. Deci un model util este acela care surprinde elementele potrivite ale realităţii cu o acurateţe acceptabilă. Un model matematic este o ecuaţie, o inegalitate sau un sistem de ecuaţii sau inecuaţii care reprezintă anumite aspecte ale sistemului fizic modelat. Modelele de acest tip sunt foarte folosite în fizică, inginerie, afaceri, economie, etc.

Un model oferă managerului (analistului) un instrument care îl ajută în studierea sistemului, fără a afecta cu nimic la nivel fizic sistemul. De exemplu să presupunem că un model matematic prevede vânzările anuale în funcţie de preţul unitar al unui produs. Dacă cunoaştem preţul unui produs putem calcula cu uşurinţă totalul vânzărilor anuale. Pentru a determina preţul de vânzare care ar aduce cele mai mari profituri se pot introduce în model diferite costuri, notându-se la fiecare cost profitul obţinut, iar în final, prin tehnica încercării şi erorii se poate determina acel cost care va aduce maximum de profit.

În mod ideal, dacă modelul este o reprezentare foarte corectă a sistemului, în final se pot obţine rezolvări la problemele sistemului real. Deci utilitatea şi aplicativitatea soluţiilor obţinute cu ajutorul modelelor depinde direct de fidelitatea cu care modelul reprezintă realitatea studiată .

Pentru a defini acele condiţii care vor conduce la găsirea soluţiei la problemele sistemului, analistul trebuie mai întâi să identifice acele criterii după care se măsoară performanţe sistemului. Criteriul cel mai frecvent utilizat este performanţa sistemului sau utilitatea lui. În aplicaţiile legate de afaceri utilitatea este deseori măsurată prin costuri sau profituri, sau în termeni de beneficiu-cost.

Ştiinţa managementului se bazează aşadar pe modele care de fapt reprezintă anumite abstractizări ale realităţii. Modelele pot fi clasificate în trei categorii:

1. Iconice
2. Analogice
3. Simbolice.

***Modelele iconic**e* sunt cel mai puțin abstracte. Ele sunt modele fizice foarte asemănătoare realității. Un model iconic poate fi modelul la scară redusă al unui avion, vapor sau al unei maşini.

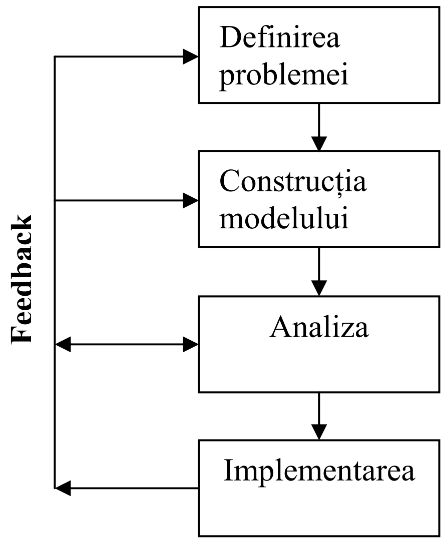
Modelele analogice sunt de asemenea modele fizice dar mult mai abstracte decât cele iconice. Modelele analogice sunt destinate pentru a facilita răspunsul la întrebarea: „Ce ar fi dacă ?". Astfel un grafic, o schiță, un termometru, un barometru, un ceas, reprezintă modele analogice.

Modelele simbolice sunt cele mai abstracte. Ele conțin numere şi simboluri algebrice care reprezintă aspecte importante ale problemelor prezentate cel mai adesea sub forma unor ecuații. Aceste numere şi simboluri sunt utilizate pentru a rezolva aspecte importante ale problemei prin găsirea valorilor unor necunoscute şi a unor variabile cheie. Ele sunt modele matematice care nu seamănă cu realitatea pe care o reprezintă. De cele mai multe ori modelele simbolice sunt superioare modelelor iconice şi celor analogice deoarece prin formalizare problemele propuse pot fi cu uşurință transferate spre rezolvare calculatoarelor. Managerul sau analistul sunt interesați în a găsi valorile variabilelor de decizie ale unui model care să le asigure un profit maxim şi doresc să cunoască previziunile anumitor acțiuni sau decizii (creşterea unor prețuri, modificarea cursului de schimb, succesul în alegeri, variația stocurilor etc). Efortul şi în acelaşi timp provocarea pentru manager şi pentru analist constau în principal în a găsi cel mai potrivit model cu cele mai adecvate obiective, constante şi variabile de decizie care să simuleze cel mai bine realitatea. Găsirea modelului, a celei mai potrivite formalizări, este o activitate chiar mai importantă decât rezolvarea ecuațiilor acestuia.

Modelele simbolice oferă multe beneficii în rezolvarea problemelor dar prezintă şi anumite riscuri. Unul din principalele avantaje este că oferă specialistului posibilitatea de a-şi concentra atenția doar asupra aspectelor esențiale ale problemei. În acelaşi timp există riscul de a neglija unele aspecte considerate în mod eronat ca secundare şi nesemnificative. Un alt avantaj constă în faptul că modelele matematice cantitative îl obligă pe specialist să cuantifice informația. Dezavantajul rezidă din faptul că poate fi cuantificată eronat mai ales informația necantitativă care este dificil sau imposibil de inclus într-un model cantitativ. Un beneficiu incontestabil al acestor modele constă în „comprimarea timpului" adică în timpul scurt de răspuns pe care îl pot oferi comparativ cu anumite experiențe reale sau studii de alt tip. Poate fi preîntâmpinat, evitat şi prevăzut efectul dezastruos sau periculos al unor experiențe fizice reale. Este evident şi pericolul simplificării care să nu mai facă posibilă corespondența între model şi realitate.

O altă clasificare a modelelor simbolice poate fi şi în funcție de problemele pe care le rezolvă. Astfel sunt modele *de optimizare* şi modele *predictive*.

Din punctul de vedere al ştiinței managementului pentru soluționarea unei probleme prin intermediul unui model se disting următoarele **faze şi relații**:



Datorită posibilităților pe care le oferă în transpunerea şi analiza modelelor prin intermediul pachetelor de programe expert, calculatoarele moderne de tip PC au devenit instrumente de bază ale managerului modern. Acum este necesar, ca managerii şi analiştii să posede cunoştințe din domeniul informatic care să le permită să utilizeze corespunzător software-ul special destinat. În prezent, fără mijlocirea calculatorului, anumite analize şi studii mai complexe devin imposibile.

După felul în care formulează şi rezolvă problemele specifice prin intermediul modelelor, ştiința managementului distinge două mari clase de modele:

1. *Modele ale programării liniare*
   - probleme rezolvate prin metode grafice
   - probleme posibil de rezolvat prin metoda Simplex
   - probleme de transport
   - probleme ale programării în numere întregi
   - probleme ale programării scopurilor
2. *Modele stochastice*
   - probleme rezolvate prin metode statistice şi teoria probabilităților
   - probleme ale previziunii şi prognozei
   - probleme de rețea
   - probleme de planificare
   - probleme ale cozilor de aşteptare
   - probleme de simulare
   - probleme rezolvate prin intermediul lanțurilor Markov.

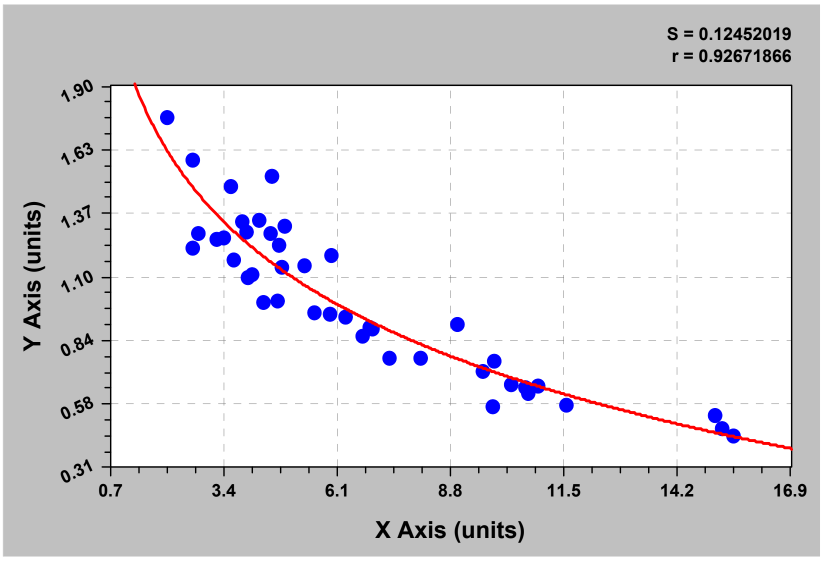
Previziune şi prognoză

Planificarea şi implementarea deciziilor sunt unele dintre principalele sarcini ale managerilor. Uneori efectele deciziilor rezultate pot avea efecte satisfăcătoare şi duc la succes, alteori nu. Adesea gradul de succes este dependent de gradul de incertitudine al evenimentelor posibile viitoare. Cu cât incertitudinea este mai mare cu atât este mai dificil de proiectat o decizie pe baza căreia să fie formulată o planificare care să ducă la rezultatele dorite. Previziunea este importantă deoarece poate reduce incertitudinea. Previziunea fundamentată ştiinţific are o importanţă vitală în planificare.

Seriile cronologice

Seriile cronologice reprezintă valori istorice ale unor variabile care au fost înregistrate la intervale periodice (ex.: cererea zilnică, săptămânală sau lunară pentru un produs, evoluţia ratei de schimb valutar etc). Seriile cronologice se mai numesc şi serii de timp sau dinamice. Ele sunt formate din două şiruri de date paralele din care primul şir arată variaţia caracteristicii timp iar cel de-al doilea arată variaţia fenomenului sau caracteristicii cercetate. Tehnicile de prognoză folosesc seriile cronologice presupunând faptul că experienţa trecută va reflecta probabil experienţa viitoare. Se consideră ca *pattern*-ul evenimentelor trecute va persista în viitor. Unele dintre cele mai comune *pattern-uri* observate cu uşurinţă în seriile de date istorice sunt tendinţele (*trend*-ul) variaţiilor ciclice, de sezonalitate.

Trend-ul reprezintă tendinţa variaţiilor crescătoare sau descrescătoare ale unei variabile, tendinţă prevăzută pentru un orizont de timp viitor, pe baza unor variaţii reale, dintr-un interval de timp cunoscut. De cele mai multe ori totuşi identificarea trend-ului este o operaţiune dificilă pentru că în realitate prezenţa în datele istorice a unor variaţii neregulate, aleatoare, face foarte greu de interpretat şi de prognozat evenimentele viitoare. Aceste variaţii pot distorsiona previziunea şi de aceea este de dorit să fie identificate şi eliminate.



Pentru manager este foarte importantă analiza care precede calculele de prognoză şi implicit fundamentează decizia. Reprezentarea grafică constituie unul din criteriile cele mai

importante pe baza cărora se va alege procedeul de extrapolare. Extrapolarea are la bază metodele şi procedeele de ajustare care conduc la micşorarea distorsiunilor prin nivelare.

O primă metodă de extrapolare este *previziunea naivă* bazată pe „bunul simț" al observației unor serii de timp.

Altă metodă de previziune prin extrapolare este cea a *mediilor mobile* care utilizează o substituire a datelor seriei dintr-un interval fix, cu mediile calculate pentru interval „din aproape în aproape".

$$M_n = \frac{\sum_{i=1}^{n} A_i}{n}$$ unde:

i = „vârsta" intervalului;
n = numărul intervalelor;
Ai = valoarea „vârstei" i.

Iată un exemplu de lucru cu mediile mobile:

Perioada	„Vârsta"	Cererea
1	5	40
2	4	44
3	3	36
4	2	42
5	1	40

$MA_3 = (36 + 42 + 40) / 3 = 39,33$

Nivelarea exponențială este o metodă de extrapolare care calculează valoarea previzionată F_t astfel:

$F_t = F_{t-1} + \alpha\,(A_{t-1} - F_{t-1})$ unde:

F_t = valoarea pentru momentul de timp prognozat t;
F_{t-1} = valoarea prognozată pentru momentul actual t-1;
α = constanta de nivelare;
A_{t-1} = valoarea reală actuală a variabilei la momentul prezent t-1.

Senzitivitatea ajustării erorii în prognoză este dată de constanta de nivelare α care poate avea valori cuprinse între 0 şi 1. Alegerea valorii constantei este foarte importantă şi trebuie făcută pe baza unor judecăți fundamentate pe cunoaşterea aproximativă a evoluției variabilei şi pe studiul erorilor rezultate din încercări succesive aplicate pe mai multe serii de date. Valorile uzuale alese pentru constantă variază între 0,05 şi 0,5. Unele pachete de programe specializate pentru astfel de calcule oferă facilități care permit modificarea automată a constantei α în funcție de valorile erorilor de prognoză rezultate.

Modelare şi trend

Trend-ul reprezintă tendința persistentă ascendentă sau descendentă a valorilor seriilor de date dinamice dintr-un orizont de timp. Foarte interesant pentru ştiința managementului şi implicit pentru manageri este găsirea unor funcții, a parametrilor acestora, care să descrie cel mai bine *trend*-ul unor serii dinamice paralele. Având un set de date (puncte) numite cel mai adesea observații, apare problema găsirii unui model care să aproximeze cel mai bine fenomenul studiat sub forma unor ecuații parametrice. Acest model poate fi un model polinomial simplu sau unul foarte complex cu mulți parametrii. Cel mai important pentru manageri şi analişti este selectarea celui mai potrivit model care să descrie cel mai bine legea de dependență existentă între variabilele studiate prin intermediul seturilor de date care le reprezintă.

Iată familiile funcțiilor de regresie şi câteva dintre ecuațiile funcțiilor acestora cel mai des utilizate. În aceste expresii de funcții y reprezintă valorile seriei variabilei studiate (prognozate), care este în funcție de valorile seriei variabilei x.

1. Regresia liniară:

- **Familia funcțiilor liniare:** y = a+bx
- **Familia funcțiilor pătratice:** y = a+bx+cx^2
- **Familia funcțiilor polinomiale:** y = a+bx+cx^2+dx^3+....

2. Regresia neliniară:

- **Familia funcțiilor exponențiale:**

Modelele exponențiale au funcții de creştere exponențiale sau logaritmice. Ele descriu în general curbe convexe sau concave, dar unele funcții pot avea un punct de inflexiune şi un punct de maxim sau minim.

- Funcția exponențială: y = a*exp(b*x)
- Funcția exponențială modificată: y = a*exp(b/x)
- Funcția logaritmică: y = a+b*ln(x)
- Funcția logaritmică reciprocă: y = 1/(a+b*ln(x))
- Funcția modelului presiunii de vaporizare: y = exp(a+b/x+c*ln(x))

- **Familia funcțiilor putere:**

Familia funcțiilor putere conține funcții de creştere de tip putere cu unul sau mai mulți parametri. Ele pot descrie evoluția unei variabile independente sau a uneia dependente a căror putere este influențată de parametrul dat. Această familie conține un set de curbe convexe sau concave fără puncte de inflexiune sau de maxim sau minim.

- Funcția putere: y= a*x^b
- Funcția putere modificată y = a*b^x
- Funcția 'Powershift': y = a*(x-b)^c

- Funcția geometrică: $y = a*x^(b*x)$
- Funcția geometrică modificată: $y = a*x^(b/x)$
- Funcția rădăcină: $y = a^(1/x)$
- Modelul Hoerl: $y = a*(b^x)*(x^c)$
- Modelul Hoerl modificat: $y = a*b^(1/x)*(x^c)$

- **Familia funcțiilor care descriu modele de tipul recoltă-densitate:**

Modelele de tipul recoltă-densitate sunt larg utilizate, în special în aplicațiile de prognoză din agricultură. Aceste modele cronologice au fost utilizate pentru a modela relațiile dintre recoltele obținute la diferite culturi în funcție de spațierea sau densitatea plantărilor. În practică, pentru relația recoltă-densitate s-au observat în special doar două tipuri de răspunsuri: "asimptotic" şi "parabolic". Astfel dacă densitatea (x) creşte, recolta obținută (y) creşte, apropiind-se în mod asimptotic de o valoare fixă. Deci peste o anumită limită relația este asimptotică. Pentru manageri este important să găsească valoarea de optim, în care relația este parabolică. Aceste tipuri de relații între cele două şiruri de date (x) şi (y) sunt foarte comune. Modelele care descriu cele mai bine aceste relații sunt:

- Modelul reciproc: y = 1 / (a + bx)
- Modelul reciproc quadratic: y = 1 / (a + bx + cx^2)
- Modelul Bleasdale: y = (a + bx) ^ (-1/c)
- Modelul Harris: y = 1 / (a + bx^c)

- **Familia funcțiilor de creştere:**

Modelele de creştere sunt caracterizate de o creştere care tinde să se plafoneze asimptotic către o valoare fixată. Aceste modele sunt comune în special ştiințelor inginereşti.

- Modelul exponențial Assoc (2): y = a*(1-exp(-bx))
- Modelul exponențial Assoc (3): y = a*(b-exp(-cx))
- Modelul creşterii saturate: y = ax / (b + x)

- **Familia funcțiilor "S-shaped" (în forma literei S):**

Procesele care produc curbe de creştere în forma literei S ("S-*shaped*" sau sigmoidale) sunt comune unei largi arii de aplicații din biologie, inginerie, agricultură şi economie. Aceste curbe încep dintr-un punct fixat şi au o creştere monotonă până la un anumit punct de inflexiune, după care în final, creşterea tinde către o valoare asimptotică. Familia de funcții "S" este un subset al familiei de funcții de creştere dar se studiază separat deoarece curbele acestor funcții au un comportament aparte, distinctiv.

- Modelul Gompertz: y = a * exp (-exp(b - cx))
- Modelul Logistic: y = a / (1 + exp (b - cx))

- Modelul Richards: y = a / (1 + exp(b - cx))^(1/d)
- Modelul MMF: y = (ab + cx^d)/(b + x^d)
- Modelul Weibull: y = a - b*exp(-cx^d)

- **Familia funcțiilor diverse:**

  Ca multe lucruri din viață, există întotdeauna unele care nu pot fi încadrate în anumite categorii specifice. Familia funcțiilor diverse este unul dintre acestea dar care totuşi descriu modele ale regresiei neliniare întâlnite în viață. Iată câteva dintre modele:

  - Modelul sinusoidal: y = a + b*cos(c*x + d)
  - Modelul Gaussian: y = a*exp((-(x - b)^2)/(2*c^2))
  - Modelul Hiperbolic: y = a + b/x
  - Modelul capacității de încălzire: y = a + bx + c/x^2
  - Modelul funcției raționale: y = (a + bx) / (1 + cx + dx^2)

Scopul aplicării unei metode de regresie este găsirea expresiei unei funcții teoretice $f(x_i)$ care să aproximeze cel mai bine valorile reale y_i obținute pentru punctele x_i culese. Deci pentru fiecare x_i real, trebuie ca valorile teoretice calculate, $f(x_i)$ să fie cât mai aproape de valorile reale y_i observate.

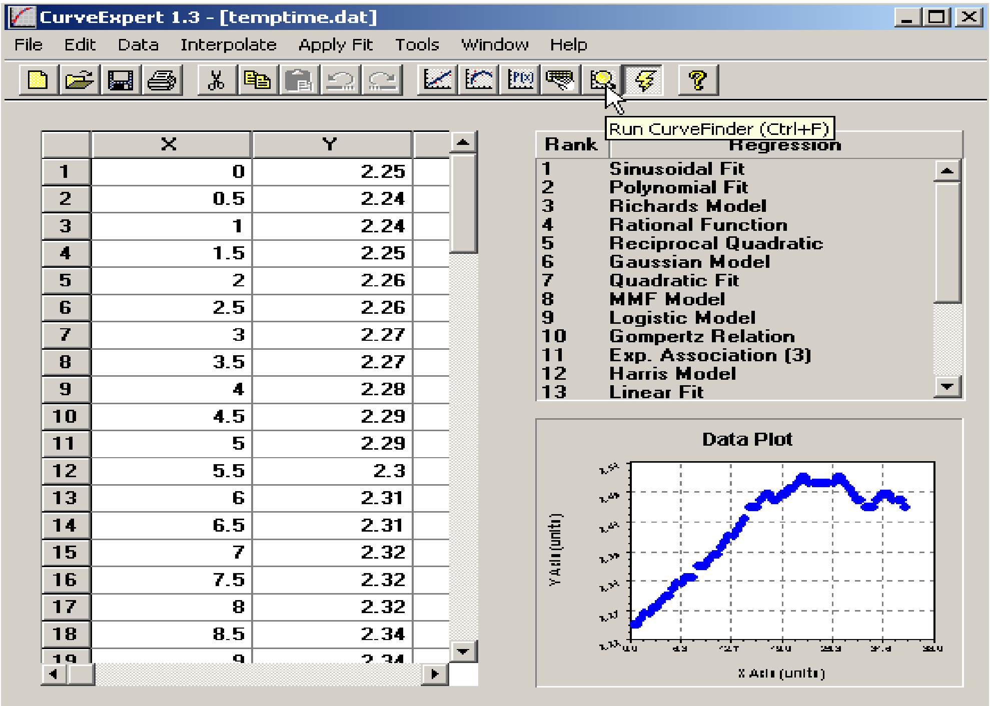
Pentru a evalua curba de regresie f(x), se utilizează:

Abaterea standard de estimare:

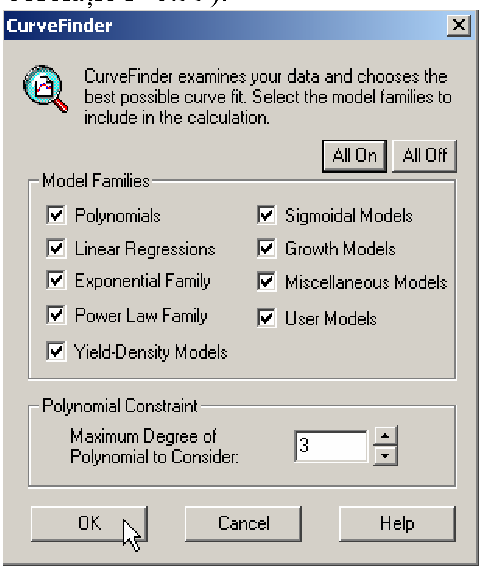
$$S_r = \sqrt{\frac{\sum_{i=1}^{n}(y_i - f(x_i))^2}{n-1}}$$

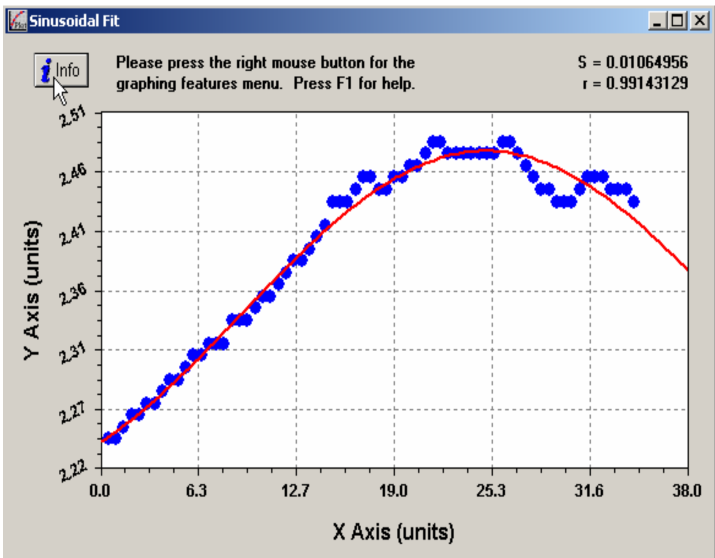
Rezolvarea problemelor de prognoză prin intermediul unui produs software specializat

Presupunând că au fost culese temperaturile reale pentru 34 de intervale de timp, reprezentate prin următoarele serii de date dinamice (X,Y - din figura de mai jos), să se determine care va fi temperatura după al 35-lea interval. Pe axa Ox sunt reprezentate intervalele de timp în care s-au cules temperaturile reale (să presupunem 0.5 zile). Pe axa Oy sunt reprezentate temperaturile corespondente pentru fiecare interval de timp.

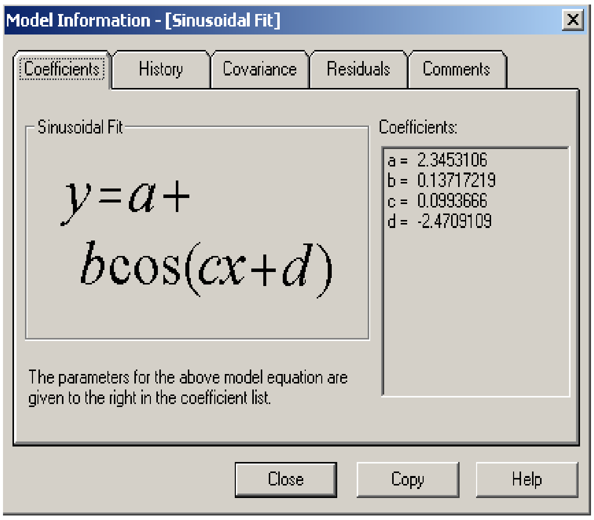


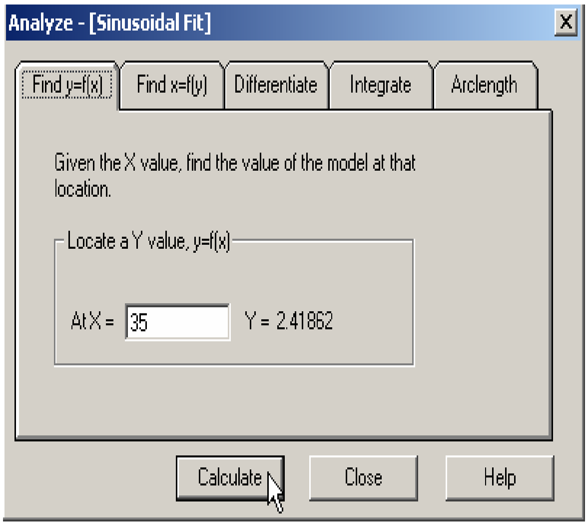
Iată în continuare în dialogul din stânga posibilitatea de a alege familiile modelelor de funcții pe care să le utilizeze programul, iar în dreapta reprezentarea grafică a funcției care aproximează cel mai bine punctele empirice (abaterea standard S=0.01, iar coeficientul de corelație r=0.99).





După apăsarea butonului <Info> apare expresia funcției propuse de program și pot fi consultate valorile coeficienților a, b, c, d (imaginea din stânga). Prin intermediul meniului contextual se poate solicita o analiză (imaginea de dialog din dreapta) care oferă posibilitatea determinării y = f(x) pentru un X ales. Dacă se introduce X=35, se va obține Y=2.41862 ceea ce reprezintă răspunsul căutat.





Programarea liniară

Prezentare generală

Pană în anii '80 toate pachetele pentru rezolvarea problemelor de programare liniară (***Linear Programming* – LP**) se bazau doar pe algoritmul *simplex*. În anul 1984 Karmarkar a publicat un nou algoritm pentru rezolvarea problemelor de programare liniară (LP), algoritm numit *interior point* care este complet diferit de algoritmul *simplex*.

Munca lui Karmarkar a adus un imens aport la rezolvarea problemelor LP atât prin metodele *interior point* cât şi prin algoritmul *simplex.*

Din anul 1984 au apărut noi produse software specializate pentru rezolvarea problemelor de programare liniară ca de exemplu:

- OSL (*Optimisation Subroutine Library*) - IBM
- Cplex (*Cplex Optimisation*)

Ambele produse pot utiliza atât algoritmi *simplex* cât şi *interior point*.

În prezent există multe produse software care pot rezolva rapid probleme de programare liniară. Calculatoarele de tip PC nu mai reprezintă un lux şi pe ele pot fi instalate cu uşurinţă asemenea produse software. Dacă dorim să rezolvăm numeric o problemă de programare liniară trebuie să ne întrebăm în primul rând dacă putem găsi un pachet software adecvat calculatorului nostru şi sistemului de operare sub care acesta lucrează şi apoi dacă acest pachet de programe are capacitatea să rezolve problema.

Pentru pachetele software care pot rezolva probleme de tip LP este foarte important să cunoaştem:

- numărul maxim de restricţii cu care pot lucra şi
- tipurile de calculatoare pe care pot fi rulate.

Pachetele software menţionate mai sus (OSL şi Cplex) au capacitatea de a lucra cu un număr maxim de 2 miliarde de variabile şi 16 milioane de restricţii. Aceste limite de capacitate depăşesc cu mult ceea ce am putea rezolva în viaţa reală. Iată în continuare câteva caracteristici ale unor probleme (LP) reale care au fost rezolvate folosind aceste programe:

Nume Program	**Număr de Restricţii**	**Număr de variabile**	**Timp de rezolvare**	**Computer**
OSL	105,000	155,000	4 ore	IBM 3090
	750	12,000,000	27 min.	IBM 3090
Cplex	145	1,000,000	6 min.	Cray YMP
	41,000	79,000	3 min.	Cray 2

Probleme complexe de programare liniară care au fost rezolvate prin intermediul unor produse software specializate

Determinarea facilităților de comunicații necesare pentru Bazinul Pacific

Aici problema pe care compania AT&T a dorit să o rezolve a fost aceea de a determina:

- unde şi cum să fie amplasate cablurile şi sateliții submarini atunci când vor fi necesari comunicațiilor din acest spațiu, şi
- numărul circuitelor necesare.

Această problemă LP a avut 28000 restricții şi 77000 variabile.

Determinarea dinamicii optime a resurselor umane din cadrul armatei USA

Problema ce trebuia rezolvată a fost aceea de a stabili promovările în cadrul armatei americane ținând cont de noile intrări şi ieşiri din armată şi de necesitățile de instruire aferente. Problema LP a avut 21000 restricții şi 43000 variabile.

Evacuarea militarilor accidentați dintr-o zonă de conflict

Aviația militară americană a avut de rezolvat o problemă de programare liniară care îşi propunea să determine fluxul de accidentați care pot fi evacuați dintr-o zonă de conflict din America continentală astfel încât aceştia să fie cât mai puțin timp în aer. Restricțiile acestei probleme sunt:

- toți pacienții care trebuie să fie transportați vor fi transportați;
- limitele de transport vor fi impuse de capacitatea spitalelor şi specializarea lor.

O astfel de problemă care acoperea 50 de zile de conflict armat a avut 79000 restricții şi 267000 variabile şi a fost rezolvată în 10 ore.

Programarea logistică militară

Departamentul de apărare al USA a avut de rezolvat o problemă de logistică cu privire la fezabilitatea susținerii operațiilor militare în timpul unei crize. Problema era de a determina dacă diferite materiale necesare pot fi transportate peste ocean în ferestre stricte de timp. Modelul LP a inclus capacitățile porturilor de îmbarcare şi debarcare, capacitățile diferitelor mijloace de transport aeriene sau navale implicate şi penalizările ce pot fi aplicate pentru neonorarea la timp a transporturilor.

O astfel de problemă care a fost simulată şi rezolvată presupunea 15 perioade de timp, 12 porturi de îmbarcare, 7 porturi de debarcare, 9 tipuri diferite de vehicule pentru 20000 de solicitări de transport. Problema a avut 20500 restricții şi 520000 variabile şi a fost rezolvată în 75 minute.

Agenda personalului din aviație

American Airlines şi-a propus rezolvarea unei probleme de tip LP prin intermediul căreia să poată stabili optim agenda personalului din aviația comercială. În urma unui studiu a rezultat că această problemă va trebui să lucreze cu un număr de aproximativ 12000000 de variabile. Trebuia să se țină seama şi de faptul că în programul liniilor aeriene pot interveni rute divizate în două parți. De exemplu ruta Chicago-Londra trebuia să treacă prin New York şi deci pentru ea trebuiau prevăzute mai multe elemente precum ora de plecare din Chicago, ora de sosire în New York respectiv în Londra precum şi disponibilitatea personalului de a însoți întreaga cursă. Zborul putea fi însoțit de un personal diferit între punctele de oprire.

Această problemă de stabilire a programului personalului de zbor trebuia să mai țină seama şi de faptul că nu orice pilot poate pilota orice tip de avion. Deci problema era aceea de a asigura pentru fiecare rută de zbor un personal corespunzător. Alte restricții mai erau legate de orele la care personalul poate lucra.

Abordarea electronică versus abordarea manuală

Rezolvarea problemelor complexe de programare liniară prin intermediul unor produse software specializate s-a impus deoarece:

- o rezolvare manuală a acestui tip de probleme (în prezent când există foarte multe produse software specializate ce pot fi rulate pe calculatoare PC relativ ieftine) este aproape inutilă, pentru că depăşeşte cu mult capacitățile umane de rezolvare clasică şi generează riscuri de eroare şi costuri foarte mari;
- o rezolvare electronică este preferată în prezent pentru că este mai rapidă, oferă o acuratețe a soluțiilor mai mare şi este mult mai ieftină comparativ cu metodele manuale.

Despre funcția obiectiv

Partea de model matematic care descrie utilitatea este denumită funcție obiectiv. Dacă funcția obiectiv trebuie să descrie măsura în care variază utilitatea produsului, atunci ea trebuie să surprindă dimensiunea utilității şi variabilele în funcție de care variază aceasta. Variabilele sistemului pot fi împărțite în variabile de decizie şi parametri. O variabilă de decizie este o variabilă care poate fi direct controlată de cel care ia deciziile. Există de asemenea unii parametri ale căror valori pot fi neclare pentru cei ce iau deciziile. Aceasta cere o analiză mai sensibilă după găsirea celei mai bune strategii. În practică este imposibil să se surprindă într-o ecuație matematică toate relațiile exacte între variabilele sistemului şi dimensiunea utilității. În schimb, analistul OR/MS trebuie să încerce să identifice si apoi să surprindă acele variabile care au cea mai mare importanță asupra dimensiunii utilității. El trebuie să le cuprindă în ecuații, sisteme de ecuații, inecuații, etc, relația matematică fiind funcția obiectiv folosită pentru a evalua performanțele sistemului studiat.

Formularea unei funcții obiectiv corectă este de obicei o sarcină foarte grea iar până la găsirea ei analistul se poate lovi de multe eşecuri. Aceste eşecuri se pot datora faptului că analistul alege un set greşit de variabile sau, chiar dacă el alege variabilele

bune, nu reuşeşte să surprindă bine relațiile dintre variabile şi dimensiunile utilității. De asemenea analistul poate încerca să găsească şi alte variabile care să îmbunătățească modelul şi să le neglijeze pe acelea care s-au dovedit a nu fi aşa de importante. În orice caz nu putem afla dacă aceşti factori îmbunătățesc într-adevăr modelul decât dacă formulăm şi testăm modele noi care conțin şi alte variabile. Tot procesul de selectare a variabilelor şi de formulare a modelului poate necesita reiterări multiple înainte de a se găsi o funcție obiectiv satisfăcătoare. Analistul speră să obțină câte o îmbunătățire a modelului la fiecare reiterare deşi aceasta nu se întâmplă întotdeauna. De cele mai multe ori succesul final aste atins după un lung şir de eşecuri şi mici succese.

La fiecare stadiu de dezvoltare a procesului, analistul trebuie să măsoare cât de adecvat sau valid este modelul. Două criterii sunt cel mai frecvent utilizate în acest tip de determinare. Primul implică experimentarea modelului: supunerea modelului la o varietate de condiții şi înregistrarea dimensiunilor utilității generate în fiecare caz. Dacă dimensiunea utilității variază într-o manieră ce diferă de aşteptări atunci există motive să se creadă că funcția obiectiv nu este corectă. De exemplu, să presupunem că un model trebuie să estimeze valoarea de piață a caselor pentru o singură familie. Modelul trebuie să indice valoarea în dolari în funcție de numărul de metri pătrați locuibili, de numărul dormitoarelor, băilor şi mărimea grădinii. După dezvoltarea modelului, analistul îl verifică aplicându-l în evaluarea unor case care au diferite valori şi caracteristici. Dacă el constată că modelul său nu surprinde corect realitatea, atunci poate concluziona că acesta nu este bun şi că mai trebuie făcute unele modificări. În schimb dacă într-adevăr valoarea caselor reflectă cele patru caracteristici nici atunci problema nu este cu siguranță rezolvată pentru că rata de creştere a valorii casei poate nu este în aceeaşi proporție cu fiecare dintre variabile şi astfel trebuie studiată importanța fiecărei variabile şi coeficientul său la valoarea casei. Al doilea stadiu în validarea modelului cere o comparație a rezultatelor modelului cu cele obținute în realitate.

Optimizare

Oamenii au căutat sau au încercat să caute mult timp metode mai bune de a-şi îmbunătăți viața zilnică. De-a lungul istoriei omenirii s-a încercat la început găsirea unor surse mai bune de hrană şi apoi găsirea de surse de materiale, energie, etc. Relativ târziu în istoria omenirii s-a început să se formuleze şi să se rezolve probleme cantitative, mai întâi în cuvinte şi apoi prin simboluri scrise. Un aspect derivat şi legat de aceste probleme a fost căutarea "optimului", a "celui mai bun". De fapt şi în prezent în cea mai mare parte a timpului managerii caută să obțină o îmbunătățire a nivelul de performanță.

Eforturi masive s-au făcut pentru a descrie situații umane şi sociale complexe. Pentru ca acestea să capete o însemnătate ele trebuiau să fie descrise printr-o ecuație matematică cu una sau mai multe variabile ale căror valori trebuiau descoperite. Întrebarea care urmează a se pune este: ce valori trebuie să capete aceste variabile astfel încât expresia matematică să aibă cele mai bune valori (cele mai mici sau cele mai mari în funcție de cum se doreşte). Acest proces general de maximizare sau minimizare este denumit optimizare. Optimizarea, denumită şi programare matematică, ajută la găsirea răspunsului care duce la cel mai bun rezultat – cel care aduce cel mai mare profit, sau acela care aduce cele mai mici costuri, pierderi sau

disconfort. Adesea aceste probleme presupun utilizarea cât mai eficientă a resurselor incluzând bani, timp, utilaje, *staff*, etc.

Problemele de optimizare sunt deseori clasificate ca fiind liniare sau neliniare în funcție de relația din problemă care este sau nu liniară în report cu variabilele. În prezent există o varietate de pachete software care rezolvă problemele de optimizare. De exemplu LINDO, QSB, LINGO şi *What's Best!* rezolvă atât modele de programare liniară cât şi neliniară.

Programarea matematică, se confruntă în general cu probleme ale determinării şi alocării optimale a resurselor limitate astfel încât să se atingă obiectivele propuse. Obiectivele trebuie să reprezinte scopurile celui care ia decizia. Resursele pot reprezenta de exemplu materiale, oameni, bani, etc. Dintre toate posibilitățile de utilizare a resurselor este de dorit să se determine aceea sau acelea care maximizează sau minimizează *calitatea numerică* precum profitul sau costul.

Scopul optimizării globale este acela de a găsi cea mai bună soluție adecvată la modelele cele mai dificile în condițiile în care există mai multe soluții posibile.

Descrierea şi formularea problemelor de programare liniară

Descriere

Programarea liniară (*Linear Programming* - LP) este o procedură care a găsit o largă aplicare practică în aproape toate domeniile de activitate. Ea poate fi utilizată în afaceri, în reclamă, în planificare, în producţie, etc. Transportul, distribuţia şi planificarea producţiei sunt problemele tipice studiate prin LP. În USA industria petrolului pare a fi cea în care LP este cel mai mult utilizată. Un manager al unei mari companii petroliere a estimat că între 5% şi 10% din timpul în care se utilizează computerele în companie este destinat analizei şi creării de modele LP.

LP rezolvă un tip de probleme (de programare) în care sunt liniare atât funcţia obiectiv care trebuie optimizată cât şi relaţiile dintre variabilele ce definesc resursele.

Acest tip de probleme a fost formulat şi rezolvat pentru prima dată la sfârşitul anilor 1940. Mai rar s-a întâlnit o tehnică matematică care să găsească o asemenea răspândire practică în: afaceri, comerţ, aplicaţii industriale şi mai rar s-a întâmplat ca o tehnică de acest fel să fie dezvoltată atât de mult şi atât de repede la nivel teoretic. Astăzi programarea liniară este utilizată cu succes în probleme de bugetare a capitalului, design, diete, conservarea resurselor, jocuri de strategie, jocuri de război, prevederea creşterilor economice, sisteme de transport, etc.

Este foarte important să se înţeleagă că programarea liniară (LP) este diferită de programarea care se referă la programarea pe calculator. În primul caz programarea (LP) înseamnă a plănui şi a organiza în timp ce în al doilea caz programarea pe calculator înseamnă scrierea de instrucţiuni pentru a realiza calcule şi programe de calculator. Cunoştinţele într-una dintre programări nu au aproape nici o importanţă pentru celălalt tip de programare. De fapt, termenul "programare liniară" a fost inventat înainte ca termenul "programare" să fie asociat cu programarea pe calculator. Această confuzie este deseori evitată prin folosirea termenilor de optimizare liniară în loc de programare liniară.
Orice problemă LP constă în existenţa unei funcţii obiectiv şi a unor restricţii.

Când se formulează o problemă de decizie ca o problemă de programare liniară trebuie verificate următoarele condiţii:

- **Funcţia obiectiv** trebuie să fie liniară. Aceasta înseamnă ca toate variabilele trebuie să fie la puterea 1 şi să fie doar adunate sau scăzute (nu înmulţite sau împărţite).
- **Obiectivul** trebuie să fie ori maximizarea ori minimizarea funcţiei liniare numită funcţia obiectiv. Obiectivul trebuie să reprezinte scopul deciziei.
- **Restricţiile** trebuie să fie de asemenea liniare. Restricţiile pot fi doar ecuaţii sau inecuaţii.

Formularea problemelor de programare liniară

Orice problemă LP este alcătuită din patru componente principale:

- un set de variabile de decizie,
- parametrii,
- funcția obiectiv şi
- setul de restricții.

Variabilele de decizie pot fi asimilate cu input-urile controlabile.
Parametrii caracterizează input-urile necontrolabile. Acestea sunt de obicei valori numerice constante date.
Obiectivul trebuie să reprezinte scopul decidentului. Funcția obiectiv arată cum este legat obiectivul de variabilele de decizie. Ea poate fi o funcție de maximizare sau de minimizare.
Restricțiile reprezintă cererile ce trebuie satisfăcute. Ele pot fi restricții de egalitate sau de inegalitate.

Exemple

În continuare se va prezenta o problemă clasică de programare liniară care va ilustra aspectele prezentate mai sus. Modul în care va fi abordată această problemă este asemănător cu modul de abordate pentru cea mai mare parte dintre problemele de programare liniară care implică luarea de deciziilor.

Problema tâmplarului nr.1

Un tâmplar produce mese şi scaune pe care le vinde în piață cu un preț de 5$ pentru o masă şi 3$ pentru un scaun. El lucrează 2 ore pentru a produce o masă şi o oră pentru a produce un scaun. Numărul total de ore de muncă pe care le poate lucra într-o săptămână este de 40 ore. Cantitățile de materiale brute necesare pentru producție sunt: 1 unitate pentru o masă şi 2 unități pentru un scaun. Cantitatea totală de material furnizat într-o săptămână este de 50 unități.
Obiectivul său este acela de a afla câte mese şi scaune trebuie să producă pe săptămână pentru a-şi maximiza venitul.

Rezolvare:

Factorii restricțiilor care de obicei vin din exterior sunt reprezentați aici de limitele muncii (care vin din partea familiei – nu mai mult de 40 de ore pe săptămână) şi resursele de material brut de care dispune într-o săptămână (livrările de material se fac după un program fix – cantitatea maximă de material care poate fi furnizat într-o săptămână este doar de 50 de unități). Astfel, formularea LP este:

Variabile:
X1 reprezintă numărul de mese ce se vor produce

X2 reprezintă numărul de scaune ce se vor produce

Funcția obiectiv:
Maximizarea venitului obținut: Max (5 X1 + 3 X2)

Restricțiile:
2 X1 + X2 $\leqslant$ 40 restricția de muncă
X1 + 2 X2 $\leqslant$ 50 restricția de material
şi ambele X1, X2 sunt ne-negative.

Acesta este un modelul matematic al problemei expuse. Variabilele de decizie, adică input-urile controlabile sunt X1 şi X2. Output-ul pentru acest model este venitul total obținut într-o săptămână adică: 5 X1 + 3 X2. Toate funcțiile folosite în model sunt liniare. Coeficienții acestor restricții sunt:

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Ei mai sunt numiți factori tehnologici şi formează matricea tehnologică.

În urma rezolvării va rezulta soluția optimă: vor fi produse într-o săptămână X1=10 mese şi X2=20 scaune. Cu această strategie optimală va putea fi obținut un venit maxim de 110 dolari.

Problema prezentată a fost reală iar soluția oferită a fost o surpriză pentru tâmplar deoarece el obişnuia să producă într-o săptămână mai multe mese decât scaune gândindu-se că acestea costau mai mult.

Problema tâmplarului nr.2

După aflarea acestei soluții având în vedere faptul că cererea de mese pe piață era foarte mare, tâmplarul a dorit să ştie dacă îşi poate permite să angajeze un ajutor astfel încât să crească cantitatea de mese produse şi să răspundă în acest fel cerințelor pieței dar fără să-şi diminueze venitul de 110$. Dacă dorea, tâmplarul îşi putea găsi un ajutor pe care să-l plătească cu 2 dolari pe oră şi care să fie disponibil mai mult de 40 de ore pe săptămână. În esență el dorea să ştie dacă ar trebui să-şi angajeze un ajutor, şi dacă da, pentru câte ore?

Rezolvare:

Variabile:
X1 reprezintă numărul de mese ce se vor produce
X2 reprezintă numărul de scaune ce se vor produce
X3 este numărul de ore suplimentare pentru care îşi va angaja un ajutor

Funcția obiectiv:
Maximizarea venitului obținut: Max (5 X1 + 3 X2 - 2 X3)

Restricțiile:
2 X1 + X2 $\leqslant$ 40 + X3 restricția de muncă cu un număr X3 necunoscut de ore, care mai poate fi scrisă şi sub forma:
2 X1 + X2 – X3 $\leqslant$ 40

X1 + 2 X2 ⩽ 50 restricția de material
şi X1, X2, X3 sunt ne-negative.

Rezolvând problema vom constata că soluția optimă este X1=50 mese, X2=0 scaune, X3=60 ore cu un venit optim pentru tâmplar de 130 dolari (5X1+3X2-3X3=250+0-120). În acest fel tâmplarul chiar va câştiga în plus 20$. Deci el ar trebui să angajeze un ajutor pentru 60 de ore.

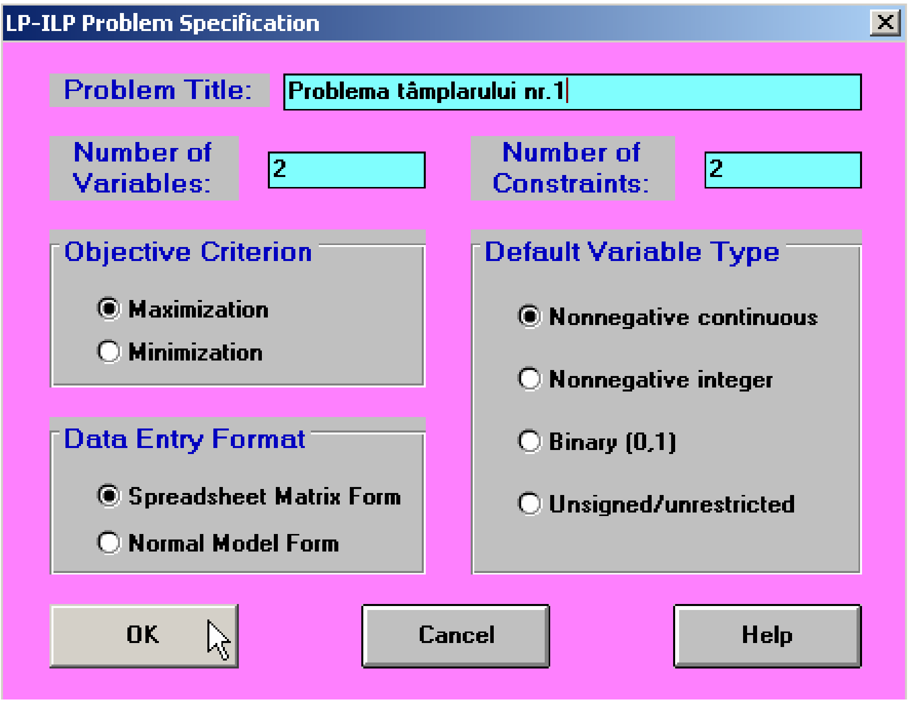
Rezolvarea problemelor utilizând un produs software specializat

Rezolvarea problemelor cu QSB

S-a ales pentru început exemplificarea pe baza produsului QSB.

Problema tâmplarului nr.1

Iată primul dialog care permite stabilirea tipului de problemă, a numărului de variabile şi a felului în care vor fi încărcate datele:



Următoarea formă matricială va fi utilizată pentru încărcarea datelor problemei:

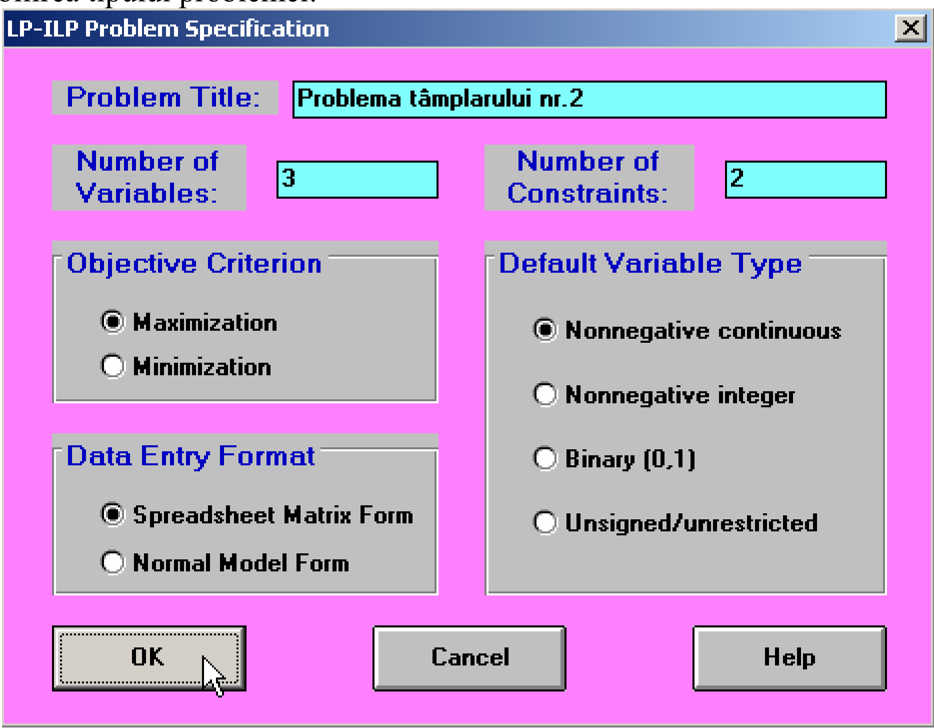
Variable -->	X1	X2	Direction	R. H. S.
Maximize	5	3		
C1	2	1	<=	40
C2	1	2	<=	50
LowerBound	0	0		
UpperBound	M	M		
VariableType	Continuous	Continuous		

După rezolvare vom obține soluția optimă:

	13:44:48		Saturday	October	19	2002		
	Decision Variable	Solution Value	Unit Cost or Profit c(j)	Total Contribution	Reduced Cost	Basis Status	Allowable Min. c(j)	Allowable Max. c(j)
1	X1	10.0000	5.0000	50.0000	0	basic	1.5000	6.0000
2	X2	20.0000	3.0000	60.0000	0	basic	2.5000	10.0000
	Objective	Function	(Max.) =	110.0000				
	Constraint	Left Hand Side	Direction	Right Hand Side	Slack or Surplus	Shadow Price	Allowable Min. RHS	Allowable Max. RHS
1	C1	40.0000	<=	40.0000	0	2.3333	25.0000	100.0000
2	C2	50.0000	<=	50.0000	0	0.3333	20.0000	80.0000

Problema tâmplarului nr.2

Stabilirea tipului problemei:



Introducerea datelor:

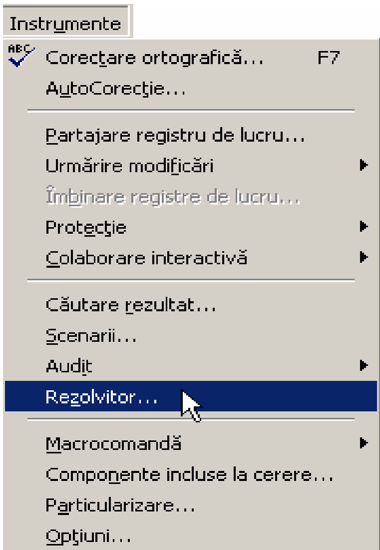
Variable -->	X1	X2	X3	Direction	R. H. S.
Maximize	5	3	-2		
C1	2	1	-1	<=	40
C2	1	2		<=	50
LowerBound	0	0	0		
UpperBound	M	M	M		
VariableType	Continuous	Continuous	Continuous		

Soluția optimă:

	14:00:22		Saturday	October	19	2002		
	Decision Variable	Solution Value	Unit Cost or Profit c(j)	Total Contribution	Reduced Cost	Basis Status	Allowable Min. c(j)	Allowable Max. c(j)
1	X1	50.0000	5.0000	250.0000	0	basic	4.5000	M
2	X2	0	3.0000	0	-1.0000	at bound	-M	4.0000
3	X3	60.0000	-2.0000	-120.0000	0	basic	-2.3333	0
	Objective	Function	(Max.) =	130.0000				
	Constraint	Left Hand Side	Direction	Right Hand Side	Slack or Surplus	Shadow Price	Allowable Min. RHS	Allowable Max. RHS
1	C1	40.0000	<=	40.0000	0	2.0000	-M	100.0000
2	C2	50.0000	<=	50.0000	0	1.0000	20.0000	M

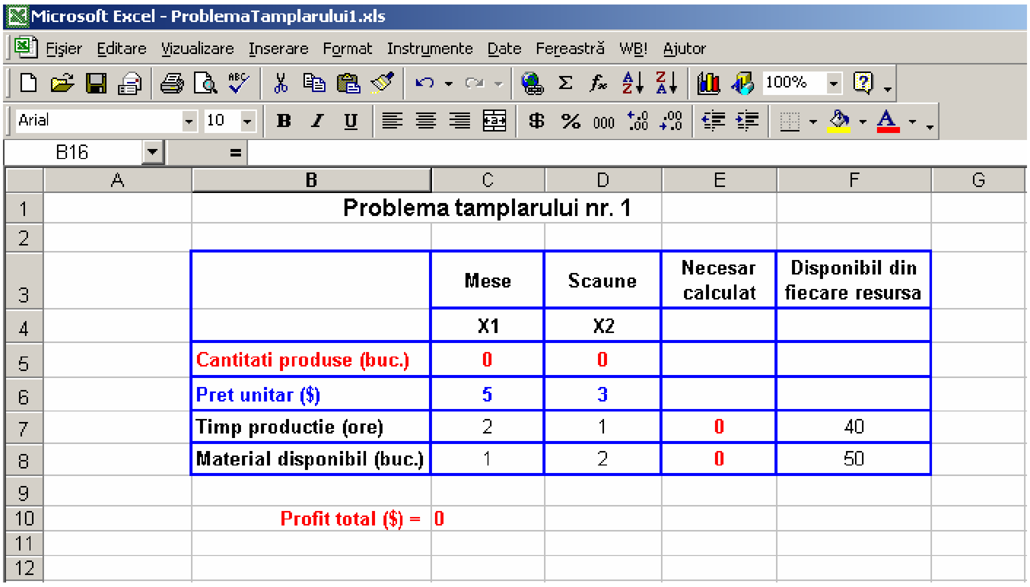
Rezolvarea problemelor cu Microsoft Excel

Programul Excel oferă de asemenea suport software pentru rezolvarea problemelor de programare liniară prin intermediului componentei *Solver* (tradus ca *Rezolvitor* în versiunile româneşti). Opțiunea care lansează *Solver*-ul se află în meniul *Tools (Instrumente)*. În mod implicit componenta nu este instalată. Ea se poate instala prin *Add-in* din cadrul aceluiaşi meniu. Iată meniul disponibil într-o versiune românească:



Opțiunea *Rezolvitor* apare doar după includerea ei prin *Componente introduse la cerere* ...

Pentru a fi posibilă rezolvarea problemei cu *Solver*-ul (*Rezolvitor*-ul) mai întâi trebuie introduse datele de intrare în celulele unei pagini Excel, apoi trebuie create câmpurile calculate.

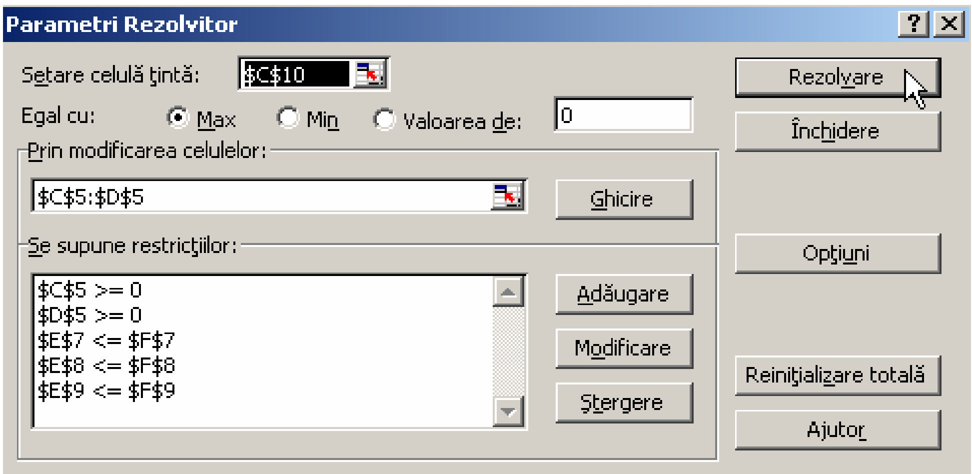


În pagina Excel de mai sus pentru o mai bună vizibilitate câmpurile calculate au fost formatate cu culoarea roşie. Ele reprezintă valorile de ieşire: X1, X2, necesarul calculat pentru fiecare resursă şi valoarea funcţiei obiectiv (profitul total). Formulele pentru necesarul celor două resurse şi pentru profit sunt:

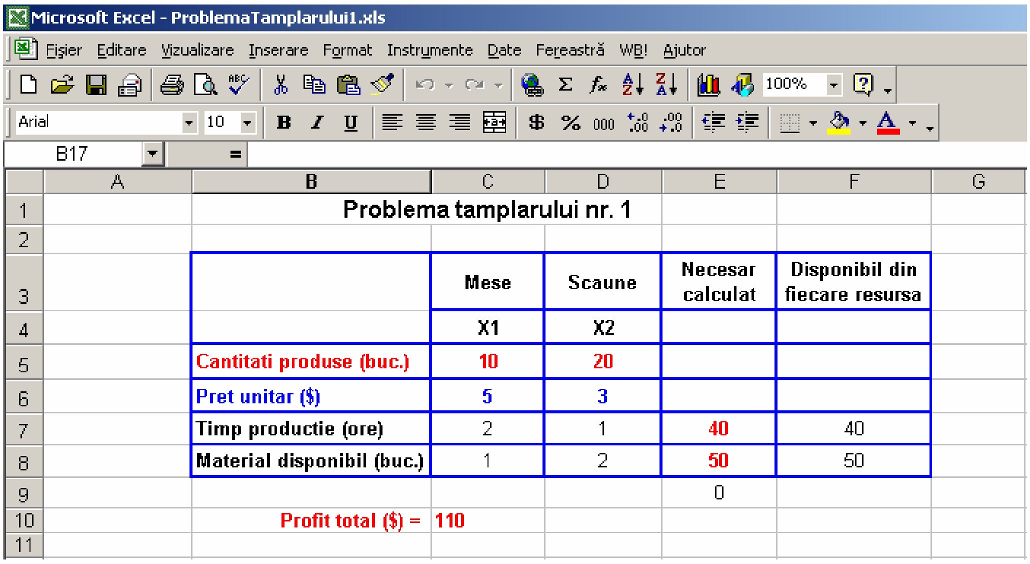
- Necesar resursa timp: E7=C7*C5+D7*D5
- Necesar resursa material: E8=C8*C5+D8*D5
- Profit total: C10=C6*C5+D6*D5

În continuare trebuie specificaţi în *Solver* (*Rezolvitor*) parametrii problemei:

- Celula ţintă (adică celula care conţine formula funcţiei obiectiv)
- Felul problemei (de maxim, de minim)
- Celulele ce se vor modifica (adică X1, X2)
- Restricţiile



După apăsarea butonului **Rezolvare** se va obține:

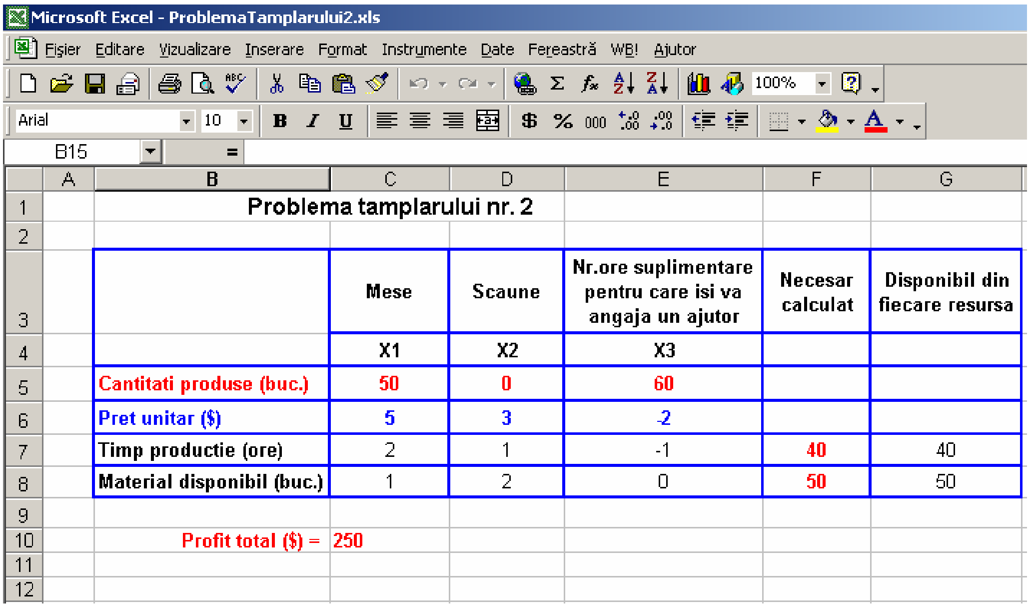


Microsoft Excel - ProblemaTamplarului1.xls

Problema tamplarului nr. 1

	Mese	Scaune	Necesar calculat	Disponibil din fiecare resursa
	X1	X2		
Cantitati produse (buc.)	10	20		
Pret unitar ($)	5	3		
Timp productie (ore)	2	1	40	40
Material disponibil (buc.)	1	2	50	50
			0	
Profit total ($) =	110			

În mod similar se rezolvă şi problema tâmplarului nr. 2:



Microsoft Excel - ProblemaTamplarului2.xls

Problema tamplarului nr. 2

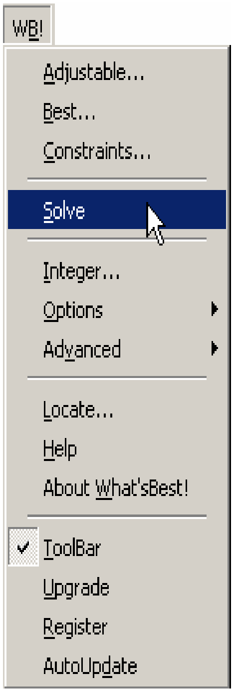
	Mese	Scaune	Nr.ore suplimentare pentru care isi va angaja un ajutor	Necesar calculat	Disponibil din fiecare resursa
	X1	X2	X3		
Cantitati produse (buc.)	50	0	60		
Pret unitar ($)	5	3	-2		
Timp productie (ore)	2	1	-1	40	40
Material disponibil (buc.)	1	2	0	50	50
Profit total ($) =	250				

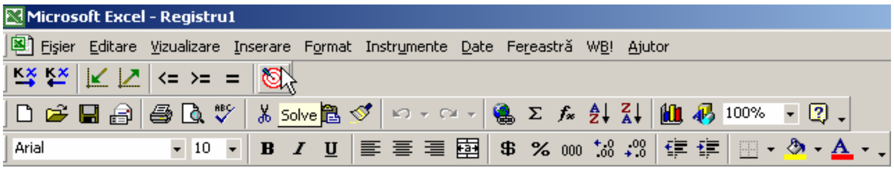
La această problemă formulele pentru necesarul celor două resurse şi pentru profit sunt:

- Necesar resursa timp: F7=C7*C5+D7*D5+E7*E5
- Necesar resursa material: F8=C8*C5+D8*D5+E8*E5
- Profit total: C10=C6*C5+D6*D5

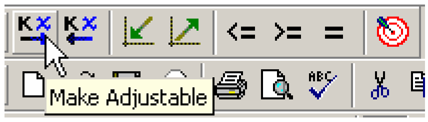
Rezolvarea problemelor cu What's Best!

Programul *What's Best!* se instalează ca o componentă Excel. După instalare sub Excel apare un nou meniu WB! Şi o nouă bară de instrumente (*What's Best!*):

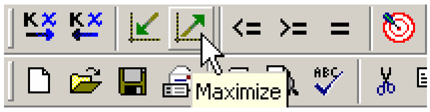




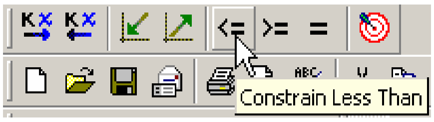
Definirea şi rezolvarea problemei se face parcurgând următorii patru paşi:



1) definirea celulelor alocate variabilelor X1, X2

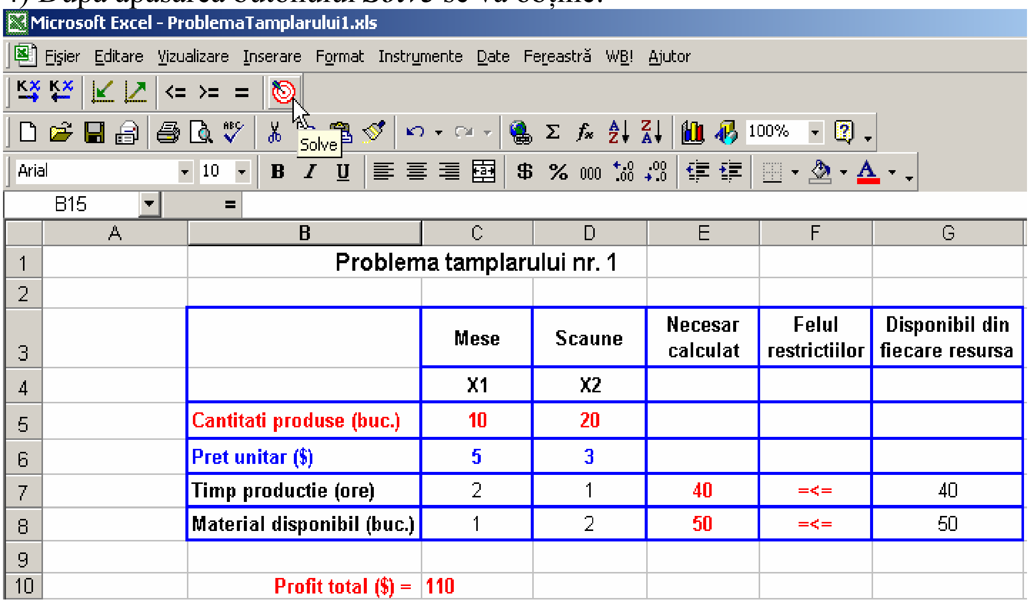


2) definirea tipului funcţiei obiectiv (Maxim sau Minim)



3) definirea tipului restricţiilor (<=, >= sau =).

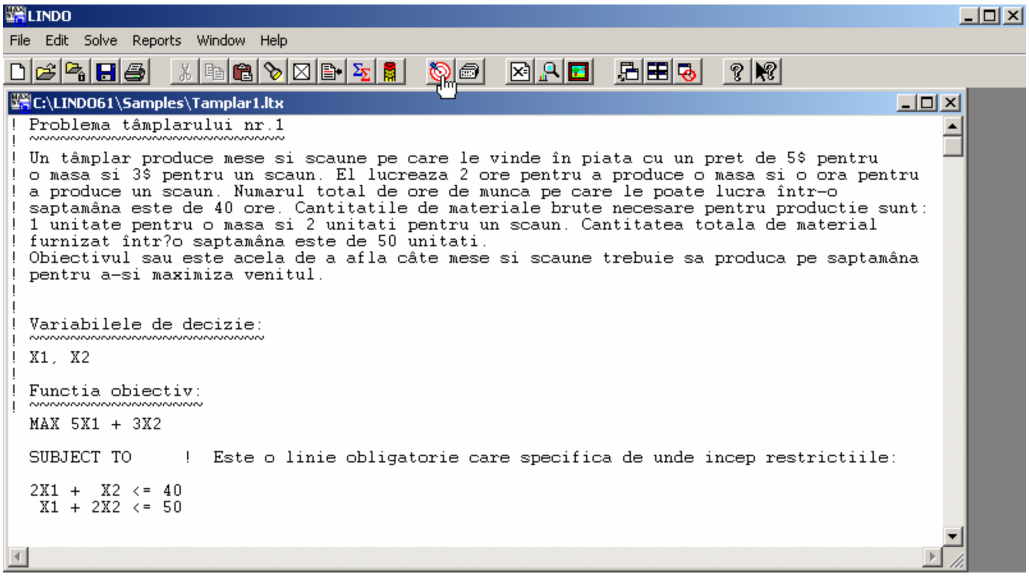
4) După apăsarea butonului ***Solve*** se va obţine:



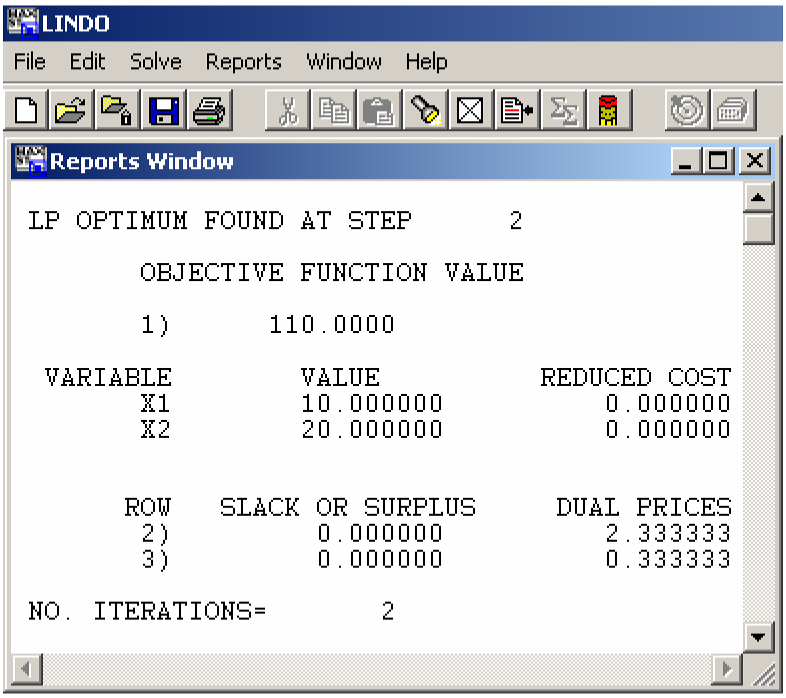
	A	B	C	D	E	F	G
1		Problema tamplarului nr. 1					
2							
3			Mese	Scaune	Necesar calculat	Felul restrictiilor	Disponibil din fiecare resursa
4			X1	X2			
5		Cantitati produse (buc.)	10	20			
6		Pret unitar ($)	5	3			
7		Timp productie (ore)	2	1	40	=<=	40
8		Material disponibil (buc.)	1	2	50	=<=	50
9							
10		Profit total ($) =	110				

Rezolvarea problemelor cu LINDO

Specific programului LINDO este faptul că oferă un editor de text în care poate fi înscrisă problema (sau adusă prin *Copy & Paste* chiar dintr-un alt editor). Atunci când se lucrează sub LINDO se poate utiliza semnul <!> drept specificator de comentariu. Întotdeauna restricțiile vor fi precedate de textul SUBJECT TO.



După apăsarea butonului (*Solve*) se va obține:



Problema duală

Construcția problemei duale şi semnificația ei

Problema duală este o problemă ce poate fi asociată oricărei probleme de programare liniară (LP).

Construcția problemei duale

Întotdeauna într-o problemă de programare liniară numită problema primală (inițială) sunt valabile următoarele reguli:

- dacă primala este o problemă de maximizare, atunci problema duală asociată ei va fi o problemă de minimizare (şi invers);
- elementele din partea dreaptă a restricțiilor (*Right Hand Side* – RHS) dintr-o problemă (primală sau duală) devin coeficienții funcției obiectiv ai celeilalte probleme (şi invers);
- coeficienții matricei restricțiilor unei probleme (primale sau duale) se obțin din transpusa matricei coeficienților restricțiilor celeilalte probleme;
- va exista câte o restricție în duală pentru fiecare variabilă din problema primală şi invers;
- valorile din partea dreaptă a restricțiilor din duală vor fi egale cu coeficienții funcției obiectiv din primală luați în ordine şi invers;
- coeficienții primei restricții din primală vor deveni coeficienții primei variabile în fiecare din restricțiile din duală ş.a;
- tipul variabilelor din duală ($\geqslant$ sau $\leqslant$) este dat sensul restricțiilor ($\geqslant$ sau $\leqslant$).

Exemple

Fiind dată următoarea problemă primală să se găsească duala ei.

Problema primală	Problema duală
Variabilele de decizie: x1, x2 **Funcția obiectiv:** Min (x1 - 2x2) **Restricțiile:** $x1 + x2 \geqslant 2$ $x1 - x2 \geqslant -1$ $x2 \geqslant 3$, şi $x1, x2 \geqslant 0$.	**Variabilele de decizie:** u1, u2, u3 **Funcția obiectiv:** Max (2u1 - u2 + 3u3) **Restricțiile:** $u1 + u2 \leqslant 1$ $u1 - u2 + u3 \leqslant -2$ $u1, u2, u3 \geqslant 0$

Problema tâmplarului

Problema primală	Problema duală
Variabilele de decizie: x1, x2 **Funcția obiectiv:** Max (5x1 + 3x2) **Restricțiile:** $2x1 + x2 \leqslant 40$ $x1 + 2x2 \leqslant 50$ $x1 \geqslant 0$ $x2 \geqslant 0$	**Variabilele de decizie:** u1, u2 **Funcția obiectiv:** Min (40u1 + 50u2) **Restricțiile:** $2u1 + u2 \geqslant 5$ $u1 + 2u2 \geqslant 3$ $u1 \geqslant 0$ $u2 \geqslant 0$

Problema duală poate fi folosită într-o mare varietate de aplicații. În unele cazuri duala poate fi mai eficientă decât primala.

Soluția dualei oferă interpretări economice importante precum prețurile umbră (*shadow prices*). **Prețurile umbră reprezintă soluția problemei duale.**

Dacă problema primală este o problemă de maximizare (a profitului) atunci prețurile umbră arată cât profit aduce fiecare unitate de resursă consumată. Concret în problema tâmplarului u1 arată cât profit aduce o oră de muncă iar u2 determină profitul obținut la o unitate de material brut consumat. Această analiză îl poate ajuta pe manager să afle care utilizare a resursei aduce cel mai mare profit.

Duala problemei tâmplarului şi interpretarea ei

Datele problemei primale:

	Masă	Scaun	Disponibil
Muncă (ore)	2	1	40
Material brut (buc)	1	2	50
Venit net ($)	5	3	

Formularea LP problemei primale:

X1 şi X2 reprezintă numărul de mese respectiv scaune care trebuie produse.

Max (5 X1 + 3 X2)

$2 X1 + X2 \leqslant 40$ restricția legată de muncă
$X1 + 2 X2 \leqslant 50$ restricția legată de material
şi ambele X1, X2 sunt nenegative.

Să presupunem că tâmplarul doreşte să-şi facă o asigurare pentru venitul net.

Din **modelul dual** putem desprinde următoarele semnificații pentru variabile:

- U1 = Suma de dolari care i se plăteşte tâmplarului pentru fiecare oră de muncă pierdută (datorată de exemplu îmbolnăvirii).
- U2 = Suma de dolari care i se plăteşte tâmplarului pentru fiecare unitate de material brut pierdută (datorată de exemplu unui incendiu).

În mod clar societatea de asigurări va încerca să minimizeze suma totală de dolari (40U1 + 50U2) care trebuie plătită tâmplarului de către compania de asigurări.

Tâmplarul va pretinde companiei de asigurări să-i despăgubească întreaga pierdere adică întregul venit net deoarece el nu va mai putea produce acea marfă.

Astfel, problema companiei de asigurări devine:

Min (40 U1 + 50 U2)

2U1 + 1U1 ⩾ 5 Venitul net de la o masă
1U1 + 2U2 ⩾ 3 venitul net de la un scaun
şi U1, U2 are nenegative.

Dacă rezolvăm problema cu un pachet de programe specializat vom obţine următoarea soluţie optimă: U1 = 2,33333$ şi U2 = 0,33333$ cu valoarea optimă de 110$ (exact suma pe care tâmplarul se aşteaptă să o primească ca asigurare).

După cum se observă din problemă tâmplarului şi din duala sa, valoarea optimă este întotdeauna aceeaşi (110$) pentru ambele situaţii. Acest fapt este cunoscut în economie ca echilibrul dintre problema primală şi cea duală.

Preţurile umbră – calculul şi semnificaţia lor

Comportamentul schimbărilor în valorile RHS ale valorii optime

RHS - *Right Hand Side* reprezintă valoarea din partea dreaptă a unei restricţii. Pentru a studia schimbările direcţionale în valoarea optimă ţinând cont de schimbările posibile din RHS, distingem următoarele două cazuri:

Cazul I: Problema de maximizare

- Pentru restricţia ⩽ schimbarea se face în aceeaşi direcţie. Deci creşterea valorii RHS nu descreşte valoarea optimă.
- Pentru restricţia ⩾ schimbarea se face în direcţia opusă. Deci creşterea valorii RHS nu duce la creşterea valorii optime. Ea descreşte sau rămâne neschimbată în funcţie de restricţie.
- Pentru restricţia = schimbarea se poate produce în ambele direcţii.

Cazul II: Problema de minimizare

- Pentru restricția ⩽ schimbarea se face în direcția opusă. Deci creşterea valorii RHS nu creşte valoarea optimă, ea descreşte sau cel mult rămâne aceeaşi.
- Pentru restricția ⩾ schimbarea se produce în aceeaşi direcție. Deci creşterea valorii RHS nu scade valoarea optimă, ea creşte sau cel mult rămâne aceeaşi.
- Pentru restricția = schimbarea se poate produce în ambele direcții.

Interpretarea prețului umbră

Prețul umbră reprezentat de o variabilă duală ne arată în ce măsură se va schimba valoarea funcției obiectiv dacă modificăm RHS-ul (adică valoarea din partea dreaptă a) restricției corespunzătoare variabilei. Aceasta se mai numeşte şi valoarea marginală, preț dual sau valoare duală. Pentru fiecare restricție, prețul umbră ne spune exact cu cât se va schimba valoarea funcției obiectiv dacă schimbăm limitele intervalului RHS (ceea ce se află în partea dreaptă a restricției).

Pentru fiecare valoare RHS, prețul umbră este rația de schimbare în valoarea optimă cauzată de orice creştere sau descreştere permisă în RHS.

$$\text{Prețul umbră al unei resurse} = \frac{\text{Schimbarea în valoarea funcției obiectiv}}{\text{Schimbarea în valoarea RHS}}$$

Din nefericire există şi concepții greşite cu privire la definiția prețului umbră. O astfel de concepție este: “În problemele de programare lineară, prețul umbră al unei restricții este diferența între valoarea optimă a funcției obiectiv şi valoarea funcției obiectiv obținută, atunci când partea dreaptă (RHS) a unei restricții este crescută cu o unitate”.

Discuție

Să considerăm următorul exemplu:

Max (X2)

X1 + X2 ⩽ 2
2.5X1 + 4X2 ⩽ 10

unde ambele variabile de decizie sunt nenegative.

Această problemă are soluția optimă pentru:

- X1 = 0, şi X2 = 2 şi
- valoare optimă rezultată = 2.

Dacă vom dori să calculăm prețul umbră al primei resurse, atunci când RHS-ul acesteia creşte cu o unitate, problema va deveni:

Max (X2)

X1 + X2 $\leqslant$ 3
2.5X1 + 4X2 $\leqslant$ 10

unde ambele variabile de decizie sunt nenegative.

Noua problemă are soluția optimă pentru:

- X1 = 0, şi X2 = 2.5 şi
- valoare optimă rezultată = 2.5.

De aceea pare că prețul umbră pentru această resursă este 2.5 – 2 = 0.5
Dar, de fapt dacă vom calcula prețul umbră prin rezolvarea corectă a problemei duale vom obține pentru această resursă prețul umbră = 1.

Prețul umbră este întotdeauna nenegativ?

Răspunsul la această întrebare depinde integral de formularea primalei şi a dualei. Ceea ce este de reținut este că prețul umbră al unui RHS dat este rata de schimbare a valorii optime ținând cont şi de schimbarea acelui RHS, schimbarea fiind între limitele senzitivității acelui RHS.

Să considerăm următorul exemplu numeric:

Max (3X1 + 5X2)

X1 + 2X2 $\leqslant$ 50
-X1 + X2 $\geqslant$ 10

X1, X2 sunt nenegative.

Ne propunem să aflăm prețul umbră al RHS2 = 10. Pentru aceasta va trebui să formulăm şi apoi să rezolvăm problema duală:

Min (50U1 + 10U2)

U1 - U2 $\geqslant$ 3
2U1 + U2 $\leqslant$ 5

Soluția dualei este U1 = 2,66, U2 = - 0,33. Deci prețul umbră corespondent pentru RHS2 = 10 este U2 = - 0,3. Aceasta înseamnă că pentru fiecare creştere/descreştere cu o unitate în valoarea RHS2 valoarea optimă pentru problema primală descreşte cu 0,33.

Prețuri umbră multiple

În acest sens întrebarea care se pune este: pentru o problemă LP care are o soluție optimă unică, este posibil să existe pentru un RHS mai mult de un preț umbră? Răspunsul este da.

Să considerăm următoarea problemă:

Min (16X1 + 24X2)

X1 + 3X2 ⩾ 6
2X1 + 2X2 ⩾ 4
X1, X2 ⩾ 0

Duala ei este:

Max (6U1 + 4U2)

U1 + 2U2 ⩽ 16
3U1 + 2U2 ⩽ 24

U1, U2 ⩾0

Această duală are mai multe soluții alternative:

- U1 = 8, U2 = 0 şi
- U1 = 4, U2 = 6.

Toate combinațiile convexe ale acestor puncte sunt şi ele soluții.
De fiecare dată când există redundanță în restricții sau dacă soluția optimă este "degenerată" ar putea exista mai mult decât un set de prețuri duale. În general, restricțiile liniare independente sunt o condiție suficientă pentru unicitatea prețurilor umbră.

Să considerăm acum următoarea problemă LP cu o restricție redundantă:

Max (10X1 + 13X2)

X1 + X2 = 1
X1 + X2 = 1
X1 + 2X2 = 2
X1, X2 sunt nenegative.

Dacă rulăm problema cu pachetul de programe LINDO vom obține:
X1 = 0, X2 = 1 cu prețurile umbră 0, 13 şi 0.
Dacă rulăm problema cu pachetul de programe QSB vom obține:
X1 = 0, X2 = 1 cu prețurile umbră 0,7,3.

În cazul redundanței, prețurile umbră obținute cu un produs software LP pot diferi de cele obținute cu un altul.

Probleme rezolvate

Problema 1

O companie care asamblează tehnică de calcul urmează să pornească producția a două tipuri de calculatoare. Fiecare din acestea necesită timp de asamblare, timp pentru testare şi spațiu de depozitare. Fiecare din aceste resurse este limitată. Managerul companiei îşi propune să determine cantitatea din fiecare tip de calculator pe care să o producă astfel încât să maximizeze profitul obținut în urma vânzării acestor calculatoare.

Informații suplimentare

Pentru a da o soluție corectă a problemei managerul a obținut de la laboratorul de producție şi financiar al companieiei următoarele informații:

	Calculator tip 1	**Calculator tip 2**
Profit unitar	60$	50$
Timp necesar pentru asamblare pe unitatea de produs	4h	10h
Timp necesar pentru testare pe unitatea de produs	2h	1h
Spațiu necesar pentru depozitarea unui produs	1m cub	1m cub

Resurse	**Disponibil (zilnic)**
Timp pentru asamblare	**100h**
Timp pentru testare	**22h**
Spațiu de depozitare	**12 m cubi**

De la compartimentul de marketing managerul află că în orice combinație se vor produce aceste tipuri de calculatoare pentru întreaga cantitate există cerere şi desfacere asigurată.

Rezolvare

Această problemă se prezintă ca o problemă de programare liniară.
În principiu soluția trebuie să fie exprimată în numere întregi însă chiar dacă rezultatele sunt numere fracționare acest lucru nu afectează în mod semnificativ soluția optimă.

Totodată se face presupunerea de nenegativitate a valorilor utilizate având în vedere că nu au sens valori negative pentru cantități, timp şi suprafețe.

Variabile

X1=numărul de calculatoare de tipul 1 care se vor produce
X2=numărul de calculatoare de tipul 2 care se vor produce

Funcția obiectiv

Max (60X1 + 50X2)

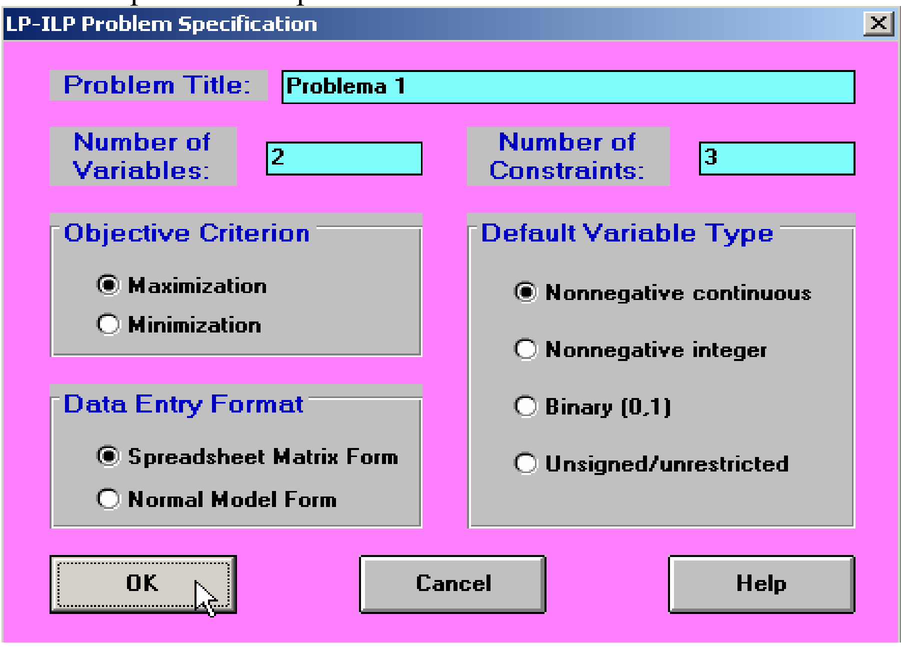
Restricții

* referitoare la timpul de asamblare:	4X1 + 10X2 <= 100
* referitoare la timpul de testare:	2X1 + X2 <= 22
* referitoare la spațiul de depozitare:	X1 + X2 <= 12
* restricții de nenegativitate:	X1 >= 0, X2 >= 0

Rezolvarea utilizând un produs software specializat

Rezolvarea problemei cu QSB

Stabilirea parametrilor problemei:



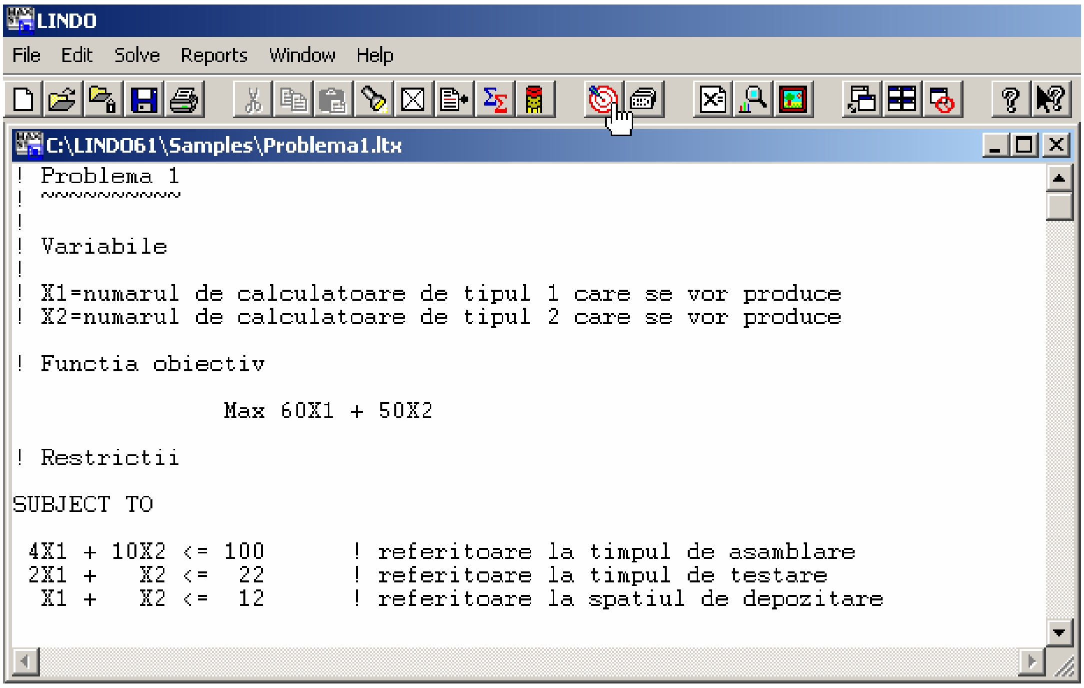
Încărcarea datelor:

Variable -->	X1	X2	Direction	R. H. S.
Maximize	60	50		
C1	4	10	<=	100
C2	2	1	<=	22
C3	1	1	<=	12
LowerBound	0	0		
UpperBound	M	M		
VariableType	Continuous	Continuous		

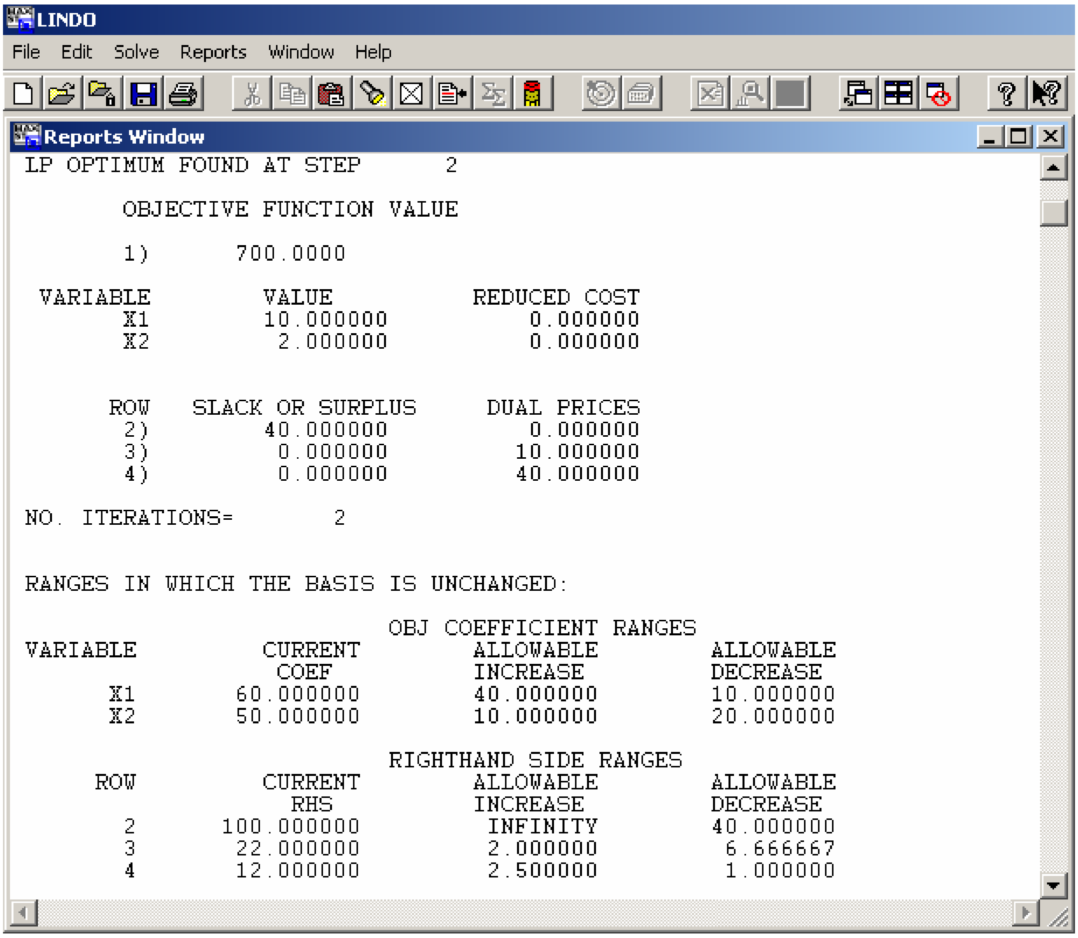
Soluția optimă:

	14:31:28		Saturday	October	19	2002		
	Decision Variable	Solution Value	Unit Cost or Profit c(j)	Total Contribution	Reduced Cost	Basis Status	Allowable Min. c(j)	Allowable Max. c(j)
1	X1	10.0000	60.0000	600.0000	0	basic	50.0000	100.0000
2	X2	2.0000	50.0000	100.0000	0	basic	30.0000	60.0000
	Objective	Function	(Max.) =	700.0000				
	Constraint	Left Hand Side	Direction	Right Hand Side	Slack or Surplus	Shadow Price	Allowable Min. RHS	Allowable Max. RHS
1	C1	60.0000	<=	100.0000	40.0000	0	60.0000	M
2	C2	22.0000	<=	22.0000	0	10.0000	15.3333	24.0000
3	C3	12.0000	<=	12.0000	0	40.0000	11.0000	14.5000

Rezolvarea problemei cu LINDO




```
! Problema 1
! ~~~~~~~~~~~
!
! Variabile
!
! X1=numarul de calculatoare de tipul 1 care se vor produce
! X2=numarul de calculatoare de tipul 2 care se vor produce

! Functia obiectiv

             Max 60X1 + 50X2

! Restrictii

SUBJECT TO

 4X1 + 10X2 <= 100      ! referitoare la timpul de asamblare
 2X1 +   X2 <=  22      ! referitoare la timpul de testare
  X1 +   X2 <=  12      ! referitoare la spatiul de depozitare
```




```
 LP OPTIMUM FOUND AT STEP       2

        OBJECTIVE FUNCTION VALUE

        1)      700.0000

  VARIABLE        VALUE          REDUCED COST
        X1        10.000000          0.000000
        X2         2.000000          0.000000


       ROW   SLACK OR SURPLUS     DUAL PRICES
        2)        40.000000          0.000000
        3)         0.000000         10.000000
        4)         0.000000         40.000000

 NO. ITERATIONS=       2

 RANGES IN WHICH THE BASIS IS UNCHANGED:

                           OBJ COEFFICIENT RANGES
 VARIABLE         CURRENT        ALLOWABLE        ALLOWABLE
                    COEF          INCREASE         DECREASE
       X1        60.000000        40.000000        10.000000
       X2        50.000000        10.000000        20.000000

                           RIGHTHAND SIDE RANGES
      ROW         CURRENT        ALLOWABLE        ALLOWABLE
                    RHS           INCREASE         DECREASE
        2       100.000000         INFINITY        40.000000
        3        22.000000         2.000000         6.666667
        4        12.000000         2.500000         1.000000
```

Problema 2

Formulați problema duală a problemei anterioare.

Pentru a formula problema duală trebuie să urmăm următorii paşi:

1. **Pentru funcția obiectiv :**

   a. Deoarece funcția obiectiv este o maximizare, în duală ea va fi o funcție de minimizare.
   b. Valorile din partea dreaptă a restricțiilor devin coeficienți ai funcției obiectiv a dualei.

   Min (100Y1 + 22Y2 + 12Y3)

S-au notat cu Y variabilele din duală pentru a le deosebi de cele din problema primală.. Va exista câte o variabilă în duală pentru fiecare restricție din primală.

2. **Pentru restricții :**

   a. Vom avea câte o restricție în duală pentru fiecare variabilă din problema primală. (Astfel, deoarece în problema primală avem două variabile, duala va avea două restricții)
   b. Valorile din partea dreaptă a restricțiilor din duală vor fi egale cu coeficienții funcției obiectiv din primală luați în ordine. Astfel, valoarea din partea dreaptă a primei restricții din duală va fi egală cu coeficientul primei variabile din funcția obiectiv a primalei iar valoarea din partea dreapta a celei de-a doua restricții din duală va fi egală cu coeficientul celei de-a doua variabile din funcția obiectiv a primalei.
   c. Coeficienții primei restricții din primală vor deveni coeficienții primei variabile în fiecare din restricțiile din duală. Astfel, coeficientul lui X1 din prima restricție din primală devine coeficientul lui Y1 în prima restricție din duală şi coeficientul lui X2 din prima restricție din primală devine coeficientul lui Y1 în cea de-a doua restricție din duală, ş.a.m.d.

   Restricțiile duale sunt :

   4Y1 + 2Y2 + 1Y3 >= 60
   10Y1 + 1Y2 + 1Y3 >= 50

Deci, dacă **problema primală** este:

Max (60X1 + 50X2)

4X1 + 10X2 <= 100
2X1 + 1X2 <= 22
1X1 + 1X2 <= 12
X1 >= 0, X2 >= 0

Atunci **problema duală** corespunzătoare va fi:

$$\text{Min } (100Y1 + 22Y2 + 12Y3)$$

$$\begin{aligned} 4Y1 + 2Y2 + 1Y3 &>= 60 \\ 10Y1 + 1Y2 + 1Y3 &>= 50 \\ Y1,Y2,Y3 &>= 0 \end{aligned}$$

Pentru a transforma o problemă primală în duala ei este mai uşor dacă toate restricțiile dintr-o problemă de maximizare sunt de forma <=, şi într-o problemă de minimizare restricțiile sunt de forma >=.

- Acest lucru se poate realiza prin înmulțirea cu (–1) a restricțiilor care nu corespund acestor criterii.
- Dacă o restricție este o egalitate, ea trebuie înlocuită cu două restricții: una de >= şi cealaltă <=.

Exemplu: 4x1 + 5x2 = 20 poate fi înlocuită cu :
4x1 + 5x2 <= 20 şi 4x1 + 5x2 >= 20

Una din aceste două restricții se înmulţeşte cu (-1) în funcție de tipul problemei (de minim sau de maxim).

Iată în continuare **soluția duală** a problemei primale – problema 1.
(*Sensitivity Analysis for RHS*):

10-19-2002 14:35:57	Constraint	Direction	Shadow Price	Right Hand Side	Allowable Min. RHS	Allowable Max. RHS
1	C1	<=	0	100.0000	60.0000	M
2	C2	<=	10.0000	22.0000	15.3333	24.0000
3	C3	<=	40.0000	12.0000	11.0000	14.5000

Explicația economică a problemei duale

Orice problemă de programare liniară poate avea două forme. Formularea inițială a problemei poartă numele de *forma primală* iar cea de-a doua *forma duală.* Forma duală este un fel de imagine în oglindă a celei primale deoarece atât în formulare cât şi în soluție valorile duale sunt versiuni „*flip-flop*" ale valorilor primalei.

Soluțiile problemei primale conțin soluțiile problemei duale şi invers. Singura problemă care se pune este interpretarea rezultatelor soluției duale. Analiza problemei duale permite managerului să evalueze impactul potențial al unui nou produs şi se foloseşte pentru a determina valorile marginale ale resurselor (restricțiilor). În legătură cu un nou produs, un manager ar putea dori să ştie ce impact ar avea adăugarea unui nou produs asupra soluțiilor cantitative şi asupra profitului; în legătură cu resursele un manager ar putea folosi soluțiile dualei pentru a determina cât profit aduce fiecare unitate de resursă consumată. Această analiză ajută managerul să aleagă acea variantă de utilizare a resursei care aduce mai mult profit.

Probleme propuse

Problema 3

Un producător de cereale studiază posibilitatea introducerii pe piață a unui nou sortiment. Costul pe kg. şi rețeta sunt prezentate în tabelul următor:

	Făina	**Orez**	**Porumb**	**Cerințe pentru o cutie de 12 oz.**
PROTEINE (G/OZ)	4	2	2	>= 27g.
Carbohidrați (g/oz.)	20	25	21	>= 240g.
Calorii / oz.	90	110	100	<= 1260 calorii
Cost / oz..	0.03$	0.05$	0.02$	

Formulați modelul acestei probleme ca o problemă de programare liniară având drept scop determinarea cantităților optime de faină, orez şi porumb pentru o cutie astfel încât să se obțină un cost de producție minim.

Problema 4

Un investitor dispune de 100000$. El decide să îi investească în obligațiuni, titluri de proprietate şi o parte să-i depună într-un cont. După consultarea unor specialişti în finanțe el asimilează ca necesară respectarea unor condiții suplimentare:

- să nu investească mai mult de 40% din suma în obligațiuni

- suma depusă în cont să fie cel puțin dublă față de suma investită în titluri de proprietate

Dobânda anuală este 8% la obligațiuni,9% la titluri şi 7% la cont.

Se presupune că întreaga sumă va fi investită şi că aceste dobânzi vor rămâne constante întreaga perioadă.

Scopul investitorului este de a maximiza profitul anual.

Formulați problema ca o problemă de programare liniară presupunând că nu se percep comisioane pentru tranzacții (deschidere cont, retrageri, răscumpărări titluri).

Problema 5

O fabrică de jucării produce 3 variante de roboți de jucărie.

Prima necesită 10 minute timp de fabricație şi ambalare şi 700g de plastic, a doua variantă necesită 12 minute şi 1050g plastic iar cea de-a treia 15 minute şi 1400g plastic.

În următorul ciclu de producție există 8 ore timp de fabricație şi ambalare disponibil pentru aceste sortimente şi 70kg de plastic.

Profitul obținut în urma comercializării unui robot de primul tip este de 1$, al doilea tip 5$, al treilea 6$.

Există o comandă anterioară care trebuie onorată din această producție, constând în 10 roboți din fiecare tip.

Formulați problema ca o problemă de programare liniară, pentru a determina cantitățile din fiecare tip ce trebuie produse pentru a asigura maximizarea profitului.

Problema 6

O companie de turism are o cerere de transport pentru 500 de persoane. Aceasta companie dispune de trei tipuri de mijloace de transport cu următoarele caracteristici:

Mijloace de transport	**Nr.de locuri**	**Consum de combustibil (l/100Km)**	**Număr de mijloace de transport disponibile**
Tip 1	30	12	6
Tip 2	50	18	5
Tip 3	45	17	4

Să se determine câte mijloace de transport din fiecare tip sunt necesare pentru transportul pasagerilor astfel încât costul transportului să fie minim.

Problema 7

O companie de producție are două utilaje cu care realizează două produse: A şi B. Fiecare din aceste produse este prelucrat pe ambele utilaje. Tabelul următor prezintă necesarul de timp de prelucrare a fiecărui produs pe cele două maşini şi disponibilul de timp al fiecărui utilaj într-o lună:

Utilaj	Produsul A (ore)	Produsul B (ore)	Timp disponibil într-o lună (ore)
1	2	3	180
2	3	2	150

Considerând că această companie are desfacerea asigurată pentru întreaga producție şi prețul este de 50$ pentru produsul A şi 60$ pentru produsul B, managerul companiei doreşte să stabilească structura producției astfel încât să realizeze o maximizare a profitului în luna următoare.

Problema 8

Un plan de nutriție cere consumarea a cel puțin 200 unități de proteine şi 180 unități de grăsimi. Analizele chimice arată că o unitate din alimentul A conține 6 unități de proteine şi 3 unități de grăsimi, iar o unitate din alimentul B conține 3 unități de proteine şi 5 unități de grăsimi. Consumul din cele doua tipuri de alimente se face numai în unități întregi. Prețul de cumpărare este de 2.5$ pentru o unitate din produsul A şi 2$ pentru o unitate din produsul B. Câte unități din fiecare aliment trebuiesc consumate astfel încât să fie satisfăcute cerințele de proteine şi grăsimi iar costul să fie minim ?

Problema 9

Un avion cargo are 3 compartimente pentru depozitarea încărcăturii: compartimentul din față, central, şi cel din spate. Acestea au următoarele limitări în ceea ce priveşte greutatea acceptată şi spațiul disponibil:

Compartiment	Greutate capacitate (tone)	Spațiu capacitate (metri cubi)
Față	10	6800
Centru	16	8700
Spate	8	5300

În plus, greutatea încărcăturii depozitate în respectivele compartimente trebuie să fie distribuită proporțional cu greutatea maxim admisă în fiecare compartiment astfel încât să se mențină echilibrul navei.

Cu următorul zbor trebuie transportate următoarele încărcături:

Încărcătura	Greutate (tone)	Volum (metri cubi/tone)	Profit ($)
C1	18	480	310
C2	15	650	380
C3	23	580	350
C4	12	390	285

Transportul acestora poate fi acceptat în orice proporții. Obiectivul este acela de a determina cât din fiecare încărcătură să fie transportat şi cum să fie distribuită încărcătura între compartimente astfel încât profitul total pe acest zbor să fie maxim.

Formulați modelul problemei de programare liniară.
Ce presupuneri se fac în formularea modelului ca o problemă de programare liniară ?
Explicați avantajele care decurg din rezolvarea problemei ca o problemă de programare liniară comparativ cu rezolvarea acesteia printr-o aproximare nefundamentată matematic.

Problema 10

O companie are două fabrici de conservare a fructelor. Există trei furnizori de fructe proaspete în următoarele cantități şi la următoarele prețuri:

- **S1**: 200t la 1100$ / t
- **S2**: 310t la 1000$ / t
- **S3**: 420t la 900$ / t

Costul transportului pe tonă este:

	La fabrica:	**A**	**B**
De la:	**S1**	300 $ / t	350 $ / t
	S2	200 $ / t	250 $ / t
	S3	600 $ / t	400 $ / t

Capacitatea de producție în tone şi costul prelucrării pentru fiecare fabrică este:

Fabrica	**A**	**B**
Capacitatea de producție	460 t	560 t
Costul prelucrării	2600 $ / t	2100 $ / t

Conservele de fructe se vând tuturor distribuitorilor companiei cu 5000 $/tonă. Compania are cerere pentru întreaga cantitate de conserve pe care o produce.

Obiectivul companiei este acela de a achiziționa fructe într-o anume proporție de la fiecare furnizor pentru a ocupa capacitatea de producție a fiecărei fabrici astfel încât să se maximizeze profitul la nivelul companiei.

- Formulați problema ca o problemă de programare liniară şi explicați-o.
- Explicați semnificația valorilor duale asociate cu restricțiile corespunzătoare (asociate) cantităților furnizate şi capacității de producție a fabricilor.
- Ce presupuneri trebuie făcute pentru a exprima problema ca o problemă de programare liniară.

Problema 11

O companie asamblează 4 tipuri de produse (1, 2, 3, 4) din componente. Profitul unitar pentru fiecare produs din cele 4 tipuri este:10$ / produs, 15$ / produs, 22$ / produs si respectiv 17$ / produs.

Comenzile pentru fiecare din cele patru produse (1, 2, 3, 4) în săptămâna următoare sunt de 50, 60, 85 şi respectiv 70 bucăți.

Fiecare produs necesită pentru asamblare 3 operații(A,B,C) care fiecare necesită un consum de om-ore pe produs diferite:

		Produs			
		1	**2**	**3**	**4**
Operație	**A**	2	2	1	1
	B	2	4	1	2
	C	3	6	1	5

Timpul disponibil în următoarea săptămână pentru fiecare operație (A,B,C) de asamblare este de: 160, 200 şi respectiv 80 om-ore.

Este admis ca muncitorii angajați pentru operația B să utilizeze maximum 20% din timpul de lucru pentru a efectua operația A (probabil cu un nivel de periculozitate ridicat sau radiații) iar cei angajați pentru a efectua operația C pot folosi până la 30% din timpul de lucru pentru a efectua operația A (altfel trebuie acordate sporuri sau trebuie să fie încadrați în alte grupe de muncă cu alt salariu).

Necesitățile de producție impun ca raportul dintre numărul de produse tip 1 asamblate şi numărul de produse de tip 4 asamblate să fie cuprins între 0.9 şi 1.15.

Formulați modelul problemei ca o problemă de programare liniară.

Planificarea programului de lucru al personalului (*Staff Scheduling*)

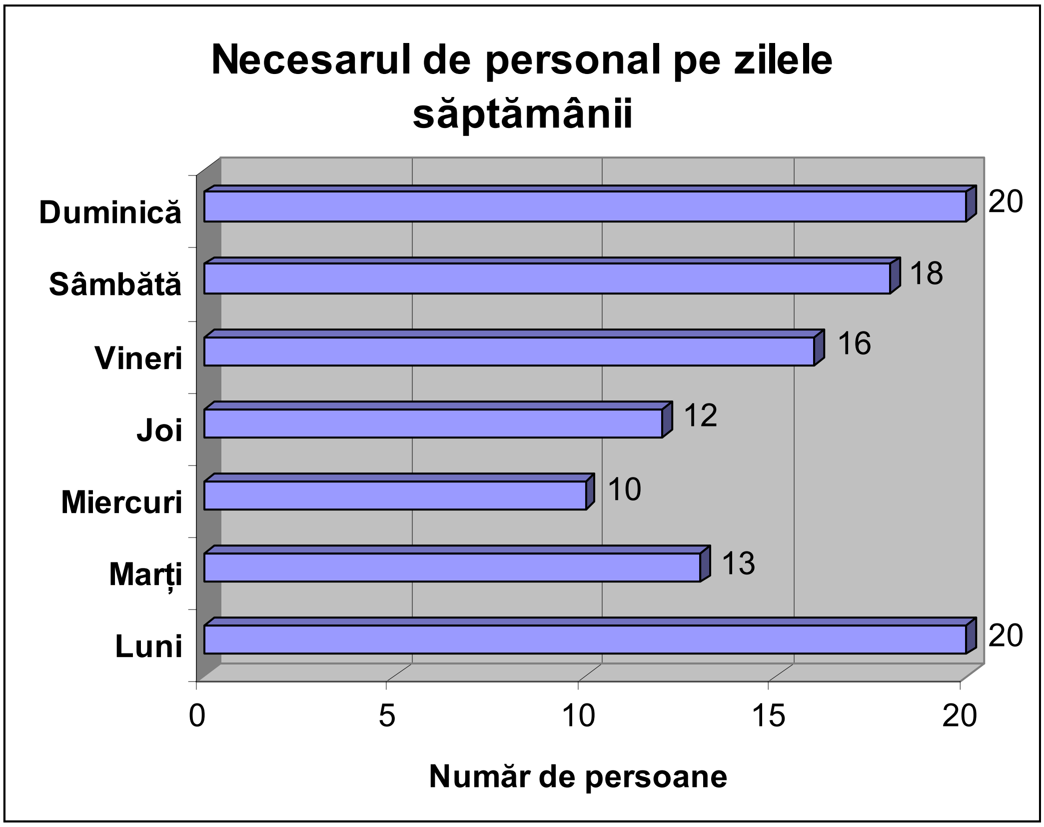
În problemele de planificare a programului de lucru al personalului se urmăreşte găsirea unei soluţii care să minimizeze costul ce trebuie plătit angajaţilor în funcţie de numărul de ore lucrate pe săptămână. Soluţia va ţine seama de un necesar minim de personal ce trebuie asigurat în fiecare zi a săptămânii şi va determina numărul de persoane care vor lucra şi care vor fi libere în fiecare zi (din săptămână).

Acest tip de planificare (a turelor) se foloseşte foarte mult în domenii cum sunt: zborurile aeriene, programarea turelor din spitale, restaurante, magazine etc.

Iată în continuare un exemplu care va uşura înţelegerea problemelor încadrate în această categorie numită *Staff Scheduling*.

Exemplul 1

În urma unui studiu efectuat la un magazin s-a constat următorul necesar de personal vânzător pentru fiecare zi a săptămânii:



Un vânzător primeşte 60 $ pentru fiecare zi lucrătoare muncită. Dacă lucrează sâmbăta el primeşte în plus 25 $ iar dacă lucrează duminica primeşte în plus 35 $.

Fiecare vânzător poate lucra doar 5 zile pe săptămâna si apoi trebuie sa aibă doua zile libere consecutiv.

Managerul magazinului doreşte o planificare optimă care să acopere necesarul de vânzători şi care să ofere soluţia pentru cea mai mică sumă totală de plată săptămânală ce trebuie plătită vânzătorilor.

Rezolvare

În următorul tabel este prezentată o planificare exaustivă (un grafic) pentru fiecare zi a săptămânii. S-au notat cu (T) zilele în care un vânzător va lucra în tură şi cu (L) zilele libere:

Planificare ziua start	Denumire variabile	Luni	Marți	Miercuri	Joi	Vineri	Sâmbătă	Duminică	Salariu săptămânal ($)
Luni	**Lu_X**	T	T	T	T	T	L	L	300
Marți	**Ma_X**	L	T	T	T	T	T	L	325
Miercuri	**Mi_X**	L	L	T	T	T	T	T	360
Joi	**Jo_X**	T	L	L	T	T	T	T	360
Vineri	**Vi_X**	T	T	L	L	T	T	T	360
Sâmbătă	**Sa_X**	T	T	T	L	L	T	T	360
Duminică	**Du_X**	T	T	T	T	L	L	T	335

Funcția obiectiv

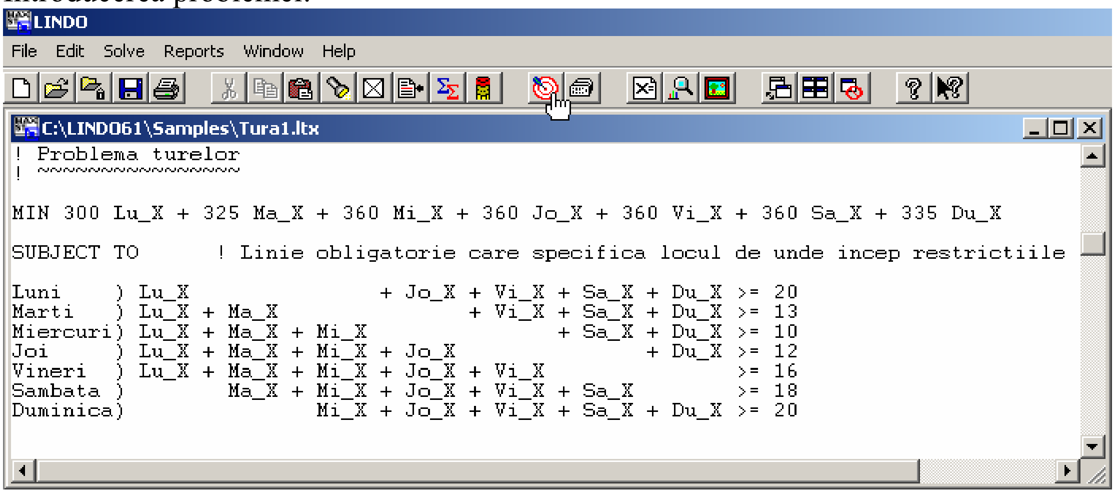
```
Min(300Lu_X + 325Ma_X + 360Mi_X + 360Jo_X + 360Vi_X +
360Sa_X + 335Du_X)
```

Restricțiile

```
Lu_X                 + Jo_X + Vi_X + Sa_X + Du_X >= 20
Lu_X + Ma_X                 + Vi_X + Sa_X + Du_X >= 13
Lu_X + Ma_X + Mi_X                 + Sa_X + Du_X >= 10
Lu_X + Ma_X + Mi_X + Jo_X                 + Du_X >= 12
Lu_X + Ma_X + Mi_X + Jo_X + Vi_X                 >= 16
       Ma_X + Mi_X + Jo_X + Vi_X + Sa_X          >= 18
              Mi_X + Jo_X + Vi_X + Sa_X + Du_X >= 20
```

Rezolvarea prin intermediul unui produs software specializat

Introducerea problemei:



Prima fereastră a soluției:

```
Reports Window

LP OPTIMUM FOUND AT STEP       8

        OBJECTIVE FUNCTION VALUE

        1)      7750.000

  VARIABLE        VALUE          REDUCED COST
      LU_X         2.000000          0.000000
      MA_X         0.000000        100.000000
      MI_X         2.000000          0.000000
      JO_X         7.000000          0.000000
      VI_X         5.000000          0.000000
      SA_X         4.000000          0.000000
      DU_X         2.000000          0.000000


       ROW   SLACK OR SURPLUS     DUAL PRICES
     LUNI)         0.000000       -100.000000
    MARTI)         0.000000          0.000000
 MIERCURI)         0.000000       -100.000000
      JOI)         1.000000          0.000000
   VINERI)         0.000000       -100.000000
  SAMBATA)         0.000000        -25.000000
 DUMINICA)         0.000000       -135.000000

 NO. ITERATIONS=       8
```

A doua fereastră a soluției:

```
LINDO
File  Edit  Solve  Reports  Window  Help

Reports Window

 RANGES IN WHICH THE BASIS IS UNCHANGED:

                           OBJ COEFFICIENT RANGES
 VARIABLE         CURRENT        ALLOWABLE        ALLOWABLE
                   COEF          INCREASE         DECREASE
     LU_X       300.000000       150.000000       300.000000
     MA_X       325.000000         INFINITY       100.000000
     MI_X       360.000000       150.000000       300.000000
     JO_X       360.000000         0.000000       100.000000
     VI_X       360.000000       150.000000         0.000000
     SA_X       360.000000       150.000000        25.000000
     DU_X       335.000000        25.000000       100.000000

                           RIGHTHAND SIDE RANGES
      ROW         CURRENT        ALLOWABLE        ALLOWABLE
                    RHS          INCREASE         DECREASE
     LUNI        20.000000         3.000000         1.500000
    MARTI        13.000000         1.000000         5.000000
 MIERCURI        10.000000         7.500000         1.500000
      JOI        12.000000         1.000000         INFINITY
   VINERI        16.000000         6.000000         1.500000
  SAMBATA        18.000000         1.000000         4.000000
 DUMINICA        20.000000         3.000000         1.500000
```

Deci vor fi necesari 22 de vânzători, iar pentru plata acestora vor trebui 7750$.
Conform planificării (graficului) turelor:
Luni vor lucra: Lu_X+Jo_X +Vi_X+Sa_X+Du_X=2+7+5+4+2=20 vânzători
Marți vor lucra: Lu_X+Ma_X+Vi_X+Sa_X+Du_X=2+0+5+4+2=13 vânzători ş.a.

Programarea INTEGER

Când formulăm o problemă de programare liniară (LP) constatăm deseori că unele variabile trebuie să ia valori întregi. Astfel de probleme le întâlnim în literatura de specialitate sub denumirea *Integer Programs* (IP) sau *Integer Linear Programming* (ILP).

Modelele IP sunt frecvent folosite în management deoarece multe decizii presupun în esență un număr finit de variante. Este cazul problemelor de decizie de tipul da / nu sau acționează / nu acționa. Alte probleme sunt astfel formulate încât nu acceptă o parte fracționară la soluții ci doar numere întregi (număr de oameni, număr de maşini număr de avioane, etc).

Există şi tipul de probleme mixte în care unele variabile trebuie să ia numai valori întregi şi altele pot lua şi valori fracționare. Acestea se încadrează în modelele numite programe *mixte-întregi* (MIP).

Problemele de programare liniară cu numere întregi constituie gama de probleme ILP (*Integer Linear Programming*) care vor fi prezentate în continuare.

Exemple

Exemplul 1

Existând patru proiecte care se vor derula pe durata a trei ani consecutivi şi având fiecare următoarele caracteristici:

Venituri returnate pe proiecte		**Cereri de capital pe ani si proiecte**		
Proiect	Venit returnat	Anul 1	2	3
1	0.2	0.5	0.3	0.2
2	0.3	1.0	0.8	0.2
3	0.5	1.5	1.5	0.3
4	0.1	0.1	0.4	0.1
Capital total disponibil:		3.1	2.5	0.4

Se cere să se stabilească ce proiecte vor fi alese pentru a se maximiza venitul final total rezultat în urma derulării acestora.

Soluția

În acest caz se observă foarte clar că soluția acestei probleme nu poate avea parte fracționară. Este exclus să obținem un răspuns de tipul: 1.72, 2.53, 3.09, 4.1. Proiectele pot fi doar alese sau respinse.

Abordăm problema în aceeaşi manieră în care am formulat problemele de tip LP, adică vom determina care sunt:

- variabilele
- restricțiile
- obiectivele

Singura modificare semnificativă în formularea IP deosebită de formularea LP este definirea variabilelor.

Variabile

Aici vom încerca să decidem dacă să garantăm un proiect sau nu. O modalitate de rezolvare este introducerea unor variabile care să ia doar valorile întregi 0 sau 1 şi care să reprezinte decizii binare:

- Decizia pozitivă (se execută se garantează) este reprezentată de 1
- Decizia negativă (nu se execută nu se garantează) este reprezentată de 0.

Aceste variabile sunt adesea numite variabile zero-unu sau variabile binare.

Vom utiliza următoarea notaţie x_j:

x_j = 1 dacă decidem să realizăm proiectul j (j=1,...,4) altfel,
x_j = 0 (dacă decidem să nu realizăm proiectul j (j=1,...,4))

Restricţii

Restricţiile referitoare la capitalul disponibil în fiecare an sunt:

$0.5x_1 + 1.0x_2 + 1.5x_3 + 0.1x_4 <= 3.1$ (anul 1)
$0.3x_1 + 0.8x_2 + 1.5x_3 + 0.4x_4 <= 2.5$ (anul 2)
$0.2x_1 + 0.2x_2 + 0.3x_3 + 0.1x_4 <= 0.4$ (anul 3)

Funcţia obiectiv

Maximizarea venitului total posibil de obţinut din proiecte.
Max $(0.2x_1 + 0.3x_2 + 0.5x_3 + 0.1x_4)$
x_j = 0 sau 1 j=1,...,4

Exemplul 2

Un album muzical care conține 9 melodii trebuie înregistrat pe o casetă matriță. Durata fiecărei melodii este:

Melodie	1	2	3	4	5	6	7	8	9
Durata (min.)	6	8	8	6	7	6	7	5	7

Se cere să se distribuie optim melodiile pe cele două fețe ale casetei astfel încât fiecare față să aibă o durată cât mai apropiată de jumătatea (30 min.) timpului total (60 min.) al celor 9 melodii.

Soluția

Variabile

$X1, X2, \ldots, X9$.
$Xi = 1$ dacă melodia <i> se înregistrează pe fata 1, altfel este 0 (dacă ea se înregistrează pe fața 2).

Funcția obiectiv

$$\text{Max } (X1 + X2 + X3 + X4 + X5 + X6 + X7 + X8 + X9)$$

Restricția

$$6X1 + 8X2 + 8X3 + 6X4 + 7X5 + 6X6 + 7X7 + 5X8 + 7X9 <= 30$$

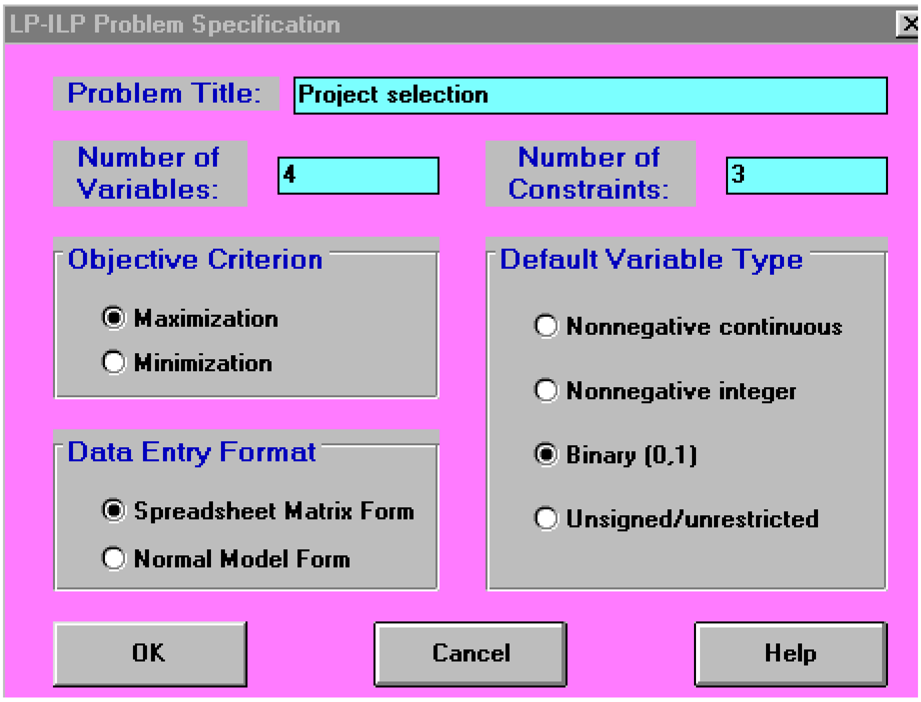
Rezolvarea prin intermediul unor produse software specializate

Rezolvarea exemplului 1 cu QSB

Pentru rezolvarea unor astfel de probleme (ILP), pachetele de programe folosesc cel mai frecvent metoda *branch and bound*.

Datele de intrare pentru problema enunțată mai sus pot fi încărcate prin intermediul următoarelor formulare:

Parametrii problemei:



Coeficienții funcției obiectiv și ai restricțiilor:

Variable -->	X1	X2	X3	X4	Direction	R. H. S.
Maximize	0.2	0.3	0.5	0.1		
C1	0.5	1.0	1.5	0.1	<=	3.1
C2	0.3	0.8	1.5	0.4	<=	2.5
C3	0.2	0.2	0.3	0.1	<=	0.4
LowerBound	0	0	0	0		
UpperBound	1	1	1	1		
VariableType	Binary	Binary	Binary	Binary		

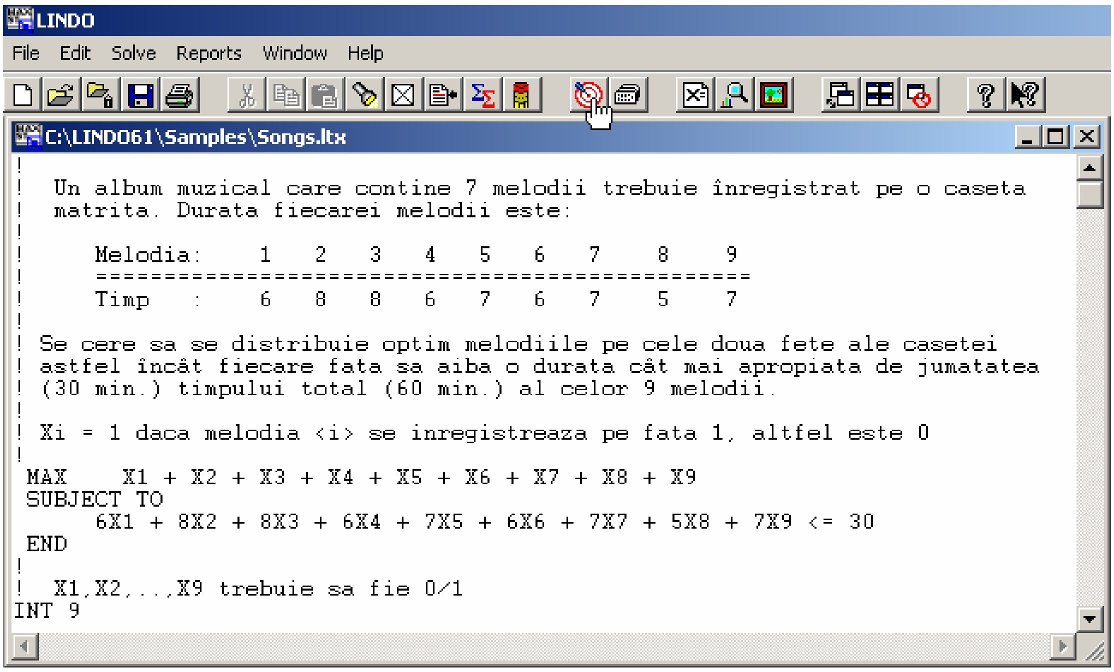
Situația finală care ne oferă soluția optimă este:

05-15-2002 09:29:22	Decision Variable	Solution Value	Unit Cost or Profit C(j)	Total Contribution	Reduced Cost	Basis Status
1	X1	0	0.2000	0	0.2000	at bound
2	X2	0	0.3000	0	0.3000	at bound
3	X3	1.0000	0.5000	0.5000	0	basic
4	X4	1.0000	0.1000	0.1000	0	basic
	Objective	Function	(Max.) =	0.6000		

Se constată că decizia optimă propusă de soluție este aceea de a alege doar proiectele 3 şi 4. Oricare altă alegere ar determina obținerea unui profit (venit) inferior sau nu ar respecta restricțiile impuse.

Rezolvarea exemplului 2 cu LINDO

Introducerea datelor:




```
!
!  Un album muzical care contine 7 melodii trebuie înregistrat pe o caseta
!  matrita. Durata fiecarei melodii este:
!
!     Melodia:    1   2   3   4   5   6   7    8    9
!     ================================================
!     Timp   :    6   8   8   6   7   6   7    5    7
!
! Se cere sa se distribuie optim melodiile pe cele doua fete ale casetei
! astfel încât fiecare fata sa aiba o durata cât mai apropiata de jumatatea
! (30 min.) timpului total (60 min.) al celor 9 melodii.
!
! Xi = 1 daca melodia <i> se inregistreaza pe fata 1, altfel este 0
!
 MAX    X1 + X2 + X3 + X4 + X5 + X6 + X7 + X8 + X9
 SUBJECT TO
      6X1 + 8X2 + 8X3 + 6X4 + 7X5 + 6X6 + 7X7 + 5X8 + 7X9 <= 30
 END
!
!  X1,X2,...,X9 trebuie sa fie 0/1
INT 9
```

```
Reports Window

LP OPTIMUM FOUND AT STEP      6
OBJECTIVE VALUE =   5.00000000


NEW INTEGER SOLUTION OF    5.00000000     AT BRANCH      0 PIVOT       6
RE-INSTALLING BEST SOLUTION...

       OBJECTIVE FUNCTION VALUE

       1)      5.000000

 VARIABLE        VALUE          REDUCED COST
       X1         1.000000         -1.000000
       X2         0.000000         -1.000000
       X3         0.000000         -1.000000
       X4         1.000000         -1.000000
       X5         0.000000         -1.000000
       X6         1.000000         -1.000000
       X7         1.000000         -1.000000
       X8         1.000000         -1.000000
       X9         0.000000         -1.000000


      ROW   SLACK OR SURPLUS     DUAL PRICES
       2)         0.000000          0.000000

NO. ITERATIONS=       6
BRANCHES=     0 DETERM.=  1.000E    0
```

Se observă că programul a propus pentru o față a casetei melodiile:
1, 4, 6, 7, 8 (adică 30 min. = 6 min. + 6 min. + 6 min. + 7 min. + 5 min.).

Programarea scopurilor (*Goal programming*)

Programarea scopurilor este o dezvoltare a programării liniare utilizată pentru probleme care implică mai multe obiective.

Modelele de programare a scopurilor diferă de modelele de programare liniară atât în ceea ce priveşte restricțiile cât şi funcția obiectiv.

Obiectivele multiple poartă numele de scopuri. Fiecare scop materializează un anumit obiectiv de atins.

În programarea scopurilor, scopurile se exprimă sub forma unor restricții. Restricțiile de scopuri sunt diferite de restricțiile prezentate la programarea liniară. Restricțiile de tipul celor prezentate la programarea liniară se numesc restricții tari (*hard*) spre deosebire de cele specifice scopurilor, care reprezintă un nivel dorit de performanță şi care se numesc restricții slabe (*soft*).

Programarea scopurilor permite folosirea exclusivă a unor restricții slabe sau o combinație de restricții tari şi slabe.

Soluția unei probleme în programarea scopurilor trebuie să satisfacă toate restricțiile tari dar nu trebuie obligatoriu să atingă nivelul de performanță fixat prin restricțiile slabe. Acest lucru se întâmplă atunci când există conflicte fie între scopuri fie între scopuri şi restricțiile tari.

De exemplu o restricție cere ca x1=10 iar altă restricție ca x1>=20. Aceste două restricții nu pot fi satisfăcute simultan şi din acest motiv în programarea liniară spunem că nu există o soluție. În astfel de cazuri putem folosi programarea scopurilor deoarece sunt permise abateri de la restricțiile slabe pentru a putea găsi o soluție acceptabilă. Astfel, în programarea scopurilor obiectivul este de a satisface restricțiile tari şi de a obține un nivel cât mai acceptabil de satisfacere a restricțiilor slabe (maleabile).

În prima fază programarea scopurilor a tratat scopurile (restricțiile slabe) ca fiind egale în importanță pentru soluția obținută. Deci, abaterea de la un scop era acceptată în mod egal cu abaterea față de la un alt scop. Mai recent s-a introdus o diferențiere în importanța scopurilor. Acest lucru se materializează în acordarea de coeficienți de importanță (prioritate) ai scopurilor. Această abordare permite obținerea unor soluții cât mai acceptabile problemelor din lumea reală.

Variabilele abatere

Pentru a măsura abaterile de la scopuri se folosesc variabile abateri care sunt incluse în restricțiile de scop. Ele reprezintă diferența dintre scopul propus şi cel posibil de atins în soluție. Există două feluri posibile de abateri de la scop: abateri sub scop şi abateri peste scop. Variabilele de abatere de la scopuri sunt incluse în fiecare din restricțiile de scop (slabe) sub numele: Ui pentru abaterile sub (*under*) scop şi Vi pentru abaterile peste (*over*) scop (i reprezintă numărul restricției de scop). Adăugând aceste două variabile abatere la o restricție de scop obținem restricții de tip egalități deoarece aceste variabile abateri măsoară tocmai diferența dintre scopul propus şi cel posibil de atins.

Exemplu: dacă un manager fixează o restricție referitoare la numărul de ore de muncă, în programarea scopurilor o transpunem sub forma unei restricții de scop de forma:

$4X1 + 2X2 + (U1 - V1) = 100$ ore

X1,X2 - reprezintă variabile de decizie
U1 - reprezintă numărul de ore de muncă neutilizate (1- prima restricție)
V1 - reprezintă numărul de ore de muncă suplimentare

Au fost introduse în restricția de scop ambele tipuri de variabile de abatere ceea ce indică faptul că sunt permise abateri în ambele sensuri.

Dacă nu este admisă abaterea într-unul dintre sensuri, abaterea corespunzătoare nu va apărea în restricția de scop. De exemplu, dacă nu este permisă prestarea de ore suplimentare variabila V1 nu va apărea în restricția de scop.

Remarcați semnele cu care sunt introduse variabilele abatere în restricția de scop: U reprezintă un minus până la scop iar V un surplus peste scop, deci U se adună şi V se scade pentru a se obține egalitatea.

Dacă $4X1 + 2X2 = 80$ => $U1 = 20$ ore nelucrate față de scopul 100
Dacă $4X1 + 2X2 = 110$ => $V1 = 10$ ore peste program față de scopul 100

În ambele aceste cazuri cealaltă variabilă abatere față de cea precizată este =0 deoarece în mod logic ele se exclud.

Dacă $4X1 + 2X2 = 100$ => $U1 = 0$ şi $V1 = 0$.

Astfel, în fiecare restricție scop cel puțin una din variabilele abatere este = 0.

Formularea modelului în programarea scopurilor

Modelul constă dintr-o funcție obiectiv şi un set de restricții. Restricțiile pot fi numai restricții de scop (slabe) sau o combinație de restricții de scop (slabe) şi restricții obişnuite (tari).

La acestea se adaugă restricțiile de nenegativitate pentru toate variabilele (de decizie şi de abatere).

În modelele care includ şi priorități ale scopurilor, funcția obiectiv măsoară care variabile abateri vor fi minimizate şi în ce ordine de prioritate (importanță)

Exemplu de model de programare a scopurilor:

Min $(P1U1 + P2V1 + P3U2)$

A) $4X1 + 2X2 \le 40$ restricții obişnuite
B) $2X1 + 6X2 \le 60$. (tari)

1) $3X1 + 3X2 + U1 - V1 = 75$ restricții de <u>scop</u>
2) $X1 + 2X2 + U2 - V2 = 50$ (slabe)

$X1, X2, U1, V1, U2, V2 \ge 0$

În funcția obiectiv sunt prezente 3 variabile abatere:U1, V1, U2. Indicii acestora arată cărei restricții de scop aparțin. P-urile (P1, P2, P3) reprezintă

prioritățile, indicii lor arată ordinea de importanță (i = 1 reprezintă cea mai mare importanță).

Observăm în funcția obiectiv că cea mai mare prioritate este aceea de a minimiza cantitatea cu care ne aflăm sub primul scop. Următoarea prioritate este aceea de a minimiza cantitatea cu care ne aflăm peste primul scop şi cea mai mică prioritate este aceea de minimiza cantitatea cu care ne aflăm sub cel de-al doilea scop.

Observăm că nu toate variabilele abatere trebuie să fie prezente în funcția obiectiv. Variabila V2 nu este inclusă deoarece nu ne propunem să minimizăm cantitatea cu care ne aflăm deasupra celui de-al doilea scop (putem presupune că acest scop se referă la profit).

Exemplul 1

O companie produce 3 tipuri de produse (X1,X2,X3). Consumul de materiale şi timp pe unitate de produs este următorul:

Produsul	**X1**	**X2**	**X3**	**Disponibil**
Materiale (unități/produs)	2	4	3	600 unități
Timp de producție (min/produs)	9	8	7	900 min.
Timp de ambalare (min/produs)	1	2	3	300 min.

Managerul a stabilit următoarele obiective (scopuri) în ordinea lor de importanță (prioritate):

1. Minimizarea orelor suplimentare de producție (prima prioritate - P1)
2. Minimizarea timpului de producție neutilizat. (a doua prioritate - P2)
3. Minimizarea atât a timpului de ambalare suplimentar cât şi a celui neutilizat. (a treia prioritate - P3)

Soluție

La o primă analiză observăm că există 3 restricții: una asupra materialelor şi două asupra timpului. Observăm că restricția asupra materialelor se exprimă ca o restricție obişnuită iar cele asupra timpului ca nişte restricții de scop deoarece ele sunt cuprinse în lista scopurilor urmărite.

Variabilele:

X1 – cantitatea de produs de tip 1
X2 – cantitatea de produs de tip 2
X3 – cantitatea de produs de tip 3
U1 – (sau X4) timpul de producție neutilizat
V1 – (sau X5) timpul de producție utilizat suplimentar
U2 – (sau X6) timpul de asamblare neutilizat
V2 – (sau X7) timpul de asamblare utilizat suplimentar

Restricțiile:

Materiale	2X1 + 4X2 + 3X3	<= 600 unități
Timp de producție	9X1 + 8X2 + 7X3 + U1 - V1 = 900 min.	
Timp de asamblare	1X1 + 2X2 + 3X3 + U2 - V2 = 300 min.	

X1, X2, X3, U1, V1, U2, V2 >= 0

Funcția obiectiv: Min (P1V1 + P2U1 + P3(U2 + V2))

Cuantificăm cele trei priorități în funcție de importanța lor:
P1 = 9, P2 = 8, P3 = 7.
Relativ la formularea scopurilor problemei, funcția obiectiv poate fi împărțită în trei obiective (scopuri):

- Min (P1V1) Scopul 1
- Min (P2U1) Scopul 2
- Min (P3(U2+V2)) Scopul 3

Exemplul 2

Managerul unei companii care produce două sortimente de stofă vrea să determine structura producției pentru o săptămână redusă de lucru (există sărbători legale). Există în acest scop suficiente stocuri de materie primă dar timpul de lucru disponibil este de 24 de ore. Fiecare metru de stofă de primul tip necesită 2 ore de muncă iar cea din tipul 2 necesită 3 ore de muncă.

Prioritățile exprimate de manager, în ordinea de importanță a acestora, sunt:

1. Minimizarea timpului de muncă neutilizat
2. Dacă sunt necesare ore suplimentare, acestea să fie de aproximativ 12 ore sau mai puțin
3. Să se încerce producerea de cel puțin 10 metri de stofă de tip 2
4. Să se evite orele suplimentare dacă acest lucru este posibil.

Soluție:

Identificăm următoarele variabile de decizie:

X1 - cantitatea de stofă de tip 1 care să se producă
X2 – cantitatea de stofă de tip 2 care să se producă
U1 – (sau X3) ore de muncă neutilizate
V1 – (sau X4) ore de muncă suplimentare
U2 – (sau X5) ore de muncă neutilizate sub cele 12 ore acceptate
V2 – (sau X6) ore de muncă suplimentare peste cele 12 ore acceptate
U3 – (sau X7) cantitatea (m) de stofă de tip 2 produsă sub cerința de 10 m
V3 – (sau X8) cantitatea (m) de stofă de tip 2 produsă peste cerința de 10 m.

Restricții:

Restricția timpului de muncă:

$$2X1 + 3X2 + (U1 - V1) = 24 \text{ ore}$$

Restricția privind orele suplimentare:

$$V1 + (U2 - V2) = 12 \text{ ore}$$

Restricția privind cantitatea de stofă de tipul 2:

$$X2 + (U3 - V3) = 10 \text{ metri}$$

Funcția obiectiv:

Relativ la scopurile exprimate de manager observăm că:

1. Minimizarea timpului de muncă neutilizat – reprezintă scopul 1 care poate fi asociat variabilei U1
2. Dacă sunt necesare ore suplimentare, acestea să fie de aproximativ 12 ore sau mai puțin – reprezintă scopul 2 care poate fi asociat variabilei V2
3. Să se încerce producerea de cel puțin 10 metri de stofă de tip 2 – reprezintă scopul 3 care poate fi asociat variabilei U3
4. Să se evite orele suplimentare dacă acest lucru este posibil – reprezintă scopul 4 care poate fi asociat variabilei V1.

$$\text{Min } (P1U1 + P2V2 + P3U3 + P4V1)$$

Cuantificăm cele patru priorități în funcție de importanța lor:
P1 = 9, P2 = 8, P3 = 7, P4 = 6.
Relativ la formularea scopurilor problemei, funcția obiectiv poate fi împărțită în patru obiective (scopuri):

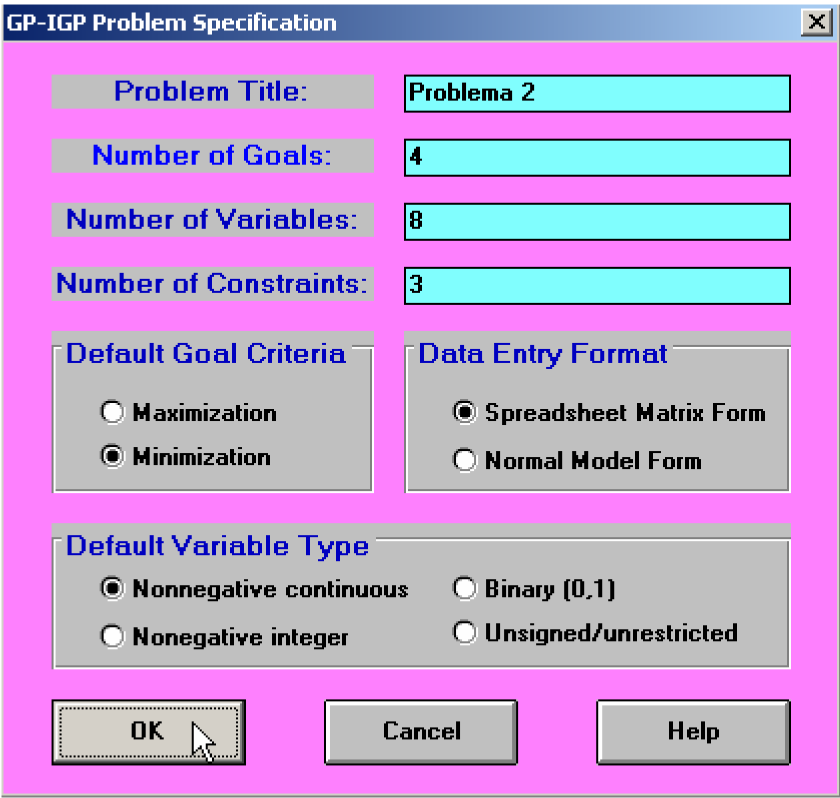
- Min (P1U1) Scopul 1
- Min (P2V2) Scopul 2
- Min (P3U3) Scopul 3
- Min (P4V1) Scopul 4

Unde: U1 = timpul de muncă neutilizat
U2 = timpul de muncă neutilizat sub cele 12 ore acceptate
V2 = timpul suplimentar de muncă peste cele 12 ore acceptate
U3 = cantitatea (metri) de stofă de tipul 2 sub cerința de 10 metri
V1 = numărul de ore suplimentare
X1, X2, U1, V1, U2, V2, U3, V3 >= 0

Rezolvarea problemelor utilizând un produs software specializat

Problema anterioară are 4 scopuri şi 8 variabile: X1, X2, U1 (notat X3), V1 (notat X4), U2 (notat X5), V2 (notat X6), U3 (notat X7) şi V3 (notat X8).

Dialogul inițial:



Formularul pentru încărcarea datelor:

Variable -->	X1	X2	X3	X4	X5	X6	X7	X8	Direction	R. H. S.
Min:G1			9							
Min:G2						8				
Min:G3							7			
Min:G4				6						
C1	2	3	1	-1					=	24
C2				1	1	-1			=	12
C3		1					1	-1	=	10
LowerBound	0	0	0	0	0	0	0	0		
UpperBound	M	M	M	M	M	M	M	M		
VariableType	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous		

Soluția finală:

	Goal Level	Decision Variable	Solution Value	Unit Cost or Profit c(j)	Total Contribution	Reduced Cost	Allowable Min. c(j)	Allowable Max. c(j)
1	G1	X1	0	0	0	0	0	M
2	G1	X2	10.00	0	0	0	0	0
3	G1	X3	0	9.00	0	9.00	0	M
4	G1	X4	6.00	0	0	0	0	0
5	G1	X5	6.00	0	0	0	0	0
6	G1	X6	0	0	0	0	0	M
7	G1	X7	0	0	0	0	0	M
8	G1	X8	0	0	0	0	0	M
9	G2	X1	0	0	0	0	0	M
10	G2	X2	10.00	0	0	0	0	0
11	G2	X3	0	0	0	0	-M	M
12	G2	X4	6.00	0	0	0	0	0
13	G2	X5	6.00	0	0	0	0	0
14	G2	X6	0	8.00	0	8.00	0	M
15	G2	X7	0	0	0	0	0	M
16	G2	X8	0	0	0	0	0	M
17	G3	X1	0	0	0	0	0	M
18	G3	X2	10.00	0	0	0	0	7.00
19	G3	X3	0	0	0	0	-M	M
20	G3	X4	6.00	0	0	0	0	2.33
21	G3	X5	6.00	0	0	0	-2.33	0
22	G3	X6	0	0	0	0	-M	M
23	G3	X7	0	7.00	0	7.00	0	M
24	G3	X8	0	0	0	0	0	M
25	G4	X1	0	0	0	12.00	-12.00	M
26	G4	X2	10.00	0	0	0	0	M
27	G4	X3	0	0	0	6.00	-M	M
28	G4	X4	6.00	6.00	36.00	0	6.00	M
29	G4	X5	6.00	0	0	0	-M	0
30	G4	X6	0	0	0	0	-M	M
31	G4	X7	0	0	0	-18.00	-M	M
32	G4	X8	0	0	0	0	0	M

	G1	Goal	Value	(Min.) =	0	(Alternate	Solution	Exists!!)		
	G2	Goal	Value	(Min.) =	0					
	G3	Goal	Value	(Min.) =	0					
	G4	Goal	Value	(Min.) =	36.00					

	Constraint	Left Hand Side	Direction	Right Hand Side	Slack or Surplus	Allowable Min. RHS	Allowable Max. RHS	ShadowPrice Goal 1	ShadowPrice Goal 2	ShadowPrice Goal 3
1	C1	24.00	=	24.00	0	18.00	30.00	0	0	0
2	C2	12.00	=	12.00	0	6.00	M	0	0	0
3	C3	10.00	=	10.00	0	8.00	12.00	0	0	0

Analize de tip rețea PERT - CPM

Introducere

Analiza de tip rețea este o tehnică specifică utilizată în planificarea şi managementul proiectelor. Specific proiectelor este faptul că:

- ele sunt temporare, au un început şi un sfârşit;
- pot fi împărțite în mai multe activități separate (*job*-uri) în care fiecare activitate are asociată o durată şi un timp de realizare (timpul cuprins între punctul de start şi cel de terminare a activității);
- există o precedență a ordinii în care se execută activitățile.

Pentru analizele de tip rețea au fost dezvoltate independent în anii 1950 două tehnici diferite:

- PERT (*Program Evaluation and Review Technique*);
- CPM (*Critical Path Management*).

PERT a fost dezvoltat de *US Navy* pentru planificarea şi controlul rachetelor *Polaris* dorindu-se a se găsi cel mai scurt timp posibil pentru realizarea proiectului.

CPM a fost dezvoltat de către Du Pont care şi-a propus să găsească un raport optim între costul proiectului şi timpul de realizare a acestuia. Se consideră că pentru unele activități poate fi micşorat timpul de realizare cheltuindu-se mai mulți bani.

În prezent există numeroase produse software care pot rezolva cu succes analizele de tip rețea atât prin tehnica PERT cât şi prin CPM. Analizele de tip rețea sunt vitale în managementul proiectelor. Ele ne dau posibilitatea ca prin intermediul unor metode cantitative corespunzătoare să putem conduce un proiect astfel încât acesta să fie realizat cu succes.

Formularea unei probleme

Pentru o mai bună înțelegere a analizelor de tip rețea, este necesară o discuție pe baza unui exemplu, care va conduce cu siguranță către o clarificare a acestei problematici.

Să presupunem că se intenționează lansarea unui nou produs pe piață. ceea ce necesită existența unui proiect compus din mai multe etape şi sarcini care vor trebui rezolvate într-o anumită ordine. O întrebare logică care apare referitor la proiect este:

În cât timp se va realiza proiectul ?

Pentru a răspunde la întrebare este necesar în primul rând crearea unei liste ce va conține toate activitățile principale ale proiectului precum şi durata de realizare pentru fiecare activitate:

Număr activitate	Denumire activitate	Timp de realizare (săptămâni)
1.	Proiectare produs	6
2.	Proiectare ambalaj	2
3.	Comanda şi recepţionarea componentelor necesare realizării produsului	3
4.	Comanda şi recepţionarea componentelor necesare realizării ambalajului	2
5.	Realizarea unui lot de produse într-o primă fază	4
6.	Realizarea ambalajului pentru acest lot de produse	1
7.	Împachetarea şi studiul metodelor de împachetare optimă	1
8.	Testarea pieţei pentru produs	6
9.	Revizuirea şi reproiectarea corespunzătoare a produsului	3
10.	Revizuirea şi reproiectarea corespunzătoare a ambalajului	1
11.	Prezentarea rezultatelor către decidenţii superiori	1

Legat de această listă se impun analize în detaliu ale activităţilor enunţate în raport cu scala timpului pe care se desfăşoară. Astfel trebuie identificate relaţiile de precedenţă ale activităţilor, indicând activităţile care datorită unor situaţii logice trebuie să fie terminate înainte ca alte activităţi să înceapă. De exemplu activitatea 1 trebuie terminată înainte ca activitatea 3 să înceapă. Activitatea 8 trebuie terminată înainte ca activitatea 9 să înceapă.

În continuare este necesară **realizarea unei liste de precedenţe imediate** a activităţilor, luând în considerare relaţiile care există între acestea. Lista va fi construită încercând să răspundem pentru fiecare activitate la următoarea întrebare:

Ce activităţi trebuie terminate înainte ca activitatea curentă să înceapă ?

Număr activitate	Observaţii	Număr activitate
Activitatea 1	trebuie terminată înainte ca activitatea	3 să înceapă
Activitatea 2	trebuie terminată înainte ca activitatea	4 să înceapă
Activitatea 3	trebuie terminată înainte ca activitatea	5 să înceapă
Activitatea 4	trebuie terminată înainte ca activitatea	6 să înceapă
Activităţile 5 şi 6	trebuie terminată înainte ca activitatea	7 să înceapă
Activitatea 7	trebuie terminată înainte ca activitatea	8 să înceapă
Activitatea 8	trebuie terminată înainte ca activitatea	9 să înceapă
Activitatea 8	trebuie terminată înainte ca activitatea	10 să înceapă
Activităţile 9 şi 10	trebuie terminată înainte ca activitatea	11 să înceapă

Este de notat că:

- activitățile 1 şi 2 nu apar în a treia coloană a tabelului deoarece nu există activități care să trebuiască să fie terminate înainte ca ele să înceapă;
- activitățile 5 şi 6 trebuie să se termine înainte ca activitatea 7 să înceapă;
- se subînțeleg din listă şi precedențele ne-imediate: astfel activitatea 1 va trebui să fie terminată înainte ca activitatea 9 să înceapă etc. Precedențele ne-imediate nu se includ în listă pentru a evita o supraîncărcare a acesteia. Ele pot fi deduse din sistemul de relații şi legături ale activităților.

Odată ce au fost completate cele două liste (lista activităților şi lista precedențelor) vom putea combina şi concentra informațiile acestora într-o diagramă numită rețea. Denumirea analizelor de tip rețea vine de fapt tocmai de la această diagramă.

Dacă revenim la prima întrebare: **În cât timp se va realiza proiectul ? (completând toate activitățile şi respectând precedența acestora)** un răspuns imediat ar putea fi: Dacă am realiza întâi activitatea 1, apoi activitatea 2, activitatea 3,…, activitatea 11 respectând precedența activităților, întregul proiect ar putea fi realizat în 30 săptămâni (suma timpilor fiecărei activități).

În acest moment o altă întrebare pe care ar putea să o pună un bun manager este:

Poate fi realizat proiectul într-un timp mai mic ?

O reformulare pragmatică a acestei întrebări este:

Care ar putea fi timpul minim de realizare a proiectului ?

În continuare vom vedea cum diagramele de tip rețea ne pot ajuta să găsim răspunsul la ultima întrebare.

Construcția diagramelor de rețea

În funcție de tipul problemelor şi de felul în care acestea sunt formulate, diagramele de rețea pot fi construite în două moduri:

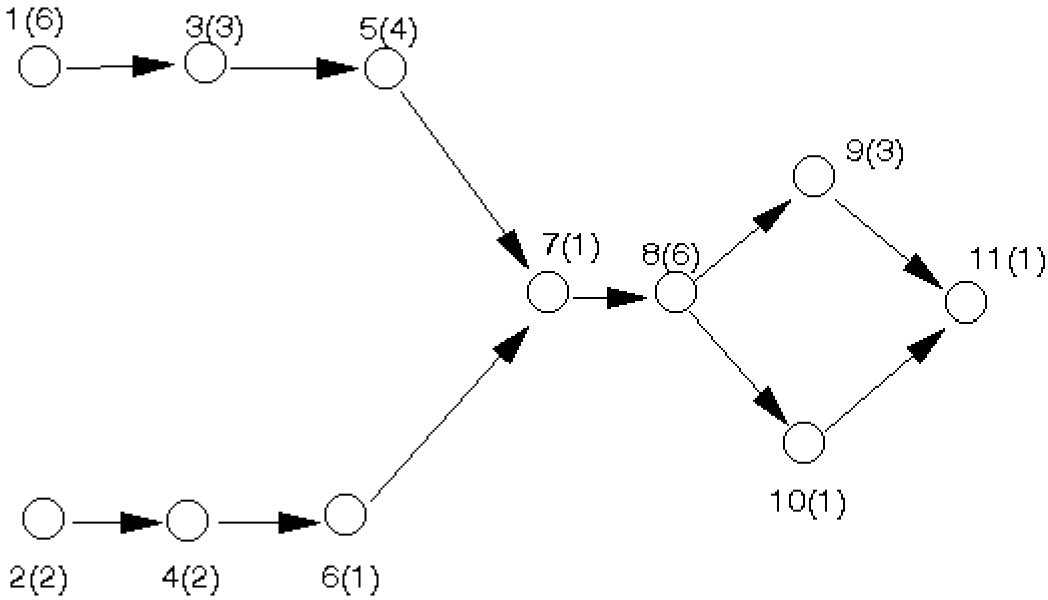
- cu activitățile trasate pe noduri (*activity on node*);
- cu activitățile trasate pe arce (*activity arc*).

Pentru exemplul în discuție s-a ales ca reprezentarea activităților să fie trasată pe noduri.

Dacă dorim rezolvarea problemei cu ajutorul unui pachet de programe specializat, atunci diagrama de rețea nu este necesară. Pentru a găsi timpul minim, respectivului program îi sunt necesare doar datele extrase din lista de precedențe.

Totuşi deoarece oamenii pot interpreta mult mai bine informația reprezentată grafic, atunci, de cele mai multe ori managerilor le sunt necesare pe lângă rezultatele numerice şi diagramele de rețea. Din acest considerent va fi prezentată în continuare şi construcția diagramelor de tip rețea.

În diagrama următoare fiecare nod (cerc) reprezintă o activitate şi conține două informații: numărul activității şi (în paranteze) timpul necesar pentru realizarea acesteia.



Aceasta este o rețea de tipul *activity on node (AON)*.
În construcția diagramei (rețelei):

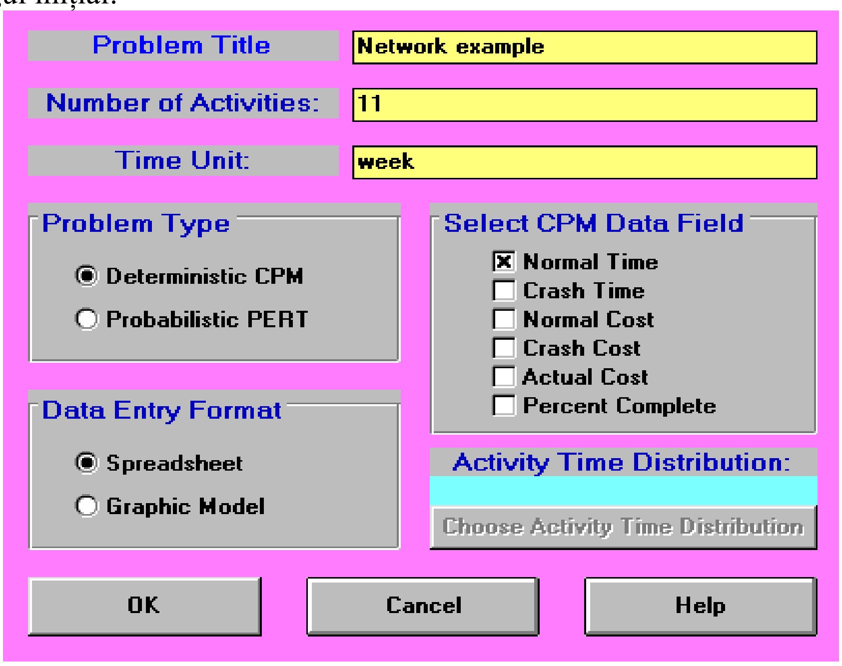
- s-a trasat un nod pentru fiecare activitate
- s-a adăugat o săgeată de la fiecare nod (activitate) (i) către un nod (activitate) (j) dacă activitatea (i) trebuie să se termine înainte ca activitatea (j) să înceapă. Toate arcele au ataşate săgeți indicând direcția în care proiectul se dezvoltă.

Pe diagramă activitățile care încep primele s-au trasat cel mai în stânga iar ultimele activități au fost plasate cel mai în dreapta. Odată trasată o rețea este relativ simplu ca ea să fie analizată. Pentru a fi găsit drumul critic poate fi folosit un algoritm de programare dinamică. În curs nu va fi tratat un asemenea algoritm ci se va insista pe rezolvarea problemelor cu ajutorul calculatorului prin utilizarea unor pachete de programe special create pentru rezolvarea unor astfel de probleme.

Rezolvarea problemelor utilizând un produs software specializat

De regulă orice pachet de programe specializat pentru rezolvarea problemelor de tip rețea prezintă întâi un dialog prin intermediul căruia solicită câteva informații de identificare a problemei iar apoi oferă un formular pentru încărcarea datelor. Iată în continuare capturile celor două ecrane aferente problemei exemplu prezentate mai sus:

Dialogul inițial:



Formularul pentru încărcarea datelor:

Activity Number	Activity Name	Immediate Predecessor (list number/name, separated by ',')	Normal Time
1	1		6
2	2		2
3	3	1	3
4	4	2	2
5	5	3	4
6	6	4	1
7	7	5,6	1
8	8	7	6
9	9	8	3
10	10	8	1
11	11	9,10	1

Situația finală cu rezolvare problemei:

01-05-2002 10:29:16	Activity Name	On Critical Path	Activity Time	Earliest Start	Earliest Finish	Latest Start	Latest Finish	Slack (LS-ES)
1	1	Yes	6	0	6	0	6	0
2	2	no	2	0	2	8	10	8
3	3	Yes	3	6	9	6	9	0
4	4	no	2	2	4	10	12	8
5	5	Yes	4	9	13	9	13	0
6	6	no	1	4	5	12	13	8
7	7	Yes	1	13	14	13	14	0
8	8	Yes	6	14	20	14	20	0
9	9	Yes	3	20	23	20	23	0
10	10	no	1	20	21	22	23	2
11	11	Yes	1	23	24	23	24	0
	Project	Completion	Time	=	24	weeks		
	Number of	Critical	Path(s)	=	1			

Studiind soluția finală observăm că timpul minim de realizare a proiectului este de 24 săptămâni. Se consideră că există suficiente resurse care să permită desfăşurarea mai multor activități simultan. Astfel activitățile 1 şi 2 se pot executa în paralel. În coloana "*Slack*" se poate vedea pentru fiecare activitate întârzierea care poate fi permisă astfel încât timpul integral de realizare a proiectului să nu fie afectat. Activitățile la care "*Slack*" (diferența) este 0 formează drumul critic. Aceste activități trebuie realizate chiar în timpul alocat. Altfel poate fi afectat timpul de realizare a proiectului. De acesta sunt numite activități critice. În cazul exemplului prezentat activitățile critice sunt: 1, 3, 5, 7, 8, 9, 11.

Notațiile prezente în situația finală sunt:

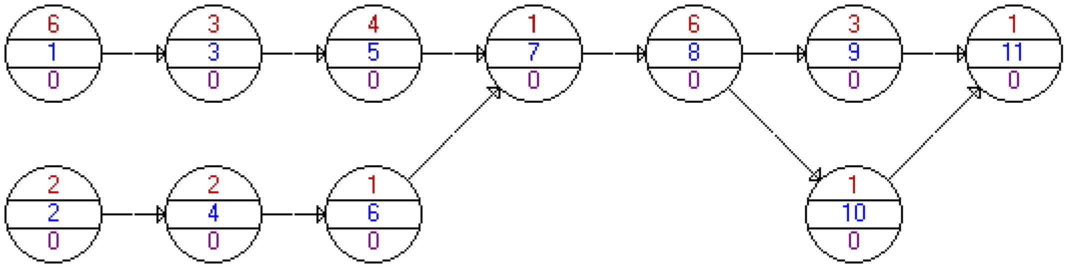
- ES (*earliest start*) - reprezintă momentul de timp cel mai devreme la care o activitate poate să înceapă.
- LS (*latest start*) - este momentul de timp cel mai târziu la care o activitate poate să înceapă astfel încât timpul de realizare a proiectului să nu fie afectat. Dacă LS=ES atunci activitatea în cauză este o activitate critică.
- EF (*earliest finish*) - reprezintă momentul de timp cel mai devreme la care o activitate se poate termina.
- LF (latest finish) - este momentul de timp cel mai târziu la care o activitate se poate termina astfel încât timpul de realizare a proiectului să nu fie afectat. LF=LS+timpul de completare a activității. O activitate este critică dacă LF=LS.
- Slack – reprezintă diferența între ES şi LS sau între LF şi EF. Slack=LS-ES=LF-EF.

Este posibil ca într-o rețea să existe mai multe drumuri critice care să aibă acelaşi rezultat.

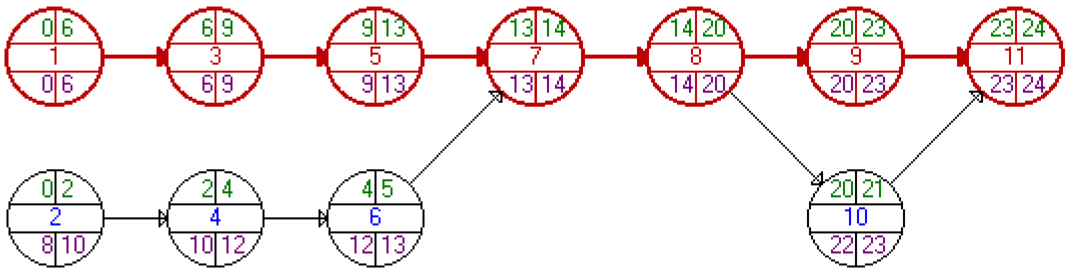
Microsoft Project

Pachetul de programe Microsoft Project oferă în plus şi alte facilităţi foarte folositoare managerului de proiect. Acestea sunt:

Trasarea automată a diagramei de reţea:

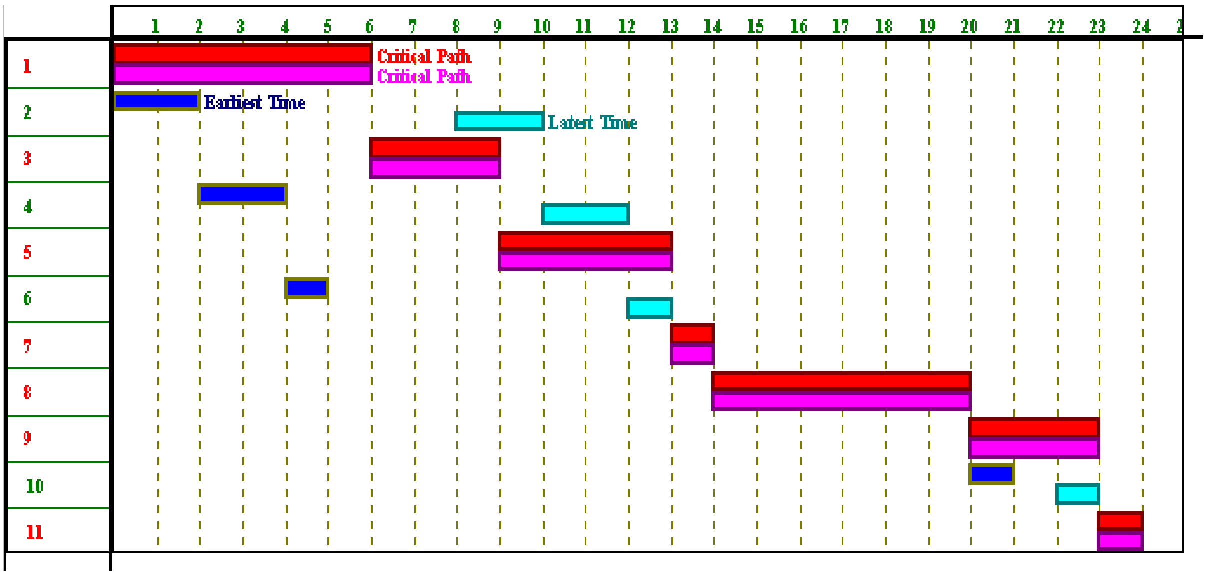


Analiza şi graficul activităţilor critice:



Graficul Gantt

Afişarea drumului critic şi **graficul Gantt**:



Graficul Gantt a fost creat de către H. L. Gantt în anul 1918. Acest grafic prezintă drumul critic şi momentele ES, LS, EF şi LF pentru fiecare activitate. Fiecare activitate este reprezentată prin două bare – una pentru ES şi cealaltă pentru LS. Pentru activitățile critice cele două bare au aceeaşi lungime.

Prin intermediul graficului Gantt se evidențiază şi pot fi studiate foarte uşor şi activitățile necritice. Astfel se observă cu uşurință că activitatea 2 poate să înceapă în oricare dintre momentele de timp 0,1,2,3,4,5,6,7,8 şi poate dura un timp corespunzător cuprins între 0 şi 8. Putem alege în care dintre aceste momente de timp activitatea 2 poate să înceapă.

Verificarea gradului de realizare a proiectului

Gradul de realizare a proiectului poate fi verificat oricând pe timpul desfăşurării acestuia, comparând starea reală cu cea planificată. De asemenea este posibilă şi o anticipare a stadiului în care se va afla proiectul după un anumit timp. De exemplu dacă se doreşte o analiză după 12 săptămâni se va obţine următoarea situaţie:

01-05-2002 12:03:12	Activity Name	On Critical Path	Activity Time	Latest Start	Latest Finish	Planned % Completion
1	1	Yes	6	0	6	100
2	2	no	2	8	10	100
3	3	Yes	3	6	9	100
4	4	no	2	10	12	100
5	5	Yes	4	9	13	75
6	6	no	1	12	13	0
7	7	Yes	1	13	14	0
8	8	Yes	6	14	20	0
9	9	Yes	3	20	23	0
10	10	no	1	22	23	0
11	11	Yes	1	23	24	0
	Overall	Project:		0	24	50

În lista de mai sus se poate observa că activităţile 1,2,3 şi 4 au fost deja completate, activitatea 5 a fost completată într-o proporţie de 75 % iar celelalte activităţi încă nu au început. S-a considerat că toate activităţile au început la momentul ES şi că au durat exact cât au fost planificate.

Probleme propuse

Problema 1

Următorul tabel prezintă activităţile unui proiect:

Activitatea	Timp de realizare (săptămâni)	Activităţi precedente
A	2	-
B	3	-
C	4	A
D	3	B,A
E	8	D,C
F	3	C
G	2	E
H	3	F,G

Trasaţi diagrama de reţea.
Calculaţi durata minimă a proiectului şi identificaţi activităţile critice.

Problema 2

Un proiect conţine 8 activităţi:

Activitatea	Timp de realizare (zile)	Activităţi precedente
A	5	-
B	7	-
C	6	-
D	3	A
E	4	B,C
F	2	C
G	6	A, D
H	5	E, F

Trasaţi diagrama de reţea.
Calculaţi durata minimă a proiectului şi identificaţi activităţile critice.
Dacă activitatea E este întârziată cu 3 zile, va fi afectată durata proiectului?
Dacă activitatea F este întârziată cu 3 zile, va fi afectată durata proiectului?

Problema de transport (*Transportation problem*)

Anumite tipuri de probleme de programare liniară pot fi rezolvate utilizând algoritmi specializați pentru aceste probleme în loc de a folosi metoda simplex (generală). Din această categorie fac parte problemele de transport, problemele de alocare, şi cele de stabilire a unui traseu optim (*problema comisului voiajor*).

Problema de transport este un tip special de problema a fluxurilor în rețea. O problemă tipică de transport include numai un set de noduri sursă şi un set de noduri destinație. Obiectivul este acela de a determina acea modalitate de transport de la surse la destinații care minimizează costurile (maximizează beneficiile) totale asociate transportului. Metoda simplex în rețea este folosită pentru a rezolva problema de transport.

Modelul de transport este aplicat în general acelor cazuri în care se pune problema distribuirii unor obiecte de la deținătorii lor la solicitanții lor, fiecare dintre aceştia (distribuitori şi solicitanți) fiind situați în amplasamente diferite.

De exemplu, o companie poate avea 10 depozite de unde sunt aprovizionate 50 magazine de desfacere. Există o multitudine de combinații în care se poate face aprovizionarea magazinelor însă cu costuri diferite.

Scopul utilizării modelării pentru rezolvarea acestei probleme este de a găsi acel plan de aprovizionare care minimizează costul de transport al produselor de la depozite la magazine şi care ține cont de necesarul de produse al fiecărui magazin şi de disponibilul din fiecare depozit.

Alte cazuri în care se poate folosi modelul problemei de transport sunt transportul produselor de la fabricile producătoare la depozite, transportul produselor între departamentele aceleiaşi companii sau alegerea între mai multe alternative de amplasare a unor fabrici sau depozite.

Formularea modelului

O problemă de transport are în general un număr de locații sursă (numite şi surse) şi un set de locații receptoare (numite şi destinații). Pentru formularea problemei mai avem nevoie de următoarele informații:

1. Cantitățile disponibile (capacitatea) fiecărei surse.
2. Cantitățile cerute la fiecare destinație.
3. Costul unitar de transport de la fiecare sursă la fiecare destinație.

Modelul problemei de transport necesită o serie de precizări:

- toate produsele sunt omogene (orice sursă poate furniza produse pentru orice destinație)
- costurile de transport cresc strict liniar în funcție de cantitatea transportată pe ruta respectivă
- într-o primă abordare se presupune că pentru întreaga cantitate disponibilă există cerere (totuşi pot fi rezolvate şi problemele care nu îndeplinesc această condiție).

Problema 1

O companie are prin contract sarcina să aprovizioneze cu lapte 3 magazine M1, M2, M3. Laptele poate fi aprovizionat din 3 ferme A, B, C.

Disponibilul de lapte în litri la fiecare fermă este:

Ferma	Disponibilul (l)
A	100
B	200
C	200

Cererea de lapte în litri la fiecare magazin este:

Magazin	Cererea (l)
M1	50
M2	150
M3	300

Costul transportului în $ pentru 1 litru de lapte pe fiecare rută este:

La: **De la ferma**	**M1**	**M2**	**M3**
A	0,4	0,2	0,8
B	0,5	0,1	0,9
C	0,7	0,6	0,3

Managerul doreşte găsirea unui plan optim de transport pentru aprovizionarea magazinelor astfel încât costul total al transportului să fie minim.

Rezolvare

Pentru rezolvare datele trebuie aranjate într-un tabel specific rezolvării problemei de transport astfel:

Variabile

	M1	**M2**	**M3**	**Disponibil**
A	4 X11	2 X12	8 X13	100
B	5 X21	1 X22	9 X23	200
C	7 X31	6 X32	3 X33	200
Cerere	50	150	300	500

S-a notat cu Xij cantitatea din cererea j care se aprovizionează de la ferma (sursa) i.

Modelul problemei de programare liniară asociat acestei probleme de transport este prezentat în continuare.

Restricții

$X11 + X12 + X13 = 100$
$X21 + X22 + X23 = 200$ restricții asupra disponibilului
$X31 + X32 + X33 = 200$

$X11 + X21 + X31 = 50$
$X12 + X22 + X32 = 150$ restricții asupra cererii
$X13 + X23 + X33 = 300$

Funcția obiectiv:

$\text{Min } (4X11 + 2X12 + 8X13 + 5X21 + 1X22 + 9X23 + 7X31 + 6X32 + 3X33)$

$X11, X12, X13, X21, X22, X23, X31, X32, X33 >= 0$

Duala problemei de transport

Formularea dualei sugerează că se încearcă transportarea unor cantități de bunuri astfel încât diferența dintre prețul unitar Ui la sursă (plecare) şi prețul unitar la destinație (sosire) Vj să nu depăşească costul unitar de transport între punctul de plecare şi cel de sosire.

Problema 2

O companie de petrol are 3 depozite şi 4 puncte de comercializare C1, C2, C3, C4. Este necesară întocmirea unui plan de transport care să minimizeze costurile totale cu transportul astfel încât să se asigure necesarul de petrol la punctele de comercializare în condițiile cunoaşterii capacității limitate a depozitelor.

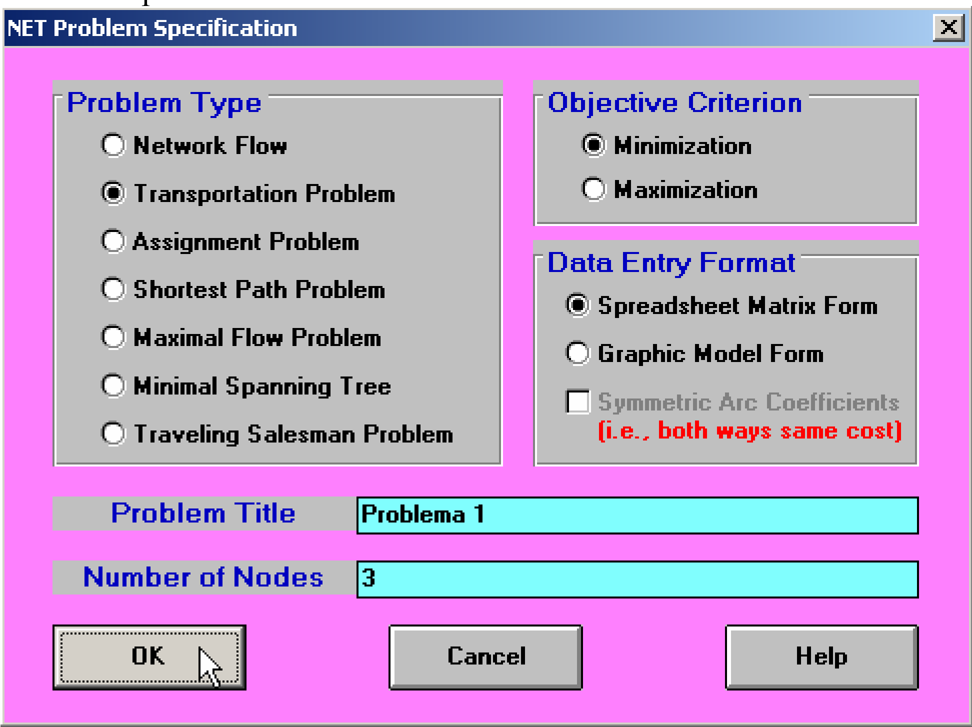
Tabelul următor prezintă valoarea cererii (în t) la punctele de comercializare, disponibilul din depozite (în t) şi costurile unitare ($/t) ale transportului de la surse la destinații:

	Puncte de comercializare:				**Capacitate depozite (t)**
	C1	C2	C3	C4	
Depozit 1	5	4	5	6	100
Depozit 2	3	3	6	6	200
Depozit 3	2	5	7	8	400
Cererea la punctele de comercializare	200	100	150	250	

Rezolvarea problemelor utilizând un produs software specializat

Problema 1

Parametrii problemei:



Introducerea datelor:

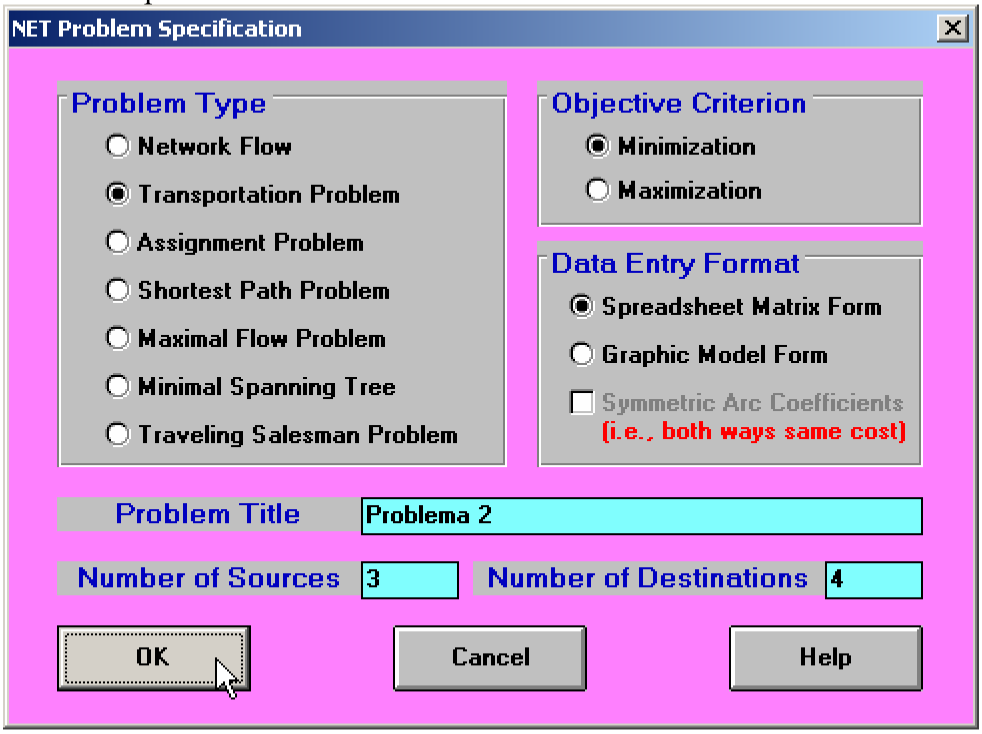
From \ To	Destination 1	Destination 2	Destination 3	Supply
Source 1	4	2	8	100
Source 2	5	1	9	200
Source 3	7	6	3	200
Demand	50	150	300	500

Soluția finală:

10-26-2002	From	To	Shipment	Unit Cost	Total Cost	Reduced Cost
1	Source 1	Destination 1	50	4	200	0
2	Source 1	Destination 3	50	8	400	0
3	Source 2	Destination 2	150	1	150	0
4	Source 2	Destination 3	50	9	450	0
5	Source 3	Destination 3	200	3	600	0
	Total	Objective	Function	Value =	1800	

Problema 2

Parametrii problemei:



Introducerea datelor:

From \ To	Destination 1	Destination 2	Destination 3	Destination 4	Supply
Source 1	5	4	5	6	100
Source 2	3	3	6	6	200
Source 3	2	5	7	8	400
Demand	200	100	150	250	

Soluția finală:

10-26-2002	From	To	Shipment	Unit Cost	Total Cost	Reduced Cost
1	Source 1	Destination 3	100	5	500	0
2	Source 2	Destination 2	100	3	300	0
3	Source 2	Destination 4	100	6	600	0
4	Source 3	Destination 1	200	2	400	0
5	Source 3	Destination 3	50	7	350	0
6	Source 3	Destination 4	150	8	1200	0
	Total	Objective	Function	Value =	3350	

Problema de alocare (*Assignment Problem*)

Problema alocării (repartizării) resurselor este un tip special de problemă de rețea sau de problemă de programare liniară în care obiectele de alocat sunt repartizate unu la unu destinațiilor. Pot face obiectul alocării anumite resurse: angajați, maşini, spații de parcare, intervale de timp, fabrici, jucători, etc.

Destinatarii (beneficiarii) repartițiilor pot fi: o activitate, un amplasament, un eveniment, o cerere, o echipă, etc. În format de rețea atât obiectele alocării cât şi beneficiarii repartițiilor sunt reprezentați ca noduri iar dacă există posibilitatea repartizării obiectului i la beneficiarul j atunci va exista un arc de la nodul i la nodul j care are asociat un cost sau un beneficiu notat Cij.

Problema alocării este considerată ca un tip special de problemă de transport având furnizori şi beneficiari şi se poate rezolva folosind metoda simplex în rețea.

În general este vorba despre un număr de candidați care pot îndeplini n sarcini şi un cost nenegativ C_{ij} implicat de atribuirea către candidatul i a sarcinii j, cost care se consideră cunoscut. Obiectivul problemei este acela de a asigna fiecărui candidat câte o sarcină astfel încât să se minimizeze costul total. Se definesc variabilele binare X_{ij} care pot lua valorile 0 sau 1. Valoarea $X_{ij} = 1$, arată că se atribuie sarcina j candidatului i. Celălalt caz, $X_{ij} = 0$ are loc atunci când nu se atribuie sarcina j candidatului i. Problemele de alocare se caracterizează prin cerința de a împerechea elemente dintr-un grup cu elemente din alt grup (unu la unu).

Un caz particular de problemă de transport este problema alocării resurselor în care toate cererile sunt 1 şi toți furnizorii sunt 1. În acest caz , integralitatea implică să existe câte un furnizor pentru fiecare cerere şi câte o cerere pentru fiecare furnizor. Costul reprezintă cheltuielile implicate de aprovizionarea unui anumit solicitant de la un anumit furnizor.

Presupunând că dorim să impunem condiția ca fie persoana i să nu îndeplinească sarcina j SAU persoana k să nu îndeplinească sarcina m, aceasta se scrie: $X_{ij}.X_{km} = 0$. Această condiție neliniară este echivalentă constrângerii liniare: $X_{ij} + X_{km} <= 1$. Această constrângere trebuie adăugată la setul de constrângeri. Cu această constrângere adăugată AP (*Assignment Problem*) devine ILP binar care poate fi rezolvat cu ajutorul pachetelor de programe LINDO sau QSB.

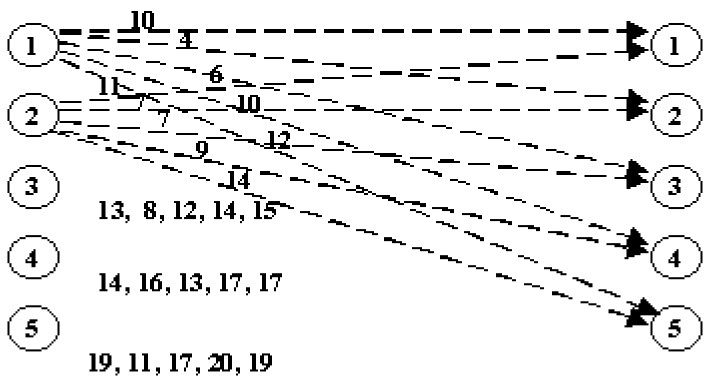
Formularea problemei duale propune ca să se încerce atribuirea de persoane pentru a îndeplini sarcini astfel încât suma valorilor U_i ale persoanei i adunate cu valoarea V_j care reprezintă costul pentru îndeplinirea sarcinii j să nu depăşească costul repartizării persoanei i pentru a îndeplini sarcina j.

De exemplu, un manager trebuie să repartizeze sarcini de execuție la 4 maşini (fiecărei maşini îi revine o sarcină de lucru) sau managerul repartizează câte un proiect fiecărei echipe. În general timpul de lucru sau costul diferă de la o maşină la alta sau de la o echipă la alta.

Scopul managerului este acela de a face acea alocare ce duce la minimizarea timpului de realizare a sarcinilor sau de minimizare a costului total.

În următoarea problemă scopul este de a asigna persoane care să rezolve anumite sarcini astfel încât să se minimizeze costul total aferent rezolvării sarcinilor. Restricțiile cer ca fiecărei sarcini să-i fie repartizată o persoană care să o rezolve şi fiecare persoană să aibă o sarcină de rezolvat.

Assignment Problem



LP Formulation:

Min $10X_{11} + 4X_{12} + 6X_{13} + 10X_{14} + 12X_{15} + 11X_{21} + 7X_{22} + 7X_{23} + 9X_{24} + 14X_{25}$
$+ 13X_{31} + 8X_{32} + 12X_{33} + 14X_{34} + 15X_{35} + 14X_{41} + 16X_{42} + 13X_{43} + 17X_{44} +$
$17X_{45} + 19X_{51} + 11X_{52} + 17X_{53} + 20X_{54} + 19X_{55}$

subject to

$$X_{11} + X_{12} + X_{13} + X_{14} + X_{15} = 1$$
$$X_{21} + X_{22} + X_{23} + X_{24} + X_{25} = 1$$
$$X_{31} + X_{32} + X_{33} + X_{34} + X_{35} = 1$$
$$X_{41} + X_{42} + X_{43} + X_{44} + X_{45} = 1$$
$$X_{51} + X_{52} + X_{53} + X_{54} + X_{55} = 1$$
$$X_{11} + X_{21} + X_{31} + X_{41} + X_{51} = 1$$
$$X_{12} + X_{22} + X_{32} + X_{42} + X_{52} = 1$$
$$X_{13} + X_{23} + X_{33} + X_{43} + X_{53} = 1$$
$$X_{14} + X_{24} + X_{34} + X_{44} + X_{54} = 1$$
$$X_{15} + X_{25} + X_{35} + X_{45} + X_{45} = 1$$
$$X_{ij} \geq 0$$

Există şi probleme în care se urmăreşte maximizarea profitului obținut.

Exemplul 1

Într-o companie, o anumită lucrare care presupune 4 operații poate fi realizată doar de 4 muncitori cu o pregătire specială. Managerul cunoaşte în urma unei analize costurile cu care fiecare din 4 cei muncitori realizează fiecare din cele 4 operații. Acestea sunt prezentate în tabelul de mai jos:

	Muncitor 1	Muncitor 2	Muncitor 3	Muncitor 4
Operația 1	15 $	20 $	18 $	24 $
Operația 2	12 $	17 $	16 $	15 $
Operația 3	14 $	15 $	19 $	17 $
Operația 4	11 $	14 $	12 $	13 $

Obiectivul urmărit de manager este acela de a minimiza costul total rezultat în urma executării celor 4 operații. Acest lucru trebuie realizat în condițiile în care cele 4 operații se execută simultan, deci fiecare din ele trebuie repartizată altui muncitor.

Cerințe pentru problemele de alocare

Situațiile în care se aplică modelul de alocare au următoarele caracteristici:

1. Există două serii de date care trebuie împerecheate unu - la - unu.
2. Obiectivul este acela de a minimiza costurile (timpul sau distanțele) sau de a maximiza profitul.
3. Se cunosc costurile sau profitul pentru fiecare pereche.

Situații speciale

Există situații care prezintă abateri față de cerințele prezentate anterior, cum ar fi:

1. Numărul de elemente din cele două serii care trebuie împerecheate nu este acelaşi (egal).
2. Obiectivul de atins este o maximizare şi nu o minimizare.
3. Anumite variante de împerechere sunt imposibile (nu sunt admise).
4. Există mai multe soluții optime.

Exemplul anterior prezintă cazul în care numărul de elemente din cele două serii de date este egal. Putem acum să considerăm o situație în care trebuiesc executate 4 operații pe 3 maşini. În acest caz una din operații va fi executată numai după ce se va termina una din celelalte operații. Pentru rezolvarea unei astfel de situații vom introduce în tabelul care prezintă costurile operațiilor pe maşini o maşină virtuală. Astfel una dintre operații va fi repartizată unei maşini inexistente, ceea ce arată că această operație nu va fi executată simultan cu celelalte, ci ulterior.

	Operația 1	Operația 2	Operația 3	Operația 4
Maşina 1	15	19	12	16
Maşina 2	23	21	18	17
Maşina 3	20	16	11	19
Maşina virtuală	0	0	0	0

Dacă obiectivul problemei este maximizarea şi nu minimizarea, trebuie făcută o operație suplimentară şi anume, se identifică cea mai mare valoare din fiecare coloană şi apoi se scad din aceasta celelalte valori din coloana respectivă. Cu tabelul modificat care s-a obținut se rezolvă problema ca şi cum ar fi o problemă de minimizare. Valorile modificate se numesc costuri de oportunitate. Setul de valori rezultate în urma minimizării costurilor de oportunitate asigură şi maximizarea valorilor inițiale.

Exemplul 2

În tabelul următor sunt prezentate valori ale profitului obținut în urma executării produselor P1, P2, P3 pe utilajele A, B, C.

	P1	P2	P3
A	14	***22***	30
B	***20***	18	40
C	11	12	***50***

Se cere maximizarea profitului total realizat în urma executării celor 3 produse.

După identificarea valorilor maxime din fiecare coloană (20, 22, 50) după scăderea celorlalte valori din coloana respectivă din valoarea maximă, se obține următorul tabel (al costurilor de oportunitate):

	P1	P2	P3
A	20-14=6	0	50-30=20
B	20-20=0	4	10
C	9	10	0

Exemplul 3

În tabelul următor sunt prezentați timpii necesari fiecăruia din 3 muncitori pentru executarea fiecăreia din 3 operații. Să se minimizeze timpul total de executare a celor 3 operații. Fiecare muncitor execută o singură operație. Ce operație va executa fiecare muncitor ? Să se formuleze modelul de programare liniară al acestei probleme.

	Operație 1	Operație 2	Operație 3
Muncitor 1	3	5	6
Muncitor 2	8	9	7
Muncitor 3	9	2	4

Xij=0 dacă muncitorul i nu va executa operația j
Xij=1 dacă muncitorul i va executa operația j

Tabelul cu variabilele de decizie devine:

	Operație 1	Operație 2	Operație 3
Muncitor 1	X11	X12	X13
Muncitor 2	X21	X22	X23
Muncitor 3	X31	X32	X33

Funcția obiectiv este:

Min (3X11 + 5X12 + 6X13 + 8X21 + 9X22 + 7X23 + 9X31 + 2X32 + 4X33)

Restricțiile sunt:

X11+X12+X13=1
X21+X22+X23=1 Pe linii
X31+X32+X33=1

X11+X21+X31=1
X12+X22+X32=1 Pe coloane
X13+X23+X33=1

Xij=0 sau Xij=1 pentru i,j=1,2,3

Exemplul 4 (extensie a problemei 3)

Dacă există 4 muncitori care pot executa fiecare oricare din cele 3 operaţii, să repartizeze optim cei 3 muncitori pe cele 3 operaţii şi să identifice muncitorul care va fi exclus. În tabelul următor sunt prezentaţi timpii necesari fiecăruia din cei 4 muncitori pentru executarea fiecăreia din cele 3 operaţii.

	Operaţie 1	Operaţie 2	Operaţie 3
Muncitor 1	3	5	6
Muncitor 2	8	9	7
Muncitor 3	9	2	4
Muncitor 4	3	2	4

Exemplul 5

O companie care asigură service pentru calculatoare are 4 clienţi notaţi C1, C2, C3 şi C4. Această companie are 4 tehnicieni (T1,T2,T3,T4) pentru a asigura *service*-ul. Datorită specializării diferite a acestora timpul necesar pentru înlăturarea defectelor reclamate de fiecare dintre clienţi este prezentat în următorul tabel.

	Clienţi			
Tehnicieni	**C1**	**C2**	**C3**	**C4**
T1	3	6	7	10
T2	5	6	3	8
T3	2	8	4	16
T4	8	6	5	9

Managerul acestei companii doreşte să afle ce tehnician să repartizeze pentru fiecare client astfel încât să minimizeze timpul total de remediere a defecţiunilor pentru cei 4 clienţi.

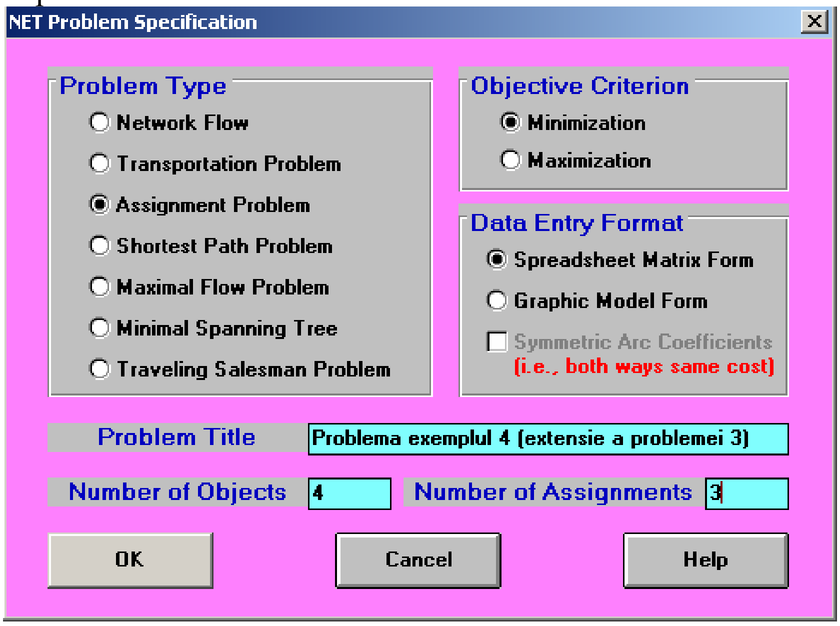
Exemplul 6

În vederea executării unor misiuni de tip comando au fost pregătiţi 5 luptători pentru 5 tipuri de misiuni speciale. Fiecare luptător poate executa oricare din cele 5 tipuri de misiuni. Performanţele lor cuantificate pe o scară de la 1 la 100 pentru fiecare tip de misiune sunt prezentate în următorul tabel:

	Misiuni speciale				
Luptători	**M1**	**M2**	**M3**	**M4**	**M5**
L1	75	70	76	74	70
L2	70	60	80	73	68
L3	68	72	80	74	65
L4	75	70	78	74	69
L5	68	61	81	75	71

Rezolvarea problemelor utilizând un produs software specializat

Introducerea parametrilor:



Introducerea datelor:

From \ To	Assignee 1	Assignee 2	Assignee 3
Assignment 1	3	5	6
Assignment 2	8	9	7
Assignment 3	9	2	4
Assignment 4	3	2	4

Soluția finală:

11-17-2002	From	To	Assignment	Unit Cost	Total Cost	Reduced Cost
1	Assignment 1	Assignee 1	1	3	3	0
2	Assignment 2	Unused_Supply	1	0	0	0
3	Assignment 3	Assignee 3	1	4	4	0
4	Assignment 4	Assignee 2	1	2	2	0
	Total	Objective	Function	Value =	9	

Muncitorul 1 va executa operația 1, muncitorul 3 operația 3 şi muncitorul 4 operația 3. Muncitorul 2 va fi exclus.

Problema comisului voiajor

Problema comis-voiajor-ului presupune existența unui număr de noduri (locații) şi arcuri care leagă toate nodurile. Obiectivul este acela de a parcurge toate nodurile (a vizita toate locațiile) o singură dată minimizând distanța parcursă. Pentru rezolvarea acestui tip de probleme există algoritmi cum sunt:

- *nearest neighbor heuristic,*
- *cheapest insertion heuristic*
- *two-way exchange improvement heuristic)*
- *branch-and-bound*

În prezent toți aceşti algoritmi sunt implementați în diferite pachete de programe printre care se numără şi QSB şi care oferă un suport software eficient pentru rezolvarea rapidă a acestor probleme.

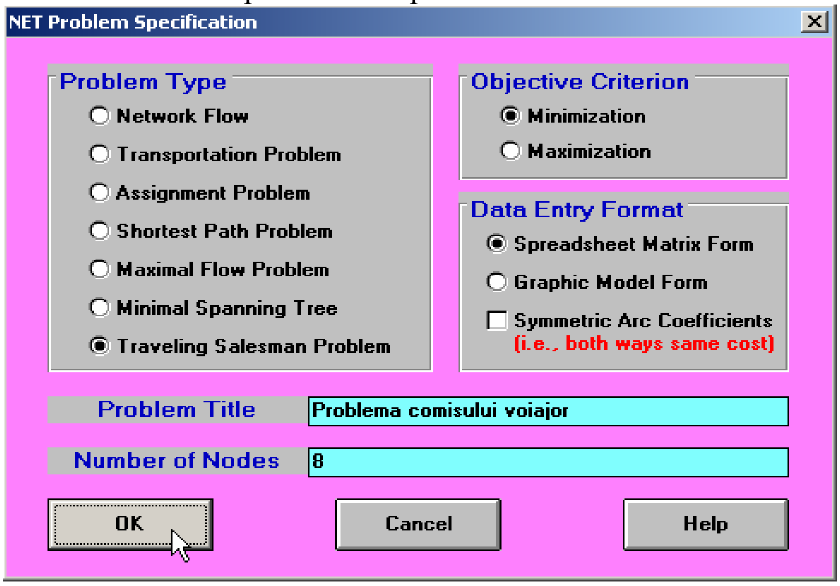
Iată în continuare un exemplu:

Cunoscându-se distanțele în km. dintre 8 oraşe prezentate în tabelul de mai jos, se cere să se determine traseul optim care să treacă o singură dată prin fiecare oraş astfel încât distanța parcursă să fie minimă.

De la oraşul	La oraşul							
	1	2	3	4	5	6	7	8
1	-	150	180	300	200	50	290	350
2	150	-	120	180	250	200	150	250
3	180	120	-	150	150	120	150	200
4	300	180	150	-	300	320	25	60
5	200	250	150	300	-	100	300	350
6	50	200	120	320	100	-	300	350
7	290	150	150	25	300	300	-	90
8	350	250	200	60	350	350	90	-

Rezolvarea problemelor utilizând un produs software specializat

Dialogul pentru introducerea parametrilor problemei anterioare este:



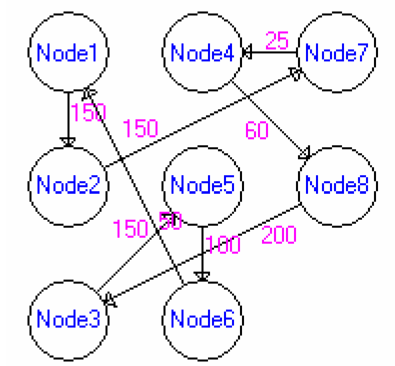
Introducerea datelor:

From \ To	Node1	Node2	Node3	Node4	Node5	Node6	Node7	Node8
Node1		150	180	300	200	50	290	350
Node2	150		120	180	250	200	150	250
Node3	180	120		150	150	120	150	200
Node4	300	180	150		300	320	25	60
Node5	200	250	150	300		100	300	350
Node6	50	200	120	320	100		300	350
Node7	290	150	150	25	300	300		90
Node8	350	250	200	60	350	350	90	

Situația finală sub formă tabelară:

11-17-2002	From Node	Connect To	Distance/Cost		From Node	Connect To	Distance/Cost
1	Node1	Node2	150	5	Node8	Node3	200
2	Node2	Node7	150	6	Node3	Node5	150
3	Node7	Node4	25	7	Node5	Node6	100
4	Node4	Node8	60	8	Node6	Node1	50
	Total	Minimal	Traveling	Distance	or Cost	=	885
	(Result	from	Branch	and	Bound	Method)	

sau în formă grafică:



Lanțuri MARKOV

Un sistem pentru a putea fi analizat cu ajutorul lanțurilor Markov trebuie să prezinte următoarele caracteristici:

- Să opereze pe un număr de intervale de timp.
- În fiecare din perioadele de timp sistemul se poate afla într-un număr finit de stări.
- Stările în care se află un sistem într-o anumită perioadă se exclud reciproc (sistemul se află într-o perioadă în una şi numai în una din stările posibile).
- Trecerea sistemului dintr-o stare în alta în perioade diferite poate fi descrisă prin probabilitățile de trecere, care rămân constante.
- Probabilitatea ca sistemul să se afle într-o anumită stare într-un anume interval depinde numai de starea sistemului în perioada anterioară şi de probabilitățile de trecere dintr-o stare în alta.

Exemple de probleme care pot fi analizate folosind lanțurile Markov sunt:

- opțiunile pe piață pentru anumite tipuri de produse sau servicii din aceeaşi gamă;
- opțiunile pentru anumite posturi TV;
- înapoierea bunurilor închiriate;
- funcționarea maşinilor şi utilajelor (trecerile succesive prin stările operațional-neoperațional), etc.

Utilizarea lanțurilor Markov este utilă pentru analiza acestor tipuri de probleme, şi ajută la înțelegerea evoluției în timp a sistemelor respective. Astfel se pot fundamenta deciziile pe termen scurt referitoare la repartizarea forței de muncă, a capacităților de producție, a stocurilor şi a fondurilor băneşti, dar şi deciziile de perspectivă privind amplasarea unor capacități de producție sau a unor puncte de asigurare a service-ului.

La rândul lor deciziile curente vor conduce la adoptarea unor strategii de modificare a prețurilor, de gestionare corectă a reclamei sau pentru acordarea unor facilități promoționale.

Probabilitățile de trecere

Probabilitățile de trecere arată tendința unui sistem (care poate fi descris ca un proces Markov) de a se modifica de la o perioadă (interval de timp) la alta.

Exemplul următor asigură o mai bună înțelegere a acestui tip de probleme:

O societate de închirieri maşini are două agenții amplasate una în centrul oraşului şi una în apropierea aeroportului. Persoanele care închiriază maşini de la una din agenții le pot înapoia la oricare din cele două (se consideră că este vorba despre un sistem închis în care toate maşinile închiriate sunt înapoiate). Acele maşini care nu îndeplinesc această condiție (nu rămân în circuitul celor înapoiate) sunt scoase din sistem şi tratate separat. Se mai presupune că toate închirierile au durata de o zi, deci la sfârşitul fiecărei zile maşinile se vor afla la una din cele două agenții.

Se presupune că urmărindu-se activitatea agenţiilor o perioadă de timp s-a constatat că 70% din maşinile închiriate de la agenţia aflată în centru au fost înapoiate tot acolo iar 30% la agenţia de la aeroport şi că 90% din maşinile închiriate de la agenţia de lângă aeroport au fost înapoiate tot acolo în timp ce 10% din acestea au fost înapoiate la agenţia din centru.

Aceste informaţii sunt prezentate în tabelul următor:

<table>
<tr><td rowspan="4">Împrumutat de la:</td><td colspan="3">Înapoiat la:</td></tr>
<tr><td></td><td>Agenţia-centru (A)</td><td>Agenţia-aeroport (B)</td></tr>
<tr><td>Agenţia-centru</td><td>0.7</td><td>0.3</td></tr>
<tr><td>Agenţia-aeroport</td><td>0.1</td><td>0.9</td></tr>
</table>

Deoarece s-a făcut presupunerea că toate maşinile închiriate sunt înapoiate la una din cele două agenţii, totalul pe fiecare linie este de 1.0 (100%) dar acest lucru nu se aplică şi pe coloane pentru că acestea se referă la modul de înapoiere a maşinilor. Probabilitatea de trecere 0.7 (70%) reprezintă proporţia (procentul) de maşini închiriate de la agenţia din centru care se aşteaptă să fie înapoiate la aceeaşi agenţie. Probabilitatea de trecere 0.3 (30%) reprezintă proporţia (procentul) de maşini închiriate de la agenţia din centru care se aşteaptă să fie înapoiate la agenţia de la aeroport.

Cunoscând că numărul total de maşini din cele două agenţii este 200 şi că distribuţia iniţială a maşinilor este: 120 maşini la agenţia din centru (A) şi 80 de maşini la agenţia din aeroport (B), managerul societăţii, în vederea luării unor decizii privind amplasarea unor unităţi de service, doreşte să afle:

1. În ce proporţie vor fi înapoiate maşinile la fiecare agenţie după o perioadă de timp (câteva zile). Această informaţie este utilă pentru planificarea muncii la fiecare din cele două agenţii.
2. În ce proporţie vor fi înapoiate maşinile la fiecare agenţie după o perioadă mai lungă de timp.

Comportarea sistemului analizat

Probabilităţile de trecere influenţează atât comportarea pe termen lung a sistemului cât şi comportarea lui pe termen scurt.

Comportarea sistemului pe termen scurt depinde de starea sistemului în perioada (intervalul de timp) actuală şi de probabilităţile de trecere.

Astfel, numărul de maşini înapoiate la fiecare din agenţii este o funcţie care depinde de probabilităţile de trecere şi de numărul maşinilor închiriate de la fiecare agenţie în perioada anterioară.

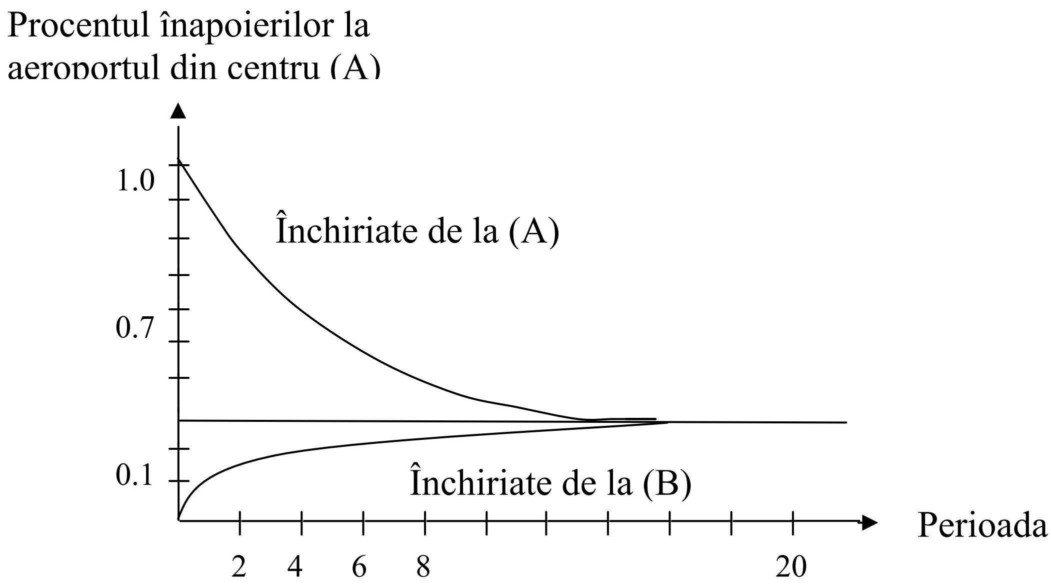
Cunoaşterea numărului de maşini existente la fiecare agenţie la un moment dat constituie punctul de pornire pentru aflarea comportării sistemului pe termen scurt. Numărul de maşini existente la fiecare agenţie în perioada (intervalul de timp) anterioară influenţează numărul acestora pentru următoarele câteva perioade.

Comportarea pe termen lung a sistemului presupune că acesta nu va fi afectat de numărul de maşini existente iniţial la fiecare agenţie, altfel spus, procentul de maşini care vor fi înapoiate la fiecare agenţie după o perioadă mai lungă de timp nu va

fi influențat de condițiile inițiale. Proporțiile (probabilitățile la care se stabilizează sistemul după o perioadă mai lungă de timp se numesc *proporții stabilizate, ferme* (*steady-state proportions or probabilities*).

Nu toate sistemele prezintă tendința de a se stabiliza, unele au tendința de a cicla iar altele de a converge către o stare absolută. Obiectul prezentării următoare îl reprezintă doar sistemele al căror comportament se stabilizează în timp.

Tendințele de evoluție pe termen scurt şi pe termen lung sunt prezentate în figura următoare:



În primele câteva intervale de timp numărul inițial de maşini de la fiecare agenție are un efect sesizabil asupra proporției înapoierilor dar după aproximativ 10 intervale de timp efectul lor este din ce în ce mai mic.

Orice lanț Markov este definit complet prin matricea sa stochastică (o matrice pătratică) şi prin distribuția inițială (care este o matrice linie). În cazul problemei noastre aceste matrici sunt:

Matricea stochastică:	(0,7	0,3)	diagonala principală a acestei matrici
	(0,1	0,9)	reprezintă coeficientul de fidelitate
Distribuția inițială este:	(0,6	0,4)	(120/200=0,6 şi 80/200=0,4)

Observație:

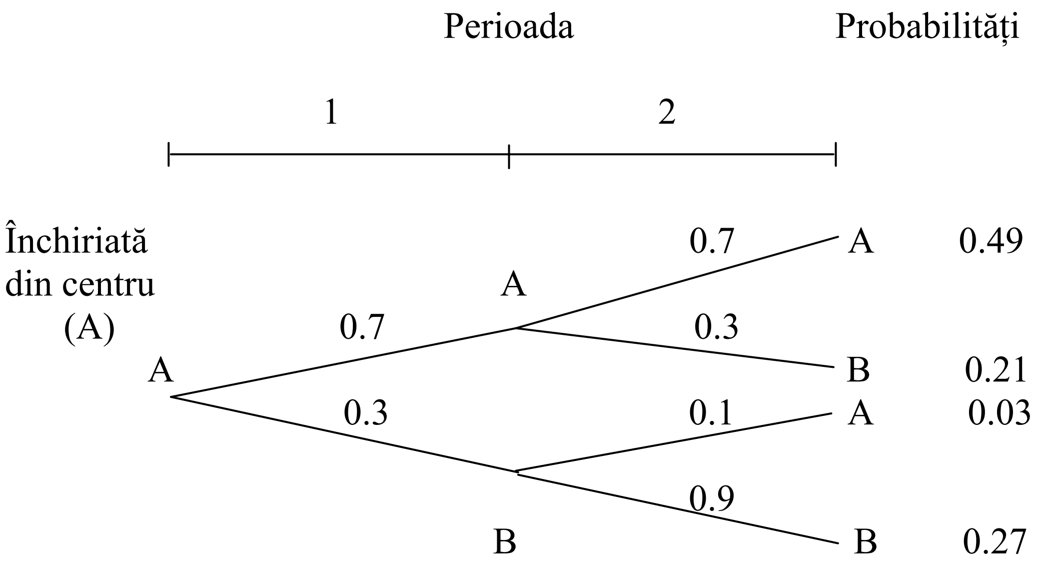
Dacă nu avem distribuția inițială, atunci ea se consideră cu probabilități egale.
Astfel pentru exemplul nostru putem avea: (0,5 0,5)
Indiferent de distribuția inițiala stabilizarea ("*steady state*") va fi aceeaşi.
După rezolvare va rezulta un "*steady state*" la : (0,25 0,75)

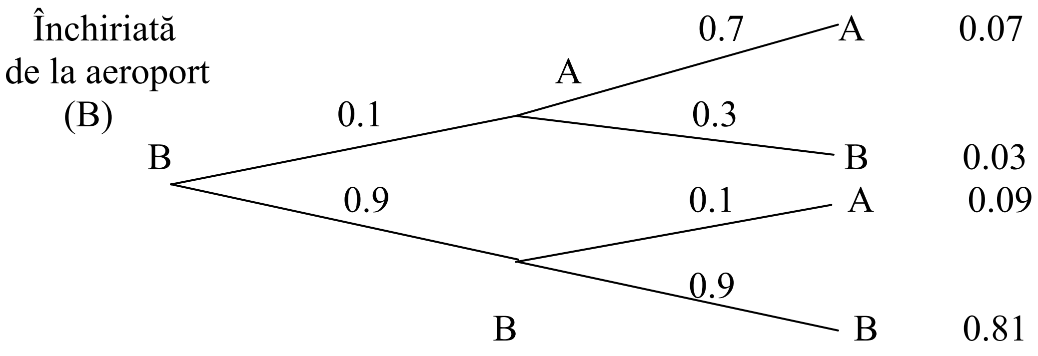
Comportarea acestor sisteme pe termen scurt poate fi descrisă prin folosirea metodei diagramei arborescente sau a matricilor de multiplicare.

Comportarea pe termen lung se poate analiza teoretic similar prin oricare din cele două metode (diagramele arborescente sau matricile de multiplicare) dar din considerente practice, atunci când se lucrează manual, aceste metode nu se aplică deoarece sunt laborioase. Cel mai adesea se utilizează (manual), metoda algebrică.

Metoda arborelui

Această metodă constă dintr-o serie de ramuri care reprezintă posibilele variante de evoluție a sistemului în fiecare perioadă în funcție de probabilitățile de trecere.





Având inițial 100 de maşini la agenția din centru (A) şi 80 la cea de la aeroport (B) rezultă la sfârşitul celei de-a doua perioade:

Maşini la A: 100*(0.49+0.03)+80*(0.07+0.09)=64.8 maşini

Maşini la B: 100*(0.21+0.27)+80*(0.03+0.81)=115.2 maşini

Această metodă are dezavantajul că devine complicată atunci când sunt mai multe elemente în sistem şi când se doreşte studiul sistemului pentru un interval de timp mai îndepărtat de cel curent (3,4….).

Metoda algebrică

Ecuațiile algebrice necesare modelării pe termen lung a comportării sistemului şi aflării probabilităților stabilizate sunt următoarele:

Matricea de trecere, reprezentată prin probabilitățile de trecere este:

	A	B
A	0.7	0.3
B	0.1	0.9

Notăm cu:

A = proporția de maşini înapoiate în centru
B = proporția de maşini înapoiate la aeroport

A=0.7A+0.1B
B=0.3A+0.9B
A+B=1

Eliminăm una din primele 2 ecuații pentru că avem 2 necunoscute şi 3 ecuații.

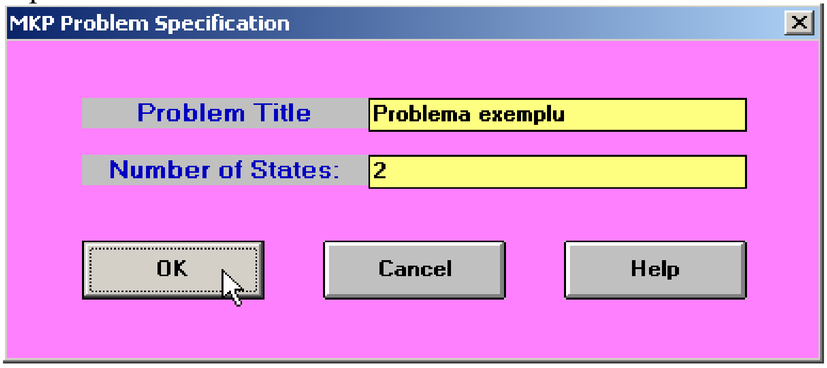
B=1-A => A=0.25   B=0.75

Dacă inițial avem 400 respectiv 300 de maşini în agențiile A respectiv B rezultă că atunci când sistemul atinge stadiul de stabilizare şi cu condițiile acceptate la început (rămân toate în sistem, probabilități de trecere constante, nu apar stări noi în care se pot afla maşinile) vom avea:

A=0.25*700=175 maşini
B=0.75*700=525 maşini

Rezolvarea problemelor utilizând un produs software specializat

Introducerea parametrilor:



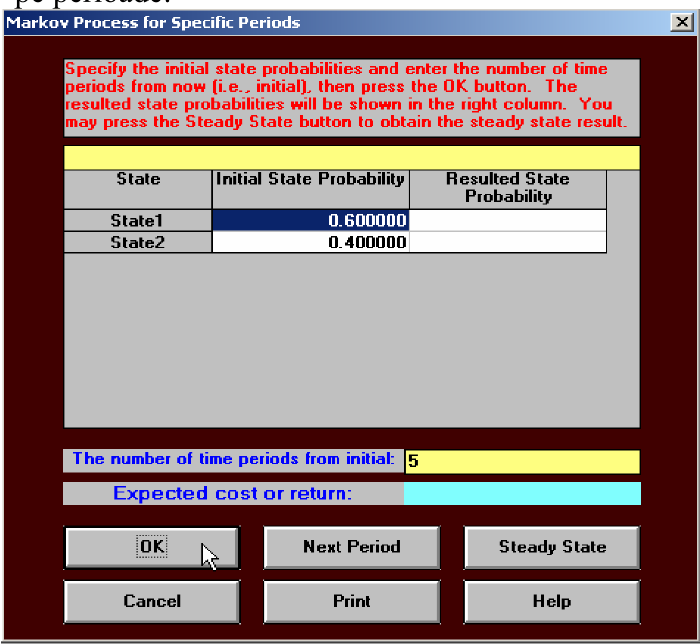
Introducerea datelor:

From \ To	State1	State2
State1	0.7	0.3
State2	0.1	0.9
Initial Prob.	0.6	0.4
State Cost		

Soluția finală:

11-18-2002	State Name	State Probability	Recurrence Time
1	State1	0.2500	4.0000
2	State2	0.7500	1.3333
	Expected	Cost/Return =	0

Analize Markov pe perioade:



Probleme propuse

Problema 1

Un studiu de piață a relevat tendințele de cumpărare ale clienților unei game de produse similare.

În urma studiului s-a stabilit că 50% din clienții care în această săptămână cumpără produsul A vor cumpăra şi săptămâna următoare acelaşi produs, 30% vor cumpăra produsul B şi 20% produsul C.

Din cei care în această săptămână cumpără produsul B 50% vor cumpăra săptămâna următoare produsul A şi 20% produsul C.

Totodată, din cei care în această săptămână cumpără produsul C, 50% vor cumpăra săptămâna următoare produsul A şi 30% vor cumpăra produsul B.

a) Scrieți matricea probabilităților de trecere
b) Care sunt probabilitățile stabilizate pe termen lung pentru cele trei produse?

Problema 2

În urma unui studiu s-a determinat probabilitatea ca o maşină să se defecteze. S-a constatat că aceasta depinde de faptul că maşina a fost reparată în ziua anterioară sau că în ziua anterioară a fost funcțională.

Probabilitățile respective sunt prezentate în tabelul următor:

		Mâine	
		În funcțiune	Defectă
Azi	În funcțiune	0.88	0.12
	Defectă	0.96	0.04

a) Având în vedere că maşina este defectă astăzi, care este probabilitatea ca ea să fie defectă şi mâine?
b) Care este probabilitatea ca maşina să fie defectă peste două zile? Dar peste trei zile?

Cozile de aşteptare

O clasă importantă de probleme în teoria deciziei o constituie cozile de aşteptare. În viața cotidiană apar probleme frecvente de aşteptare la stațiile de benzină, la magazine, restaurante, etc. Aceste probleme se pun de asemenea în transporturile aeriene când aeronavele aşteaptă aprobarea de aterizare de la turnul de control, în cazul camioanelor care aşteaptă să fie încărcate / descărcate, al pasagerilor care aşteaptă taxiuri la aeroport. Exemple de cozi de aşteptare se pot întâlni la ghişeele din bănci sau oficii poştale, în fabrici pentru anumite procese care trebuie parcurse, maşini care trebuiesc reparate, documente care trebuiesc emise, angajați care trebuie să-şi introducă fişa (cartela) de pontaj sau să fie serviți la cantină.

Acest gen de sisteme se caracterizează prin sosiri în sistem foarte variabile şi prin rate de deservire. Astfel, deşi capacitatea totală de deservire a sistemului depăşeşte necesitățile de prelucrare, cozile de aşteptare se formează din când în când datorită variației intrărilor în sistem. Pe de altă parte, această variație a intrărilor determină, în anumite perioade, surplusul de stații de deservire sau de angajați, determinate de absența temporară a intrărilor în sistem (a cererilor). Deci, cozile de aşteptare tind să se formeze datorită supraîncărcărilor create fie de rata de deservire, fie de rata intrărilor în sistem.

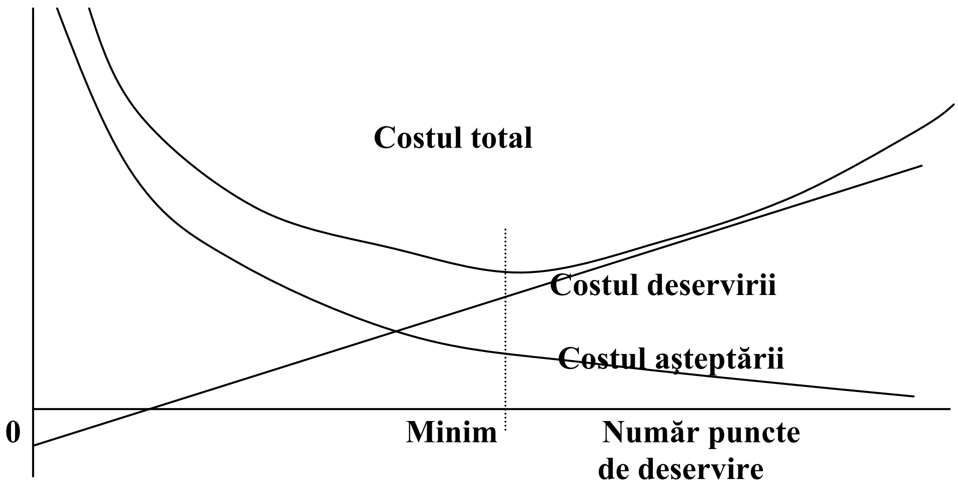
Analizele şi modelarea sistemelor în care apar cozi de aşteptare sunt necesare managerilor care trebuie să stabilească o capacitate optimă a acestora.

Scopuri în proiectarea modelelor de cozi de aşteptare

Aceste modele sunt modele predictive deoarece ele constituie un proiect al comportării sistemelor în care pot apărea cozi de aşteptare. Sistemul analizat poate fi unul în funcțiune dar care nu are performanțe satisfăcătoare, în acest caz se pune problema de a decide ce trebuie modificat pentru a creşte performanțele. În alte situații poate fi cazul unor sisteme aflate în faza de proiectare pentru care trebuie stabilite caracteristicile astfel încât să atingă performanțele dorite.

Un scop principal al proiectării sistemelor în care pot apărea cozi de aşteptare este acela de a găsi un echilibru între costul determinat de creşterea punctelor de deservire şi costul pe care-l implică aşteptarea clienților la cozi.

Aceste două costuri sunt invers proporționale: o scădere a costului determinat de faptul că unii clienți trebuie să aştepte să fie deserviți se realizează prin creşterea costului deservirii (prin creşterea numărului de puncte de deservire sau a ratei de deservire). Funcția care combină cele două costuri (de deservire, de aşteptare) numită funcția costului total are formă de U. Costul total este minimizat în punctul cel mai de jos al acestei curbe. Găsirea punctului de minim implică adesea o creştere în trepte a capacității de deservire. Deoarece creşterea capacității de deservire în cele mai multe cazuri este discretă (una, două stații de deservire) este mai indicat să reprezentăm costul deservirii prin nişte paliere decât printr-o linie continuă. Astfel curba costului total va fi mai puțin lină decât cea prezentată în figura următoare.

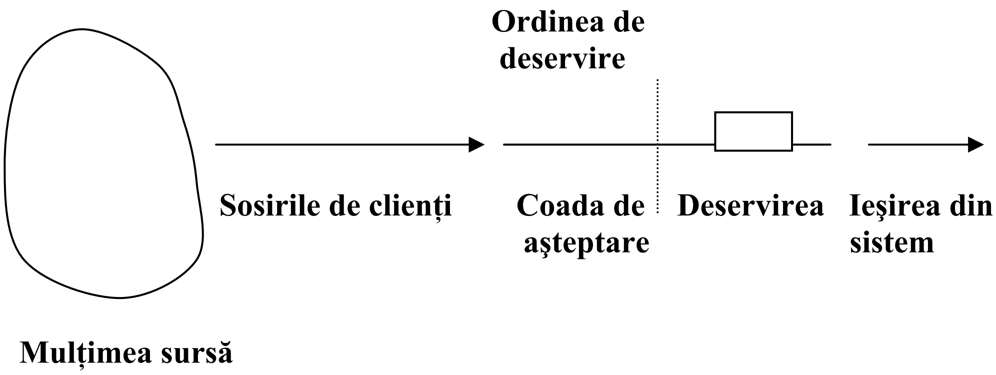


În proiectarea sistemelor în care apar cozi de aşteptare pot fi specificate şi anumite cerinţe particulare. De exemplu managerul unei bănci poate cere ca numărul mediu de clienţi care aşteaptă la coadă să nu fie mai mare de 7 sau managerul unei policlinici doreşte să ştie câte locuri să prevadă în sala de aşteptare pentru pacienţi.

Elemente şi caracteristici ale sistemelor în care pot apărea cozi de aşteptare

Aceste sisteme se diferenţiază prin diferite caracteristici, cum ar fi numărul de puncte de deservire sau faptul că accesul în sistem este liber sau restricţionat. Cunoaşterea acestor caracteristici ajută la modelarea sistemului prin folosirea celei mai potrivite metode.

Elementele componente ale acestor sisteme sunt: mulţimea din care provin clienţii (mulţimea sursă), sosirile de clienţi, coada de aşteptare, ordinea de deservire (procesare), deservirea şi ieşirea din sistem.



Mulțimea sursă

Se referă la populația care furnizează potențialele sosiri în sistem, deci sursa de clienți. Dacă această sursă este suficient de mare astfel încât probabilitatea sosirilor nu este în mod semnificativ influențată de faptul că clienții trebuie să aştepte la coadă, spunem că *mulțimea sursă este infinită*.

Exemple de astfel de sisteme sunt: stațiile de benzină, supermarket-urile, băncile, oficiile poştale, agențiile de voiaj, etc. Deşi mulțimile sursă pentru aceste sisteme nu sunt de fapt infinite, faptul că unii clienți aşteaptă nu împiedică pe alți clienți să intre în sistem.

Pe de altă parte există *sisteme care au un acces limitat* la serviciile oferite de ele. De exemplu într-o fabrică un operator răspunde de încărcarea / descărcarea a 5 prese sau un depanator este responsabil de menținerea în stare de funcționare a unui număr fix de utilaje. Dacă numărul de unități sau clienți care necesită să fie deserviți influențează probabilitatea ca să se producă o nouă intrare în sistem să descrească (deoarece procentul populației care furnizează clienți se reduce) se spune că avem o *mulțime sursă finită sau limitată*.

Sosirile de clienți

Clienții sunt unități care necesită să fie deservite. În unele sisteme clienții sunt oameni, în altele pot fi automobile care trebuie reparate, vase maritime care trebuie descărcate, avioane care trebuie să aterizeze, apeluri telefonice care trebuiesc procesate, etc .

În legătură cu acestea se analizează o serie de probleme, cum ar fi:
Dacă clienții sosesc în sistem câte unul sau în grupuri. De exemplu maşinile sosesc la service câte una pe când consumatorii sosesc la restaurant în grupuri. Toate modelele prezentate în continuare consideră că sosirile în sistem sunt individuale.

Coada de aşteptare

Coada de aşteptare este alcătuită din clienți care au fost admişi în sistem şi aşteaptă să fie deserviți. De exemplu, maşinile care aşteaptă să fie spălate la o spălătorie de maşini. Se pun o serie de probleme în legătură cu clienții care refuză să intre în sistem datorită faptului că trebuie să stea la coadă, clienții care renunță după ce au stat un timp la coadă, sau clienții care se mută la o altă coadă (la supermarket se mută la o altă casă de marcat unde coada este mai mică). Modelele prezentate presupun că odată ce un client s-a aşezat la o coadă rămâne acolo până când va fi servit. În plus se presupune că există o singură coadă de aşteptare de unde clienții, atunci când le vine rândul sunt îndrumați către unul din punctele de deservire libere.

Ordinea de deservire

Regula obişnuită de deservire este *primul venit-primul servit FIFO*. În unele cazuri pentru a se păstra o ordine corectă de servire se dau bonuri de ordine. Există însă şi cazuri în care este necesară o clasificare a clienților intrați în sistem *după anumite criterii de prioritate*. Acest lucru poate fi util la camera de gardă a unui spital unde pacienții trebuie tratați în ordinea sosirii dar ținând seama şi de gravitatea

afecţiunilor, în reţelele de calculatoare unde timpii şi ordinea de procesare ţine cont de anumite priorităţi sau la companii unde clienţii cei mai importanţi sunt serviţi preferenţial.

În continuare vor fi prezentate ambele sisteme: primul venit - primul servit şi sistemul de deservire preferenţial.

Deservirea

Aceasta este caracterizată în principal prin numărul de puncte de deservire, prin numărul de faze necesare de parcurs în procesul de deservire şi prin timpul necesar acestora.

Un sistem poate avea un singur punct de deservire sau mai multe. Se consideră că atunci când există mai multe puncte de deservire acestea funcţionează independent, deci pot fi serviţi simultan atâţia clienţi câte puncte de deservire avem. Dacă aceste puncte nu lucrează independent vor fi tratate ca şi cum ar fi vorba despre un singur punct de deservire. Ambele modele: cu un punct de deservire şi cu mai multe puncte de deservire vor fi tratate în continuare.

Procesul de deservire poate consta în una sau mai multe faze sau operaţii. De exemplu o operaţiune bancară poate consta în constituirea unui depozit şi încasarea unui cec (proces de deservire multifazic). Dacă ar fi vorba numai de constituirea depozitului am spune că este vorba despre un proces de deservire cu o singură fază. Un alt exemplu de proces de deservire multifazic este înscrierea unui candidat pentru susţinerea examenului de admitere.

De regulă distribuţia timpului de deservire poate fi considerată *Poisson*. Această distribuţie implică faptul că în majoritatea cazurilor este necesar un timp redus de procesare, o anumită proporţie necesită timp mediu de procesare şi o proporţie foarte redusă necesită timp mare de procesare.

Ieşirea din sistem

Se pune aici problema ce fac clienţii după ce ies din sistem. Variantele posibile sunt: intrarea în mulţimea sursă a sistemului imediat, după un anumit timp sau niciodată. De exemplu o maşină care a fost spălată nu va reveni imediat pentru o altă spălare, un pacient vaccinat pentru o anumită boală nu va mai reveni la spital pentru acea boală, etc. Modelele prezentate în continuare presupun o intrare imediată în mulţimea sursă pentru sistem (în cazurile cu mulţimi sursă limitate) fie sunt prezente şi celelalte variante dar mulţimea sursă este suficient de mare încât nu au o influenţă măsurabilă asupra ratei intrărilor în sistem.

PSIHOLOGIA AŞTEPTĂRII

Scopul acestui subcapitol este de a stabili ce decizii trebuie luate privind caracteristicile unui sistem în care pot apărea cozi de aşteptare astfel încât să se asigure funcţionarea lor performantă.

Totuşi există sisteme de acest fel, în funcţiune, în care nu se pot face modificări astfel încât să se reducă timpii de aşteptare sub anumite limite.

Două costuri sunt asociate aşteptării. Unul este un cost economic asociat spaţiului necesar pentru aşteptare şi utilajelor care aşteaptă să fie reparate, altul determinat de plata angajaţilor şi utilajelor care aşteaptă clienţi pentru a-i deservi. Aceste costuri pot fi estimate cu un grad înalt de corectitudine.

Alt cost este asociat cu pierderile potenţiale determinate de renunţarea clienţilor la serviciile oferite din cauza cozilor pe care le percep ca prea mari. Acesta are un impact negativ asupra profitului. Acest cost mai este numit şi cost psihologic şi este mult mai greu de estimat.

De exemplu, dacă clienţii consideră că timpul lor este irosit prin aşteptare şi mai ales dacă au impresia că ar putea exista o organizare mai bună care să reducă timpul de aşteptare, atunci ei vor renunţa la serviciul respectiv dacă există alternative. În consecinţă, dacă clienţii ar avea timpul de aşteptare ocupat într-un mod constructiv sau printr-o formă de distracţie ar suporta mult mai uşor acest timp inevitabil deşteptare.

Exemple de cazuri în care se încearcă prin metode psihologice să se atenueze impactul neplăcut al aşteptării sunt: ziarele şi revistele din sălile de aşteptare de la dentist, coafor sau chiar din mijloacele de transport, muzica şi filmele care se difuzează în timpul zborurilor aeriene, muzica pe care o pun automat la telefon unele companii pe perioada aşteptării legăturii cerute, amplasarea de oglinzi în locurile de aşteptare a lifturilor în hoteluri şi magazine, etc.

În unele cazuri este chiar de dorit un timp moderat de aşteptare. Astfel la supermarket-uri se observă amplasarea unor produse nu foarte costisitoare dar care în general fac plăcere cum ar fi gumă de mestecat, bomboane, ţigări, reviste. În timpul de aşteptare se dă ocazia clienţilor să mai adauge mici cumpărături de plăcere la cele deja achiziţionate.

Notații folosite în sistemele cozilor de aşteptare:

A/B/C/D/E

A: specifică natura procesului sosirilor în sistem. Notațiile standard pentru tipurile de sosiri în sistem sunt:

- M: Timpul dintre sosiri este o variabilă aleatoare independentă, cu o distribuție exponențială. Este acelaşi lucru cu a spune că rata sosirilor are o distribuție Poisson.
- D: Timpul dintre sosiri este constant (valoare fixă).
- E(k):Timpul dintre sosiri este o variabilă aleatoare cu distribuție Erlang funcție de parametrul k.
- G: Timpul dintre sosiri este o variabilă aleatoare cu o distribuție generală având media şi abaterea (variația) cunoscută.

B: specifică natura timpului de deservire. Notațiile standard pentru tipurile de deservire sunt:

- M: Timpul dintre sosiri este o variabilă aleatoare independentă, cu o distribuție exponențială. Este acelaşi lucru cu a spune că rata sosirilor are o distribuție Poisson.
- D: Timpul de deservire este constant (valoare fixă).
- E(k):Timpul de deservire este o variabilă aleatoare cu distribuție Erlang funcție de parametrul k.
- G: Timpul de deservire este o variabilă aleatoare cu o distribuție generală având media şi abaterea (variația) cunoscută.

C: specificarea numărului s de stații de deservire. Se presupune că acestea sunt identice şi paralele.

D: specificarea numărului maxim de clienți admişi în sistem. Este egal cu (Q+s), unde Q este capacitatea cozii de aşteptare şi s este numărul de stații de deservire.

E: specificarea dimensiunii mulțimii (populației) sursă a sistemului.

Atunci când D sau E sunt omise, înseamnă că D sau E sunt infinite.

Modele posibile utilizate în rezolvarea problemelor cozilor de aşteptare:
M/M/1, M/G/1, D/M/1, M/M/s, M/M/s/N, M/M/s/N/K, M/M/s/s, M/G/∞ etc

s: numărul de stații de deservire (sau canale)

μ: Rata de deservire (miu) pentru un server.
Timpul mediu de deservire pentru un client este $1/\mu$. În general rata de deservire sau timpul de deservire are o anumită comportare sau distribuție a probabilităților. QA

necesită cunoaşterea distribuției timpului de deservire pentru a clasifica şi rezolva problema respectivă.

λ: Rata sosirilor (lambda).
Timpul mediu între sosirile clienților este $1/\lambda$. În general rata sosirilor sau timpul dintre două sosiri succesive are o anumită comportare sau distribuție a probabilităților. QA necesită cunoaşterea distribuției timpului dintre sosiri pentru a clasifica şi rezolva problema respectivă. De remarcat că atunci când mulțimea (populația) sursă este infinită, λ reprezintă rata sosirilor mulțimii (populației), iar dacă mulțimea sursă este limitată (<10000), λ reprezintă rata sosirilor individuale.

Q: Capacitatea cozii (spațiul de aşteptare maxim), care reprezintă numărul maxim de clienți care pot aştepta să fie deserviți.

N: Populația de clienți (mulțimea sursă) care numărul de clienți potențiali. QA consideră că mulțimea (populația) este infinită atunci când există mai mult de 10000 de clienți potențiali.

Cs: Costul pe unitatea de timp generat de ocuparea stației de deservire

Ci: Costul pe unitatea de timp generat de staționarea (neocuparea) stației de deservire

Cw: Costul pe unitatea de timp generat de faptul că clienții trebuie să aştepte

Cu: Costul pe unitatea de timp generat de deservirea clienților

Ch: Costul generat de renunțarea clienților

Cq: Costul unitar al cozii

n: Numărul de clienți în sistem, în curs de deservire şi aşteptând la coadă

K: Numărul de clienți admişi în sistem. $K=Q+s$.

α: Coeficientul de presiune a deservirii. În general α este un coeficient nenegativ cu rolul de a mări viteza de deservire atunci când sistemul este ocupat (încărcat).

β: Coeficientul de descurajare a sosirilor. Este în general un coeficient nenegativ care are rolul de a descuraja clienții să intre în sistem atunci când acesta este ocupat (încărcat = toate punctele de deservire sunt ocupate).

B: Numărul mediu de clienți care renunță pe unitatea de timp.
$B=\lambda$ - Rata efectivă a sosirilor.

b: Mărimea grupului. Aceasta poate fi de un anume tip sau cu o anume distribuție a probabilitățiii. Implicit este 1.

γ: Intensitatea traficului = λ/μ şi este numărul mediu de clienți deserviți

Lq: Numărul mediu de clienți la coadă

L: Numărul mediu de clienți în sistem. $L=Lq + \gamma$

Lb: Numărul mediu de clienți la coadă într-un sistem ocupat (încărcat). $Lb = Lq/Pw$

Wq: Timpul mediu de aşteptare a clienților la coadă. $Wq = Lq/\lambda$

W: Timpul mediu pe care un client îl consumă în sistem.

Wb: Timpul mediu de aşteptare la coadă de către un client într-un sistem ocupat (încărcat). $Wb = Wq/Pw$

ρ: Factorul de utilizare a sistemului $= \lambda/(s\mu)$

P0: probabilitatea ca toate stațiile de deservire să fie nefolosite

P(n) Probabilitatea ca să fie în sistem n clienți

Pw sau Pb, probabilitatea ca un client sosit să aştepte, sistemul să fie ocupat (toate stațiile de deservire să fie ocupate). $Pw = \sum P(n)$ pentru $n \geq s$.

Rezolvarea problemelor utilizând un produs software specializat

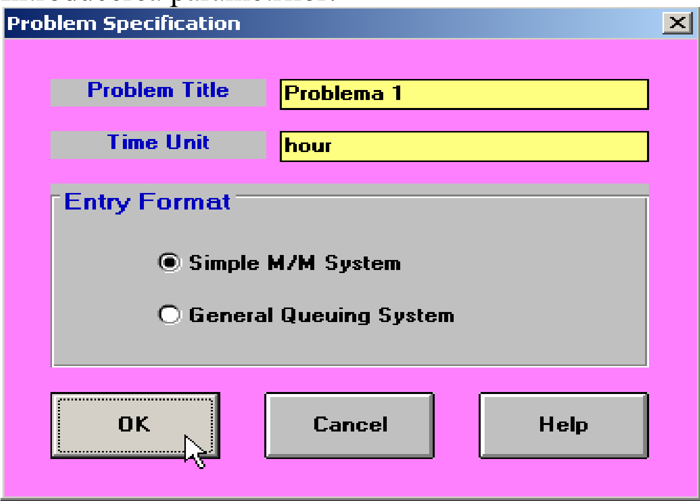
Problema 1

Proprietarul unei spălătorii de maşini intenţionează să deschidă încă o spălătorie în altă parte a oraşului. Bazându-se pe experienţa sa şi pe observaţiile făcute el estimează că rata sosirilor va fi de 20 de maşini pe oră iar rata de deservire va fi de 25 de maşini pe oră. Rata de deservire este variabilă deoarece maşinile sunt spălate manual. Maşinile sunt spălate una câte una (deci este vorba de un sistem cu o singură staţie de deservire şi un singur canal).

Determinaţi:

1. Numărul mediu de maşini în sistem (la coadă şi în curs de spălare).
2. Timpul mediu de stat la coadă.
3. Timpul mediu pe care o maşină îl consumă în sistem (la coadă şi pentru deservire).
4. Utilizarea sistemului.

Introducerea parametrilor:



Introducerea datelor:

Data Description	ENTRY
Number of servers	1
Service rate (per server per hour)	25
Customer arrival rate (per hour)	20
Queue capacity (maximum waiting space)	M
Customer population	M
Busy server cost per hour	
Idle server cost per hour	
Customer waiting cost per hour	
Customer being served cost per hour	
Cost of customer being balked	
Unit queue capacity cost	

Situaţia finală:

11-18-2002	Performance Measure	Result
1	System: M/M/1	From Formula
2	Customer arrival rate (lambda) per hour =	20.0000
3	Service rate per server (mu) per hour =	25.0000
4	Overall system effective arrival rate per hour =	20.0000
5	Overall system effective service rate per hour =	20.0000
6	Overall system utilization =	80.0000 %
7	Average number of customers in the system (L) =	4.0000
8	Average number of customers in the queue (Lq) =	3.2000
9	Average number of customers in the queue for a busy system (Lb) =	4.0000
10	Average time customer spends in the system (W) =	0.2000 hours
11	Average time customer spends in the queue (Wq) =	0.1600 hours
12	Average time customer spends in the queue for a busy system (Wb) =	0.2000 hours
13	The probability that all servers are idle (Po) =	20.0000 %
14	The probability an arriving customer waits (Pw or Pb) =	80.0000 %
15	Average number of customers being balked per hour =	0
16	Total cost of busy server per hour =	$0
17	Total cost of idle server per hour =	$0
18	Total cost of customer waiting per hour =	$0
19	Total cost of customer being served per hour =	$0
20	Total cost of customer being balked per hour =	$0
21	Total queue space cost per hour =	$0
22	Total system cost per hour =	$0

Probleme propuse

Problema 2

Rata sosirii clienților la o agenție de bilete cu un singur ghişeu este de 3 clienți pe minut şi rata deservirii de 4 clienți pe minut. Calculați:

1. Utilizarea sistemului
2. Numărul mediu de clienți la coadă
3. Numărul mediu de clienți în sistem
4. Timpul mediu de aşteptare la coadă
5. Timpul mediu petrecut în sistem
6. Probabilitatea ca în sistem să fie 0 clienți
7. Timpul mediu de aşteptare pentru o sosire care nu este deservită imediat.

Problema 3

Proprietarul unui magazin alimentar urmează să deschidă o nouă aripă. El estimează pentru aceasta o rată a sosirilor de 1.2 clienți pe minut. Pentru această zonă vor fi angajați 2 vânzători fiecare având o rată de deservire de 1 client pe minut. Calculați indicatorii de la exemplul anterior.

Probleme de aranjare (*Facility layout*)

Vor fi prezentate în continuare două tipuri caracteristice de probleme din categoria problemelor de aranjare:

1. Amplasarea compartimentelor (*departamental layout*)
2. Echilibrarea liniilor de asamblare (*assembly line balancing*)

Pentru o mai bună înțelegere aceste tipuri de probleme vor fi prezentate prin intermediul unor exemple.

Amplasarea compartimentelor (*departamental layout, functional layout*)

Vom considera o instituție ale cărei compartimente sunt dispuse ca în imaginea de mai jos. Există în total 11 compartimente numerotate: 1, 2, 3,…,9, A, B. Vom presupune că instituția se află pe un singur nivel (are doar parter) şi că podeaua instituției este împărțită în linii orizontale şi verticale adică în rânduri şi coloane ale căror intersecții formează pătrate, (celule de dimensiuni egale), similar unei table de şah. În cazul exemplului nostru există 16 rânduri şi 14 coloane.

```
                       Coloane
                1 2 3 4 5 6 7 8 9 0 1 2 3 4
                ===========================
Rânduri   1     1 1 1 1 1 1 2 3 3 3 3 3 3 3
          2     1 1 1 1 1 1 2 3           3
          3     8 8 8 8 8 8 8 3           3
          4     8           8 3 3 3 3 3 3 3
          5     8           8 5 5 4 4 4 4 4
          6     8           8 5 5 4       4
          7     8           8 6 6 4       4
          8     8           8 6 6 4       4
          9     8           8 7 7 4       4
         10     8           8 7 7 4 4 4 4 4
         11     8           8 A A A A B B B
         12     8 8 8 8 8 8 8 A     A B   B
         13     9 9 9 9 9 9 9 A     A B   B
         14     9           9 A     A B   B
         15     9           9 A     A B   B
         16     9 9 9 9 9 9 9 A A A A B B B
```

Dimensiunile celor 11 compartimente pot fi deduse din figura de mai sus. De exemplu, amplasamentul compartimentului 9 ține de la rândul 13 coloana 1, la rândul 16 coloana 7 inclusiv. Amplasamentul compartimentului 5 acoperă o suprafață cuprinsă între rândul 5 coloana 8 şi rândul 6 coloana 9 care într-o altă exprimare poate fi scris ca (5,8)-(6,9) .

Managerul instituției doreşte să schimbe modul de amplasare al compartimentelor (prezentate mai sus) astfel încât activitatea instituției să fie cât mai eficientă.

Schimbarea modului de aranjare a compartimentelor trebuie să țină seamă de următoarele reguli:

- suprafața totală a instituției nu se poate modifica;
- un compartiment nu poate fi împărțit în mai multe părți disparate. Amplasamentul său trebuie să cuprindă un lot compact de celule.

Ceea ce poate fi modificat este:

- modul cum sunt amplasate compartimentele unul față de celălalt;
- forma diferitelor compartimente.

Pentru a amplasa cât mai eficient compartimentele este nevoie să cunoaştem fluxurile sau gradul de interacțiune dintre compartimente (fluxuri de materiale, de persoane, de documente). În mod logic, compartimentele între care există o interacțiune (un flux) mai mare ar trebui amplasate mai aproape unele de altele decât cele între care interacțiunea (fluxul) este mai redusă.

De exemplu dacă instituția este o fabrică vom amplasa secțiile fabricii astfel încât să minimizăm suma tuturor produselor dintre distanța de la secția i la secția j (d_{ij}) şi cantitatea care reprezintă fluxul de la secția i la secția j (c_{ij}):

$$\textbf{Min} \sum d_{ij} \times c_{ij} \quad i,j=1,n$$

Aceste produse dintre cantitate şi distanță se mai numesc şi distanțe parcurse cu încărcătura.

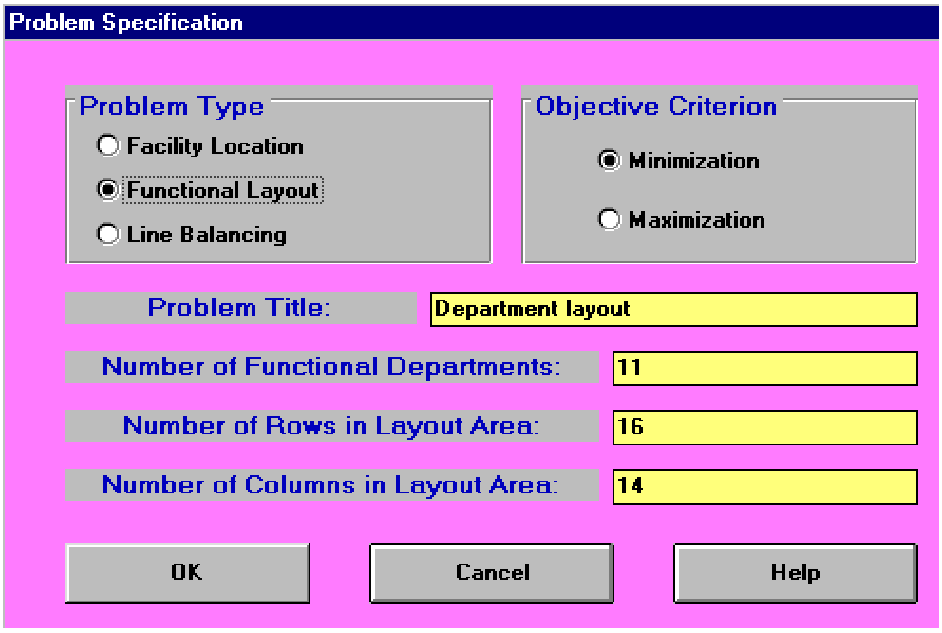
Cuantificările fluxurilor (interacțiunilor) dintre compartimente pentru exemplul nostru reprezintă de asemenea (ca şi amplasarea inițială) date de intrare pentru model şi sunt prezentate în tabelul de mai jos:

		La:										
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**A**	**B**
De la:	**1**	-	-	-	-	30	-	20	-	-	-	-
	2	-	-	40	-	-	-	-	10	-	-	-
	3	100	-	-	-	-	-	-	40	50	30	-
	4	-	-	-	-	-	5	-	-	-	-	150
	5	-	40	-	-	-	80	-	30	90	-	-
	6	10	30	30	20	10	-	90	-	-	10	-
	7	-	-	40	100	-	-	-	-	-	-	20
	8	10	-	-	30	-	40	-	-	50	60	-
	9	10	-	-	50	-	40	-	30	-	40	-
	A	15	20	-	80	70	30	-	10	-	55	-
	B	-	-	-	10	-	20	-	-	-	-	-

Rezolvarea problemelor utilizând un produs software specializat

Problema enunțată poate fi rezolvată pe calculator, cu ajutorul produselor software specializate. Iată felul în care se pot introduce datele de intrare dacă va fi utilizat pachetul de programe QSB:

Parametrii problemei:



Introducerea datelor:

)epartmen	)epartmen	Location Fixed	To Dep.	To Dep.	To Dep.	To Dep.	To Dep.	To Dep.	To Dep.	To Dep.	To Dep.	To Dep.	To Dep.	Initial Layout in
1	1	No					30		20					(1,1)-(2,6)
2	2	No			40					10				(1,7)-(2,7)
3	3	No	100							40	50	30		(1,8)-(4,14)
4	4	No						5					150	(5,10)-(10,14)
5	5	No		40				80		30	90			(5,8)-(6,9)
6	6	No	10	30	30	20	10		90			10		(7,8)-(8,9)
7	7	No			40	100							20	(9,8)-(10,9)
8	8	No	10			30		40			50	60		(3,1)-(12,7)
9	9	No	10			50		40		30		40		(13,1)-(16,7)
10	A	No	15	20		80	70	30		10		55		(11,8)-(16,11)
11	B	No				10		20						(11,12)-(16,14)

Se poate observa că poziția inițială a fiecărui compartiment este dată de rândul și coloana (poziția) colțurilor opuse de pe diagonala principală. De exemplu compartimentul 1 ține de la (1,1) – (rândul 1, coloana 1) la (2,6) – (rândul 2, coloana 6). Este de reținut faptul că pachetul de programe QSB cere folosirea parantezelor rotunde ”( )” pentru introducerea pozițiilor amplasamentului inițial, și nu a parantezelor pătrate “[]”.

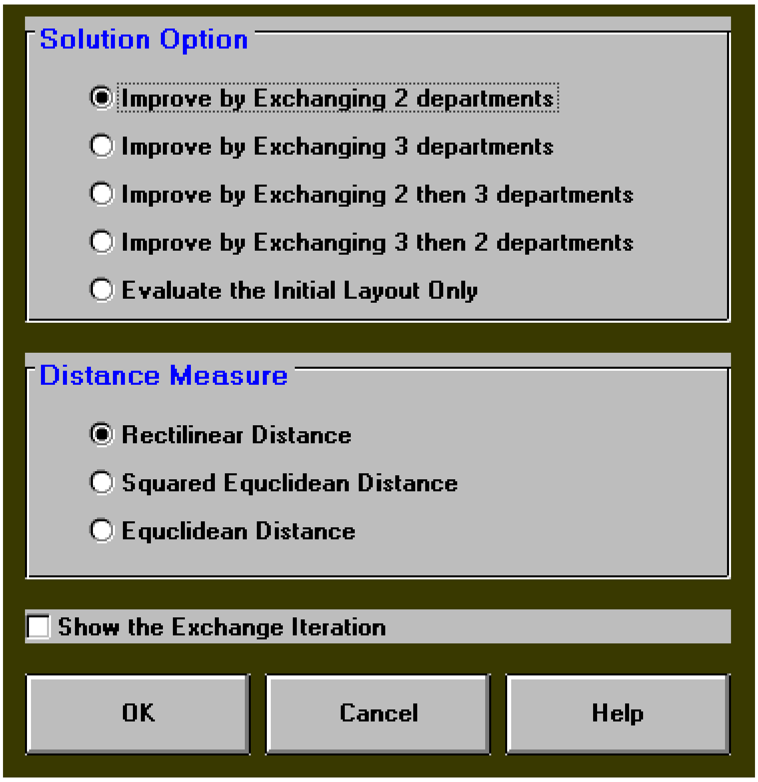
Remarcăm că în rezolvarea acestei probleme este necesar să stabilim modalitatea de calcul a distanței dintre două locații. Dacă (x_i,y_i) și (x_j,y_j) reprezintă coordonatele a două locații: i și j atunci distanța dintre ele poate fi calculată folosind modelul:

- **rectiliniar**, ceea ce presupune calculul distanței dintre i și j ca:
  - $|x_i-x_j|+|y_i-y_j|$
- **Euclidian**, ceea ce presupune calculul distanței dintre i și j ca:
  - $[(x_i-x_j)^2 + (y_i-y_j)^2]^{0.5}$
- **pătrat Euclidian**, ceea ce presupune calculul distanței dintre i și j ca:
  - $(x_i-x_j)^2 + (y_i-y_j)^2$

Modelul rectiliniar este folosit mai ales atunci când studiem modul de aranjare a secţiilor fabricilor, a oraşelor americane, etc. în general atunci când aranjarea se face după o grilă de dreptunghiuri. Din acest motiv modelul de măsurare a distanţei se mai numeşte *sistemul de măsurare Manhattan*.

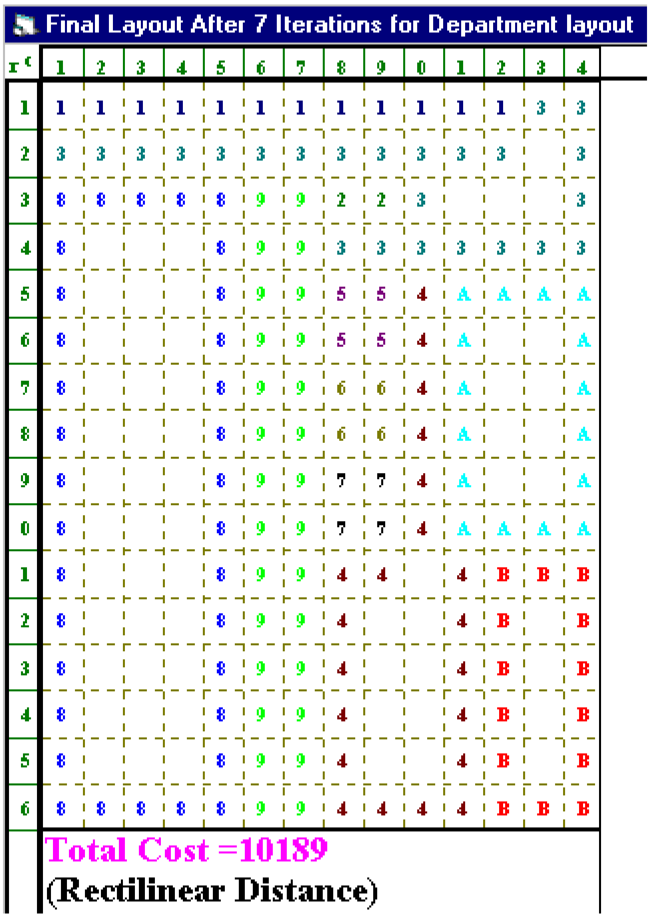
Modelul Euclidian se foloseşte acolo unde se poate trasa o linie dreaptă, la fel ca şi modelul pătratic Euclidian, numai că acesta din urmă realizează şi o descurajare a distanţelor foarte mari (ridicarea la pătrat a unor numere mari, care reprezintă distanţe, are ca rezultat numere şi mai mari care fac parte din funcţia obiectiv pe care dorim să o minimizăm).

Rezultatul obţinut prin rularea pachetului de programe QSB, folosind modelul rectiliniar de măsurare a distanţei dintre secţii, este prezentat mai jos. Menţionăm că s-a ales ca opţiune de rezolvare a problemei schimbarea simultană a amplasării a două compartimente (secţii).



După fixarea opţiunilor şi apăsarea butonului **<OK>** calculatorul caută o aranjare mai bună schimbând poziţia a două departamente. Se încearcă toate variantele posibile şi se alege acea variantă care îmbunătăţeşte cel mai mult funcţia obiectiv prin schimbarea poziţiei a două compartimente.

Soluţia finală:



Noul amplasament al compartimentelor propus de program este prezentat mai jos:

	1	2	3	4	5	6	7	8	9	0	1	2	3	4
1	1	1	1	1	1	1	1	1	1	1	1	1	3	3
2	3	3	3	3	3	3	3	3	3	3	3	3		3
3	8	8	8	8	8	9	9	2	2	3				3
4	8				8	9	9	3	3	3	3	3	3	3
5	8				8	9	9	5	5	4	A	A	A	A
6	8				8	9	9	5	5	4	A			A
7	8				8	9	9	6	6	4	A			A
8	8				8	9	9	6	6	4	A			A
9	8				8	9	9	7	7	4	A			A
10	8				8	9	9	7	7	4	A	A	A	A
11	8				8	9	9	4	4	4	4	B	B	B
12	8				8	9	9	4			4	B		B
13	8				8	9	9	4			4	B		B
14	8				8	9	9	4			4	B		B
15	8				8	9	9	4			4	B		B
16	8	8	8	8	8	9	9	4	4	4	4	B	B	B

Cu un cost de 10189 acest nou aranjament al compartimentelor îmbunătăţeşte considerabil costul, care iniţial avea o valoare de 14147.5. Acest cost de 10189 se obţine în condiţiile în care se utilizează modelul rectiliniar de calcul a distanţei dintre compartimente şi fiecare compartiment se consideră a fi amplasat în centrul ei. Astfel, compartimentul 1 are centrul pe linia 1) deoarece el este amplasat în întregime în linia 1) şi pe coloana 6.5 (deoarece secţia se întinde de la coloana 1 la coloana 12

inclusiv). Se alege coloana 6.5 deoarece se consideră că această coloană are 6 coloane în stânga şi 6 coloane în dreapta.

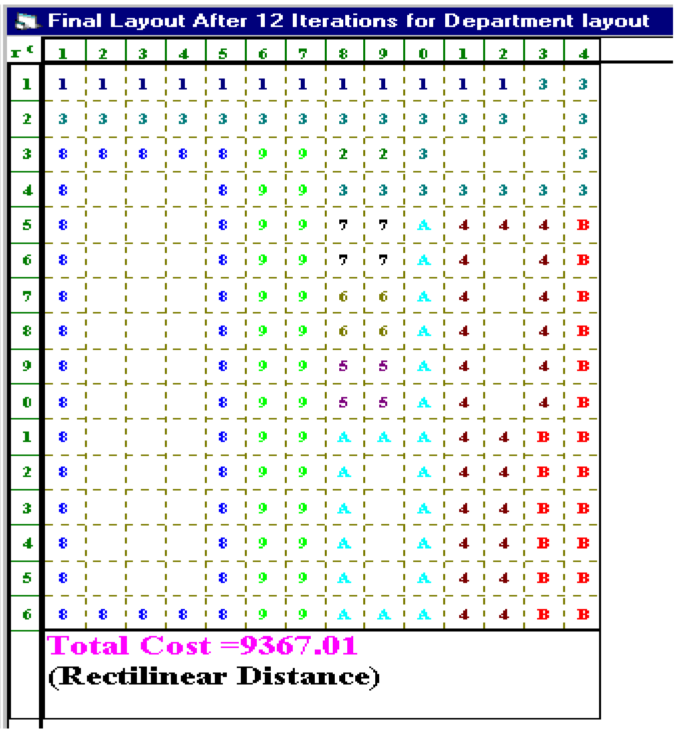
Tabelul de mai jos prezintă centrul calculat pentru fiecare din compartimentele prezentate în exemplul nostru.

11-09-2002 12:31:32	Department Name	Center Row	Center Column	Flow To All Departments	Cost To All Departments
1	1	1	6.50	50	405
2	2	3	8.50	50	180.00
3	3	2.61	9.61	220	1,745.00
4	4	12.30	9.60	155	719.50
5	5	5.50	8.50	240	1085
6	6	7.50	8.50	200	758
7	7	9.50	8.50	160	880
8	8	9.50	3	190	1567
9	9	9.50	6.50	170	965
10	A	7.50	12.50	280	1,628.50
11	B	13.50	13	30	256
	Total			1745	10189
	Distance	Measure:	Rectilinear		

O soluție mai bună a problemei putem obține alegând altă opțiune inițială, cum ar fi:

- schimbarea a 3 compartimente
- schimbarea a 2 şi apoi a 3 compartimente
- schimbarea a 3 şi apoi a 2 compartimente

Pentru exemplul nostru, aşa cum se vede în imaginea de mai jos, opțiunea “Schimbarea a 3 şi apoi a 2 compartimente” conduce la o soluție îmbunătățită a problemei.



Final Layout After 12 Iterations for Department layout

r c	1	2	3	4	5	6	7	8	9	0	1	2	3	4
1	1	1	1	1	1	1	1	1	1	1	1	1	3	3
2	3	3	3	3	3	3	3	3	3	3	3	3		3
3	8	8	8	8	8	9	9	2	2	3				3
4	8				8	9	9	3	3	3	3	3	3	3
5	8				8	9	9	7	7	A	4	4	4	B
6	8				8	9	9	7	7	A	4		4	B
7	8				8	9	9	6	6	A	4		4	B
8	8				8	9	9	6	6	A	4		4	B
9	8				8	9	9	5	5	A	4		4	B
0	8				8	9	9	5	5	A	4		4	B
1	8				8	9	9	A	A	A	4	4	B	B
2	8				8	9	9	A		A	4	4	B	B
3	8				8	9	9	A		A	4	4	B	B
4	8				8	9	9	A		A	4	4	B	B
5	8				8	9	9	A		A	4	4	B	B
6	8	8	8	8	8	9	9	A	A	A	4	4	B	B

Total Cost =9367.01
(Rectilinear Distance)

În ambele amplasamente prezentate mai sus compartimentul 1 este amplasat de-a lungul primului rând. În cazul în care considerăm din anumite motive că ar fi mai potrivit amplasamentul inițial al compartimentului, pe primele 2 rânduri avem posibilitatea să stabilim că poziția acestui compartiment este fixă şi dacă folosim opțiunea de îmbunătățire a soluției prin schimbarea a 3 şi apoi a 2 compartimente vom obține:

r c	1	2	3	4	5	6	7	8	9	0	1	2	3	4
1	1	1	1	1	1	1	3	3	3	3	3	3	3	3
2	1	1	1	1	1	1	3							3
3	8	8	8	8	8	9	9	3					3	3
4	8				8	9	9	3	3	3	3	3	2	2
5	8				8	9	9	7	7	A	4	4	4	B
6	8				8	9	9	7	7	A	4		4	B
7	8				8	9	9	6	6	A	4		4	B
8	8				8	9	9	6	6	A	4		4	B
9	8				8	9	9	5	5	A	4		4	B
0	8				8	9	9	5	5	A	4		4	B
1	8				8	9	9	A	A	A	4	4	B	B
2	8				8	9	9	A		A	4	4	B	B
3	8				8	9	9	A		A	4	4	B	B
4	8				8	9	9	A		A	4	4	B	B
5	8				8	9	9	A		A	4	4	B	B
6	8	8	8	8	8	9	9	A	A	A	4	4	B	B

Total Cost =10585.23
(Rectilinear Distance)

Echilibrarea liniilor de producție (*Assembly line balancing*)

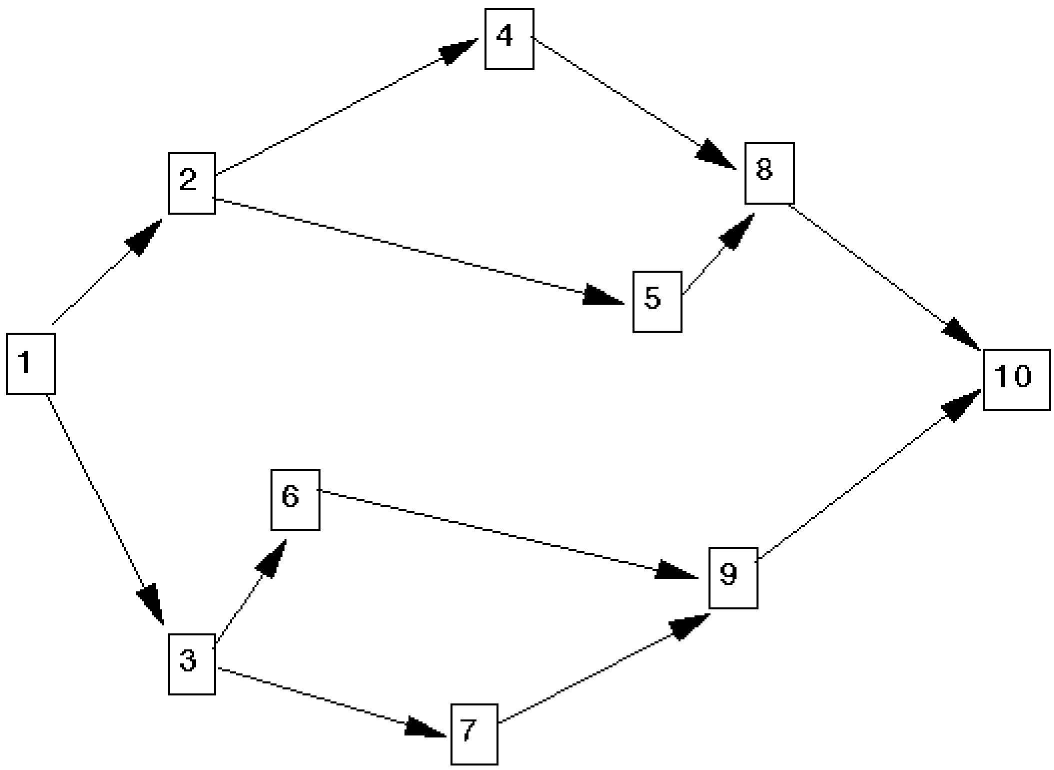
În problemele de echilibrare a liniilor de producție (de asamblare) ne propunem să reunim operațiile individuale (*tasks*) în grupuri de operații executate la un punct de lucru (*workstations*) astfel încât aceste puncte de lucru să funcționeze cât mai puțin în gol (să fie cât mai puțin timp în lipsă de lucrări).

Pentru a ilustra acest tip de problemă, vom considera o firmă de confecții care doreşte să-şi optimizeze producția unui anumit tip de costum. Linia de asamblare pentru producerea acestui costum presupune realizarea a 10 operații individuale, durata fiecărei operații şi operația imediat următoare sunt prezentate în tabelul de mai jos:

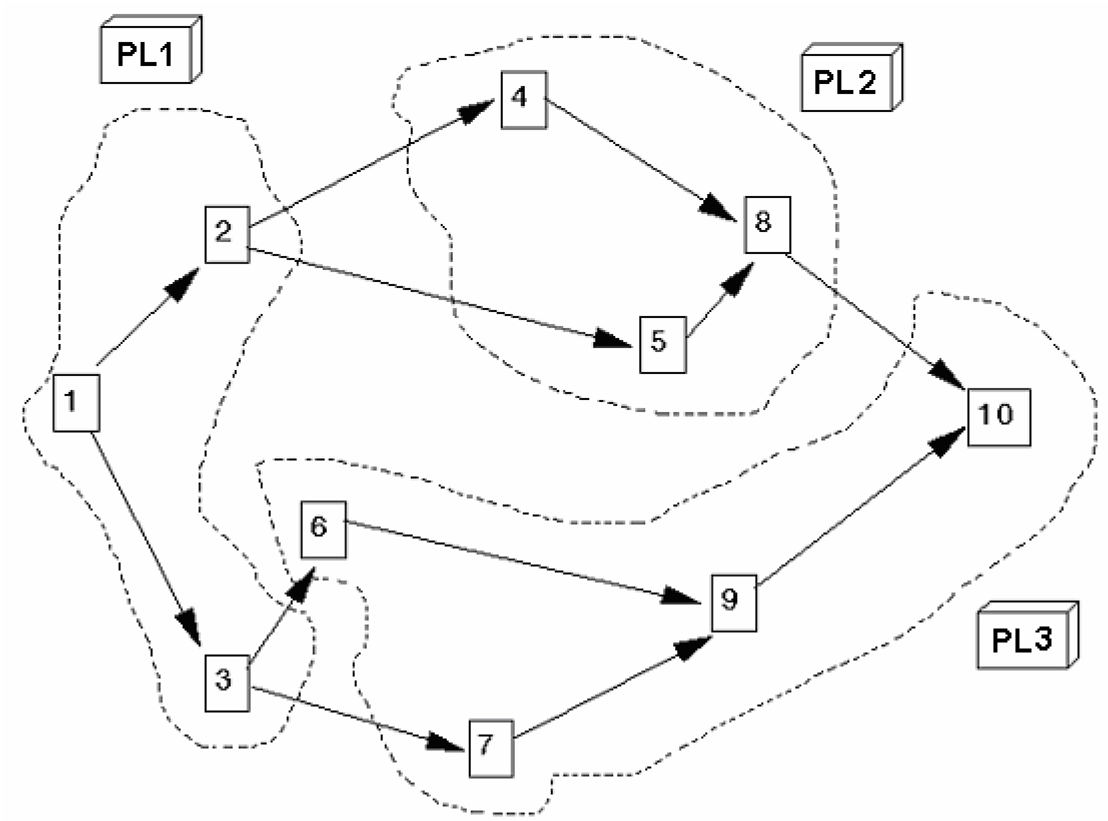
Operația elementară	Durata (secunde)	Operația care urmează
1	40	2,3
2	30	4,5
3	50	6,7
4	40	8
5	6	8
6	25	9
7	15	9
8	20	10
9	18	10
10	30	-

Se observă că operațiile 2 şi 3 nu pot începe până nu se termină executarea operației 1, operațiile 4 şi 5 nu pot începe până nu se termină executarea operației 2 ş.a.

Reprezentarea grafică a acestui proces de producție sub formă de diagramă este:



În imaginea de mai jos este prezentată o posibilă grupare a operațiilor în 3 puncte de lucru (PL1, PL2, PL3). Considerăm că operațiile care sunt executate la acelaşi punct de lucru sunt executate succesiv (una după cealaltă), de către un singur muncitor şi că activitatea se desfăşoară în acelaşi timp la toate punctele de lucru.



Gruparea operațiilor în puncte de lucru prezentată mai sus, face ca durata operațiilor executate la punctul de lucru PL1 să fie de 120 secunde, la punctul de lucru PL2 să fie 66 secunde iar la punctul de lucru PL3 să fie 88 secunde.

La un anumit moment, de la punctul de lucru PL1 se transferă materialul prelucrat la punctul de lucru PL2, de la acesta la punctul de lucru PL3 iar apoi se trece materialul prelucrat la o altă fază a procesului de producție. Această modalitate de transfer a materialului prelucrat de la un punct de lucru la altul presupune ca transferul să aibă loc numai în momentul în care în cadrul fiecărui punct de lucru au fost executate toate operațiile individuale. Ea (această modalitate) are avantajul că nu necesită existența unor puncte intermediare de depozitare şi că prelucrarea începe imediat ce materialul a fost transferat, punctul de lucru fiind liber. În cele mai multe cazuri punctele de lucru sunt amplasate unul lângă celălalt, deci nu se mai pierde timp suplimentar cu transportul.

Modalitatea descrisă mai sus poartă numele de "proces în paşi", adică sistemul avansează cu viteza celui mai încet punct de lucru.

Deoarece soluția de grupare a operațiilor în puncte de lucru respectă toate condițiile tehnologice precizate în enunț putem spune că am stabilit o variantă de linie de asamblare cu 3 puncte de lucru, viteza liniei de asamblare fiind de 120 de secunde. Problema care se pune este aceea de a echilibra această linie de asamblare în sensul de a afla care este cea mai bună grupare a operațiilor în stații de lucru astfel încât să obținem o anumită viteză a sistemului (să se producă un număr de unități într-o perioadă dată de timp).

Presupunând că dorim să producem 60 de unități de produs pe oră, vom avea nevoie practic să producem o unitate de produs la fiecare 60 (3600/60) de secunde.

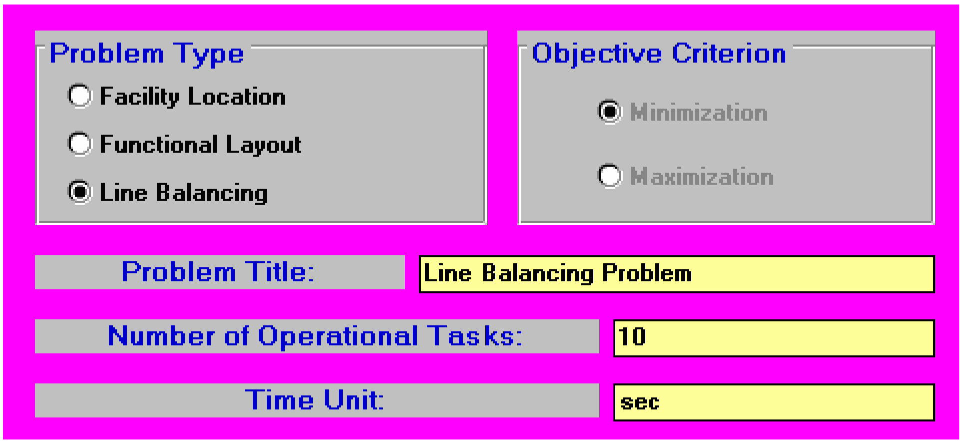
Având în vedere că o linie de asamblare este alcătuită din puncte de lucru înseamnă că la fiecare 60 de secunde trebuie să se producă schimbarea materialului prelucrat între punctele de lucru.

Timpul total de prelucrare este de 40+30+…+18+30=274, ceea ce presupune că avem nevoie de cel puţin 274/60=4.6, deci 5 puncte de lucru pentru a asigura ca o unitate de produs să fie prelucrată la fiecare 60 de secunde. Având însă în vedere constrângerile tehnologice legate de succesiunea operaţiilor, nu există nici un fel de garanţie că cele 5 puncte de lucru vor asigura un flux de 60 de unităţi pe oră.

Rezolvarea problemelor utilizând un produs software specializat

Problema echilibrării liniei de asamblare poate fi rezolvată cu ajutorul produselor software specializate, furnizându-se datele iniţiale şi datele de intrare conform dialogului şi machetei matriciale de mai jos.

Observăm că avem posibilitatea să stabilim şi anumite operaţii izolate, adică operaţii care pot fi executate oricând în cadrul procesului, şi care nu au predecesori sau succesori.



Task Number	Task Name	Task Time in sec	Task Isolated (Y/N)	Immediate Successor (task number separated by ,)
1	Task 1	40	No	2,3
2	Task 2	30	No	4,5
3	Task 3	50	No	6,7
4	Task 4	40	No	8
5	Task 5	6	No	8
6	Task 6	25	No	9
7	Task 7	15	No	9
8	Task 8	20	No	10
9	Task 9	18	No	10
10	Task 10	30	No	

În continuare este prezentată soluţia furnizată de pachetul de programe, utilizându-se ca metodă de găsire a soluţiei "*Optimizing Best Bud Search*", aceasta asigurând găsirea unei soluţii cu un număr minim de puncte de lucru care să satisfacă cerinţele tehnologice de respectare a succesiunii operaţiilor.

Solution Method

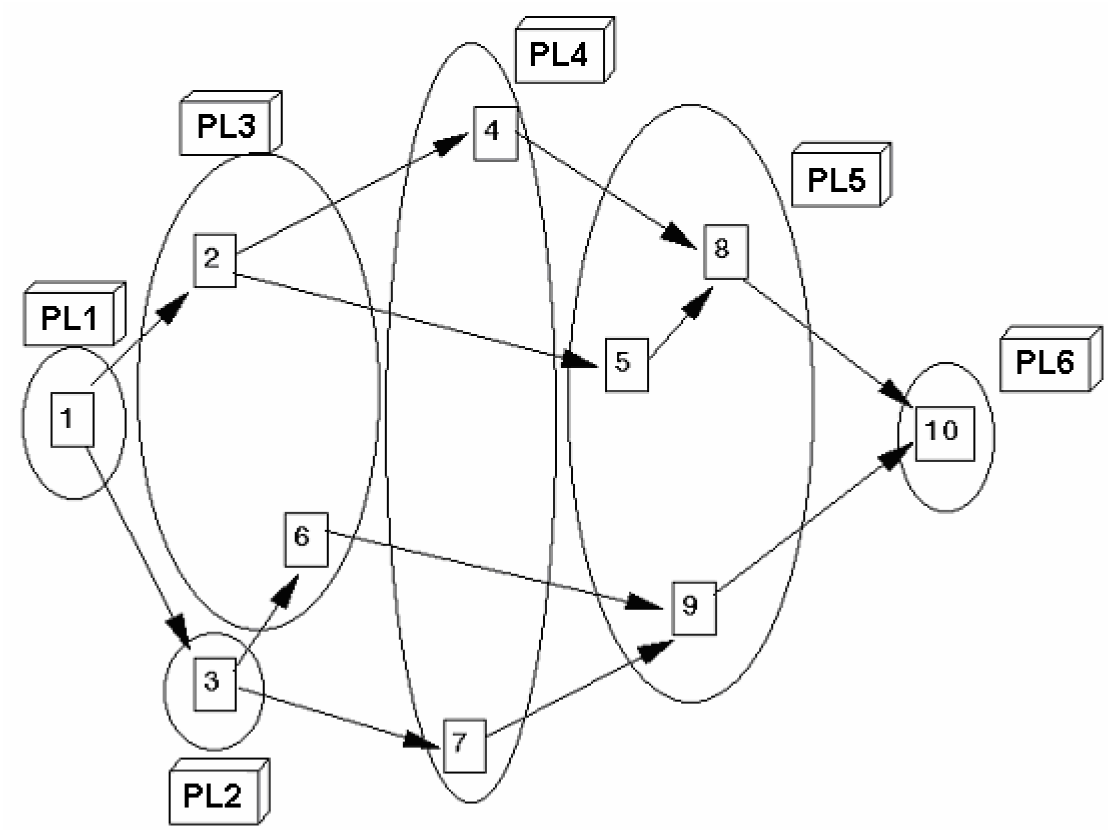
- ( ) Heuristic Procedure
- (•) Optimizing Best Bud Search
- ( ) COMSOAL Type Random Generation

Cycle time in sec	60
Time length in sec	3600
Desired production quantity:	60

11-09-2002 17:27:20	Line Station	Number of Operators	Task Assigned	Task Name	Task Time	Time Unassigned	% Idleness
1	1	1	1	Task 1	40	20	33.33%
2	2	1	3	Task 3	50	10	16.67%
3	3	1	2	Task 2	30	30	50.00%
4			6	Task 6	25	5	8.33%
5	4	1	4	Task 4	40	20	33.33%
6			7	Task 7	15	5	8.33%
7	5	1	9	Task 9	18	42	70.00%
8			5	Task 5	6	36	60.00%
9			8	Task 8	20	16	26.67%
10	6	1	10	Task 10	30	30	50.00%
	Solved by	Best Bud	Search				

11-09-2002	Item	Result
1	Desired Cycle Time in sec	60
2	Number of Line Stations	6
3	Number of Required Operators	6
4	Total Available Time in sec	360
5	Total Task Time in sec	274
6	Total Idle Time in sec	86
7	Balance Delay (%)	23.89%
	Solution has been obtained by	
	Best Bud Search	

În continuare sunt prezentate sub formă de diagramă punctele de lucru şi operațiile, aşa cum rezultă din soluția furnizată de calculator. Observăm că ordinea de parcurgere a punctelor de lucru este: PL1, unde se execută operația 1, apoi PL2, unde se execută operația 3, apoi PL3, unde se execută operațiile 2 şi 6, etc.



Pachetul de programe furnizează o soluție cu 6 puncte de lucru, fiecare având un disponibil de 60 de secunde pentru executare operațiilor repartizate lor. În total rezultă un timp disponibil de 60*6=360 de secunde.

Timpul necesar pentru executarea tuturor operațiilor este de 274 de secunde, deci rezultă un timp de staționare de 360-274=86 secunde pentru un ciclu de producție. Astfel, rezultă că (274/360)*100=76.11% din timpul total disponibil este folosit pentru producție, ceea ce mai este cunoscut sub numele de eficiență a ciclului de producție şi (86/360)*100=23.89% din timpul total disponibil este neutilizat, care mai este cunoscut sub numele de întârziere a ciclului de producție.

Într-un proces ideal timpul neutilizat ar trebui să fie 0, deci la nici unul dintre punctele de lucru să nu existe timp neproductiv. Pentru a ne apropia cât mai mult de acest deziderat se pot încerca mai multe variante de cicluri de timp în care să fie furnizat un anumit număr de unități de produs.

În continuare este prezentată soluția furnizată de calculator problemei noastre în condițiile în care se doreşte producerea a 65 de unități de produs pe oră (un ciclu de producție de 55 de secunde).

11-09-2002 18:06:32	Line Station	Number of Operators	Task Assigned	Task Name	Task Time	Time Unassigned	% Idleness
1	1	1	1	Task 1	40	15	27.27%
2	2	1	3	Task 3	50	5	9.09%
3	3	1	2	Task 2	30	25	45.45%
4			6	Task 6	25	0	0.00%
5	4	1	4	Task 4	40	15	27.27%
6			7	Task 7	15	0	0.00%
7	5	1	9	Task 9	18	37	67.27%
8			5	Task 5	6	31	56.36%
9			8	Task 8	20	11	20.00%
10	6	1	10	Task 10	30	25	45.45%
	Solved by	Best Bud	Search				

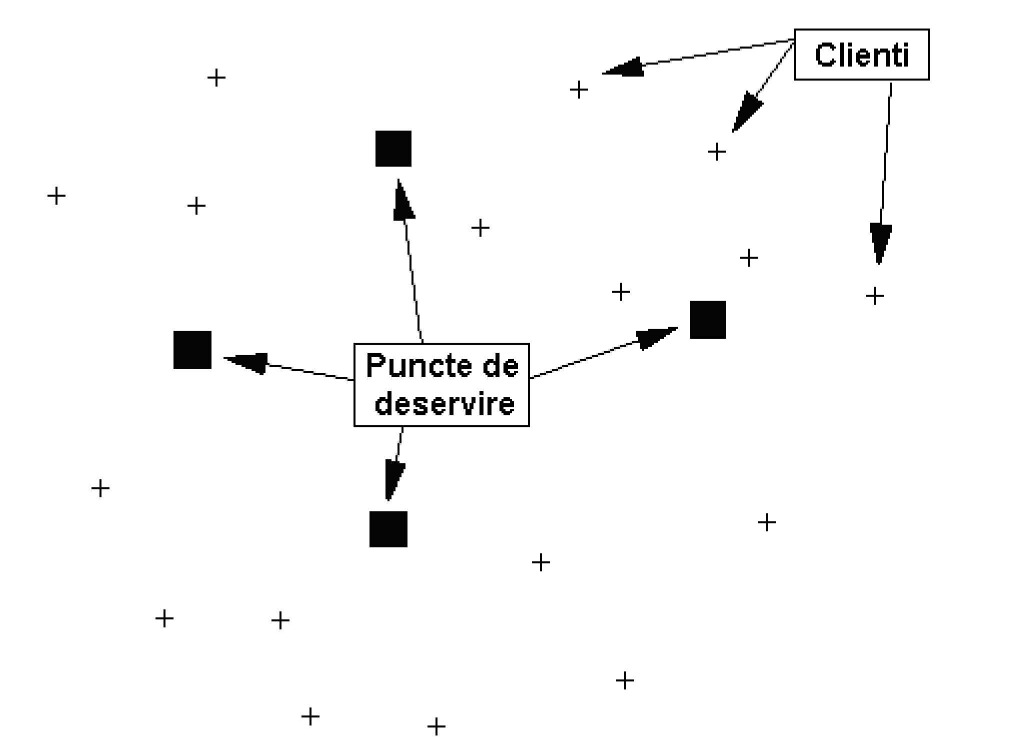
11-09-2002	Item	Result
1	Desired Cycle Time in sec	55
2	Number of Line Stations	6
3	Number of Required Operators	6
4	Total Available Time in sec	330
5	Total Task Time in sec	274
6	Total Idle Time in sec	56
7	Balance Delay (%)	16.97%
	Solution has been obtained by	
	Best Bud Search	

Amplasarea facilităţilor (punctelor de deservire) (*Facility location*)

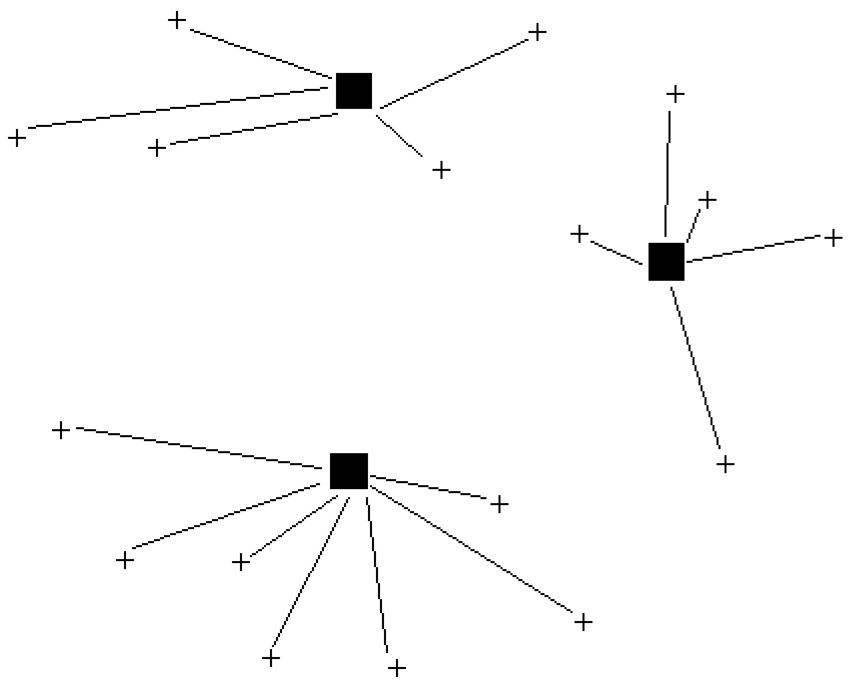
La modul general, problema facilităţilor (punctelor de deservire) se enunţă astfel: dându-se un număr de facilităţi şi un număr de clienţi care folosesc aceste facilităţi, care dintre aceste facilităţi vor fi folosite şi ce clienţi vor fi deserviţi de fiecare dintre ele astfel încât să se minimizeze costul deservirii tuturor clienţilor.

Vom considera că facilităţile (punctele de deservire) sunt fie deschise (deservesc un client), fie închise şi există un cost fix care este implicat de deschiderea unui punct de deservire. Ceea ce trebuie stabilit este care dintre punctele de deservire să fie deschise şi care dintre ele închise.

În continuare poate fi văzută o reprezentare grafică a problemei:



O posibilă soluţie este prezentată mai jos:



Alți factori care influențează acest proces mai sunt:

- clienții au asociat un anumit nivel al cererii iar punctele de deservire au asociată o anumită capacitate limită de deservire;
- clienții pot fi deserviți de la unul sau mai multe puncte de deservire.

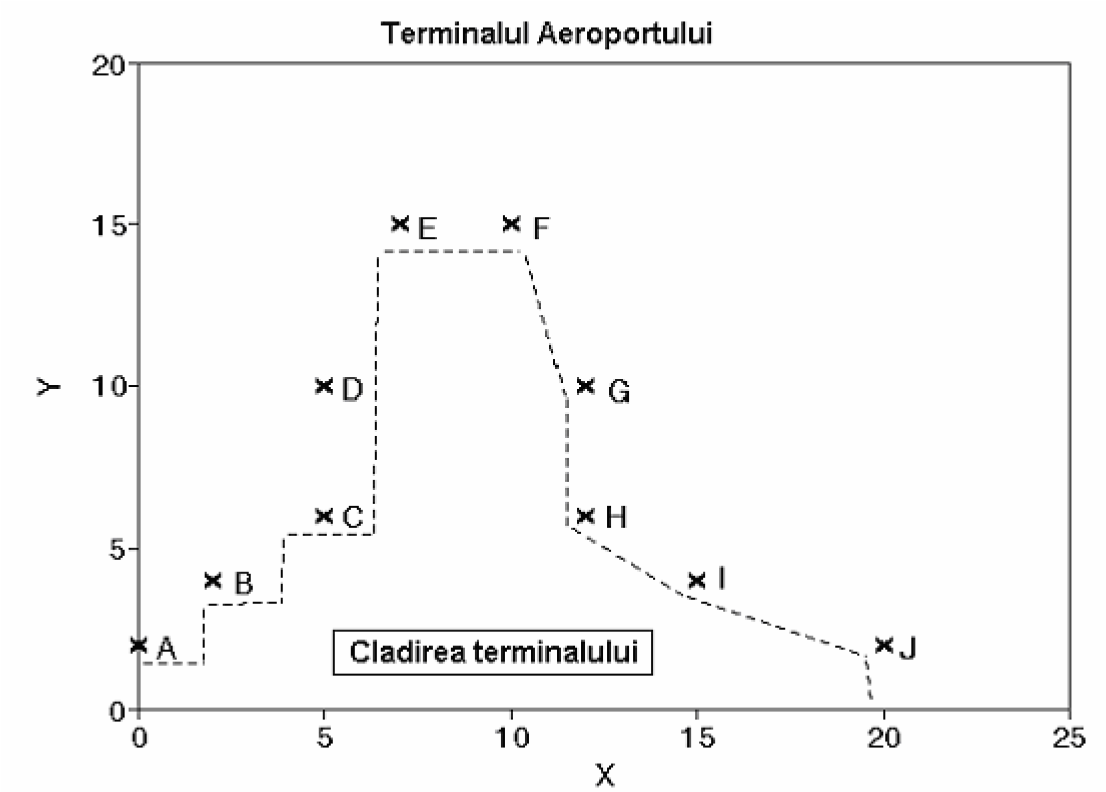
Problemele dimensionării punctelor de deservire şi ale amplasării acestora sunt strâns legate între ele şi vor fi abordate simultan. De fapt programul prezentat aici disociază din motive de simplificare cele două probleme şi se ocupă numai de amplasarea facilităților.

Exemplu

Vom ilustra problema amplasării punctelor de deservire prin intermediul următorului exemplu:

Terminalul unui aeroport are 10 porți (notate de la A la J). O reprezentare grafică a terminalului precum şi coordonatele amplasamentului porților sunt prezentate în continuare:

Poarta	coordonata x	coordonata y
A	0	2
B	2	4
C	5	6
D	5	10
E	7	15
F	10	15
G	12	10
H	12	6
I	15	4
J	20	2



Bagajele de la zborurile care sosesc pe aeroport sunt descărcate la porțile de sosire şi sunt transportate de aici la un punct general de predare a bagajelor.

Se estimează că numărul de unități de bagaje care sosesc la fiecare din porți într-o zi este de: 3600, 2500, 1800, 2200, 1000, 4500, 5600, 1400, 1800 şi respectiv 3000.

Unde ar trebui amplasat punctul general de predare a bagajelor astfel încât să se minimizeze fluxul bagajelor transportate?

Soluție

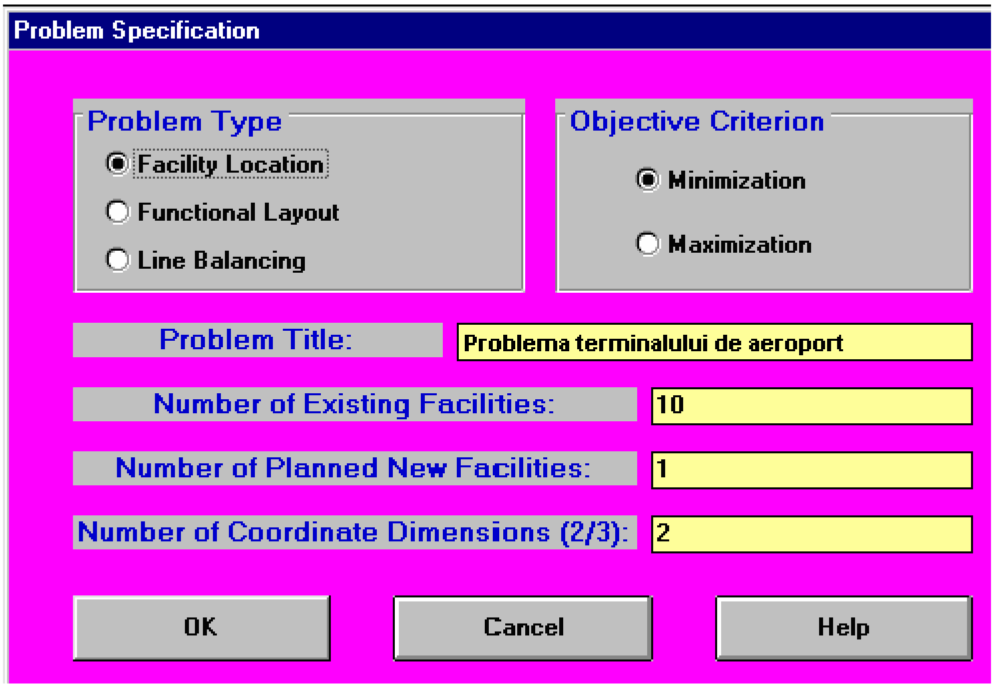
Pentru a poziționa punctul general de predare a bagajelor trebuie să ținem seama de cantitatea de bagaje care trebuie transportată de la fiecare poartă la punctul de predare. În mod logic o poartă la care soseşte un număr mai mare de bagaje ar trebui să fie mai aproape de punctul de predare a bagajelor decât una la care sosesc mai puține bagaje. Altfel spus, vom amplasa punctul de predare a bagajelor astfel

încât să minimizăm suma tuturor produselor dintre distanţa de la fiecare poartă la punctul de predare a bagajelor(d_i) şi numărul bagajelor de la fiecare poartă (c_i).

$$\textbf{Min} \sum d_i \times c_i \quad i=1,n$$

Rezolvarea problemelor utilizând un produs software specializat

Datele iniţiale ale problemei enunţate anterior care vor fi furnizate pachetului de programe sunt:



Terminologia utilizată în cadrul pachetului are următoarea semnificaţie:

- ***existing facilities*** – se referă la acele puncte care sunt deja fixate în spaţiu şi ale căror coordonate le cunoaştem
- ***new facilities*** – se referă la acele facilităţi ale căror poziţii în spaţiu dorim să le determinăm cu ajutorul pachetului de programe.

Pentru a introduce datele despre porţile de sosire, putem ignora în ecranul de introducere a datelor coloanele care se referă la fluxurile dintre facilităţile existente, deoarece, în cazul nostru, toate fluxurile sunt între fiecare din porţile de sosire şi punctul de predare a bagajelor.

Facility Number	Facility Name	To New 1 Flow/Unit Cost	Location X Axis	Location Y Axis
Existing 1	A	3600	0	2
Existing 2	B	2500	2	4
Existing 3	C	1800	5	6
Existing 4	D	2200	5	10
Existing 5	E	1000	7	15
Existing 6	F	4500	10	15
Existing 7	G	5600	12	10
Existing 8	H	1400	12	6
Existing 9	I	1800	15	4
Existing 10	J	3000	20	2
New 1	NEW			

De remarcat că în rezolvarea acestei probleme este necesar să stabilim modalitatea de calcul a distanței dintre două locații. Noi nu cunoaştem la momentul inițial amplasarea punctului de predare a bagajelor, deci nici distanța dintre el şi fiecare poartă de sosire. În mod similar cu problema *Functional Layout* (prezentată anterior), dacă (x_i,y_i) şi (x_j,y_j) reprezintă coordonatele a două locații: i şi j atunci distanța dintre ele poate fi calculată folosind modelul:

- rectiliniar, ceea ce presupune calculul distanței dintre i şi j ca:
  - $|x_i-x_j|+|y_i-y_j|$
- Euclidian, ceea ce presupune calculul distanței dintre i şi j ca:
  - $[(x_i-x_j)^2 + (y_i-y_j)^2]^{0.5}$
- pătratic Euclidian, ceea ce presupune calculul distanței dintre i şi j ca:
  - $(x_i-x_j)^2 + (y_i-y_j)^2$

Rezultatele obținute prin rularea pachetului de programe QSB, folosind fiecare din modelele de măsurare a distanțelor prezentate anterior sunt.

Modelul rectiliniar

11-09-2000 14:33:37	Facility Name	X Axis	Y Axis	Flow To All Facilities	Cost To All Facilities
1	A	0	2	3600	50400
2	B	2	4	2500	25000
3	C	5	6	1800	9000
4	D	5	10	2200	19800
5	E	7	15	1000	12000
6	F	10	15	4500	40500
7	G	12	10	5600	33600
8	H	12	6	1400	2800
9	I	15	4	1800	12600
10	J	20	2	3000	42000
11	NEW	10	6	0	0
	Total			27400	247700
	Distance	Measure:	Rectilinear		

Modelul Euclidian

11-09-2000 14:34:36	Facility Name	X Axis	Y Axis	Flow To All Facilities	Cost To All Facilities
1	A	0	2	3600	44,266.76
2	B	2	4	2500	23,819.75
3	C	5	6	1800	10,667.54
4	D	5	10	2200	11,483.13
5	E	7	15	1000	6,776.21
6	F	10	15	4500	27,073.98
7	G	12	10	5600	11,965.45
8	H	12	6	1400	4,938.54
9	I	15	4	1800	12,556.59
10	J	20	2	3000	36,299.11
11	NEW	10.12	8.98	0	0
	Total			27400	189,847.06
	Distance	Measure:	Euclidean		

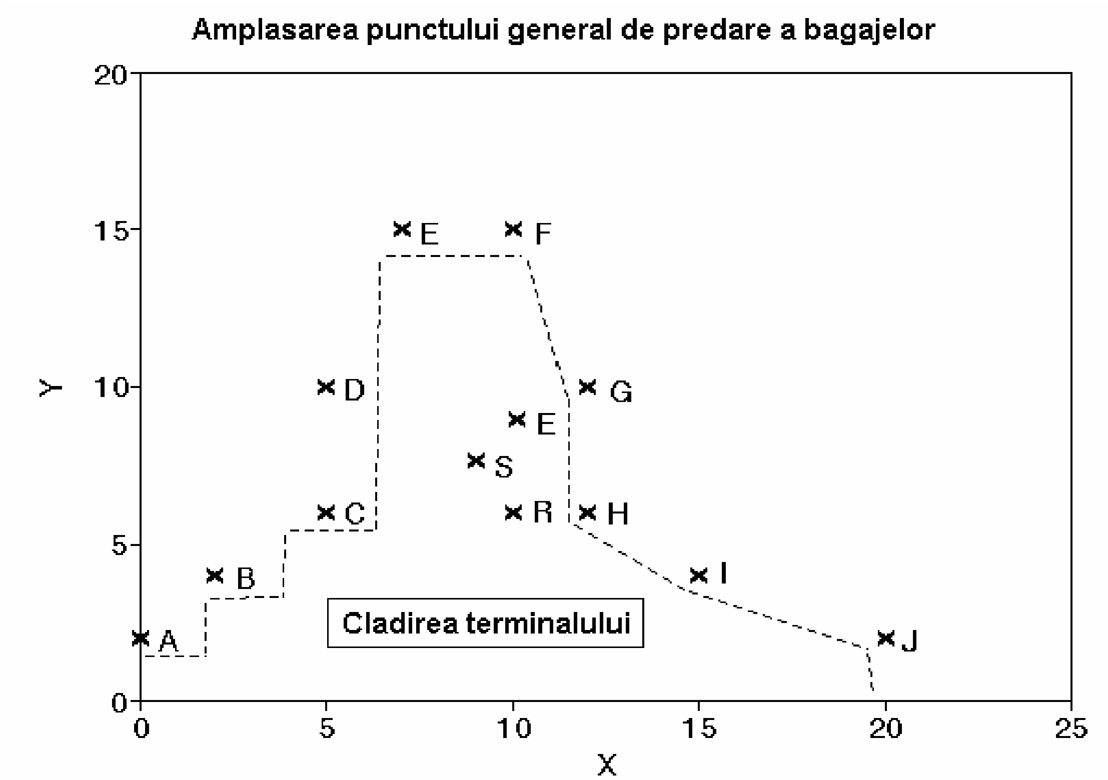
Modelul pătratic Euclidian

11-09-2000 14:35:13	Facility Name	X Axis	Y Axis	Flow To All Facilities	Cost To All Facilities
1	A	0	2	3600	410570
2	B	2	4	2500	157,928.28
3	C	5	6	1800	34,547.77
4	D	5	10	2200	48,070.31
5	E	7	15	1000	57,966.92
6	F	10	15	4500	245,971.59
7	G	12	10	5600	79,154.93
8	H	12	6	1400	16,069.03
9	I	15	4	1800	87,917.11
10	J	20	2	3000	456,010.25
11	NEW	9.05	7.67	0	0
	Total			27400	1,594,206.25
	Distance	Measure:	Squared	Euclidean	

Aşa cum vedem din soluția furnizată de pachetul de programe, coordonatele la care ar trebui amplasat punctul general de predare a bagajelor sunt diferite în funcție de modelul folosit pentru măsurarea distanțelor, astfel:

Modelul	coordonata x	coordonata y
Rectiliniar	10	6
Euclidian	10.12	8.98
Pătratic Euclidian	9.05	7.67

Acelaşi lucru este reprezentat grafic în imaginea de mai jos, folosindu-se notațiile R,E şi S pentru cele trei modele de măsurare a distanței: rectiliniar, Euclidian, pătratic Euclidian.



De remarcat că alegerea inițială a punctului de amplasare a facilității este cel mai frecvent între soluția dată de modelul rectiliniar şi cel Euclidian (ținând cont de modul în care se desfăşoară fluxul bagajelor între porți şi punctul de predare a bagajelor). Dacă bagajele circulă în linie dreaptă, atunci este de preferat soluția obținută cu modelul pătratic Euclidian deoarece acesta descurajează distanțele foarte mari.

De remarcat că nu ne propunem să determinăm o poziție exactă până la nivel de milimetri ci folosim pachetul de programe pentru a determina o regiune în care să fie amplasat punctul de predare a bagajelor, costurile fiind aproximativ egale pentru acea regiune. Pachetul de programe foloseşte şi el o serie de aproximații în cadrul algoritmului de calcul..

O extensie a acestei probleme rezolvabilă tot cu ajutorul pachetului de programe QSB este aceea de a amplasa mai multe puncte de predare a bagajelor şi de a stabili pentru fiecare dintre porțile de sosire la ce punct de predare se vor transporta bagajele ei.

Elemente de analiză statistică. Programul SPSS

Înainte de a demara prezentarea programului SPSS şi a principalelor aplicații ale acestuia sunt necesare clarificarea unei terminologii statistice cu care se va opera şi înțelegerea unor elemente fundamentale de statistică care vor include: caracteristicile statistice, variabile şi scale de măsurare, mărimile medii, indicatorii variației şi asimetriei, valorile medii de poziție şi de structură, scorurile standard, curba normală, ipotezele statistice şi coeficientul de corelație.

Caracteristicile statistice, variabile şi scale de măsurare

Caracteristicile statistice reprezintă însuşirile fenomenelor studiate. Astfel putem deosebi caracteristici variabile ca formă de manifestare sau ca nivel de dezvoltare.

Caracteristicile statistice pot fi clasificate după:

- conținutul caracteristicii: de timp, de spațiu sau atributive;
- natura variației: cu variație continuă, discontinuă;
- modul de obținere: primare, derivate;
- forma de manifestare: alternative, nealternative.

Caracteristicile statistice care pot avea mai mult decât o singură valoare cu care variază în funcție de o serie de factori se mai numesc şi **variabile statistice** iar formele de manifestare ale acestor caracteristici se numesc variante.

Variabilele pot fi:

- dependente dacă sunt spuse influenței altor variabile;
- independente dacă sunt variabile ce influențează alte variabile;
- continue dacă au un număr infinit al nivelurilor de măsurare;
- discrete dacă au un număr finit al nivelurilor de măsurare.

Dacă variabila se referă la caracterstica supusă măsurării, **scalele de măsurare** se referă la modalitatea de măsurare. Programul SPSS utilizează următoarele scale de măsurare:

- **Scală** – valorile numerice ale datelor se reprezintă pe un interval sau printr-un raport. Exemplu: vârsta, venitul, temperatura, lungimea, timpul de răspuns. Variabilele reprezentate şi măsurate pe scală trebuie să aibă valori numerice.
- **Nominale** – se utilizează atunci când valorile datelor unei variabile reprezintă valori de ordin neintrinsec în funcție de existența sau inexistența unei caracteristici. Exemplu: apartenența la o anumită categorie de funcții de încadrare sau sexul: 1 – masculin, 2 – feminin.
- **Ordinale** (de ordin sau de rang) – se folosesc atunci când valorile datelor reprezintă categorii de ordin intrinsec care pot fi puse într-o anumită ordine de măsurare. Exemplu nivelul de pregătire: scăzut, mediu, înalt.

Chiar dacă o variabilă (caracteristică) poate fi măsurată pe oricare dintre tipurile de scale prezentate este foarte important în prelucrările statistice ca pentru ea să fie aleasă scala cea mai potrivită.

Mărimile medii, indicatorii variaţiei şi asimetriei

Mărimile medii

În analiza statistică se utilizează foarte frecvent mărimile medii deoarece pe baza lor se poate exprima într-o anumită măsură tendinţa unor fenomene. Aceste mărimi fac parte din cadrul indicatorilor derivaţi şi au un caracter abstract.

Pentru a determina valoarea tendinţei centrale a unei serii statistice se utilizează: media aritmetică, media armonică, media pătratică şi media geometrică. Toate aceste medii se pot calcula ca medii simple sau ca medii ponderate.

Media aritmetică

Formula matematică de exprimare a mediei aritmetice simple este:

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

unde:

$\overline{X}$ **este media aritmetică simplă**;

x_i reprezintă termenii individuali ai seriei (x_1, x_2, … x_n) iar suma acestora este: $x_1 + x_2 + \ldots + x_n = \sum_{i=1}^{n} x_i$

n reprezintă numărul de termeni pentru care se calculează media.

În urmatorul exemplu sunt prezentate notele obţinute la teste şi la interviu de către 5 studenţi care s-au prezentat pentru promovarea unui examen:

Nume studenţi	**Note la testul 1**	**Note la testul 2**	**Note la interviu**
Student 1	10	9	8
Student 2	7	7	8
Student 3	7	8	9
Student 4	9	9	8
Student 5	10	9	7
-	$\overline{X}$ = 43/5 = 8,60	$\overline{X}$ = 42/5 = 8,40	$\overline{X}$ = 40/5 = 8,00

În ultima linie a tabelului s-au calculat mediile aritmetice simple pentru teste şi interviu.

Mărimile medii ponderate se pot calcula atunci când fiecărei variante a unei caracteristici i se poate ataşa o frecvenţă.

Formula de calcul a mediei aritmetice ponderate este:

$$\overline{Xp} = \frac{\sum_{i=1}^{m} x_i f_i}{\sum_{i=1}^{m} f_i}$$

unde:

$\overline{Xp}$ este **media aritmetică ponderată**;

x_i reprezintă termenii variantelor ($x_1, x_2, \ldots x_n$);

f_i reprezintă frecvenţa sau ponderea fiecărei variante;

m reprezintă numărul variantelor pentru care se calculează media.

Dacă vom organiza datele astfel încât să evidenţiem frecvenţele pentru notele obţinute în două variante (varianta 1 corespunzătoare intervalului de note cuprinse între 9 şi 10 şi varianta 2 pentru intervalul de note cuprinse între 7 şi 8) atunci termenii variantelor vor trebui calculaţi ca fiind mijlocul fiecărui interval. În acest caz vom obţine următorul tabel:

Variante	**Mijlocul de interval (x_i)**	**Frecvenţă note test 1 $(f_i)_1$**	**Frecvenţă note test 2 $(f_i)_2$**	**Frecvenţă note interviu $(f_i)_3$**
9 - 10	9,5	3	3	1
7 - 8	7,5	2	2	4
Media ponderată	-	$\overline{Xp}$ = 8,7	$\overline{Xp}$ = 8,7	$\overline{Xp}$ = 7,9

O altă aplicaţie a mediei ponderate este atunci când se utilizează ponderi (coeficienţi de importanţă) pentru termenii individuali ai seriei sau pentru intervale. Pentru exemplificare vom presupune că se doreşte ca notele obţinute la proba interviu să fie ponderate astfel:

Nota	**Pondere**
10	0,4
9	0,3
8	0,2
7	0,1
≤ 6	0

Dacă se lucrează cu ponderi subunitare, atunci întotdeauna suma ponderilor acordate trebuie să fie 1. În cazul utilizării valorilor procentuale, suma ponderilor va trebui să fie 100.

Revenind la primul tabel, conform ponderilor dorite vom avea:

Note interviu	**Ponderi**
8	0,2
8	0,2
9	0,3
8	0,2
7	0,1
$\overline{Xp}$ = 7,5	-

Media armonică

Media armonică a termenilor unei serii se defineşte ca fiind acea valoare a cărei mărime inversă este media aritmetică rezultată din valorile inverse ale termenilor aceleiaşi serii.

Dacă $x_1, x_2, \dots x_n$ reprezintă termenii seriei, atunci valorile lor inverse vor fi: $\frac{1}{x_1}, \frac{1}{x_2}, \dots, \frac{1}{x_n}$.

Formula de calcul a mediei armonice simple este:

$$\overline{X}_a = \frac{n}{\sum_{i=1}^{n} 1/x_i}$$

unde:

$\overline{X}a$ este **media armonică simplă**;

x_i reprezintă termenii individuali ai seriei $(x_1, x_2, \dots x_n)$;

n reprezintă numărul de termeni pentru care se calculează media.

Pentru exemplul anterior mediile armonice calculate vor fi:

Nume studenți	Note la testul 1	Note la testul 2	Note la interviu
Student 1	10	9	8
Student 2	7	7	8
Student 3	7	8	9
Student 4	9	9	8
Student 5	10	9	7
-	$\overline{X}a$ = 8,38	$\overline{X}a$ = 8,32	$\overline{X}a$ = 7,95

În tabel se observă că valorile mediilor armonice calculate sunt mai mici decât cele ale mediilor aritmetice. În acest sens trebuie cunoscut faptul că, întotdeauna dacă valorile termenilor sunt pozitive atunci valorile mediilor armonice calculate vor fi mai mici decât cele ale mediilor aritmetice.

Expresia mediei armonice ponderată este:

$$\overline{X}_{ap} = \frac{\sum_{i=1}^{m} f_i}{\sum_{i=1}^{m} f_i/x_i}$$

unde:

$\overline{X}ap$ este **media armonică ponderată**;

x_i reprezintă termenii variantelor $(x_1, x_2, \dots x_n)$;

f_i reprezintă frecvența sau ponderea fiecărei variante;

m reprezintă numărul variantelor pentru care se calculează media.

Media pătratică

Media pătratică se obține în mod similar cu media aritmetică dar cu deosebirea că valorile termenilor seriei se ridică la pătrat, iar apoi se extrage radicalul de ordinul 2 din suma acestora raportată la numărul de termeni pentru care se calculează.

$$x_1^2 + x_2^2 + \dots + x_n^2 = \frac{\sum_{i=1}^{n} x_i^2}{n}$$

Formulele de calcul ale mediei pătratice simple ($\overline{X}pa$) şi ale mediei pătratice ponderate ($\overline{X}pap$) sunt:

$$\overline{X}pa = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} \qquad \overline{X}pap = \sqrt{\frac{\sum_{i=1}^{m} x_i^2 f_i}{\sum_{i=1}^{m} f_i}}$$

unde:

$\overline{X}pa$ este **media pătratică simplă**;

$\overline{X}pap$ este **media pătratică ponderată**;

x_i reprezintă termenii seriei sau ai variantelor (x_1, x_2, … x_n);

f_i reprezintă frecvențele sau ponderile variantelor;

m reprezintă numărul variantelor pentru care se calculează media;

n reprezintă numărul de termeni pentru care se calculează media.

Pentru exemplul prezentat mediile pătratice calculate vor fi:

Nume studenți	**Note la testul 1**	**Note la testul 2**	**Note la interviu**
Student 1	10	9	8
Student 2	7	7	8
Student 3	7	8	9
Student 4	9	9	8
Student 5	10	9	7
-	$\overline{X}pa$ = 8,70	$\overline{X}pa$ = 8,43	$\overline{X}pa$ = 8,02

Este de reținut faptul că întotdeauna media pătratică este mai mare decât media aritmetică, indiferent de semnul termenilor. Cu cât valorile termenilor sunt mai mari cu atât ei vor influența mai mult valoarea mediei pătratice. Din acest motiv media pătratică se va folosi atunci când se doreşte să se acorde o importanță mai mare nivelurilor mai ridicate ale termenilor seriei.

Media geometrică

Media geometrică se calculează ca radical de ordinul n din produsul celor n termeni ai seriei.

$$x_1 . x_2 \ldots x_n = \prod_{i=1}^{n} x_i$$

Expresiile mediei geometrice simple ($\overline{X}g$) şi ale mediei geometrice ponderate ($\overline{X}gp$) sunt:

$$\overline{X}g = \sqrt[n]{\prod_{i=1}^{n} x_i} \qquad \overline{X}gp = \sqrt[\sum_{i=1}^{m} f_i]{\prod_{i=1}^{m} x_i^{f_i}}$$

unde:

$\overline{X}g$	este media geometrică simplă;
$\overline{X}gp$	este media geometrică ponderată;
x_i	reprezintă termenii seriei sau ai variantelor (x_1, x_2, … x_n);
$\prod_{i=1}^{n} x_i$	este produsul termenilor seriei;
f_i	reprezintă frecvenţele sau ponderile variantelor;
m	reprezintă numărul variantelor pentru care se calculează media;
n	reprezintă numărul de termeni pentru care se calculează media.

Este evident că prin proceduri manuale calculul mediei geometrice este foarte laborios. Pentru a putea fi calculată media geometrică este necesar ca toţi termenii seriei să fie pozitivi şi mai mari decât 0. Este de reţinut faptul că întotdeauna media geometrică va fi mai mică decât media aritmetică.

Relaţia între mediile prezentate este:

$$\overline{X}a < \overline{X}g < \overline{X} < \overline{X}pa$$

Valorile medii de poziție şi de structură

Cele mai frecvent utilizate valori medii de poziție şi de structură sunt mediana, modulul, cuartilele şi decilele.

Mediana

Mediana (Me) reprezintă valoarea centrală a unei serii statistice care împarte termenii unei seriei în două părți egale. Jumătatea inferioară va conține termenii ale căror valori sunt mai mici decât valoarea medianei iar jumătatea superioară va conține termenii care au o valori mai mari decât valoarea medianei.

Locul medianei (LMe) într-o serie de n termeni este:

$$LMe = \frac{n+1}{2}$$

Locul medianei (LMef) într-o serie de distribuție de frecvențe în care $n = \sum_{i=1}^{m} f_i$ este:

$$Lmef = \frac{\left(\sum_{i=1}^{m} f_i\right)+1}{2}$$

Intervalul median va fi considerat acela în care frecvențele cumulate depăşesc locul medianei în serie.

Formula de calcul a medianei într-o serie de distribuție de frecvențe este:

$$Me = x_0 + d\frac{Lmef - \sum f_p}{f_m}$$

unde:

- Me este **mediana**;
- x_0 este limita inferioară a intervalului median;
- d reprezintă mărimea intervalului median;
- Lmef reprezintă locul medianei;
- f_m este frecvența intervalului median;
- f_p este frecvența intervalului precedent celui median.

Pentru o bună înțelegere a semnificației şi calculului medianei vom considera următorul exemplu care prezintă rezultatele în puncte obținute de 9 studenți la un test.

Nume studenți	**Puncte obținute**
Student 1	230
Student 2	310
Student 3	250
Student 4	310
Student 5	150
Student 6	180
Student 7	80
Student 8	350
Student 9	220

Se doreşte organizarea datelor într-un alt tabel care să evidențieze frecvențele pentru următoarele intervale de puncte obținute de studenți: < 100, între 100 şi 200, între 200 şi 300, > 300. Pentru aceasta vom realiza următorul tabel în care vor fi înscrise şi frecvențele cumulate:

Grupe de puncte obținute	**Număr de studenți (f_i)**	**Frecvențe cumulate**
< 100	1	1
100 - 200	2	3
200 - 300	3	6
> 300	3	9
Total	$\sum f_i = 9$	

Locul medianei (LMef) va fi:

Lmef = $\frac{\left(\sum_{i=1}^{m} f_i\right)+1}{2}$ = (9 + 1) / 2 = 5 rezultă că intervalul medianei este 200 – 300 deoarece aici frecvențele cumulate depăşesc locul medianei în serie. Mediana (Me) va fi cuprinsă între 200 şi 300 adică: 200 < Me < 300. Deci limita inferioară a intervalului median (x_0) este 200. Mărimea intervalului median (d) este 100. Din tabel se observă că suma cumulată a frecvențelor precedente intervalului median ($\sum f_p$) este 3. De asemenea frecvența intervalului median (f_m) este tot 3. Conform formulei de calcul a medianei avem:

$$Me = x_0 + d\frac{Lmef - \sum f_p}{f_m} = 200 + 100\,\frac{5 - 3}{3} = 266 \text{ puncte}$$

Modulul

Modulul (Mo) reprezintă valoarea termenului dintr-o serie care are frecvența maximă.

Revenind la exemplul anterior în care se prezintă rezultatele în puncte obținute de 9 studenți la un test observăm că doi studenți (studentul 2 şi studentul 4) au obținut 310 puncte. Deci frecvența maximă este 2 şi ea corespunde termenului 310, ceea ce înseamnă că Mo = 310 puncte. Pentru acest exemplu media aritmetică este 231,11. Dacă vom sorta tabelul după punctele obținute vom observa că mediana Me = 230:

Nume studenți	Puncte obținute
Student 7	80
Student 5	150
Student 6	180
Student 9	220
Student 1	230
Student 3	250
Student 2	310
Student 4	310
Student 8	350

Modulul poate fi calculat de asemenea şi pentru o serie de intervale de distribuție de frecvențe. În acest caz intervalul modal este intervalul corespondent intervalului corespondent celei mai mari frecvențe. Pentru analize se va folosi media aritmetică ponderată.

Deoarece într-o serie de distribuție de frecvențe pot exista mai multe intervale corespunzătoare frecvenței maxime rezultă deci că pot exista mai multe module şi mai multe intervale modale. În aceste cazuri seria este plurimodală.

Formula de calcul a modulului într-o serie de intervale este:

$$Mo = x_0 + d \frac{\Delta_1}{\Delta_1 + \Delta_2}$$

unde:

Mo este **modulul**;
x_0 reprezintă limita inferioară a intervalului modal;
d reprezintă mărimea intervalului modal;
$\Delta_1 = f_m - f_{m-1}$;
$\Delta_2 = f_m - f_{m+1}$;
f_m este frecvența intervalului modal;
f_{m-1} este frecvența intervalului precedent celui modal;
f_{m+1} este frecvența intervalului următor celui modal.

Pentru exemplificare vom considera că într-o companie există următoarea distribuție de frecvențe a veniturilor angajaților:

Venit ($)	Mijlocul de interval (x_i)	Număr angajați (f_i)
0 - 500	250	3
500 - 1000	750	5
1000 - 1500	1250	6
1500 - 2000	1750	4
2000 - 2500	2250	4
2500 - 3000	2750	2
3000 - 3500	3250	1

Se observă că frecvența cea mai mare (f_m) este 6 şi corespunzător acesteia intervalul modal este cel cuprins între 1000$ şi 1500$. Deci f_{m-1} = 5 iar f_{m+1} = 4. Limita inferioară a intervalului modal (x_0) este 1000$ iar d = 1500 – 1000 = 500.

$$Mo = 1000 + 500 \frac{6 - 5}{(6-5) + (6-4)} = 1167\$$$

Pentru analize se recomandă şi calculul mediei aritmetice ponderate:
(m = numărul de intervale = 7)

$$\overline{Xp} = \frac{\sum_{i=1}^{m} x_i f_i}{\sum_{i=1}^{m} f_i} = (250x3 + 750x5 + \ldots + 3250x1) / 2 = 36750 / 25 = 1470\$$$

Deoarece valoarea modulului este mai mică decât valoarea mediei rezultă concluzia că în exemplul prezentat frecvențele termenilor mai mici sunt mai numeroase decât frecvențele termenilor mai mari.

În cazul unei distribuții perfect simetrice media, mediana şi modulul au valori egale.

Mediana poate fi utilizată pe scale ordinale şi pe scale de interval (*scale* în SPSS).

Modulul poate fi utilizat pentru orice tip de scală dar este singurul indicator pentru scala nominală.

Cuartile, decile, percentile

Pentru seriile cu asimetrie mare şi care au o amplitudine mare a variației se calculează şi alți indicatori de poziție cum sunt: quartilele, decilele, centilele şi percentilele.

Quartilele sunt acele valori ale termenilor care separă seria în patru părți egale. Deosebim astfel quartila inferioară (Q1) care delimitează sfertul inferior (25%) al termenilor, quartila a doua (medie) (Q2) care este egală cu mediana deoarece împarte termenii în două părți egale (50%) şi quartila superioară (Q3) care delimitează sfertul superior (75%). În mod similar decilele împart seria în 10 părți egale iar centilele în 100 de părți egale. Rezultă că vor exista 9 decile şi 99 centile.

Pentru exemplificare vom considera următorul tabel care prezintă punctele obținute de studenți la un test, transformate în note:

Nume studenți	**Puncte obținute**	**Nota**
Student 1	230	8
Student 2	310	10
Student 3	250	9
Student 4	310	10
Student 5	150	7
Student 6	180	7
Student 7	80	6
Student 8	350	10
Student 9	220	8

Pe baza acestui tabel vom realiza tabelul sintetic al frecvențelor:

Nota	**Frecvența**			
	Absolută		**Relativă**	
	absolută (fi)	**cumulată (fc)**	**relativă (fr=fi/9)**	**cumulată (frc)**
6	1	1	0,11	0.11
7	2	3	0,22	0.33
8	2	5	0,22	0,55
9	1	6	0,11	0,66
10	3	9	0,34	1
Total	**9**		**1**	

Frecvența relativă cumulată procentuală (adică înmulțită cu 100) se numeşte **rang percentil**. **Percentilele** reprezintă valoarea dintr-o distribuție care corespunde unui anumit rang percentil. Quartilele sunt percentilele corespunzătoare rangurilor percentile 25%, 50% şi 75%.

Astfel în exemplul prezentat pentru nota 8 corespunde rangul percentil 55% ceea ce înseamnă că 55% dintre studenți au obținut la test o notă mai mică sau egală cu 8. Putem spune şi că rangului percentil 55% îi corespunde percentila 8.

Indicatorii variației şi asimetriei

Indicatorii simpli ai variației

Amplitudinea absolută a variației (R de la ***R****ange*) se calculează ca diferență între nivelul maxim (x_{max}) şi nivelul minim (x_{min}) al caracteristicii:

$R = x_{max} - x_{min}$

Amplitudinea relativă a variației (R%) se calculează ca raport între amplitudinea absolută a variației şi nivelul mediu al variației:

$R\% = 100 (R / \overline{X})$

Abaterile individuale absolute (d de la ***d****eviation*) se calculează ca diferențe între fiecare variantă înregistrată şi media aritmetică a acestora.

$d_i = x_i - \overline{X}$ Este evident că întotdeauna suma abaterilor absolute este nulă: $\sum(x_i - \overline{X}) = 0$. În procedurile statistice se recomandă calcularea acestei sume în modul: $\sum| x_i - \overline{X} |$.

Abaterile individuale relative (d%) se calculează ca raport între abaterile absolute la nivelul mediu al caracteristicii:

$d_i \% = 100 (d_i / x_i)$

Abaterea quartilă (RQ) se calculează ca diferență între quartila 3 şi quartila 1:

$RQ = Q3 – Q1$

Abaterea semi-interquartilă are următoarea expresie:

$RSQ = (Q3 – Q1) / 2$

Indicatorii sintetici ai variației

Indicatorii sintetici ai variației (împrăştierii) sunt: abaterea medie liniară, abaterea medie pătratică, dispersia şi coeficientul de variație.

Abaterea medie liniară (dm) se calculează ca o medie aritmetică a termenilor seriei de la media lor.

$dm = (\sum| x_i - \overline{X} |) / n$

Abaterea standard (s) se calculează ca o medie pătratică din abaterile tuturor variantelor seriei de la media lor aritmetică:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n}}$$

Coeficientul de variație (cv) se calculează ca raport între abaterea medie pătratică şi nivelul mediu al seriei:

$cv = s / \overline{X}$

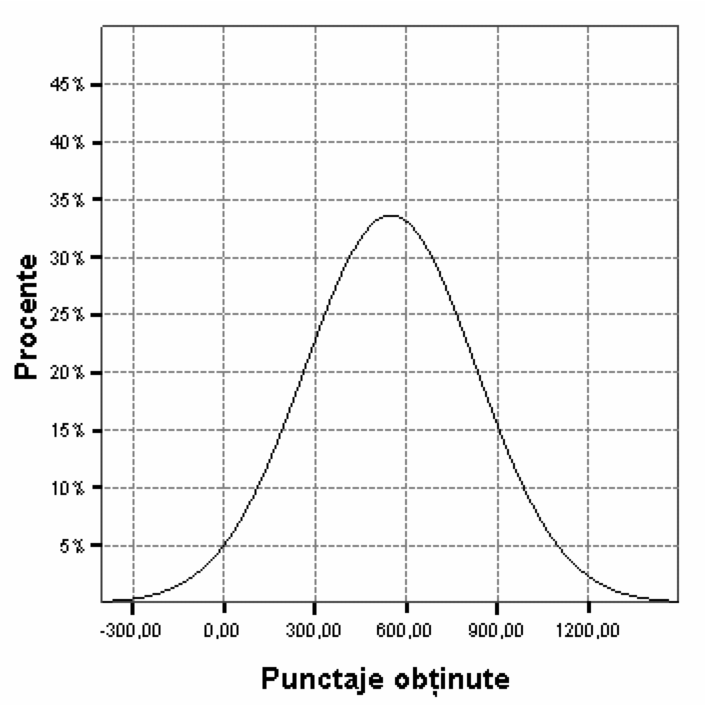
Dispersia (varianța s^2) se calculează ca o medie aritmetică a pătratelor abaterilor termenilor față de media lor

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n}$$

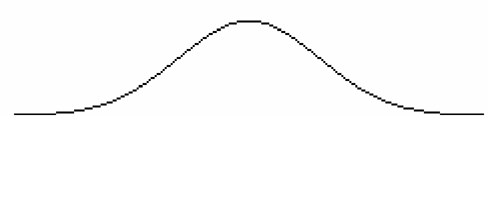
S-a demonstrat că valorile pentru abaterea standard şi pentru dispersie calculate după formulele de mai sus pentru un eşantion conțin o imprecizie (*bias*). Cu cât eşantionul este mai mic cu atât dispersia va fi influențată mai mult de valoarea de la numitor. Se poate introduce astfel o corecție n – 1 (la numitor). Adică în ambele formule în loc de n putem avea n – 1. Valoarea n – 1 de la numitor este denumită **numărul gradelor de libertate**.

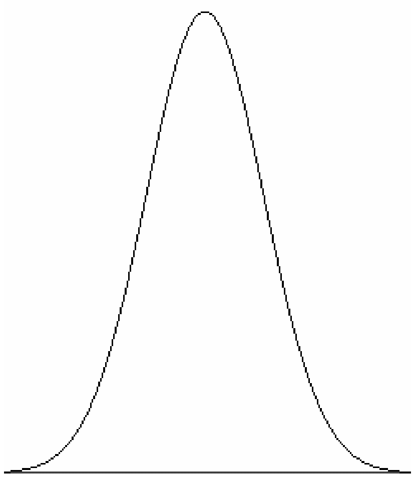
Indicatori ai formei distribuției

Forma grafică a unei distribuții normale poate avea față de mediana (Me) o abatere de simetrie pe orizontală orientată către stânga (simetrie negativă) sau către dreapta acesteia (simetrie pozitivă). Atunci când pentru o distribuție Me = Mo = $\overline{X}$ spunem despre aceasta că este perfect simetrică. Următorul grafic ilustrează forma acestei distribuții în care: Me = Mo = $\overline{X}$ = 550

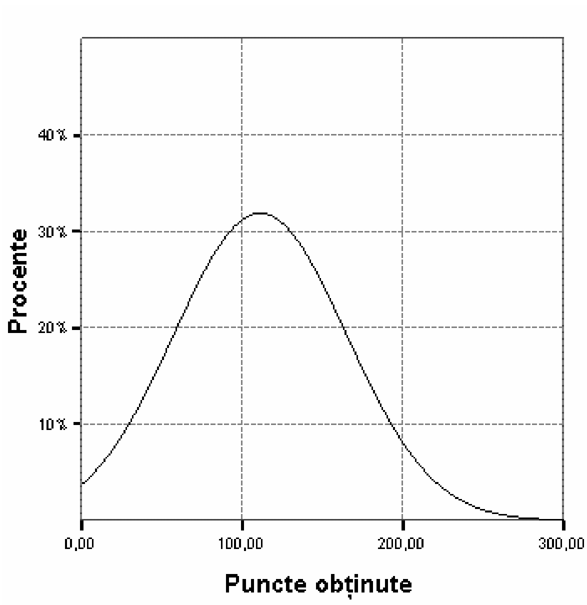


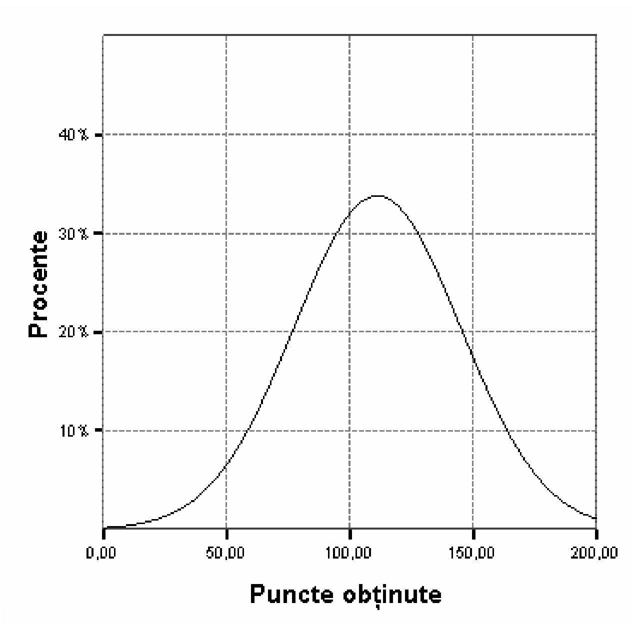
În funcție de distribuție pot fi obținute şi altfel de grafice ale curbei, prea aplatizate sau prea înalte. Pentru a corecta aceste cazuri se utilizează indicele de boltire (*kurtosis*).





De asemenea, aşa cum am specificat anterior putem întâlni două forme de grafice cu simetrie negativă sau cu simetrie pozitivă:





Pentru corecta simetria negativă sau pozitivă se utilizează indicele de simetrie sau de oblicitate (*skewness*).

Indicele de boltire (*kurtosis*) se calculează în mod similar cu abaterea standard dar prin ridicarea la puterea a patra. El variază în jurul valorii 0. Indicele de boltire cu valori pozitive indică o curbă înaltă iar cel cu valori negative o curbă aplatizată.

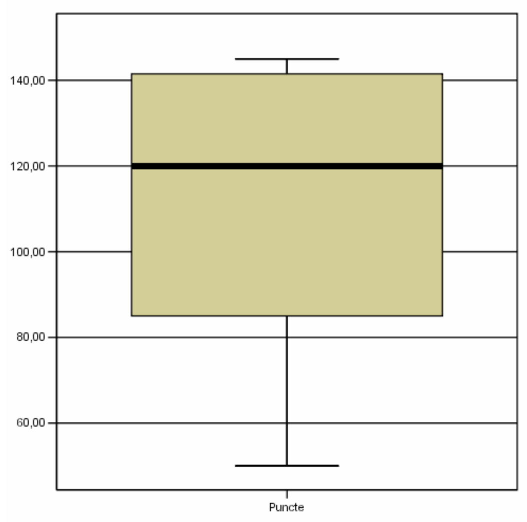
Indicele de simetrie sau de oblicitate (*skewness*) se calculează în mod similar cu abaterea standard dar prin ridicarea la puterea a treia. Pentru valoarea 0 el indică o simetrie perfectă. Valorile pozitive sau negative indică o asimetrie stânga / dreapta mai mult sau mai puțin pronunțată.

Valorile extreme ale unei distribuții pot fi foarte bine evidențiate prin intermediul unui grafic numit *Box and Whisker Plot* sau pe scurt *Box Plot*, creat de Tukey. **Graficul *Box Plot*** se prezintă sub forma unui dreptunghi care are baza în dreptul quartilei 25 şi latura superioară în dreptul quartilei 75. Deci înălțimea acestui dreptunghi cuprinde 50% din valorile distribuției. Mediana este reprezentată prin intermediul unei linii ce apare în interiorul dreptunghiului.

Pentru exemplificare vom considera următoarea distribuție care conține punctele obținute la un test:

50, 60, 70, 80, 90, 100, 110, 120, 130, *140, 141, 142, 143, 144, 145*

Se observă (evidențiat cu caractere înclinate) aglomerarea mai multor valori în partea superioară a distribuției (în intervalul 140 – 145). Iată în continuare graficul *Box Plot* corespunzător acestei distribuții, realizat cu SPSS:



Graficul ***Steam and Leaf Plot*** (tulpină şi frunză) creat tot de către Tukey prezintă de asemenea într-o formă sugestivă explorarea unei distribuţii scoţând în evidenţă frecvenţele grupate. Iată în continuare graficul *Steam and Leaf Plot* corespunzător aceleiaşi distribuţii şi realizat cu SPSS:

```
Puncte Stem-and-Leaf Plot

 Frequency    Stem &  Leaf

      ,00        0 .
     5,00        0 .  56789
    10,00        1 .  0123444444

 Stem width:    100,00
 Each leaf:       1 case(s)
```

Se observă că sunt evidenţiate toate cele 15 valori ale distribuţiei dintre care 5 se află sub 100 ele fiind 50, 60, 70, 80, 90 având corespondentă în grafic linia a doua: (0 . 56789). De asemenea 10 valori sunt mai mari sau egale cu 100, dintre care 6 aparţin domeniului 140. Aceste valori sunt evidenţiate în linia a treia: (1 . 0123444444).

Scoruri standard, curba normală, ipoteze statistice, coeficientul de corelație

De multe ori în analiza statistică limitarea doar la valorile unei caracteristici, la indicatorii tendinței centrale sau la indicatorii variației nu oferă răspunsuri satisfăcătoare pentru cercetător. De exemplu dacă vom cunoaşte doar nota obținută (sau numărul de puncte obținute) de un student la un test nu putem trage prea multe concluzii deoarece nu cunoaştem forma distribuției pe care se plasează în raport cu ceilalți studenți participanți la test (mai aplatizată sau nu, deplasată spre dreapta sau spre stânga), gradul de împrăştiere a notelor şi abaterea de la medie. De asemenea nu putem răspunde unor întrebări (ipoteze) mai nuanțate cum ar fi: există o legătură între performanța realizată de studenții care au urmat anterior un curs pregătitor şi cei care nu au urmat un astfel de curs? A îmbunătățit acest curs performanța studenților? Performanța realizată este dependentă de vârstă, sex sau categoria socială căreia îi aparțin studenții? Performanța realizată de studenți la test este superioară sau nu performanței medii realizată pe parcursul mai multor ani de către studenții din universitate care au susținut acest test? Şirul acestor întrebări ar putea continua cu aplicații şi din alte domenii de activitate: există o legătură între numărul voturilor pentru un anumit candidat şi vârsta votanților? Există o relație între atitudinea angajaților dintr-o companie față de o anumită decizie şi salariile acestora?

Pentru a putea da răspunsuri pertinente unor astfel de întrebări se utilizează scorurile (notele) standard, ipotezele şi testele statistice.

Scorurile standard şi curba normală

Scorurile z (scoruri standardizate) se calculează pe baza valorilor caracteristicii, mediei şi a abaterii standard:

$$z_i = (x_i - \overline{X}) / s$$

unde:
- z_i este **scorul standardizat** corespunzător termenului i;
- x_i este valoarea termenului i al seriei;
- $\overline{X}$ este media;
- s reprezintă abaterea standard.

În continuare va fi prezentat un exemplu de transformare a punctelor obținute la un test, în scoruri z. Media şi abaterea standard sunt 231,11 respectiv 85,94.

Nume studenți	**Puncte obținute**	**Scorul (nota) z**
Student 1	230	-0,012
Student 2	310	+0,917
Student 3	250	+0,219
Student 4	310	+0,917
Student 5	150	-0,943
Student 6	180	-0,594
Student 7	80	-1,758
Student 8	350	+1,383
Student 9	220	-0,129
$\overline{X}$ =231,11; s=85,94;		Suma = 0

Convertirea valorilor unei distribuții în scoruri z nu modifică forma acesteia. Suma scorurilor într-o distribuție z este întotdeauna 0.

Scorurile t (*Thrustone*) se calculează pe baza scorurilor z după următoarea relație:

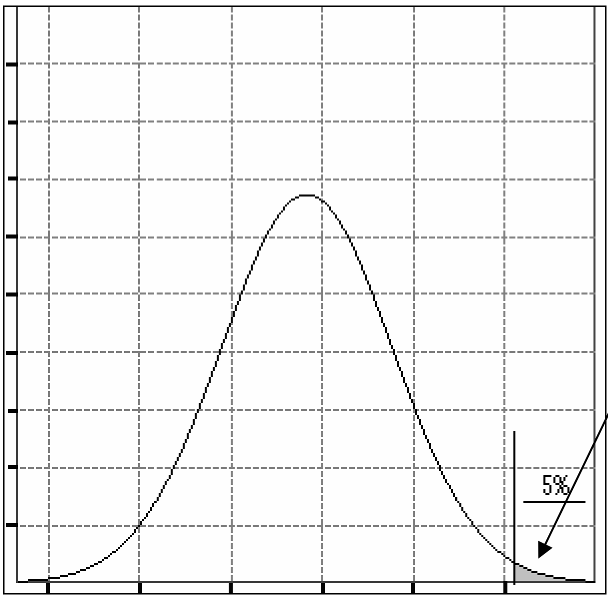
$t_i = 50 + 10 z_i$

SAT (*Scholastic Assessement Test*) se calculează similar tot pe baza notelor z:

$SAT_i = 500 + 100 z_i$

Cel mai frecvent se utilizează scorurile z şi t. Similar se calculează pentru eşantioane scorurile z şi t. În acest caz x_i se înlocuieşte cu media eşantionului, $\overline{X}$ cu media populației din care s-a extras eşantionul iar s cu eroarea standard a mediei de eşantionare. Corespunzător scorurilor (z şi t) există testele z şi t cu ajutorul cărora pot fi confirmate sau infirmate ipotezele. Pentru a lucra cu aceste teste întotdeauna trebuie să existe **ipoteza cercetării** (de exemplu: parcurgerea unui curs anterior influențează performanțele studenților la test) şi o **ipoteză de nul** care infirmă ipoteza cercetării (de exemplu: parcurgerea unui curs anterior nu influențează performanțele studenților la test).

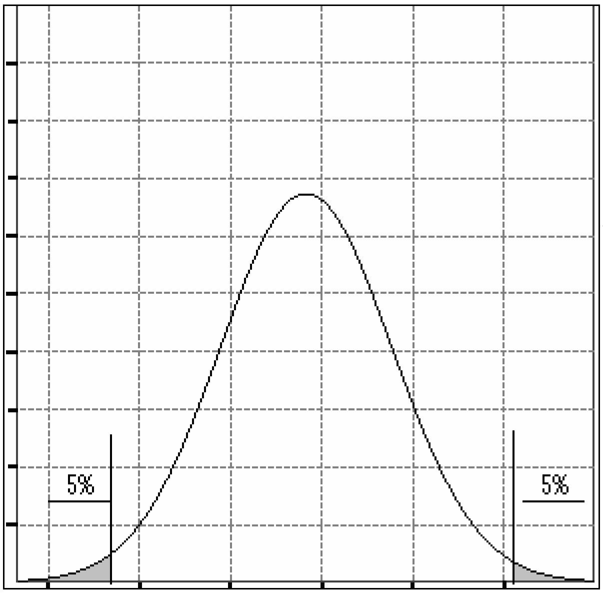
Pentru fiecare scor z sau t se calculează o probabilitate p corespunzătoare (ce poate fi găsită în tabelele z sau t). Deoarece nu există o garanție 100% pentru acceptarea sau respingerea unei ipoteze, atunci trebuie să fie asumat un risc de eroare. Nivelul convențional minim acceptat pentru acesta este de 5%. Nivelul de risc poate fi exprimat şi prin opusul lui (95% aşa cum lucrează SPSS). În acest caz 95% nu reprezintă eroarea maximă acceptată ci nivelul minim de încredere în rezultatul experimentului. O explicitare grafică a erorii maxim acceptate este:



Dacă scorul eşantionului se află în zona haşurată de 5% atunci ipoteza de nul se respinge şi se acceptă ipoteza cercetării. Cu alte cuvinte *dacă $p < 0,05$ atunci se acceptă ipoteza cercetării.*

Evident problema se judecă similar şi pentru zona de 5% din stânga graficului. Acesta este un test statistic unilateral (***one tailed***).

Pentru a verifica ipoteza cercetării simultan pe ambele laturi (stânga – dreapta) ale distribuției se aplică testul z sau t bilateral (***two tailed***):



Dacă scorul eşantionului se află în oricare dintre zonele haşurate de 5% atunci ipoteza de nul se respinge şi se acceptă ipoteza cercetării. Cu alte cuvinte *dacă p < 0,05 atunci se acceptă ipoteza cercetării.*

Se recomandă utilizarea testului z pentru un eşantion de minim 30. Atunci când volumul eşantionului este < 30 sau chiar > 30 se poate utiliza testul t.

Distribuția t mai este denumită şi distribuția Student. Ea are toate caracteristicile unei distribuții normale, are aceeaşi formă de clopot dar care de această dată depinde de numărul gradelor de libertate df (***d**egrees of **f**reedom*). Cu cât df este mai mic cu atât forma curbei devine mai aplatizată. Pentru un test statistic t aplicat unui sigur eşantion expresia de calcul a df este:

$df = N - 1$

unde:
df reprezintă numărul gradelor de libertate;
N este volumul eşantionului.

Dacă testul t se aplică pentru două eşantioane primul de volum N1 iar al doilea de volum N2 atunci:

$df = N1 + N2 - 2$

Pentru a putea aplica testul t trebuie ca cele două eşantioane să fie omogene. Se recomandă ca volumul cele două eşantioane să fie egal (N1 = N2) iar dispersiile acestora să fie apropiate. În acest sens se poate aplica un **test de omogenitate *Levene's Test*** care verifică valorile dispersiilor şi calculează un indicator F numit **raportul Fisher**. Dacă valorile F sunt mari atunci chiar dacă probabilitatea p calculată este mai mică decât 0,05 aceasta indică faptul că dispersiile sunt heterogene. Rezultă astfel că ipoteza nu poate fi acceptată. SPSS oferă şi o altă posibilitate de calcul a valorilor t şi pentru dispersii heterogene (*Unequal variances*). Această alternativă asigură o acuratețe mai mare a rezultatelor şi ajută cercetătorul să poată discerne cu siguranță dacă acceptă sau respinge o ipoteză.

Trebuie reținut că în toate testele statistice criteriul de bază pentru semnificația statistică ***Sig.*** (*Significance*) este probabilitatea (p) calculată pentru testul bilateral (***two tailed*** sau ***2-tailed significance***).

Atunci când se lucrează cu mai mult de două eşantioane odată se poate utiliza testul ANOVA (***AN**alisys **Of** **VA**raiance*).

Coeficientul de corelație Pearson

Coeficientul de corelație Pearson se va utiliza atunci când se doreşte măsurarea valorilor a două variabile din acelaşi eşantion pentru a se afla dacă între acestea există o relație şi care este intensitatea relației.

Dacă relația există vom deosebi două feluri de corelație: pozitivă şi negativă.

Corelația este **pozitivă** atunci când creşterea valorilor unei variabile determină creşterea valorilor celeilalte variabile.

Corelația negativă apare atunci când creşterea valorilor unei variabile determină scăderea valorilor pentru a doua variabilă.

Felul corelației se exprimă prin semnul coeficientului de corelație Pearson (r) iar **intensitatea legăturii** dintre cele două variabile se exprimă prin valoarea acestuia. Cu alte cuvinte atunci când avem o valoare pozitivă a lui r spunem că între variabile există o corelație pozitivă şi invers. Cu cât valoarea lui r este mai mare cu atât legătura dintre variabile este mai puternică.

Expresia coeficientului de corelație este:

$$r = \frac{\sum_{i=1}^{N} z_x z_y}{N}$$

unde:

r	este coeficientul de corelație pentru variabilele x şi y;
z_x	este scorul z al variabilei x;
z_y	este scorul z al variabilei y;
N	reprezintă este volumul eşantionului.

Este de reținut faptul că valorile lui r pot varia doar în intervalul [-1, +1].

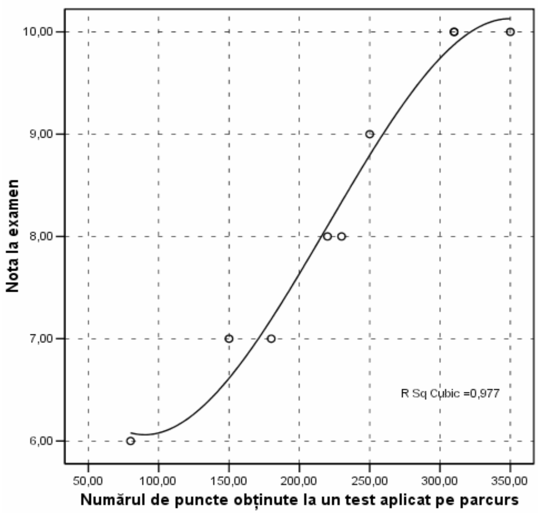
Numărul gradelor de libertate df = N – 2. Atunci când se analizează corelația se va alege şi un nivel de risc acceptat (în mod convențional 5%) sau un nivel minim de încredere (în mod convențional 95%).

Valorile lui r se exprimă pe o scală ordinală. Pentru a putea compara doi coeficienți de corelație aceştia se ridică la pătrat. $\mathbf{r^2}$ se numeşte **coeficientul de determinare**.

O reprezentare grafică interesantă a corelației poate fi obținută în SPSS prin intermediul unui grafic de tip *Scatter – Plot*.

Iată în continuare un exemplu elocvent: se doreşte să se afle dacă rezultatele obținute de 9 studenți la un test aplicat pe parcurs influențează notele lor finale la examen.

Număr de puncte obținute la test (variabila x)	Nota la examen (variabila y)
230,00	8
310,00	10
250,00	9
310,00	10
150,00	7
180,00	7
80,00	6
350,00	10
220,00	8



Coeficientul de corelație Pearson calculat pentru cele două variabile este r = 0,977. Valoarea acestuia arată că între cele două variabile există o legătură puternică şi o corelație pozitivă. Graficul *Scatter* redă foarte sugestiv aceste informații.

În concluzie se poate afirma că studenții care au obținut un număr mare de puncte la test au obținut note mari la examenul final.

Chiar dacă în analiza statistică cel mai utilizat este coeficientul de corelație Pearson totuşi se mai poate lucra şi cu alți coeficienți de corelație cum sunt Kendall şi Spearman. Semnificația acestora este similară cu cea a coeficientului de corelație Pearson. Iată pentru exemplul prezentat valorile celor doi coeficienți calculați de către SPSS pentru ***2-tailed significance***:

Kendall's tau_b Correlation Coefficient Sig. (2-tailed) = 0,941
Spearman's rho Correlation Coefficient Sig. (2-tailed) = 0,979

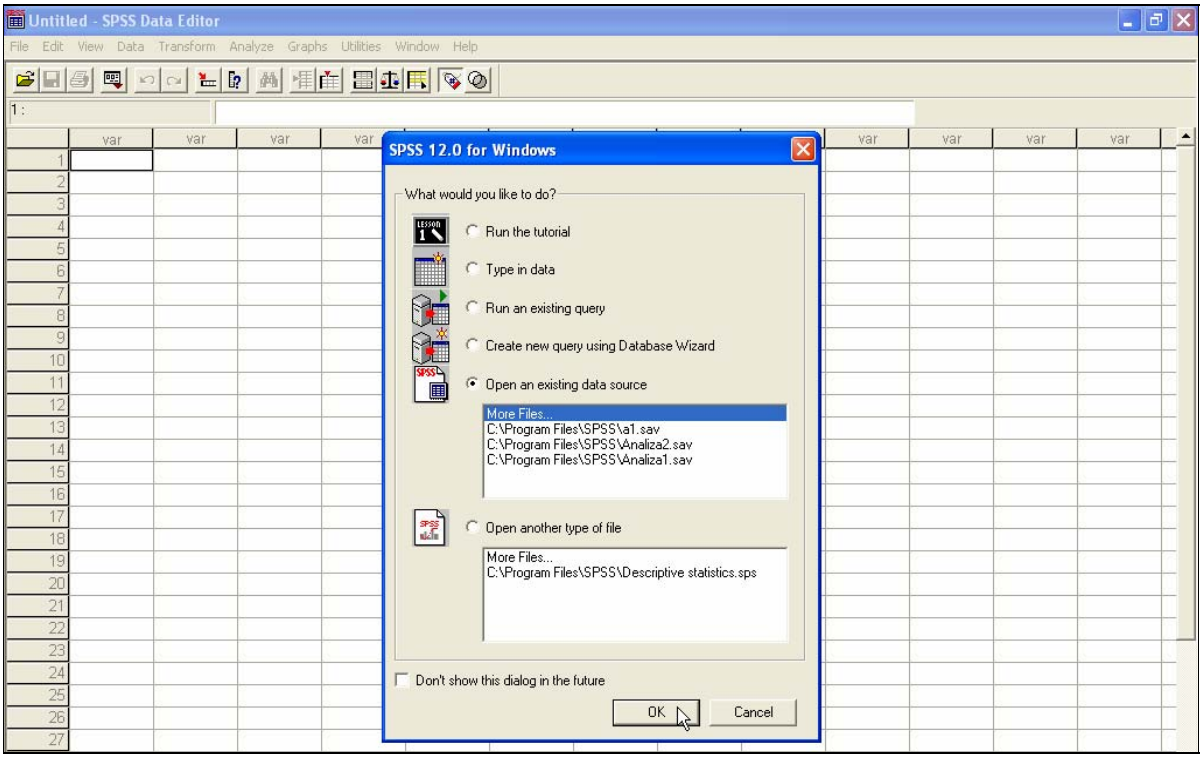
Prezentarea în detaliu a acestora nu face obiectul acestui curs. Informații teoretice suplimentare atât despre testele statistice cât şi despre corelații pot fi găsite în cursurile avansate de statistică.

Mai multe informații despre utilizarea practică a testelor statistice şi a coeficientului de corelație vor fi date în următorul capitol în paralel cu prezentarea programului SPSS şi cu aplicațiile acestuia redate prin intermediul unor studii de caz.

Programul SPSS. Analize şi aplicaţii

SPSS este în prezent unul dintre cele mai populare programe utilizate în analiza statistică. Versiunea 12 a acestui produs software oferă o puternică interfaţă interactivă şi include foarte multe facilităţi de achiziţie a datelor, de procesare şi analiză statistică, de prezentare a rezultatelor obţinute.

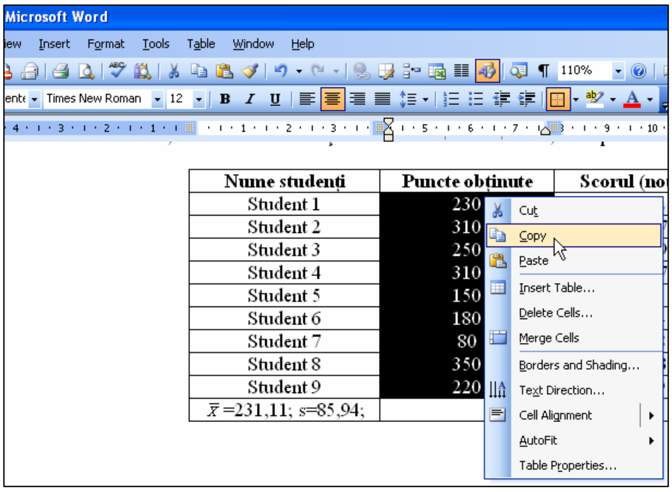
După instalarea versiunii Windows, produsul plasează un *shortcut* în *desktop* şi creează un submeniu corespunzător *SPSS for Windows* în cadrul meniului *Start* al acestui sistem de operare. La lansarea în execuţie (de exemplu printr-un clic dublu dat pe *shortcut*-ul din *Desktop*) SPSS va prezenta fereastra corespunzătoare bazei de date şi o casetă de dialog prin intermediul căreia vor putea fi realizate mai multe acţiuni cum sunt: afişarea unui tutorial, posibilitatea de încărcare a datelor prin tastare, încărcarea unui set de date care există creat anterior în urma unei interogări (*query*) într-o altă bază de date, crearea unei interogări cu ajutorul unui „vrăjitor” (asistent software), deschiderea unui fişier (cu extensia .sav) în care au fost salvate date în format SPSS.

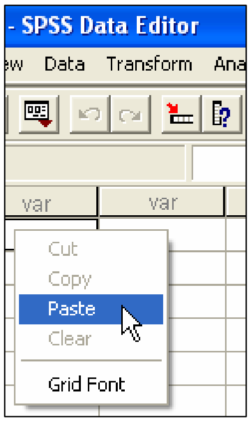


Dacă se va selecta a doua opţiune (*Type in data*) şi se va apăsa apoi butonul **Ok** sau dacă se va răspunde cu ***Cancel*** acestui dialog se va intra în interfaţa de lucru cu baza de date (vizibilă în fundal – *SPSS Data Editor*).

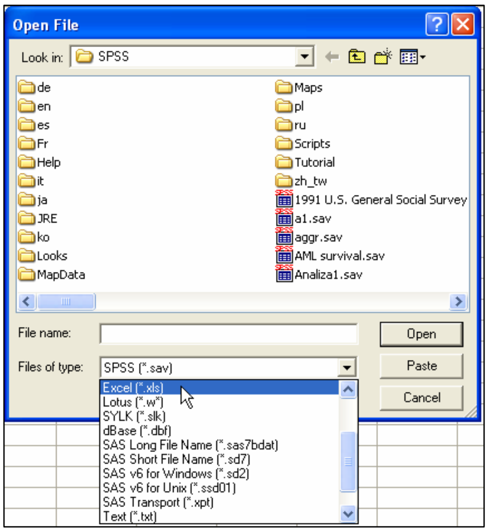
SPSS a fost proiectat să funcţioneze ca un produs software de tip deschis care să permită cu uşurinţă schimbul de date cu alte baze de date sau produse software. Astfel prin intermediul meniului ***File*** este posibil oricând un import / export de date, crearea unor date noi sau salvarea datelor existente. Aceste operaţiuni pot fi realizate cu ajutorul opţiunilor *New, Open, Open Database, Read Text Data, Save, Save As …* din cadrul meniului ***File***.

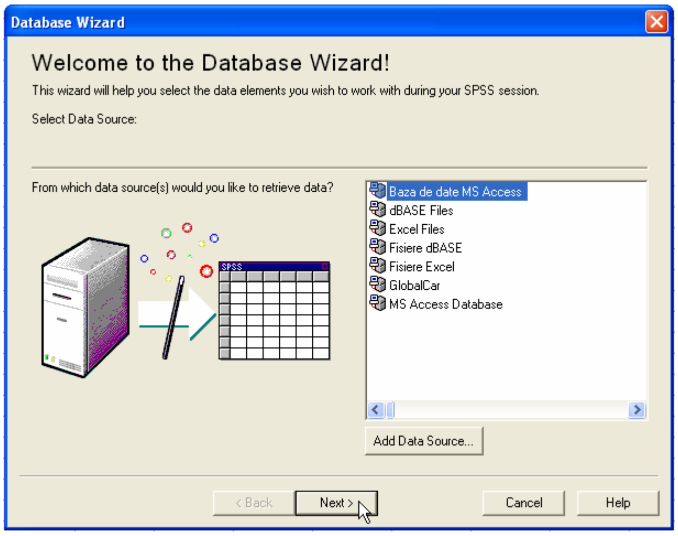
Schimbul de date sau obiecte (tabele, grafice etc) între SPSS şi alte programe se poate realiza foarte elegant prin intermediul mecanismului *Clipboard* al sistemului de operare (*Copy, Paste, Cut*).





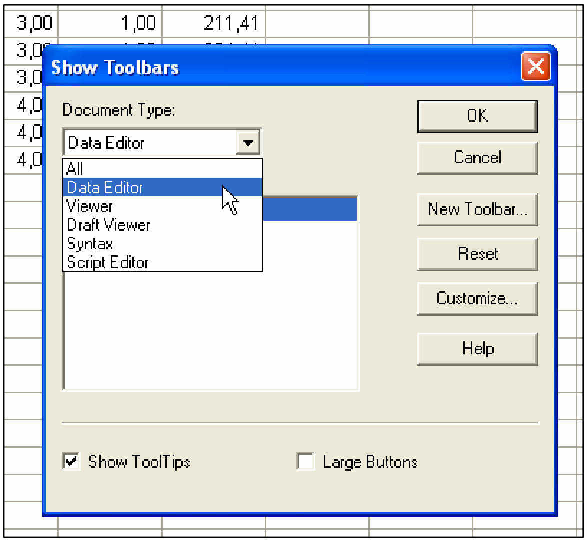
SPSS poate încărca direct din hard disc date din fişiere care se află în format text, Excel, Lotus, dBase sau se poate conecta prin ODBC la alte baze de date cum sunt: MS Access, MS SQL, Paradox, Postgres, Oracle etc pentru a încărca date din tabelele acestora sau date rezultate în urma unor interogări.

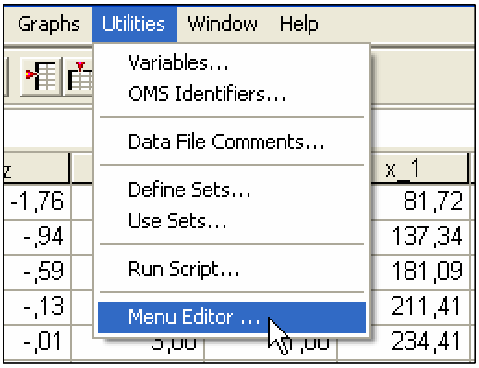




SPSS este un instrument flexibil care permite lucrul intuitiv cu meniuri şi butoane ce pot fi configurate de către utilizatori conform preferinţelor acestora. Iată modul în care pot fi configurate barele de butoane (*Toolbars*) şi meniurile produsului:

- din opţiunea *Toolbars* a meniului *View* poate fi lansată o casetă de dialog care permite configurarea barelor de butoane iar prin intermediul meniului *Utilities* poate fi lansat editorul de meniuri:

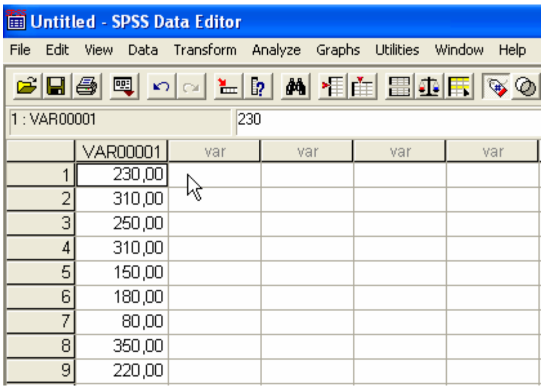




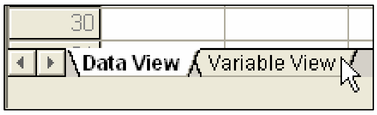
Încărcarea, editarea şi transformarea datelor. Componenta *Data Editor*

În SPSS tabela bazei de date este împărţită în linii şi coloane. Liniile reprezintă înregistrările acesteia sau cazurile (*cases)* cum sunt denumite în SPSS iar coloanele (câmpurile) reprezintă variabilele distribuţiei. În celulele de intersecţie se află valorile corespunzătoare variabilelor pentru fiecare caz (*case*).

În continuare vom presupune că am încărcat în *Data Editor*, prin tastare, următorul set de date care conţine punctele obţinute de 9 studenţi în urma aplicării unui test. Acest set de date a mai fost utilizat pentru exemplificări şi în capitolul anterior. În fereastra editorului vom avea:

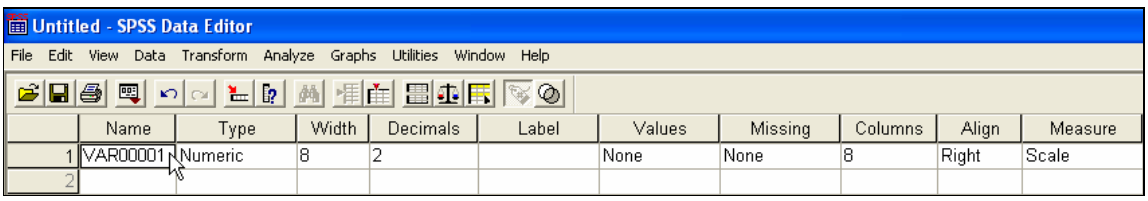


Observăm că SPSS a denumit generic prima variabilă VAR00001. În subsolul acestei ferestre identificăm posibilitatea de a comuta în fereastra variabilelor (*Variable View*).

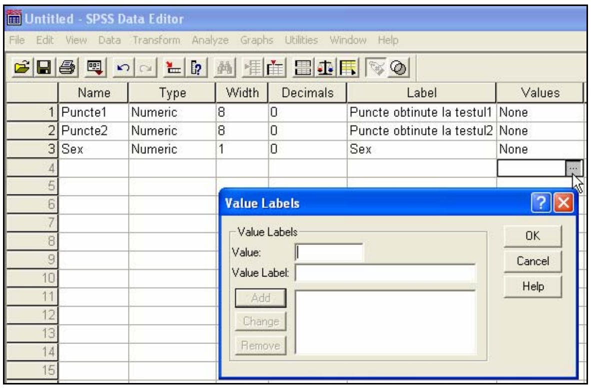


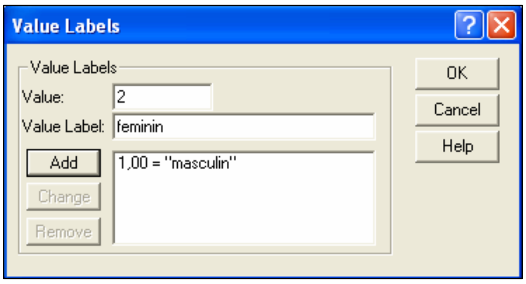
După un clic în *Variable View* va fi afişată fereastra variabilelor.

Observăm că există deja o formatare implicită pentru VAR00001. Aceasta este declarată ca fiind de tip numeric, cu o lungime de 8 caractere din care 2 zecimale ce se va afişa aliniată la dreapta şi care va fi măsurată pe o scală de tip *Scale* (raport / interval).

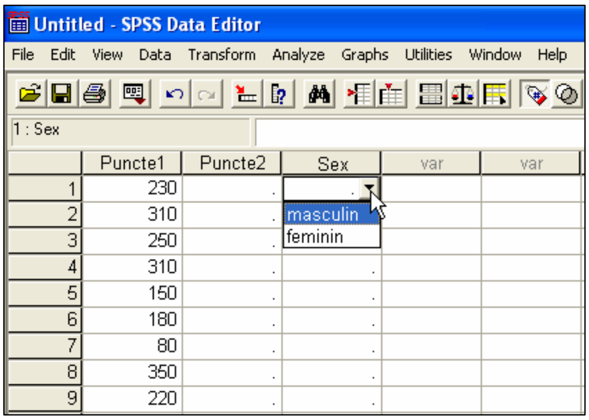


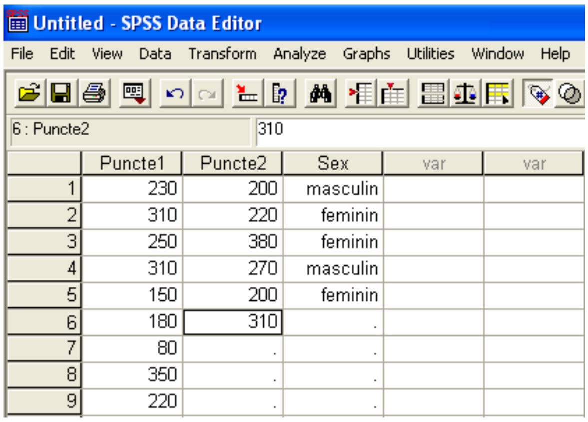
După ce vom selecta variabila (VAR00001) vom putea modifica numele acesteia prin tastare. Noul nume ales va fi **Puncte1**. Sub SPSS fiecare variabilă poate avea o etichetă care se va afişa în capul de tabel al datelor sau în rapoartele finale. Presupunem că am editat similar eticheta acestei variabile - **Puncte obţinute la testul1**. În acelaşi mod se va crea o nouă variabilă **Puncte2** cu eticheta **Puncte obţinute la testul2** şi o altă variabilă nominală **Sex** care va conţine doar două valori: 1 pentru masculin şi 2 pentru feminin. Pentru a se seta aceste valori s-a dat clic pe butonul **…** :





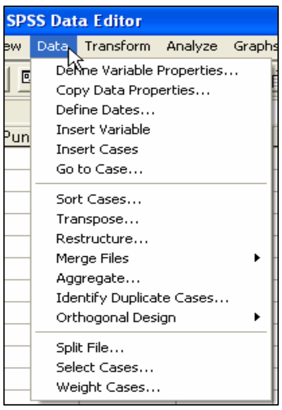
Toate aceste modificări se vor reflecta instantaneu în fereastra de date:





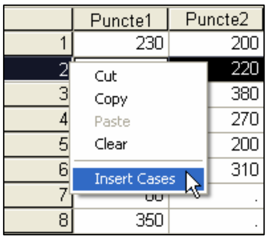
Este de remarcat faptul că SPSS marchează datele neîncărcate (sau lipsă) cu un punct. Acestea sunt numite *Missing Values*. Atunci când procesează datele SPSS va exclude cazurile care au valori lipsă sau în funcţie de model le va aproxima.

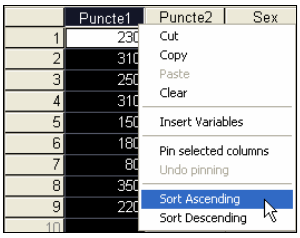
Dacă meniurile *Edit* şi *View* ne oferă funcţii comune şi utile pentru copierea, mutarea, ştergerea şi afişarea datelor şi variabilelor (*Copy, Paste, Cut, Clear, Fonts, Grid Lines* etc), nu acelaşi lucru se poate afirma despre **meniul *Data*** care merită o tratare specială.

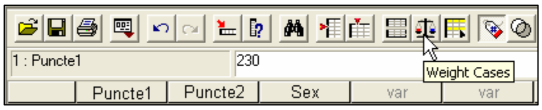


Prin intermediul acestui meniu pot fi definite proprietăţile variabilelor şi ale datelor, pot fi inserate cazuri şi variabile noi, seriile de date pot fi sortate, se pot adăuga date din alte fişiere externe (*Merge Files*), liniile pot fi transformate în coloane şi coloanele în linii (*Transpose*), pot fi identificate cazurile duplicate, se poate împărţi baza de date în subgrupuri (*Split File*), se poate lucra cu cazuri cu „greutăţi” specifice (*Weight Cases*) etc.

Unele dintre acțiunile enumerate anterior pot fi realizate atât în fereastra datelor cât şi în cea a variabilelor prin intermediul meniurilor contextuale sau a barei de unelte:



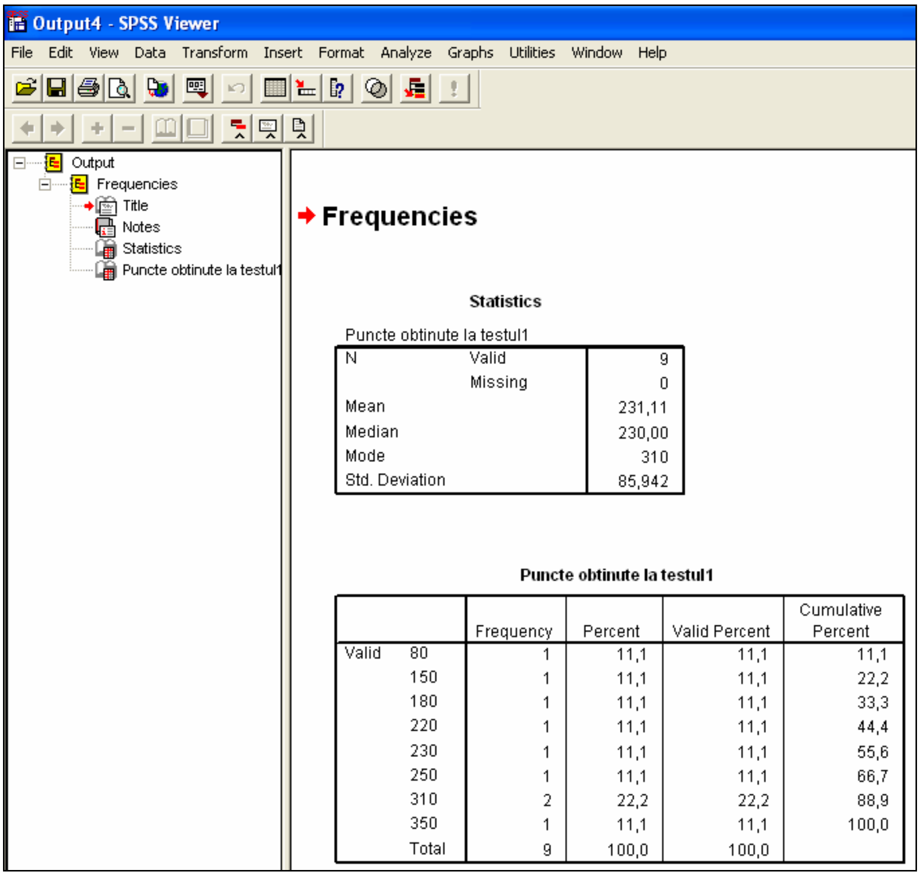




Deoarece nu întotdeauna semnificațiile ultimelor trei opțiuni din meniul *Data* (*Split File, Select Cases şi Weight Cases*) sunt înțelese corect de către utilizatorii SPSS începători, acestea vor fi prezentate şi explicate în continuare prin intermediul unor exemple referitoare la utilizarea lor.

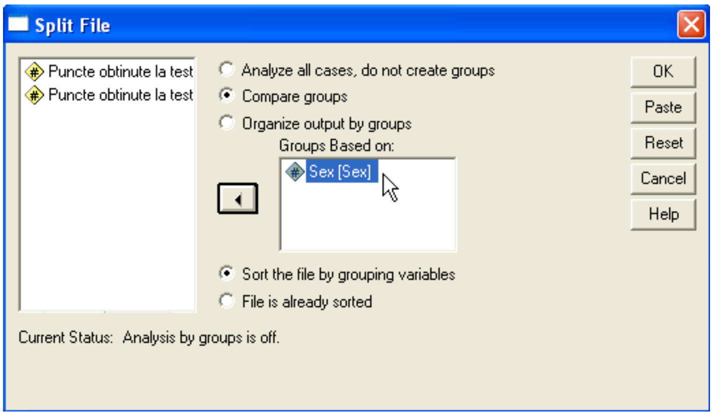
Dacă se doreşte împărțirea bazei de date în grupuri pentru o analiză detaliată în fereastra de raportare (afişare) a rezultatelor (*Output SPSS Viewer*), atunci va putea fi utilizată **opțiunea *Split File*** al cărui efect va fi vizibil doar în fereastra de afişare a rezultatelor (nu şi în *SPSS Data Editor*).

De exemplu dacă se doreşte afişarea mediei, medianei, modulului şi a abaterii standard pentru punctele obținute la testul 1, fereastra *Output SPSS Viewer* va arăta astfel:

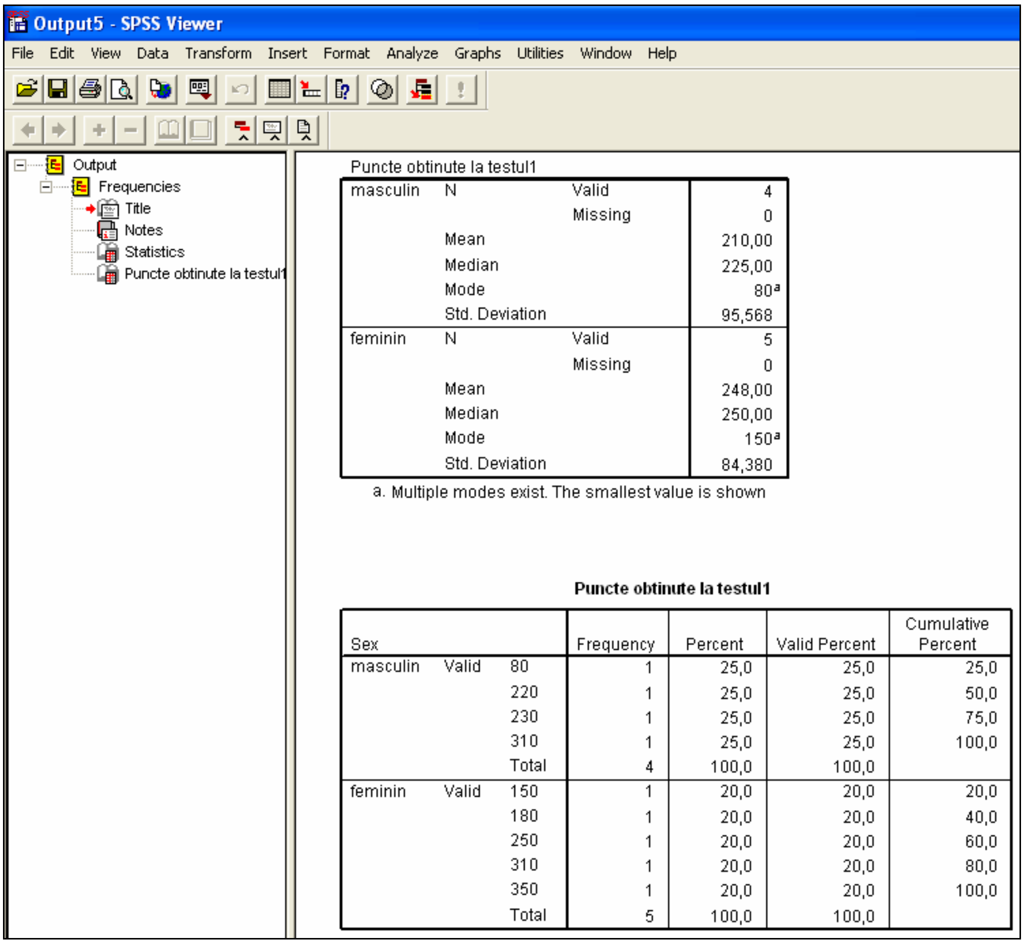


Notă
Felul în care a fost obținut acest raport (*Frecuencies*) va fi prezentat atunci când va fi descris meniul (*Analyze*).

Pentru a împărți datele în două grupuri conform sexului (masculin, feminin) în vederea unor analize separate vom utiliza opțiunea ***Split File*** şi vom specifica acest lucru în următorul dialog care va apărea:



În acest caz atunci când va fi afişat raportul solicitat, el va conține următoarele informații:



Puncte obtinute la testul1

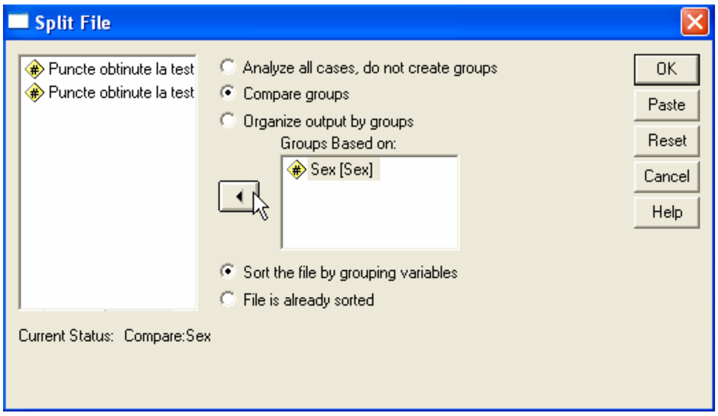
masculin	N	Valid	4
		Missing	0
	Mean		210,00
	Median		225,00
	Mode		80a
	Std. Deviation		95,568
feminin	N	Valid	5
		Missing	0
	Mean		248,00
	Median		250,00
	Mode		150a
	Std. Deviation		84,380

a. Multiple modes exist. The smallest value is shown

Puncte obtinute la testul1

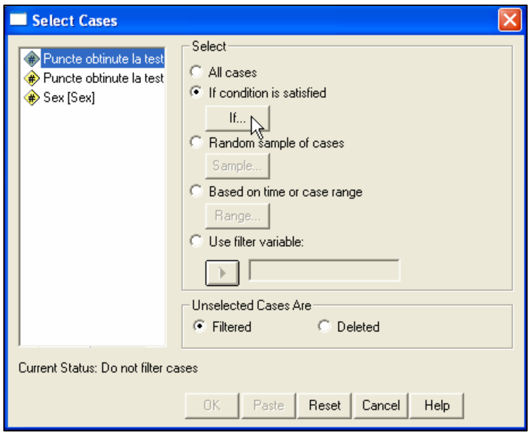
Sex			Frequency	Percent	Valid Percent	Cumulative Percent
masculin	Valid	80	1	25,0	25,0	25,0
		220	1	25,0	25,0	50,0
		230	1	25,0	25,0	75,0
		310	1	25,0	25,0	100,0
		Total	4	100,0	100,0	
feminin	Valid	150	1	20,0	20,0	20,0
		180	1	20,0	20,0	40,0
		250	1	20,0	20,0	60,0
		310	1	20,0	20,0	80,0
		350	1	20,0	20,0	100,0
		Total	5	100,0	100,0	

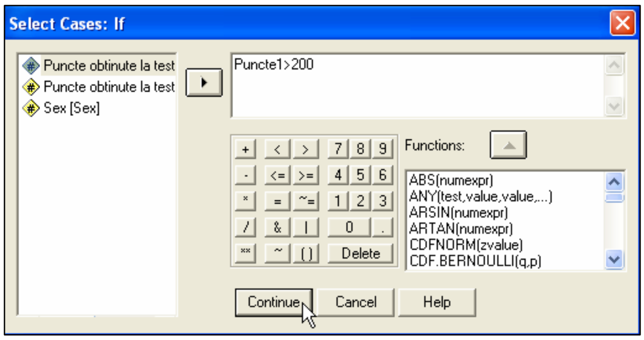
Dacă se doreşte anularea efectului *Split File* atunci se solicită din nou această opţiune din meniul *Data* şi se trece variabila de grupare (Sex) înapoi în subfereastra variabilelor (din partea stângă), după care se închide dialogul:



Selectarea unor cazuri pentru aplicarea procedurilor de analiză separat, doar pentru acestea se poate realiza prin intermediul opţiunii ***Select Cases*** din acelaşi meniu *Data*.

Pentru exemplificare vom presupune că dorim să selectăm pentru analize doar cazurile în care studenţii au obţinut la testul 1 un număr de puncte mai mare de 200 (Puncte1>200). În acest sens după solicitarea *Select Cases* se va specifica în dialogul care va apărea variabila dorită (Puncte1) după care se va înscrie condiţia în următorul dialog *If* (tradus *dacă* şi disponibil după apăsarea butonului *If*):





În urma acestei acţiuni fereastra *Data Editor* va prezenta cazurile neselectate marcate cu o linie diagonală în capul de linie. În acelaşi timp SPSS va crea o variabilă nouă *filter_$* care va specifica pentru fiecare caz dacă a fost selectat sau nu. Evident, în mod asemănător va putea fi realizată cu uşurinţă şi o altă selecţie de exemplu doar pentru sexul masculin (Sex=1):

	Puncte1	Puncte2	Sex	filter_$
1	230	200	masculin	Selected
2	310	220	feminin	Selected
3	250	380	feminin	Selected
4	310	270	masculin	Selected
5	150	200	feminin	Not Select
6	180	310	feminin	Not Select
7	80	120	masculin	Not Select
8	350	170	feminin	Selected
9	220	220	masculin	Selected

	Puncte1	Puncte2	Sex	filter_$
1	230	200	masculin	Selected
2	310	220	feminin	Not Select
3	250	380	feminin	Not Select
4	310	270	masculin	Selected
5	150	200	feminin	Not Select
6	180	310	feminin	Not Select
7	80	120	masculin	Selected
8	350	170	feminin	Not Select
9	220	220	masculin	Selected

Opţiunea ***Weight Cases*** dă cazurilor o „greutate” diferită creată prin simularea unei replicări a acestora. De exemplu dacă avem seria 1, 2, 4, 5, 6, 7, 3, 5 atunci 1 va fi considerat cu greutatea cea mai mică, el va avea o singură replicare. Pentru 2 chiar dacă apare o singură dată în serie se vor crea două replicări, şa. Iată în continuare analiza de frecvenţe afişată în fereastra *Output SPSS Viewer* (stânga) atunci când *Weight Cases* nu este activat (modul normal de lucru) şi în (dreapta) atunci când *Weight Cases* este activat:

Frequencies

Statistics

VAR00001

N	Valid	8
	Missing	0
Mean		4,13
Median		4,50
Mode		5
Sum		33

VAR00001

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	12,5	12,5	12,5
	2	1	12,5	12,5	25,0
	3	1	12,5	12,5	37,5
	4	1	12,5	12,5	50,0
	5	2	25,0	25,0	75,0
	6	1	12,5	12,5	87,5
	7	1	12,5	12,5	100,0
	Total	8	100,0	100,0	

Frequencies

Statistics

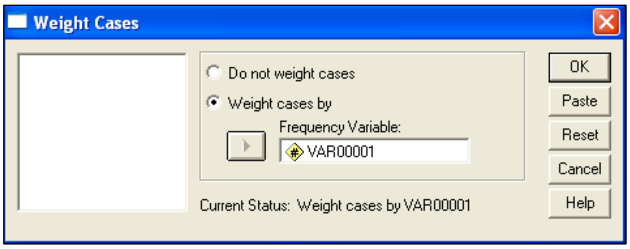
VAR00001

N	Valid	33
	Missing	0
Mean		5,00
Median		5,00
Mode		5
Sum		165

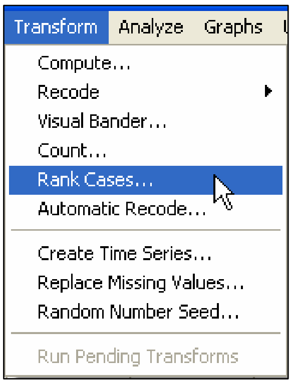
VAR00001

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	3,0	3,0	3,0
	2	2	6,1	6,1	9,1
	3	3	9,1	9,1	18,2
	4	4	12,1	12,1	30,3
	5	10	30,3	30,3	60,6
	6	6	18,2	18,2	78,8
	7	7	21,2	21,2	100,0
	Total	33	100,0	100,0	

Atenţie! Dacă s-a selectat *Weight Cases*, efectul acestuia rămâne activat permanent. În datele salvate cu *Weight Cases* activ se va salva automat şi această setare. Pentru revenirea la modul de lucru normal este necesară deselecarea (prin bifarea opţiunii *Do not weight cases*).



Meniul *Transform* este de asemenea un meniu important al programului SPSS.

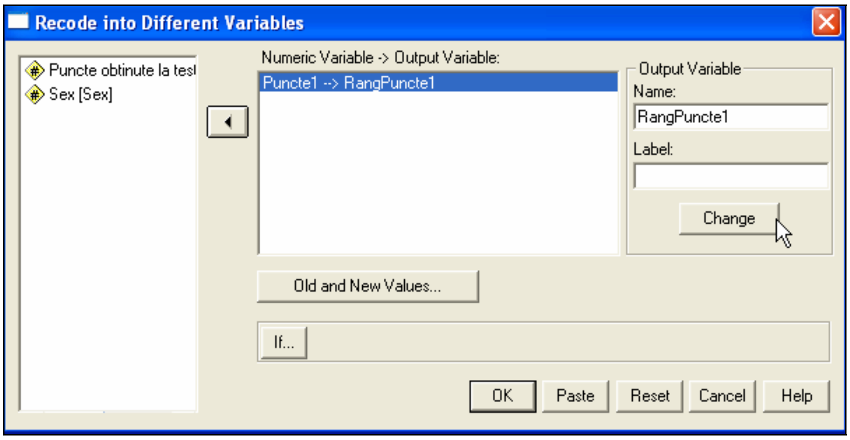


Aşa cum sugerează numele, acest meniu este important pentru transformarea şi crearea datelor. Astfel pot fi generate variabile derivate (câmpuri calculate), recodificate şi contorizate datele, se pot genera ranguri etc. În continuare va fi prezentat lucrul cu opţiunile meniului *Transform* fiind folosite în acest sens câteva exemple.

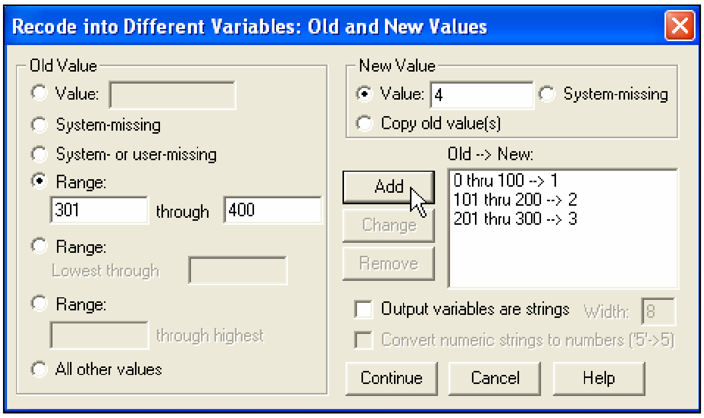
Vom presupune că se doreşte crearea unor ranguri corespunzătoare valorilor variabilei Puncte1 astfel:

- pentru $0 \le$ Puncte1 ≤ 100 variabila RangPuncte1 = 1
- pentru $101 \le$ Puncte1 ≤ 200 variabila RangPuncte1 = 2
- pentru $201 \le$ Puncte1 ≤ 300 variabila RangPuncte1 = 3
- pentru $301 \le$ Puncte1 ≤ 100 variabila RangPuncte1 = 4

Pentru aceasta va trebui să alegem din meniul *Transform* **opțiunea *Recode***. Mai întâi vom crea o variabilă RangPuncte1 (prin înscrierea numelui în *Output variable*):



După apăsarea butonului *Old and New Values* vom specifica în următorul dialog intervalele de clasă pentru ranguri:

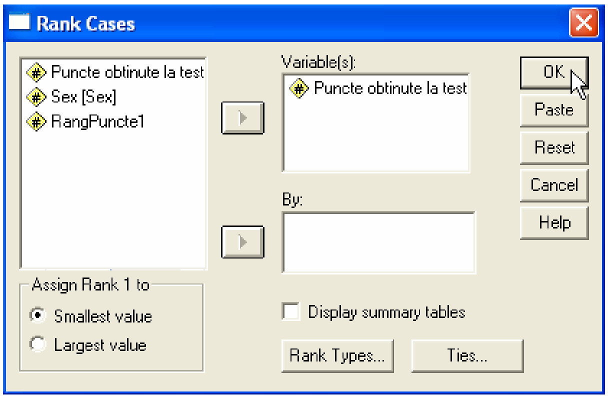


În urma acestei acțiuni în fereastra *Data Editor* va apărea corespunzător rangurilor alese, noua variabilă RangPuncte1:

1 : Puncte1 230

	Puncte1	Puncte2	Sex	RangPuncte1
1	230	200	masculin	3
2	310	220	feminin	4
3	250	380	feminin	3
4	310	270	masculin	4
5	150	200	feminin	2
6	180	310	feminin	2
7	80	120	masculin	1
8	350	170	feminin	4
9	220	220	masculin	3

O altă modalitate de a ierarhiza datele conform unor ranguri este utilizarea opțiunii ***Rank Cases***. După apelarea acestei opțiuni va apărea un dialog prin intermediul căruia se va specifica variabila pentru care se doreşte crearea rangurilor:



În acest exemplu s-a ales aceeaşi variabilă (Puncte obținute la testul 1). După apăsarea butonului OK în fereastra editorului de date va apărea o variabilă nouă RPuncte1. Sortând datele după această variabilă vom avea:

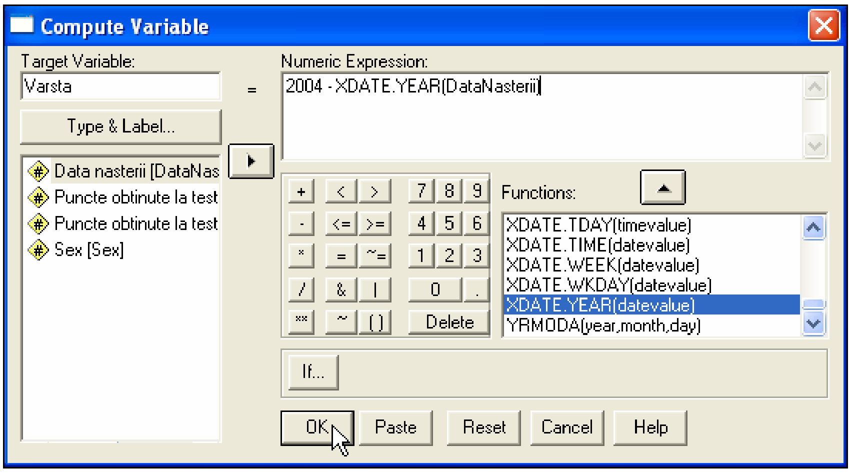
	Puncte1	Puncte2	Sex	RangPuncte1	RPuncte1
1	80	120	masculin	1	1,000
2	150	200	feminin	2	2,000
3	180	310	feminin	2	3,000
4	220	220	masculin	3	4,000
5	230	200	masculin	3	5,000
6	250	380	feminin	3	6,000
7	310	220	feminin	4	7,500
8	310	270	masculin	4	7,500
9	350	170	feminin	4	9,000

Iată în continuare două exemple care vor ajuta la înțelegerea utilității opțiunii ***Compute*** din cadrul meniului *Transform*.

Deoarece dorim să realizăm analize care să țină seama şi de vârstă, am inserat în baza de date o variabilă nouă numită DataNasterii. După încărcarea datelor corespunzătoare acestei coloane fereastra editorului de date conține:

	DataNasterii	Puncte1	Puncte2	Sex
1	20.10.1984	230	200	masculin
2	15.01.1983	310	220	feminin
3	01.02.1983	250	380	feminin
4	19.09.1984	310	270	masculin
5	05.07.1982	150	200	feminin
6	02.12.1981	180	310	feminin
7	08.11.1981	80	120	masculin
8	16.05.1982	350	170	feminin
9	28.06.1984	220	220	masculin

Va trebui să creăm şi să calculăm o variabilă Vârsta, derivată din DataNasterii. În acest sens vom avea: Varsta = 2004 – Anul_Nasterii. S-a considerat anul curent ca fiind 2004. În acest sens SPSS pune la dispoziţie funcţia XDATE.YEAR(*DateValue*) care va întoarce anul dintr-o valoare tip dată calendaristică. De exemplu XDATE.YEAR(02.12.1981) = 1981. Deci după selectarea opţiunii *Compute* va apărea următorul dialog în care am înscris în câmpul *Target Variabile* numele variabilei ce va fi creată (Varsta) şi apoi prin redactare şi cu ajutorul generatorului de formule am compus expresia (din *Numeric Expression*):
2004 - XDATE.YEAR(DataNasterii).



În urma acestei acţiuni în *SPSS Data Editor* a apărut ca un câmp calculat vârsta studenţilor (Varsta):

1 : DataNasterii 20.10.1984

	DataNasterii	Puncte1	Puncte2	Sex	Varsta
1	20.10.1984	230	200	masculin	20,00
2	15.01.1983	310	220	feminin	21,00
3	01.02.1983	250	380	feminin	21,00
4	19.09.1984	310	270	masculin	20,00
5	05.07.1982	150	200	feminin	22,00
6	02.12.1981	180	310	feminin	23,00
7	08.11.1981	80	120	masculin	23,00
8	16.05.1982	350	170	feminin	22,00
9	28.06.1984	220	220	masculin	20,00

Dacă se doreşte obţinerea unei noi variabile PunctajMediu al celor două teste care să se calculeze după formula: PunctajMediu = (Puncte1 + Puncte2) / 2 atunci în dialogul *Compute Variable* va trebui să se completeze în câmpul *Target Variabile*

numele variabilei (PunctajMediu) iar în *Numeric Expression* (Puncte1 + Puncte2) / 2. După apăsarea butonului OK fereastra editorului va conține:

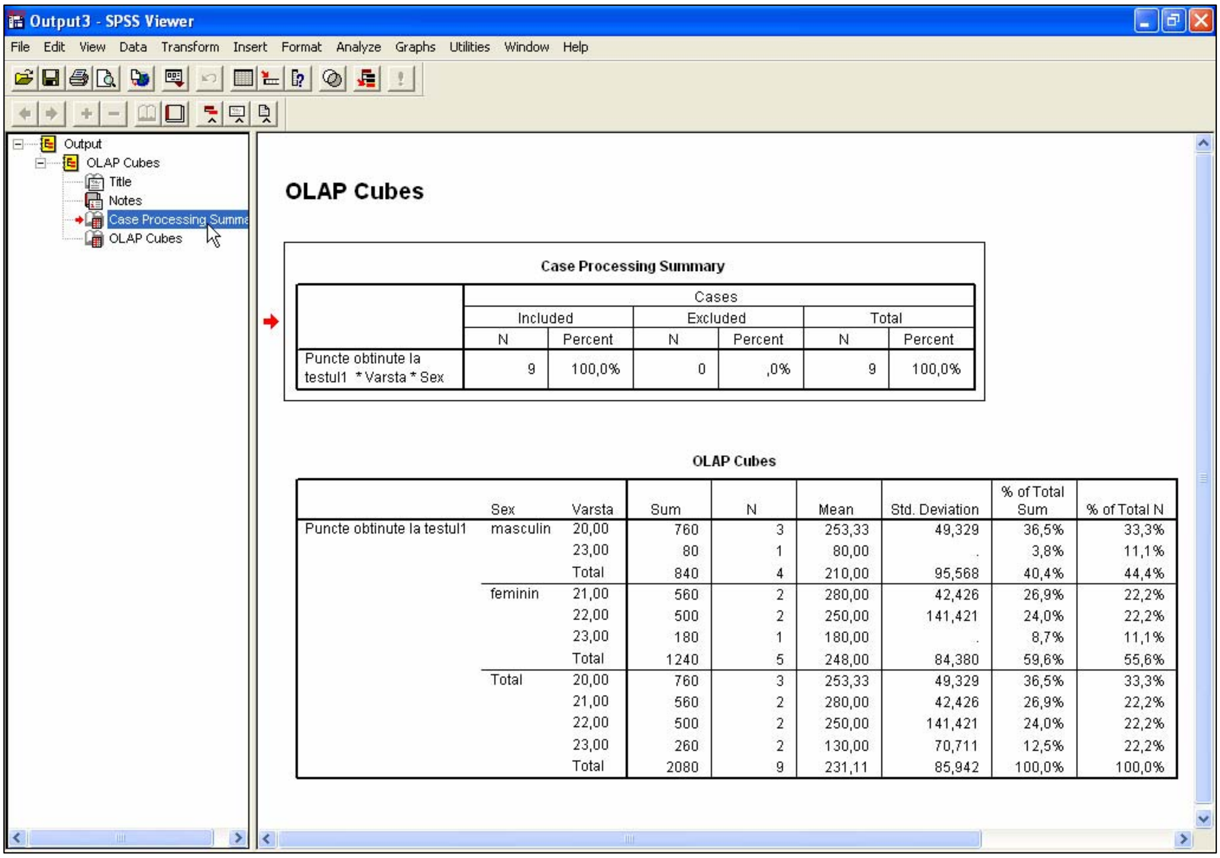
1 : DataNasterii	20.10.1984					
	DataNasterii	Puncte1	Puncte2	Sex	Varsta	PunctajMediu
1	20.10.1984	230	200	masculin	20,00	215,00
2	15.01.1983	310	220	feminin	21,00	265,00
3	01.02.1983	250	380	feminin	21,00	315,00
4	19.09.1984	310	270	masculin	20,00	290,00
5	05.07.1982	150	200	feminin	22,00	175,00
6	02.12.1981	180	310	feminin	23,00	245,00
7	08.11.1981	80	120	masculin	23,00	100,00
8	16.05.1982	350	170	feminin	22,00	260,00
9	28.06.1984	220	220	masculin	20,00	220,00

Componenta *Output Viewer*

Meniurile *Analyze* şi *Graphs* constituie partea „de forţă” a programului SPSS care întăreşte convingerea unei mari părţi a comunităţii ştiinţifice, că în prezent SPSS este unul dintre cele mai performante programe ce pot fi utilizate pentru analize statistice. Opţiunile acestor meniuri asigură proceduri care fac posibilă procesarea datelor existente în *Data Editor* astfel încât să poată fi rezolvate cu uşurinţă probleme de statistică descriptivă, de analiză dispersională, de corelaţie, de validare a unor ipoteze prin intermediul testelor statistice de vizualizare şi interpretare a datelor şi a rezultatelor procesărilor pe baza unor reprezentări grafice sugestive. În prezentările făcute nu se va insista pe posibilităţile de personalizare şi pe setările suplimentare pe care SPSS le oferă şi care presupun cunoştinţe avansate se Statistică. Cel mai frecvent se vor accepta variantele implicite presetate de către SPSS.

Rezultatele, situaţiile finale, rapoartele oferite de opţiunile celor două meniuri vor fi afişate într-o fereastră specială numită *SPSS Output Viewer* care constituie de asemenea ca şi *Data Editor* o componentă importantă a programului. În plus *Output Viewer* oferă o interfaţă interactivă ce permite modificarea formei de afişare a rezultatelor prin alegerea celor mai potrivite tabele şi grafice de raportare, a fonturilor, culorilor, tipurilor de linii, a formatului numerelor şi textelor şi a conţinutului acestora. Astfel de exemplu, rapoartele pot fi modificate pentru a fi afişate în limba română. De asemenea componentele (obiectele) conţinute în rapoarte pot fi copiate în alte programe (de exemplu *Word*) prin mecanismul *Clipboard al Windows*-ului (*Copy & Paste*). Rapoartele SPSS pot fi salvate în fişiere specifice cu extensia (.spo) în acest fel fiind posibilă şi o utilizare ulterioară a acestora.

Fereastra *Output Viewer* este împărţită în două secţiuni:



Secţiunea din stânga în care apar pe o structură tip „arbore” itemii de selecţie formaţi din titlurile componentelor fiecărui raport, şi secţiunea din dreapta care

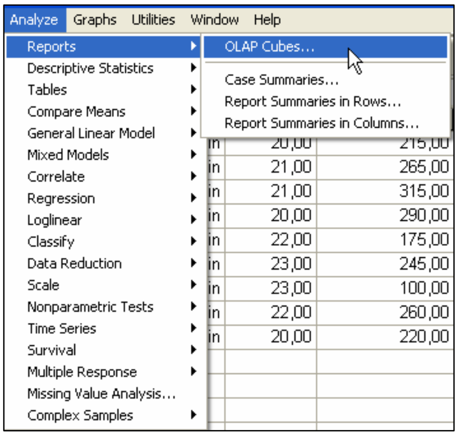
conține titlurile, tabelele, valorile, textele ajutătoare etc ale soluțiilor furnizate de SPSS şi care constituie de fapt situația finală de raport. Un clic într-un item de selecție activează componenta corespondentă aflată în secțiunea din dreapta. În imaginea anterioară este prezentat un raport rezultat al opțiunii *Reports - OLAP Cubes* din meniul *Analyze*. Vom denumi această imagine *Output3 – SPSS Viewer*.

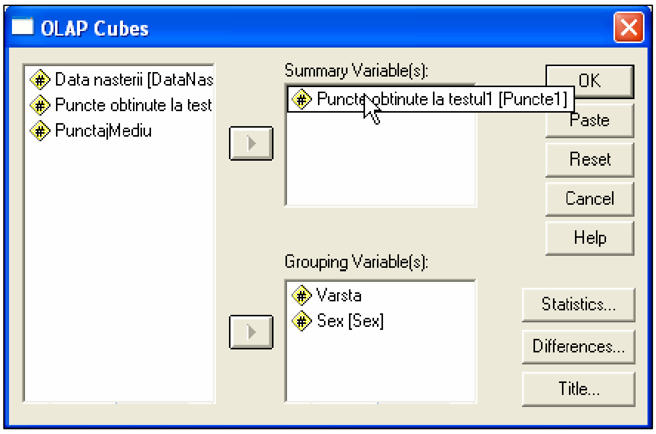
Datele pe baza cărora s-a realizat acest raport, existente în *Data Editor* sunt:

1 : DataNasterii 20.10.1984

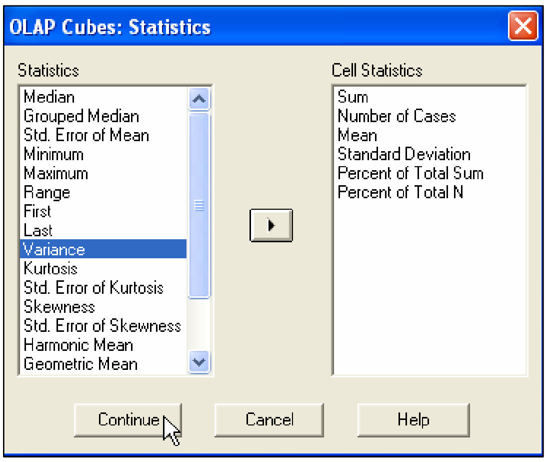
	DataNasterii	Puncte1	Puncte2	Sex	Varsta	PunctajMediu
1	20.10.1984	230	200	masculin	20,00	215,00
2	15.01.1983	310	220	feminin	21,00	265,00
3	01.02.1983	250	380	feminin	21,00	315,00
4	19.09.1984	310	270	masculin	20,00	290,00
5	05.07.1982	150	200	feminin	22,00	175,00
6	02.12.1981	180	310	feminin	23,00	245,00
7	08.11.1981	80	120	masculin	23,00	100,00
8	16.05.1982	350	170	feminin	22,00	260,00
9	28.06.1984	220	220	masculin	20,00	220,00

Meniul *Analyze* oferă toate opțiunile pe baza cărora pot fi realizate analize statistice sub SPSS. Prima opțiune este *Reports*. După un clic dat pe opțiunea OLAP (***On**line **A**nalytical **P**rocessing) Cubes* a submeniului *Reports* a apărut dialogul *OLAP Cubes* prin intermediul căruia s-au selectat variabilele pentru care se doreşte o analiză OLAP.

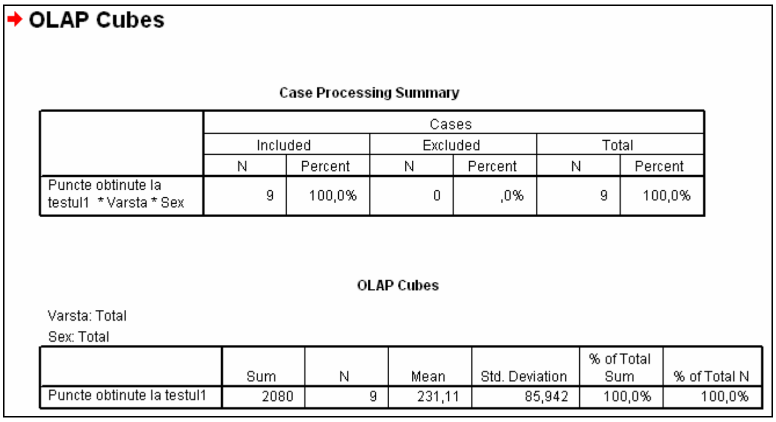




Apăsarea butonului *Statistics* oferă posibilitatea alegerii indicatorilor statistici ce se doreşte a fi calculați în cadrul acestei analize (*OLAP Cubes Statistics*):



Forma inițială furnizată de SPSS pentru acest raport în *Output Viewer* a fost:



OLAP Cubes

Case Processing Summary

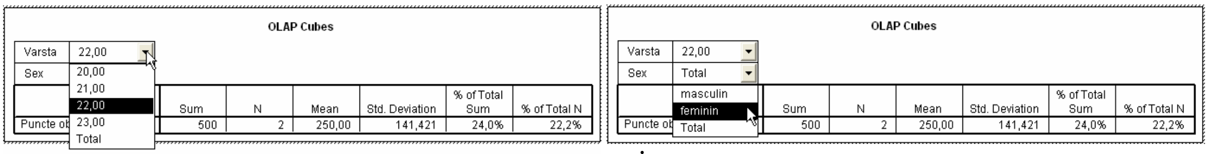
	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
Puncte obtinute la testul1 * Varsta * Sex	9	100,0%	0	,0%	9	100,0%

OLAP Cubes

Varsta: Total
Sex: Total

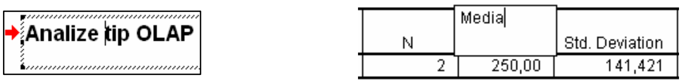
	Sum	N	Mean	Std. Deviation	% of Total Sum	% of Total N
Puncte obtinute la testul1	2080	9	231,11	85,942	100,0%	100,0%

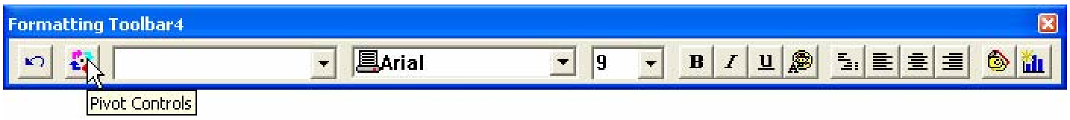
Raportul *OLAP Cubes* (la fel ca şi alte rapoarte) oferă posibilitatea studierii într-o formă interactivă a indicatorilor statistici doriți. Astfel un clic dublu dat în tabelul *OLAP Cubes* determină un studiu în detaliu pe baza selecțiilor disponibile după vârstă



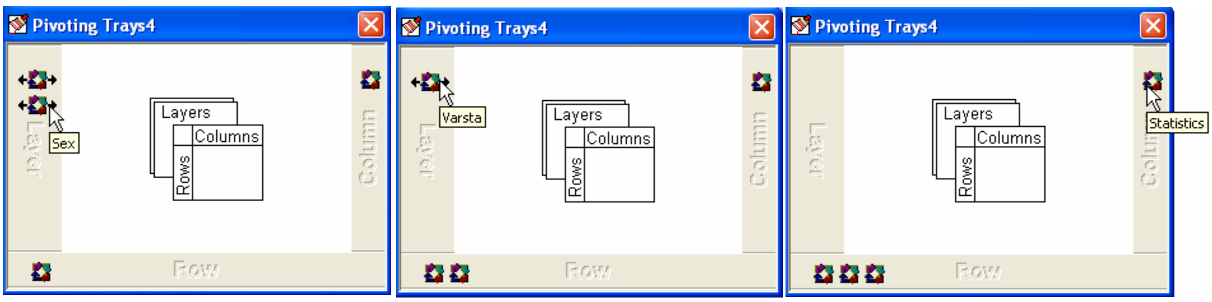
şi sex:

De asemenea un clic dublu dat în titlul rapoartelor sau în oricare text existent oferă posibilitatea schimbării şi formatării acestora prin intermediul unei bare de unelte (*Formatting Toolbar*):



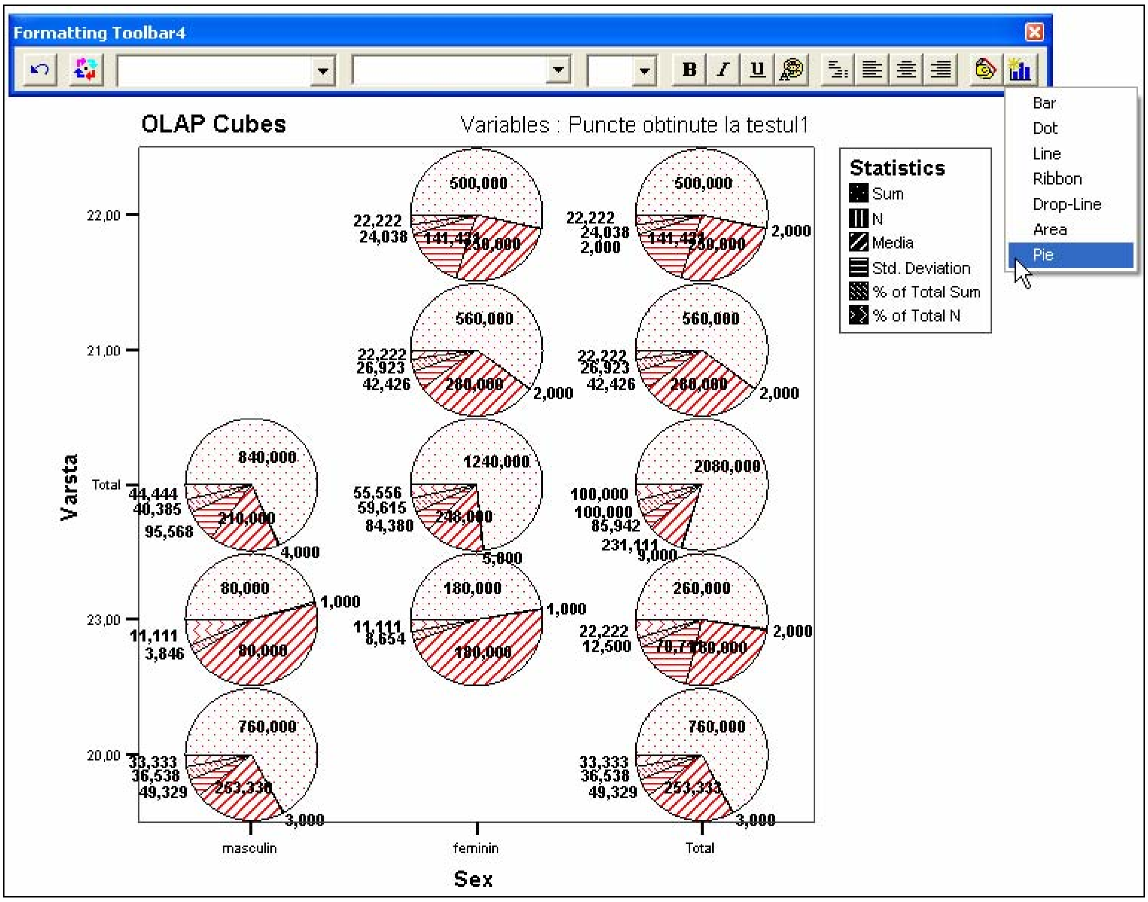


O facilitate foarte importantă pe care această bară o oferă este *Pivot Controls* (vezi butonul marcat din imaginea barei). Această facilitate este disponibilă pentru configurarea într-un raport a oricărui tabel (cu una sau mai multe intrări). Rândurile (*rows*) şi coloanele (*columns*) pot fi create interactiv prin *drag & drop* (tragerea lor cu ajutorul *mouse*-ului). În exemplul următor s-a dorit obținerea unui raport OLAP detaliat pe sexe şi pe vârste.



Prin „tragerea" variabilei Sex şi a variabilei Varsta pe rânduri (*Rows*) s-a obținut tabelul prezentat în prima imagine a acestui capitol numită *Output3 – SPSS Viewer*.

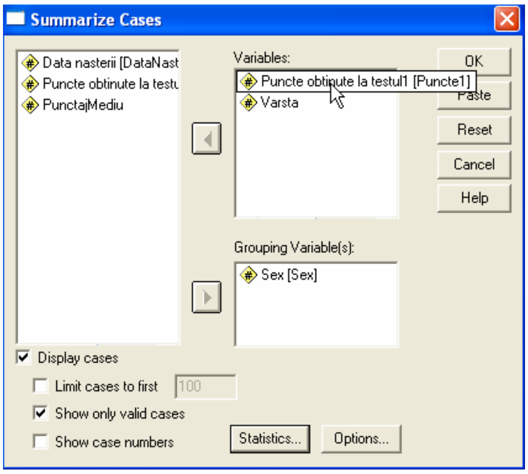
O altă facilitate interesantă a barei de unelte de formatare este şi cea oferită pentru realizarea reprezentării grafice a conţinutului oricărui tabel dintr-un raport. Iată în continuare un grafic de tip *Pie* (plăcintă) pe vârste care a fost realizat pentru tabelul din aceeaşi imagine *Output3 – SPSS Viewer*. (Vezi această imagine şi datele din *Data Editor* pe baza cărora a fost construit raportul din imagine).

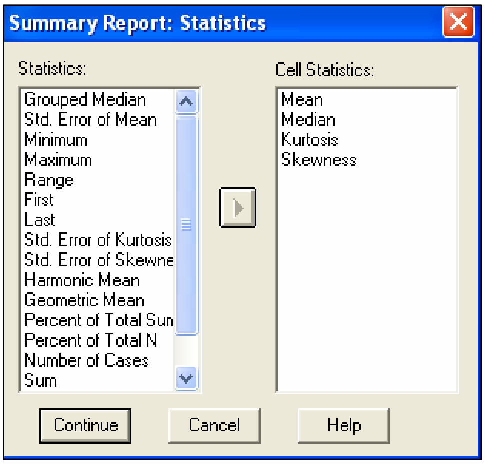


În continuare vor fi prezentate cele mai importante şi frecvent utilizate opţiuni ale meniului *Analyze* utilizându-se exemplificări pe baza unor aplicaţii practice relevante. În subcapitolul anterior în paralel cu componenta *Output – SPSS Viewer* a fost prezentată deja prima opţiune – *OLAP Cubes,* din cadrul meniului *Analyze – Reports*.

Indicatorii descriptivi

Opţiunea ***Case Summaries*** a submeniului *Reports* oferă o imagine de ansamblu asupra valorii indicatorilor statistici descriptivi ai unor variabile ce pot fi prezentaţi grupaţi pe baza altor variabile. Pentru exemplificare s-a utilizat acelaşi set de date ca şi pentru *OLAP Cubes*. Primul dialog care apare după lansarea opţiunii este *Summarize Cases*. Prin intermediul acestui dialog se pot selecta variabilele pentru analiză. În cazul nostru s-au ales Puncte1 şi Varsta iar pentru grupare Sex. La apăsarea butonului *Statistics* apare următorul dialog *Summary Report: Statistics* prin intermediul căruia vor fi selectaţi indicatorii statistici doriţi:



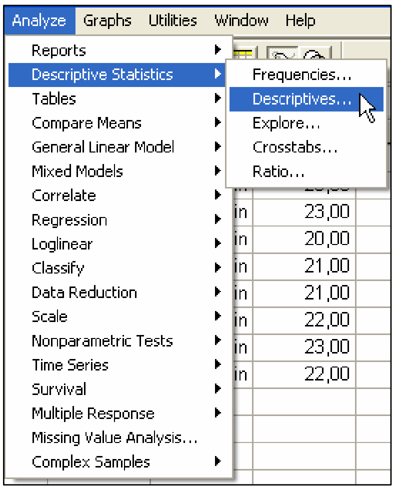


După un clic dat pe butonul OK, SPSS va afişa în *Output – SPSS Viewer* următorul raport:

Case Summaries

				Puncte obtinute la testul1	Varsta
Sex	masculin	1		230	20,00
		2		310	20,00
		3		80	23,00
		4		220	20,00
		Total	Mean	210,00	20,7500
			Median	225,00	20,0000
			Kurtosis	1,916	4,000
			Skewness	-,907	2,000
	feminin	1		310	21,00
		2		250	21,00
		3		150	22,00
		4		180	23,00
		5		350	22,00
		Total	Mean	248,00	21,8000
			Median	250,00	22,0000
			Kurtosis	-2,165	-,612
			Skewness	,030	,512
	Total	Mean		231,11	21,3333
		Median		230,00	21,0000
		Kurtosis		-,403	-1,556
		Skewness		-,367	,233

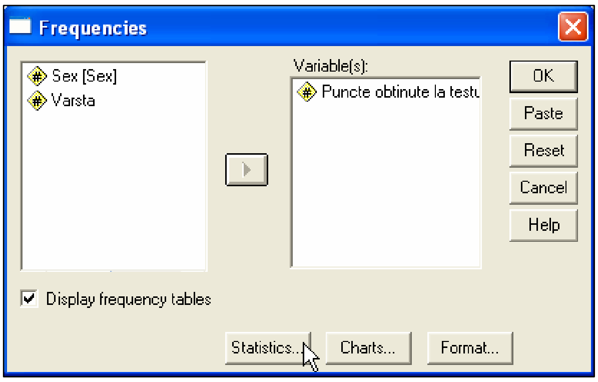
Meniul *Descriptives Statistics* oferă posibilitatea calculului şi analizei indicatorilor statistici descriptivi.

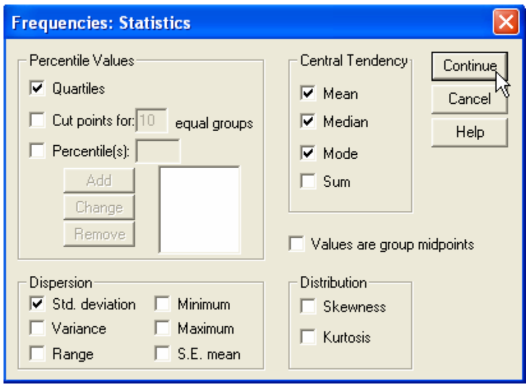


Pentru toate exemplificările acestui meniu vom lua în considerare următorul set de date:

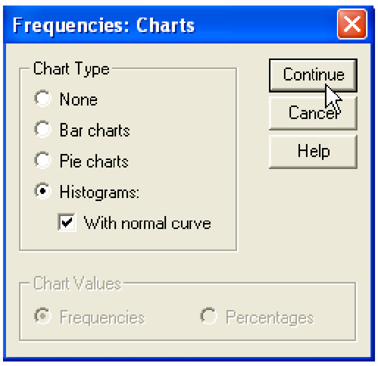
	Puncte1	Sex	Varsta
1	220	masculin	20,00
2	310	masculin	20,00
3	80	masculin	23,00
4	220	masculin	20,00
5	310	feminin	21,00
6	250	feminin	21,00
7	150	feminin	22,00
8	180	feminin	23,00
9	350	feminin	22,00
10	180	masculin	23,00
11	310	feminin	20,00

Prima opţiune a meniului este ***Frequencies***, singura care permite analiza de frecvenţe. Dialogul *Frequencies* oferă posibilitatea selectării variabilei (variabilelor) pentru analiză iar dialogul *Statistics* prezintă indicatorii statistici ce pot fi calculaţi pentru raport.





Butonul *Charts* disponibilizează un alt dialog prin intermediul căruia se poate specifica dacă se doreşte ca raportul să conţină o reprezentare grafică şi se poate selecta tipul de grafic dorit.



După apăsarea butonului OK, raportul prezentat în *Output – SPSS Viewer* va conţine:

Frequencies

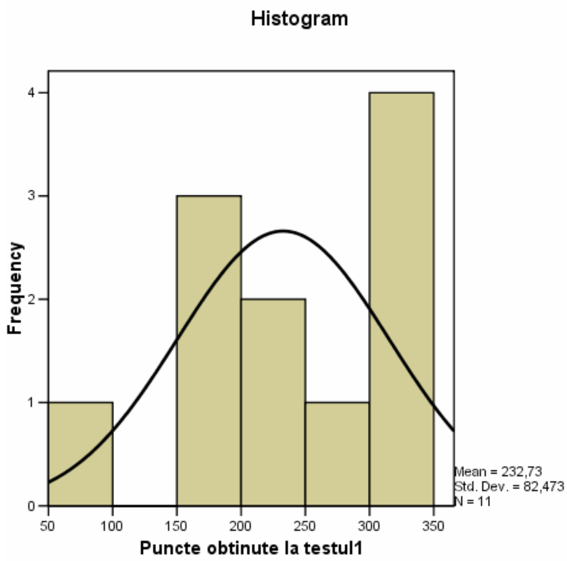
Statistics

Puncte obtinute la testul1

N	Valid	11
	Missing	0
Mean		232,73
Median		220,00
Mode		310
Std. Deviation		82,473
Percentiles	25	180,00
	50	220,00
	75	310,00

Puncte obtinute la testul1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	80	1	9,1	9,1	9,1
	150	1	9,1	9,1	18,2
	180	2	18,2	18,2	36,4
	220	2	18,2	18,2	54,5
	250	1	9,1	9,1	63,6
	310	3	27,3	27,3	90,9
	350	1	9,1	9,1	100,0
	Total	11	100,0	100,0	



Opțiunea *Descriptives*, după selectarea variabilei pentru analiză (Puncte obținute la testul 1) şi alegerea indicatorilor statistici (minim, maxim, media, abaterea standard şi dispersia) va oferi următorul raport:

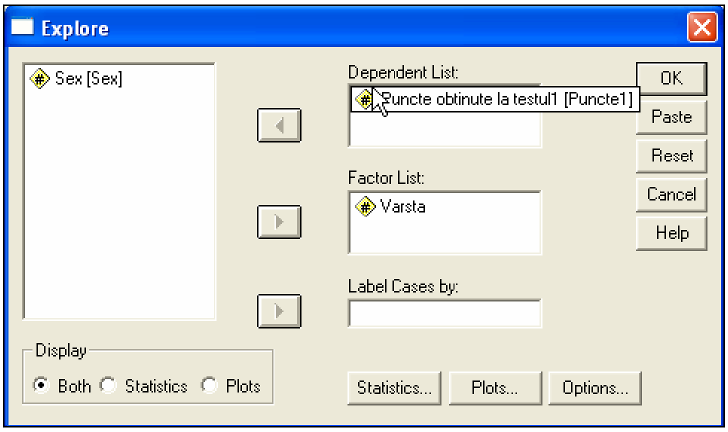
Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Puncte obtinute la testul1	11	80	350	232,73	82,473	6801,818
Valid N (listwise)	11					

Opțiunea *Explore* este utilă atunci când se doreşte un studiu complet al indicatorilor statistici descriptivi pentru una sau mai multe variabile considerate a fi dependente (*Dependent List*) de una sau mai multe variabile categoriale (*Factors*). Ca atare în prima casetă de dialog care va apărea după lansarea acestei opțiuni se vor selecta variabilele de analizat în *Dependent List* iar (opțional) în *Factor List* se vor selecta variabilele categoriale în funcție de care se doreşte analiza. În acest dialog apare şi un câmp *Label cases by* (rar utilizat) care permite etichetarea cazurilor la afişare.

În exemplul care urmează s-a utilizat acelaşi set de date (folosit anterior), alegându-se variabila Puncte1 pentru analiză şi variabila Varsta ca variabilă categorială:



Raportul corespunzător afişează (în mod implicit) un sumar al cazurilor procesate, indicatorii statistici descriptivi şi reprezentările grafice *Stem-and-Leaf Plots* structurate pe vârste.

Explore

Varsta

Case Processing Summary

		Cases					
		Valid		Missing		Total	
	Varsta	N	Percent	N	Percent	N	Percent
Puncte obtinute la testul1	20,00	4	100,0%	0	,0%	4	100,0%
	21,00	2	100,0%	0	,0%	2	100,0%
	22,00	2	100,0%	0	,0%	2	100,0%
	23,00	3	100,0%	0	,0%	3	100,0%

Puncte obtinute la testul1

Stem-and-Leaf Plots

```
Puncte obtinute la testul1 Stem-and-Leaf Plot for
Varsta= 20,00

 Frequency    Stem &  Leaf

     2,00        2 .  22
      ,00        2 .
     2,00        3 .  11

 Stem width:       100
 Each leaf:       1 case(s)
```

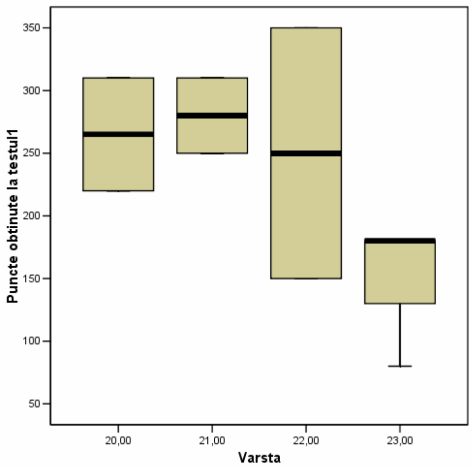
```
Puncte obtinute la testul1 Stem-and-Leaf Plot for
Varsta= 21,00

 Frequency    Stem &  Leaf

     1,00        2 .  5
     1,00        3 .  1

 Stem width:       100
 Each leaf:       1 case(s)
```

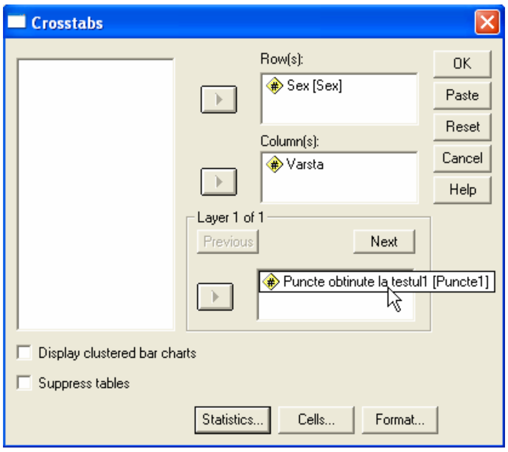
```
Puncte obtinute la testul1 Stem-and-Leaf Plot for
Varsta= 23,00

 Frequency    Stem &  Leaf

     1,00        0 .  8
     2,00        1 .  88

 Stem width:       100
 Each leaf:       1 case(s)
```

```
Puncte obtinute la testul1 Stem-and-Leaf Plot for
Varsta= 22,00

 Frequency    Stem &  Leaf

     2,00        0 .  13

 Stem width:       1000
 Each leaf:        1 case(s)
```

Descriptives

	Varsta			Statistic	Std. Error
Puncte obtinute la testul1	20,00	Mean		265,00	25,981
		95% Confidence Interval for Mean	Lower Bound	182,32	
			Upper Bound	347,68	
		5% Trimmed Mean		265,00	
		Median		265,00	
		Variance		2700,000	
		Std. Deviation		51,962	
		Minimum		220	
		Maximum		310	
		Range		90	
		Interquartile Range		90	
		Skewness		,000	1,014
		Kurtosis		-6,000	2,619
	21,00	Mean		280,00	30,000
		95% Confidence Interval for Mean	Lower Bound	-101,19	
			Upper Bound	661,19	
		5% Trimmed Mean		.	
		Median		280,00	
		Variance		1800,000	
		Std. Deviation		42,426	
		Minimum		250	
		Maximum		310	
		Range		60	
		Interquartile Range		.	
		Skewness		.	.
		Kurtosis		.	.
	22,00	Mean		250,00	100,000
		95% Confidence Interval for Mean	Lower Bound	-1020,62	
			Upper Bound	1520,62	
		5% Trimmed Mean		.	
		Median		250,00	
		Variance		20000,000	
		Std. Deviation		141,421	
		Minimum		150	
		Maximum		350	
		Range		200	
		Interquartile Range		.	
		Skewness		.	.
		Kurtosis		.	.
	23,00	Mean		146,67	33,333
		95% Confidence Interval for Mean	Lower Bound	3,24	
			Upper Bound	290,09	
		5% Trimmed Mean		.	
		Median		180,00	
		Variance		3333,333	
		Std. Deviation		57,735	
		Minimum		80	
		Maximum		180	
		Range		100	
		Interquartile Range		.	
		Skewness		-1,732	1,225
		Kurtosis		.	.





Opţiunea *Crosstab* permite realizarea unor tabele cu dublă intrare.

Caseta de dialog corespunzătoare ne oferă un suport interactiv pentru selectarea variabilelor reprezentate pe rânduri (*Rows*), pe coloane (*Columns*) şi a valorilor din celulele de intersecţie ale acestora. În mod implicit valorile celulelor vor reprezenta numărul de cazuri (*Counts*) ale fiecărei intersecţii.

Iată în continuare raportul ***Crosstabulation*** afişat în *Output – SPSS Viewer*:

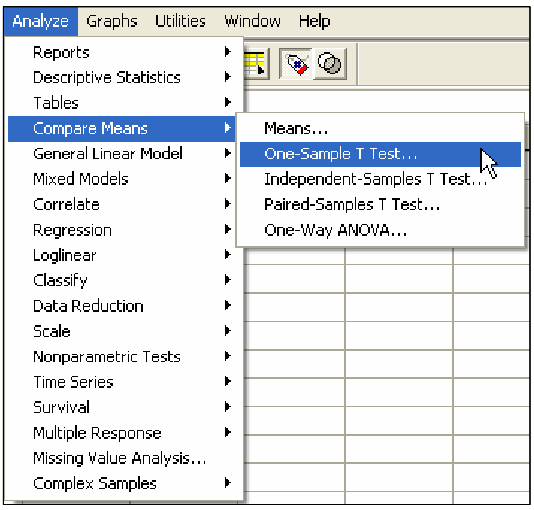
Sex * Varsta * Puncte obtinute la testul1 Crosstabulation

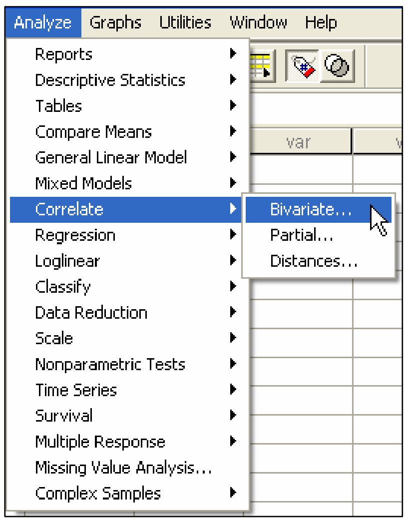
Count

Puncte obtinute la testul1			Varsta 20,00	21,00	22,00	23,00	Total
80	Sex	masculin				1	1
	Total					1	1
150	Sex	feminin			1		1
	Total				1		1
180	Sex	masculin				1	1
		feminin				1	1
	Total					2	2
220	Sex	masculin	2				2
	Total		2				2
250	Sex	feminin		1			1
	Total			1			1
310	Sex	masculin	1	0			1
		feminin	1	1			2
	Total		2	1			3
350	Sex	feminin			1		1
	Total				1		1

Teste statistice şi analiza corelaţiei

În acest subcapitol vor fi prezentate aplicaţii rezolvate pe baza suportului software SPSS dedicat pentru testele statistice şi studiul corelaţiei. În acest sens se va insista pe facilităţile oferite de opţiunile meniurilor *Compare Means* şi *Corelate*.





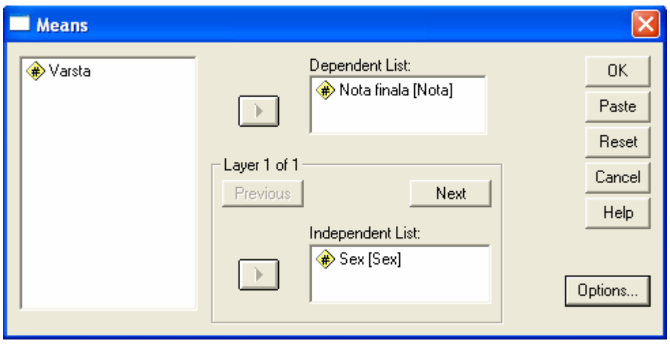
Aşa cum s-a arătat anterior testul T este util pentru testarea diferenţelor între valorile medii ale unei variabile care pot fi măsurate între diferite grupuri la momente diferite de timp, sau în comparaţie cu o populaţie statistică ale cărei valori pentru respectiva variabilă sunt cunoscute. Cele mai frecvente aplicaţii ale testului T sunt legate de testarea diferenţei între grupuri dependente sau independente.

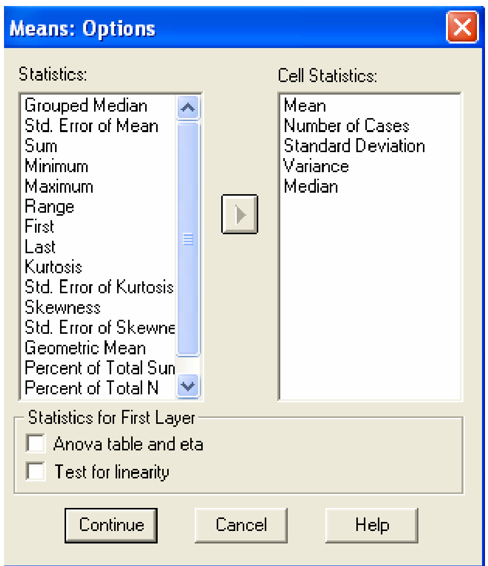
Testul T pentru grupuri independente este utilizat atunci când aceeaşi variabilă a fost măsurată între două asemenea grupuri şi cercetătorul vrea să ştie dacă diferenţa mediilor între grupuri este semnificativă statistic. Prin grupuri independente înţelegem acele grupuri care conţin subiecţi diferiţi.

Vom considera în continuare un set de date pe baza căruia dorim să verificăm dacă notele finale ale unui grup de 14 studenţi compus din 7 băieţi şi 7 fete diferă semnificativ în funcţie de sexul acestora. Setul de date este:

Nota finală	Sex	Vârsta
8	masculin	20
9	masculin	20
6	masculin	23
8	masculin	20
9	feminin	21
8	feminin	21
7	feminin	22
7	feminin	23
10	feminin	22
7	masculin	23
9	feminin	20
6	masculin	22
7	feminin	23
8	masculin	22

Opţiunea *Means* (medii) a meniului *Compare Means* ar trebui să fie prima consultată, deoarece ea ne poate oferi o informaţie interesantă relativ la notele medii ale celor două sexe. Prin intermediul primului dialog al opţiunii (*Means*) vom selecta variabila dependentă (Nota) şi variabila independentă (Sex). Un clic pe butonul *Options* va deschide un nou dialog prin care vom putea alege indicatorii statistici ce vor fi afişaţi.





Raportul final (Report) care ne va prezenta cele două medii 7,43 pentru băieţi şi 8,14 pentru fete este:

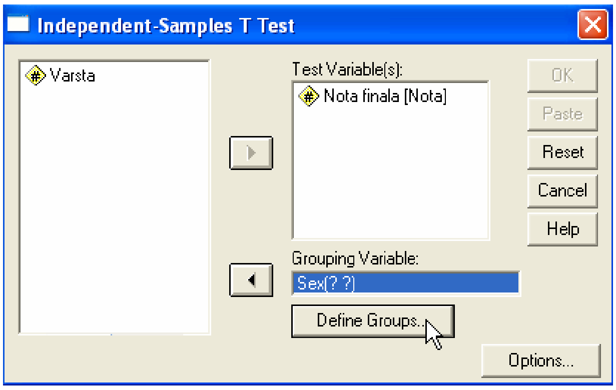
Report

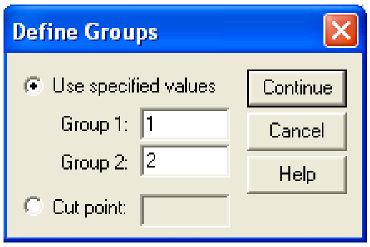
Nota finala

Sex	Mean	N	Std. Deviation	Variance	Median
masculin	7,43	7	1,134	1,286	8,00
feminin	8,14	7	1,215	1,476	8,00
Total	7,79	14	1,188	1,412	8,00

Pe baza acestui raport s-ar putea trage următoare concluzie: notele fetelor sunt mai mari, deci notele depind de sex. Dar să nu ne grăbim cu această afirmaţie care încă nu este decât o ipoteză pe care va trebui să o verificăm prin intermediul unui test T pentru grupuri independente.

Opţiunea *Independent-Samples T Test* ne dă această posibilitate. Primul dialog (*Independent-Samples T Test*) ne oferă suportul pentru a putea alege variabilele dorite. Am selectat ca variabilă testată Nota şi ca variabilă de grupare Sex. Deoarece grupurile nu sunt definite (Sex (? ?)) un clic în butonul *Define Groups* ne va ajuta în acest sens. Pentru că variabila de grupare este Sex şi pentru că atunci când a fost definită i s-au atribuit două valori: 1=masculin şi 2=feminin, vom specifica acest lucru în dialogul *Define Groups*.





După apăsarea butonului OK va apărea următorul raport:

Group Statistics

	Sex	N	Mean	Std. Deviation	Std. Error Mean
Nota finala	masculin	7	7,43	1,134	,429
	feminin	7	8,14	1,215	,459

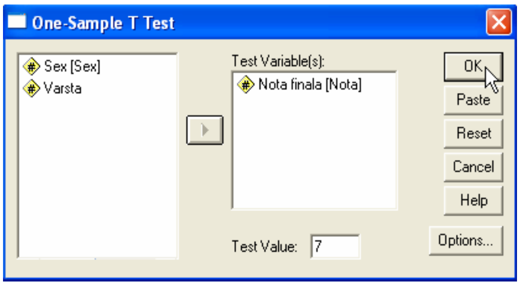
Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Nota finala	Equal variances assumed	,090	,769	-1,137	12	,278	-,714	,628	-2,083	,654
	Equal variances not assumed			-1,137	11,943	,278	-,714	,628	-2,084	,655

Se observă în ultimul tabel testul *Levene* care a fost aplicat pentru verificarea omogenității. Reamintim faptul că o valoare mică a raportului F (Ficher) şi o probabilitate (*Sig.*) < 0,05 confirmă faptul că omogenitatea dispersiilor este asigurată. Dacă omogenitatea dispersiilor nu există, atunci citirea rezultatelor trebuie făcută pe linia a doua (*Equal variances not assumed*). Observăm că pentru acest test nota t calculată este t = -1,137, numărul gradelor de libertate df = N – 2 = 12 şi probabilitatea *Sig.(2-tailed)* = 0,278. Deoarece 0,278 > 0,05 atunci ipoteza cercetării este infirmată şi este acceptată ipoteza de nul: notele nu depind de sex.

Opțiunea *One-Sample T Test* se poate aplica atunci când se verifică media unui eşantion față de media unei populații căreia aparține acel eşantion. În cazul exemplului nostru vom presupune că nota medie obținută la respectiva disciplină (probabil obligatorie) de către toți studenții absolvenți ai universității, pe parcursul mai multor ani este cunoscută, ea fiind 7. Dorim să verificăm ipoteza conform căreia notele studenților din eşantion se încadrează în această tendință.

În acest sens vom lansa opțiunea *One-Sample T Test* şi vom specifica nota 7 în câmpul *Test Value*.



Raportul obținut după apăsarea butonului OK va fi:

One-Sample Test

	Test Value = 7					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Nota finala	2,474	13	,028	,786	,10	1,47

Se observă că *Sig.* = 0,028 < 0,05. Rezultă că ipoteza cercetării se acceptă.

Sub SPSS **testul T poate fi aplicat şi pentru grupuri dependente**. Acest test necesită două măsurări ale aceleiaşi variabile la două momente diferite de timp. Fiecare caz va corespunde unei singure persoane.

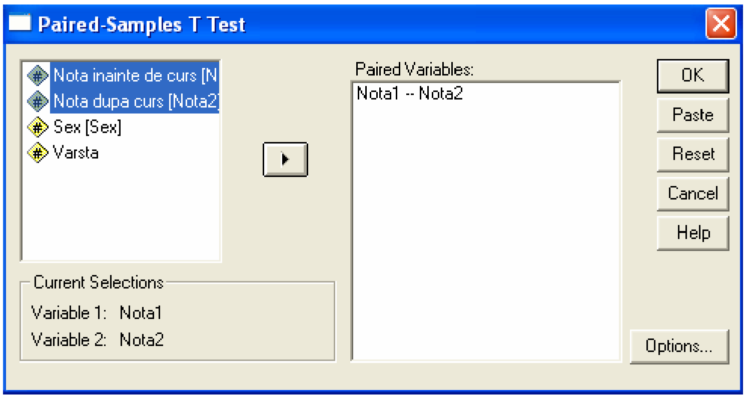
Astfel poate fi verificat de exemplu dacă absolvirea unui curs plătit de o companie pentru anumit un grup de angajați contribuie sau nu la creşterea performanțelor acestora. În acest sens este necesară măsurarea performanțelor angajaților care au urmat cursul, înainte şi după absolvire. Pe baza diferențelor care apar între medii se va accepta sau se va respinge ipoteza bazată pe presupunerea că respectivul curs a contribuit la creşterea performanțelor angajaților.

Opțiunea *Paired-Samples T Test* (testul T pentru grupuri pereche) ne oferă posibilitatea să aplicăm testul T pentru două grupuri dependente. În cazul exemplului enunțat variabila măsurată va fi nota care cuantifică performanța angajaților înainte şi după absolvirea cursului. Vom defini două variabile Nota1 şi Nota2. Variabila Nota1 conține notele angajaților înainte de curs şi variabila Nota2 notele angajaților după curs.

Setul de date care include cele două note ale unui grup de 16 angajați ce au urmat cursul este:

	Nota1	Nota2
1	8,50	8,30
2	9,10	8,80
3	6,00	7,00
4	8,00	9,50
5	9,00	9,60
6	8,30	8,00
7	7,00	7,80
8	7,00	8,70
9	10,00	9,80
10	7,50	7,80
11	9,10	8,80
12	6,00	6,80
13	7,30	7,00
14	8,00	8,80
15	9,00	8,40
16	7,50	8,60

După lansarea opțiunii *Paired-Samples T Test* din cadrul meniului *Compare Means*, va apărea următorul dialog prin intermediul căruia vor fi selectate cele două variabile supuse studiului (Nota1 şi Nota2):



Raportul din *Output – SPSS Viewer* va fi:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Nota inainte de curs	7,9563	16	1,14657	,28664
	Nota dupa curs	8,3563	16	,91504	,22876

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Nota inainte de curs & Nota dupa curs	16	,773	,000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Nota inainte de curs - Nota dupa curs	-,40000	,72847	,18212	-,78817	-,01183	-2,196	15	,044

Se observă *Sig.(2-tailed)* = 0,044 care este < 0,05. Rezultă că ipoteza cercetării se acceptă. Cursul a contribuit la creşterea performanţelor angajaţilor.

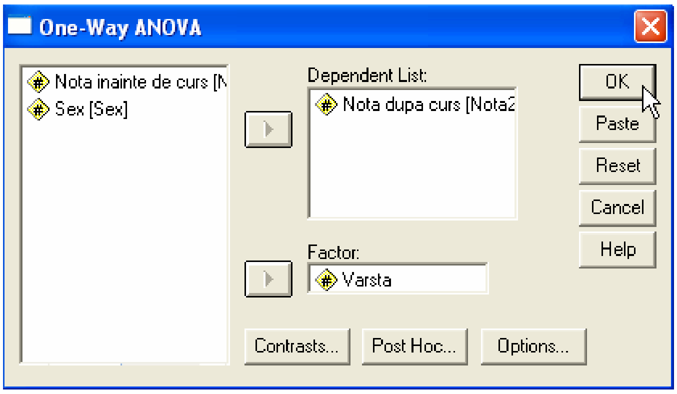
Analiza dispersională ANOVA (*ANalisys Of VAriance*) constituie un alt instrument statistic foarte flexibil cu ajutorul căruia pot fi comparate diferenţele dintre mai mult de două valori medii ale unor variabile. Există două modele ANOVA ce pot fi utilizate. Primul este ***One-Way*** **ANOVA** care este similar cu testul T cu excepţia că pot fi testate simultan mai mult de două diferenţe între medii. Al doilea model este ***Two-Way Factorial*** **ANOVA** şi reprezintă o extensie a primului model. Diferenţa constă în posibilitatea utilizării unei analize factoriale care poate avea mai mult decât o singură variabilă de grupare.

Pentru a exemplifica lucrul cu cele două modele ANOVA vom mai adăuga tabelului de date din (*Data Editor*) două variabile Varsta şi Sex. Setul de date va fi:

	Nota1	Nota2	Sex	Varsta
1	8,50	8,30	masculin	20
2	9,10	8,80	masculin	20
3	6,00	7,00	masculin	23
4	8,00	9,50	masculin	20
5	9,00	9,60	feminin	21
6	8,30	8,00	feminin	21
7	7,00	7,80	feminin	22
8	7,00	8,70	feminin	23
9	10,00	9,80	feminin	22
10	7,50	7,80	masculin	23
11	9,10	8,80	feminin	20
12	6,00	6,80	masculin	22
13	7,30	7,00	feminin	23
14	8,00	8,80	masculin	22
15	9,00	8,40	masculin	21
16	7,50	8,60	feminin	20

Vom presupune că cercetătorul doreşte să cunoască dacă vârsta angajaţilor influenţează creşterea performanţelor realizate de către aceştia după absolvirea unui curs. Deci variabila dependentă va fi Nota2 iar variabila factor Varsta. Din tabelul de date se observă că vârsta angajaţilor este cuprinsă între 20 şi 23 de ani.

Opţiunea ***One-Way*** **ANOVA** disponibilă de sub acelaşi meniu *Means* va lansa un dialog prin intermediul căruia va fi setată semnificaţia celor două variabile:



Raportul oferit de către SPSS va fi:

Oneway

Descriptives

Nota dupa curs

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
20	5	8,8000	,44159	,19748	8,2517	9,3483	8,30	9,50
21	3	8,6667	,83267	,48074	6,5982	10,7351	8,00	9,60
22	4	8,3000	1,29099	,64550	6,2457	10,3543	6,80	9,80
23	4	7,6250	,80984	,40492	6,3364	8,9136	7,00	8,70
Total	16	8,3563	,91504	,22876	7,8687	8,8438	6,80	9,80

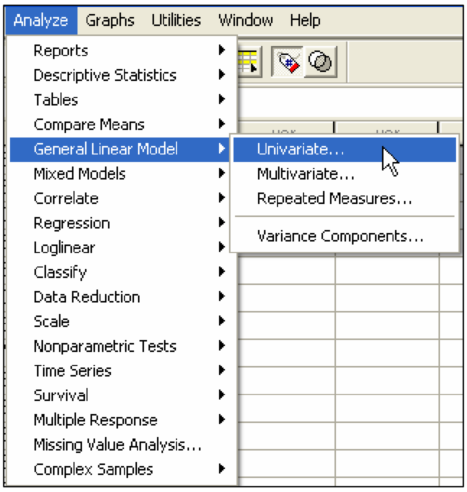
ANOVA

Nota dupa curs

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3,425	3	1,142	1,500	,265
Within Groups	9,134	12	,761		
Total	12,559	15			

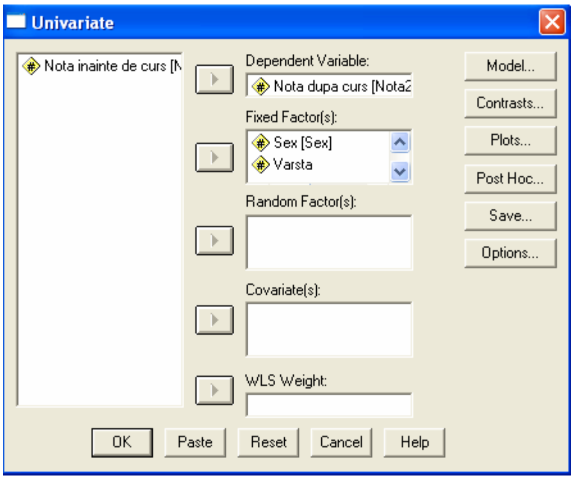
Conform rezultatelor din raport *Sig.* = 0,265 > 0,05 concluzionăm că se respinge ipoteza cercetării conform căreia vârsta influențează performanțele angajaților care au absolvit cursul.

Meniul *General Linear Model* oferă posibilitatea unor analize ANOVA (factoriale) realizate simultan pentru una sau mai multe variabile ce pot fi influențate de mai mulți factori. Vor fi prezentate în acest sens opțiunile *Univariate* și *Multivariate*.



Mergând pe același exemplu, vom presupune că cercetătorul dorește o analiză complexă prin care să studieze efectul separat și împreună a doi factori - vârsta și sexul - asupra variabilei de performanță Nota2. Ipoteza cercetării este: vârsta și sexul influențează performanțele angajaților măsurate prin valorile variabilei Nota2.

Opțiunea *Univariate* permite realizarea unor astfel de analize. Prin intermediul dialogului cu același nume se pot selecta variabilele tip factor și variabila dependentă:



Rezultatele întoarse de raport vor fi:

Univariate Analysis of Variance

Between-Subjects Factors

		Value Label	N
Sex	1	masculin	8
	2	feminin	8
Varsta	20		5
	21		3
	22		4
	23		4

Tests of Between-Subjects Effects

Dependent Variable: Nota dupa curs

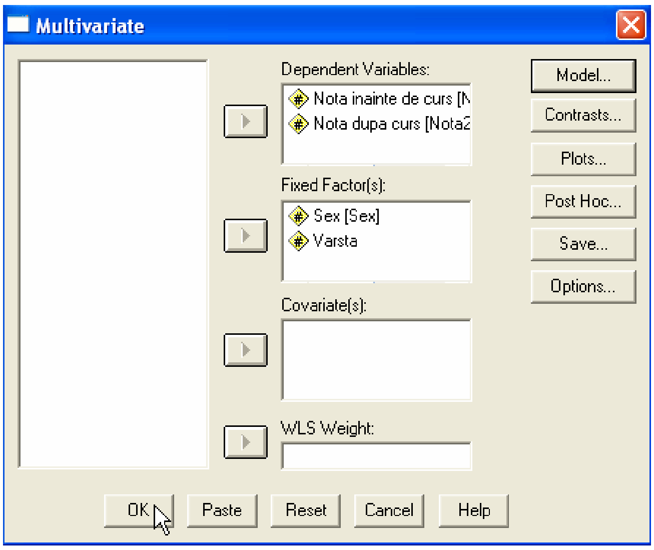
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4,768[a]	7	,681	,699	,675
Intercept	1024,103	1	1024,103	1051,486	,000
Sex	,654	1	,654	,671	,436
Varsta	3,172	3	1,057	1,086	,409
Sex * Varsta	,748	3	,249	,256	,855
Error	7,792	8	,974		
Total	1129,790	16			
Corrected Total	12,559	15			

a. R Squared = ,380 (Adjusted R Squared = -,163)

Rezultă că ipoteza cercetării nu poate fi acceptată nici separat pentru fiecare factor nici împreună pentru cei doi factori (Varsta şi Sex). Toate valorile *Sig.* (0,436, 0,409, 0,855) sunt mai mari decât 0,05.

Opțiunea *Multivariate* permite lucrul cu mai multe variabile care pot fi dependente simultan de mai mulți factori.

Pentru exemplificare conform setului de date utilizat în exemplele anterioare vom presupune că se doreşte studiul a două variabile dependente (Nota1 şi Nota2) în funcție de doi factori (Varsta şi Sex). În acest sens dialogul *Multivariate* va fi:



Raportul corespunzător va prezenta următoarele tabele:

Between-Subjects Factors

		Value Label	N
Sex	1	masculin	8
	2	feminin	8
Varsta	20		5
	21		3
	22		4
	23		4

Descriptive Statistics

	Sex	Varsta	Mean	Std. Deviation	N
Nota inainte de curs	masculin	20	8,5333	,55076	3
		21	9,0000	.	1
		22	7,0000	1,41421	2
		23	6,7500	1,06066	2
		Total	7,7625	1,21059	8
	feminin	20	8,3000	1,13137	2
		21	8,6500	,49497	2
		22	8,5000	2,12132	2
		23	7,1500	,21213	2
		Total	8,1500	1,12504	8
	Total	20	8,4400	,69857	5
		21	8,7667	,40415	3
		22	7,7500	1,70783	4
		23	6,9500	,66583	4
		Total	7,9562	1,14657	16
Nota dupa curs	masculin	20	8,8667	,60277	3
		21	8,4000	.	1
		22	7,8000	1,41421	2
		23	7,4000	,56569	2
		Total	8,1750	,92698	8
	feminin	20	8,7000	,14142	2
		21	8,8000	1,13137	2
		22	8,8000	1,41421	2
		23	7,8500	1,20208	2
		Total	8,5375	,92727	8
	Total	20	8,8000	,44159	5
		21	8,6667	,83267	3
		22	8,3000	1,29099	4
		23	7,6250	,80984	4
		Total	8,3563	,91504	16

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	Nota inainte de curs	9,918[a]	7	1,417	1,156	,417
	Nota dupa curs	4,768[b]	7	,681	,699	,675
Intercept	Nota inainte de curs	941,788	1	941,788	768,676	,000
	Nota dupa curs	1024,103	1	1024,103	1051,486	,000
Sex	Nota inainte de curs	,400	1	,400	,327	,583
	Nota dupa curs	,654	1	,654	,671	,436
Varsta	Nota inainte de curs	7,244	3	2,415	1,971	,197
	Nota dupa curs	3,172	3	1,057	1,086	,409
Sex * Varsta	Nota inainte de curs	2,060	3	,687	,560	,656
	Nota dupa curs	,748	3	,249	,256	,855
Error	Nota inainte de curs	9,802	8	1,225		
	Nota dupa curs	7,792	8	,974		
Total	Nota inainte de curs	1032,550	16			
	Nota dupa curs	1129,790	16			
Corrected Total	Nota inainte de curs	19,719	15			
	Nota dupa curs	12,559	15			

a. R Squared = ,503 (Adjusted R Squared = ,068)

b. R Squared = ,380 (Adjusted R Squared = -,163)

Se observă că toate valorile *Sig.* (atât pentru Nota1 cât şi pentru Nota2) sunt mai mari decât 0,05. Rezultă că pentru toate cazurile ipoteza cercetării nu va fi acceptată. Se poate concluziona: rezultatele obţinute atât înainte de participarea la curs cât şi cele obţinute după absolvirea cursului nu cresc semnificativ nici pe seama factorului vârsta nici pe seama factorului sex nici simultan pe baza celor doi factori.

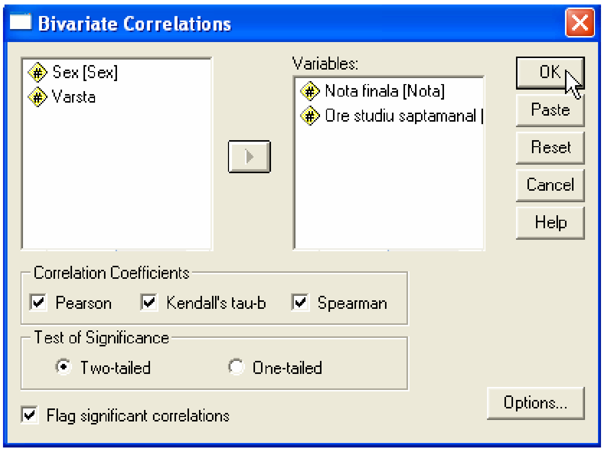
Meniul *Corelate* ne oferă posibilitatea de a putea studia corelaţia între două variabile prin intermediul coeficienţilor de corelaţie. S-a arătat într-un capitol anterior că atunci când se doreşte măsurarea valorilor a două variabile din acelaşi eşantion

pentru a se afla dacă între acestea există o relație şi care este intensitatea relației se utilizează coeficienții de corelație. Cel mai frecvent se utilizează coeficientul de corelație Pearson. Felul corelației se exprimă prin semnul coeficientului de corelație (r) iar **intensitatea legăturii** dintre cele două variabile se exprimă prin valoarea acestuia care va fi cuprinsă în intervalul [-1, +1].

Să considerăm că dorim să studiem dacă între notele finale obținute de un grup de studenți şi numărul de ore alocate săptămânal de către aceştia pentru studiu există o legătură (o relație). Vom realiza un sondaj prin care le vom solicita studenților participanți să specifice aceste date. Ipoteza cercetării este: între notele finale şi numărul de ore alocate săptămânal studiului există o corelație (legătură) semnificativă. Presupunem că datele sondajului aplicat unui grup de 14 studenți şi înscrise în *Data Editor* sunt:

	Nota	Ore
1	8	24
2	9	40
3	6	10
4	8	32
5	9	35
6	8	30
7	7	30
8	7	24
9	10	45
10	7	28
11	9	42
12	6	15
13	7	28
14	8	28

Opțiunea *Bivariate* din cadrul meniului *Corelate* ne oferă posibilitatea să realizăm acest studiu de corelație. Dialogul *Bivariate Corelations* ne ajută să selectăm variabilele pentru care dorim calculul coeficienților de corelație şi să specificăm care coeficienți de corelație dorim să fie calculați şi ce test de semnificație vom utiliza.



Apăsarea butonului OK va determina afişarea raportului corespunzător care va conține:

Correlations

Correlations

		Nota finala	Ore studiu saptamanal
Nota finala	Pearson Correlation	1	,910**
	Sig. (2-tailed)	.	,000
	N	14	14
Ore studiu saptamanal	Pearson Correlation	,910**	1
	Sig. (2-tailed)	,000	.
	N	14	14

**. Correlation is significant at the 0.01 level (2-tailed).

Nonparametric Correlations

Correlations

			Nota finala	Ore studiu saptamanal
Kendall's tau_b	Nota finala	Correlation Coefficient	1,000	,784**
		Sig. (2-tailed)	.	,000
		N	14	14
	Ore studiu saptamanal	Correlation Coefficient	,784**	1,000
		Sig. (2-tailed)	,000	.
		N	14	14
Spearman's rho	Nota finala	Correlation Coefficient	1,000	,869**
		Sig. (2-tailed)	.	,000
		N	14	14
	Ore studiu saptamanal	Correlation Coefficient	,869**	1,000
		Sig. (2-tailed)	,000	.
		N	14	14

**. Correlation is significant at the 0.01 level (2-tailed).

Observăm că între cele două variabile există o corelație pozitivă foarte intensă (r=0,910). Rezultă că ipoteza cercetării va fi acceptată.

BIBLIOGRAFIE

Anderson, R., David	Introduction to Management Science, Quantitative Approach to Decision Making, West Publishing Company, College & School Division, 1993
Boldur-Lățescu, Săcuiu, I., Țigănescu, E.	Cercetare operațională cu aplicații în economie, Editura didactică şi pedagogică, Bucureşti, 1979
Dantzig, G., B.	Linear Programming and Extensions, Princeton University Press, New York, 1963
Daren, George Paul, Mallery	SPSS for Windows Step by Step. A simple Guide and Reference, 11.0
Fred R., David	Strategic Management, Prentice Hall PTR, 2001
Heller, Robert Hindle, Tim	Making Decision, Publisher Dorling Kindersley, Incorporated, 1997
Hammond J., R. Keeney, H. Raiffa	Smart Choices: A Practical Guide to Making Better Decisions, Harvard Business School Press., 1998.
Kaufmann, A.	Metode şi modele ale cercetării operaționale, vol.I, II Editura ştiințifică, Bucureşti, 1967
Lawrence, A. John, Pasternack, Barry	Applied Management Science, John & Sons, Incorporated, Wiley, 2002
Lawrence, L., William, D.	Quantitative Decision Making with Spreadsheet Applications, Publisher Brooks / Cole, 2001
Marija, Norusis	SPSS 12.0 Guide to Data Analysis,
Schonberger, Richard	Operations Management, Irwin, 1988, Homewood, IL 60430, Boston, MA 02116, Library of Congress
Stevenson, William, J.	Management Science,, Irwin, 1989, Homewood, IL 60430, Boston, MA 02116, Library of Congress
Tummala V.	Decision Analysis with Business Applications, Educational Publishers, 1973.
Yih-Long, Chang, Sullivan, Robert	QSB+ Quantitative Systems for Business Plus, Prentice Hall, Englewood Cliffs, NJ 07632, 1991
Williams H.,	Model Building in Mathematical Programming, Wiley, 1999.
***	Documentație utilizare SPSS v12.0.