

Curs 11

Analiza și prognoza seriilor temporale (II)

Modelarea și analiza datelor pentru decizii de management

Conf.dr.ing. Lăcrimioara GRAMA

Anul 2 master · Tehnologii, Sisteme și Aplicații pentru eActivități

Facultatea de Electronică, Telecomunicații și Tehnologia Informației
Universitatea Tehnică din Cluj-Napoca



Cursul precedent

Analiza și prognoza seriilor temporale (I)

- ✓ Analiza previzională
 - ✓ Metode cantitative de prognoză
- ✓ Tipare ale seriilor temporale
 - ✓ Tipar orizontal
 - ✓ Tipar de trend
 - ✓ Tipar de sezonabilitate
 - ✓ Tipar de trend și sezonabilitate
 - ✓ Tipar ciclic
 - ✓ Selectarea unei metode de prognoză
- ✓ Acuratețea prognozei
- ✓ Metode de prognoză pentru serii cu tipar orizontal
 - ✓ Media alunecătoare
 - ✓ Media alunecătoare ponderată
 - ✓ Netezirea exponențială



La finalul acestui curs, studenții vor putea să:

- diferențieze modelele de trend liniar, trend neliniar, sezonalitate și trend cu sezonalitate;
- formuleze modele de prognoză bazate pe trend liniar, quadratic și exponențial;
- estimeze parametrii modelelor prin metoda celor mai mici pătrate și cu ajutorul Excel Solver;
- construiască modele de sezonalitate utilizând variabile fictive;
- combine efectul trendului cu efectele sezoniere într-un model unic de prognoză;
- utilizeze modelele estimate pentru prognozarea valorilor viitoare ale seriei temporale

Obiectivele cursului



Cuprins

Proгноza utilizând metode de trend

- Trend liniar
- Trend neliniar
 - Model quadratic
 - Model exponențial

Proгноza utilizând metode de sezonality

- Sezonality fără trend
- Sezonality și trend
- Metode bazate pe date lunare

Rezumat

Trend liniar: metoda celor mai mici pătrate

Metoda celor mai mici pătrate (*least squares method*), utilizată și în analiza de regresie, determină dreapta de trend care minimizează suma erorilor pătratice dintre valorile observate și cele estimate

- Perioada de timp este variabila independentă, iar valoarea observată efectiv în seria temporală este variabila dependentă

Modelul de trend liniar

$$T_t = b_0 + b_1 t \quad (1)$$

- T_t – valoarea estimată prin modelul de trend pentru perioada t
- b_0 – ordonata la origine a dreptei de trend liniar
- b_1 – panta dreptei de trend liniar
- t – perioada de timp ($t = \overline{1, n}$; $t = 1$ corespunde primei observații, iar $t = n$ celei mai recente)

Obiectiv: determinarea parametrilor b_0 și b_1 care minimizează suma erorilor pătratice (SSE)

$$\min_{b_0, b_1} \sum_{t=1}^n (A_t - T_t)^2$$

Minimizarea SSE este echivalentă cu minimizarea MSE, deoarece n este fix

Estimarea parametrilor

$$b_0 = \bar{A} - b_1 \bar{t}$$

$$b_1 = \frac{\sum_{t=1}^n (t - \bar{t})(A_t - \bar{A})}{\sum_{t=1}^n (t - \bar{t})^2}$$

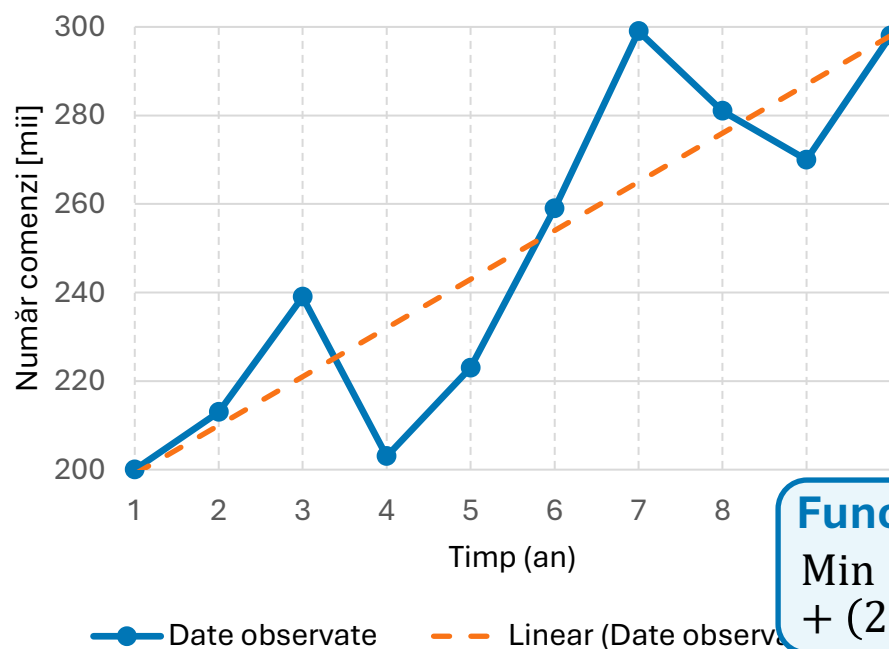
- $\bar{A}$ – media valorilor realizate
- $\bar{t}$ – media perioadelor de timp
- A_t – valoarea realizată a seriei temporale în perioada t
- n – numărul de observații

Trend liniar: exemplu

An	1	2	3	4	5	6	7	8	9	10
Număr comenzi [mii]	200	213	239	203	223	259	299	281	270	298

- Revenim la seria temporală privind numărul anual de comenzi procesate de un magazin online, analizată în cursul anterior

Graficul seriei temporale



- Deși seria prezintă atât creșteri, cât și scăderi în cei 10 ani analizați, dreapta de trend liniar surprinde în mod rezonabil evoluția sa pe termen lung

- Indicele temporal ia valorile $t = 1, \dots, 10$, unde $t = 1$ corespunde primei observații, iar $t = 10$ celei mai recente

Obiectiv: determinarea parametrilor b_0 și b_1 care minimizează suma pătratelor erorilor dintre valorile observate și cele estimate prin modelul de trend

Variable de decizie: b_0 și b_1

Variable auxiliare: $T_1, T_2, \dots, T_{10}$

Funcția obiectiv

$$\text{Min } (200 - T_1)^2 + (213 - T_2)^2 + (239 - T_3)^2 + (203 - T_4)^2 + (223 - T_5)^2 + (259 - T_6)^2 + (299 - T_7)^2 + (281 - T_8)^2 + (270 - T_9)^2 + (298 - T_{10})^2$$

Constrângeri: calculul valorilor estimate

$$T_1 = b_0 + 1 \cdot b_1$$

$$T_3 = b_0 + 3 \cdot b_1$$

$$T_5 = b_0 + 5 \cdot b_1$$

$$T_7 = b_0 + 7 \cdot b_1$$

$$T_9 = b_0 + 9 \cdot b_1$$

$$T_2 = b_0 + 2 \cdot b_1$$

$$T_4 = b_0 + 4 \cdot b_1$$

$$T_6 = b_0 + 6 \cdot b_1$$

$$T_8 = b_0 + 8 \cdot b_1$$

$$T_{10} = b_0 + 10 \cdot b_1$$

Trend liniar: calcul analitic al parametrilor

$$b_0 = \bar{A} - b_1 \bar{t}$$

$$b_1 = \frac{\sum_{t=1}^n (t - \bar{t})(A_t - \bar{A})}{\sum_{t=1}^n (t - \bar{t})^2}$$

$$\bar{t} = \frac{55}{10} = 5.5$$

$$\bar{A} = \frac{2485}{10} = 248.5$$

$$b_1 = \frac{907.5}{82.5} = 11$$

$$b_0 = 248.5 - 11 \cdot 5.5 = 188$$

Ecuatia trendului liniar

$$T_t = 188 + 11t$$

Panta $b_1 = 11$ indică o creștere medie estimată de **11 mii de comenzi pe an**

Dacă tendința observată în ultimii 10 ani este relevantă pentru viitor, modelul de trend poate fi utilizat pentru prognoză

Pentru anul 11:

$$T_{11} = 188 + 11 \cdot 11 = 309$$

➤ Se prognozează aproximativ **309 mii de comenzi** pentru anul 11

An (t)	$t - \bar{t}$	$(t - \bar{t})^2$	A_t [mii]	$A_t - \bar{A}$	$(t - \bar{t})(A_t - \bar{A})$
1	-4.5	20.25	200	-48.5	218.25
2	-3.5	12.25	213	-35.5	124.25
3	-2.5	6.25	239	-9.5	23.75
4	-1.5	2.25	203	-45.5	68.25
5	-0.5	0.25	223	-25.5	12.75
6	0.5	0.25	259	10.5	5.25
7	1.5	2.25	299	50.5	75.75
8	2.5	6.25	281	32.5	81.25
9	3.5	12.25	270	21.5	75.25
10	4.5	20.25	298	49.5	222.75
55		82.50	2485		907.50

Trend liniar: validarea soluției cu Excel Solver

Metoda de rezolvare în Excel Solver: *GRG Nonlinear* (Generalized Reduced Gradient)

Ecuatia trendului liniar

$$T_t = 188 + 11t$$

Eroarea medie pătratică

$$MSE = 307$$

Rezultatul coincide cu
soluția analitică

b0	188.00					MSE	307.00	
b1	11.00							
An (t)	t - $\bar{t}$	(t - $\bar{t}$) ²	Comenzi [mii] (At)	At - $\bar{y}$	(t - $\bar{t}$)(At - $\bar{y}$)	Valoare estimată (Tt)	Eroare de prognost	Eroare pătratică
1	-4.50	20.25	200.00	-48.50	218.25	199.00	1.00	1.00
2	-3.50	12.25	213.00	-35.50	124.25	210.00	3.00	9.00
3	-2.50	6.25	239.00	-9.50	23.75	221.00	18.00	324.00
4	-1.50	2.25	203.00	-45.50	68.25	232.00	-29.00	841.00
5	-0.50	0.25	223.00	-25.50	12.75	243.00	-20.00	400.00
6	0.50	0.25	259.00	10.50	5.25	254.00	5.00	25.00
7	1.50	2.25	299.00	50.50	75.75	265.00	34.00	1156.00
8	2.50	6.25	281.00	32.50	81.25	276.00	5.00	25.00
9	3.50	12.25	270.00	21.50	75.25	287.00	-17.00	289.00
10	4.50	20.25	298.00	49.50	222.75	298.00	0.00	0.00

Trend liniar: modelul general de optimizare

Pentru o serie temporală cu trend liniar, formată din n observații A_t , valorile optime ale parametrilor b_0 și b_1 se obțin prin minimizarea sumei erorilor pătratice

$$\text{Min} \quad \sum_{t=1}^n [A_t - T_t]^2 \quad (\text{funcția obiectiv})$$

$$\text{s. t.} \quad T_t = b_0 + b_1 t, \quad t = \overline{1, n} \quad (\text{rel. de definire})$$

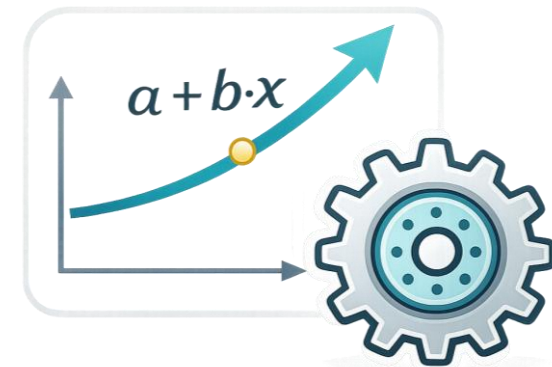
Variabile de decizie: b_0, b_1

Variabile auxiliare: $T_1, T_2, \dots, T_n$

Relații de definire: n , câte una pentru fiecare observație

Generalizare

Modelul poate fi extins la alte forme de trend prin înlocuirea relației liniare pentru T_t cu relația corespunzătoare noului model



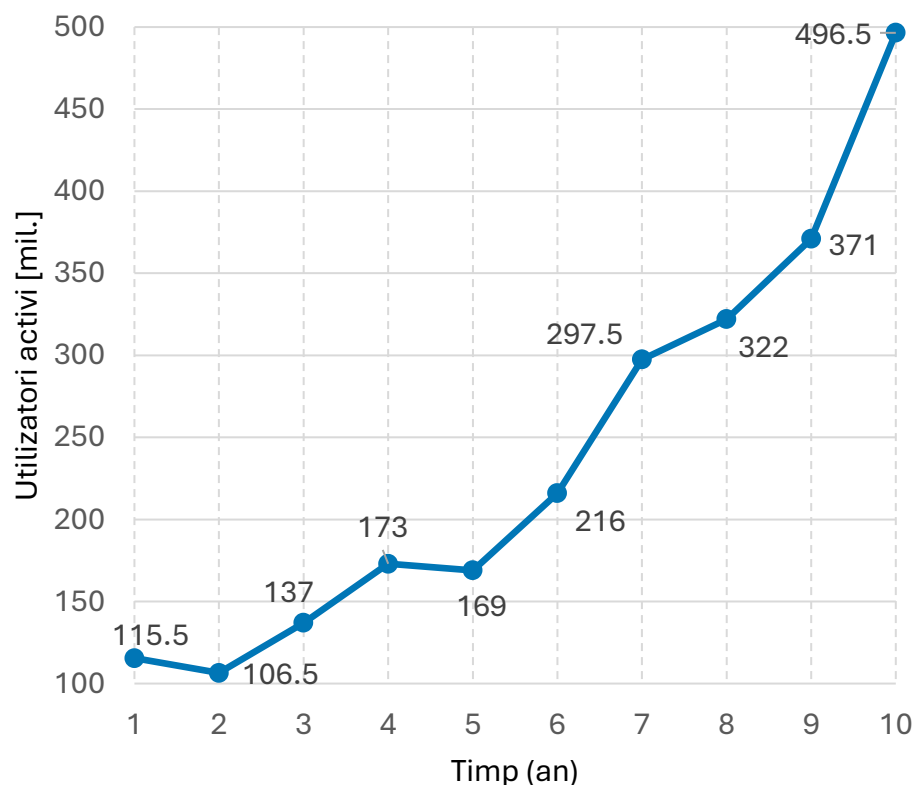
Instrumente software

Excel Solver, Minitab și SAS pot estima parametrii utilizând metode de regresie sau de ajustare a curbei (*curve fitting*)

Trend neliniar: identificarea tiparului

- Revenim la seria temporală privind evoluția anuală a numărului de utilizatori activi ai unei platforme digitale globale, analizată în cursul anterior

Graficul seriei temporale



Observație

Graficul evidențiază o **tendință generală ascendentă**

Concluzie

Spre deosebire de seria privind numărul anual de comenzi, această evoluție nu este descrisă adecvat printr-un trend liniar

Ritmul de creștere se modifică în timp, sugerând necesitatea unui **model de trend neliniar**

Modele candidate

Pentru modelarea unui trend neliniar pot fi utilizate, între altele:

- modelul quadratic
- modelul exponențial

Trend neliniar: modelul quadratic

An	1	2	3	4	5	6	7	8	9	10
Utilizatori activi [mil.]	115.5	106.5	137	173	169	216	297.5	322	371	496.5

Modelul de trend quadratic

$$T_t = b_0 + b_1 t + b_2 t^2 \quad (2)$$

- T_t – valoarea estimată prin modelul de trend pentru perioada t
- b_0 – termenul constant al modelului
- b_1 – coeficientul termenului liniar
- b_2 – coeficientul termenului quadratic, care stabilește sensul și intensitatea curburii trendului
- t – indicele temporal ($t = \overline{1, n}$)

Funcția obiectiv – model general:

determinarea parametrilor b_0 , b_1 și b_2 care minimizează suma erorilor pătratice (SSE)

$$\min_{b_0, b_1, b_2} \sum_{t=1}^n (A_t - T_t)^2$$

Pentru un număr fix de observații, minimizarea SSE este echivalentă cu minimizarea MSE

Variabile de decizie: b_0 , b_1 și b_2

Variabile auxiliare: $T_1, T_2, \dots, T_{10}$

Funcția obiectiv

$$\text{Min } (115.5 - T_1)^2 + (106.5 - T_2)^2 + (137 - T_3)^2 + (173 - T_4)^2 + (169 - T_5)^2 + (216 - T_6)^2 + (297.5 - T_7)^2 + (322 - T_8)^2 + (371 - T_9)^2 + (496.5 - T_{10})^2$$

Relații de calcul ale valorilor estimate

$$T_1 = b_0 + 1 \cdot b_1 + 1 \cdot b_2$$

$$T_2 = b_0 + 2 \cdot b_1 + 4 \cdot b_2$$

$$T_3 = b_0 + 3 \cdot b_1 + 9 \cdot b_2$$

$$T_4 = b_0 + 4 \cdot b_1 + 16 \cdot b_2$$

$$T_5 = b_0 + 5 \cdot b_1 + 25 \cdot b_2$$

$$T_6 = b_0 + 6 \cdot b_1 + 36 \cdot b_2$$

$$T_7 = b_0 + 7 \cdot b_1 + 49 \cdot b_2$$

$$T_8 = b_0 + 8 \cdot b_1 + 64 \cdot b_2$$

$$T_9 = b_0 + 9 \cdot b_1 + 81 \cdot b_2$$

$$T_{10} = b_0 + 10 \cdot b_1 + 100 \cdot b_2$$

Trend neliniar: soluția modelului quadratic

Metoda de rezolvare în Excel Solver: *GRG Nonlinear* (Generalized Reduced Gradient)

Ecuția trendului quadratic

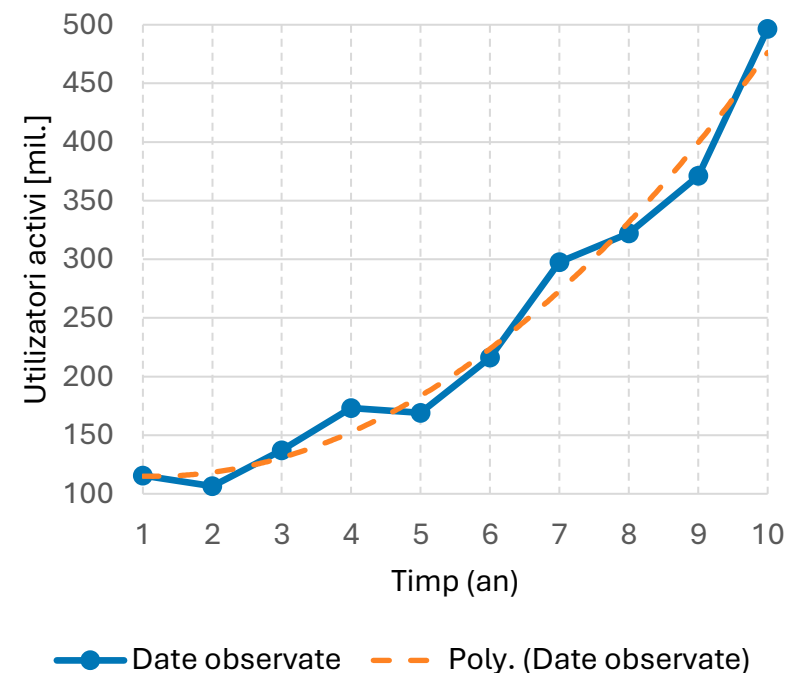
$$T_t = 120.91 - 10.53t + 4.61t^2$$

Eroarea medie pătratică

$$MSE = 276.62$$

b0	120.91			
b1	-10.53		MSE	276.62
b2	4.61			
An	Utilizatori [mil.]	Valoare estimată	Eroare de prognoză	Eroare pătratică
1	115.5	114.99	0.51	0.26
2	106.5	118.28	-11.78	138.78
3	137	130.79	6.21	38.56
4	173	152.52	20.48	419.57
5	169	183.46	-14.46	209.04
6	216	223.62	-7.62	58.00
7	297.5	272.99	24.51	600.79
8	322	331.58	-9.58	91.74
9	371	399.38	-28.38	805.59
10	496.5	476.40	20.10	403.86
			Total	2766.20

Date observate și model de trend



Dacă tendința observată în ultimii 10 ani se menține, modelul de trend poate fi utilizat pentru prognozarea valorilor viitoare

Pentru anul 11: utilizând coeficienții cu precizia completă furnizați de Solver, rezultă $T_{11} = 562.64$

➤ Se prognozează aproximativ **562.64 mil de utilizatori activi** pentru anul 11

Trend neliniar: modelul exponențial

Modelul de trend exponențial

$$T_t = b_0(b_1)^t \quad (3)$$

- T_t – valoarea estimată prin modelul de trend pentru perioada t
- b_0 – termenul constant al modelului
- b_1 – factorul de creștere sau de scădere de la o perioadă la alta
- t – indicele temporal ($t = \overline{1, n}$)

Funcția obiectiv – model general:

determinarea parametrilor b_0 și b_1 care minimizează suma erorilor pătratice

$$\min_{b_0, b_1} \sum_{t=1}^n (A_t - T_t)^2$$

Pentru un număr fix de observații, minimizarea SSE este echivalentă cu minimizarea MSE

- Continuăm analiza aceleiași serii temporale privind numărul de utilizatori activi

An	1	2	3	4	5	6	7	8	9	10
Utilizatori activi [mil.]	115.5	106.5	137	173	169	216	297.5	322	371	496.5

Variabile de decizie: $b_0 > 0$ și $b_1 > 0$

Variabile auxiliare: $T_1, T_2, \dots, T_{10}$

$b_1 > 1$: trend exponențial crescător

$0 < b_1 < 1$: trend exponențial descrescător

$b_1 = 1$: nivel constant

Funcția obiectiv

$$\text{Min } (115.5 - T_1)^2 + (106.5 - T_2)^2 + (137 - T_3)^2 + (173 - T_4)^2 + (169 - T_5)^2 + (216 - T_6)^2 + (297.5 - T_7)^2 + (322 - T_8)^2 + (371 - T_9)^2 + (496.5 - T_{10})^2$$

Relații de calcul ale valorilor estimate

$$T_1 = b_0 b_1^1$$

$$T_2 = b_0 b_1^2$$

$$T_3 = b_0 b_1^3$$

$$T_4 = b_0 b_1^4$$

$$T_5 = b_0 b_1^5$$

$$T_6 = b_0 b_1^6$$

$$T_7 = b_0 b_1^7$$

$$T_8 = b_0 b_1^8$$

$$T_9 = b_0 b_1^9$$

$$T_{10} = b_0 b_1^{10}$$

Trend neliniar: soluția modelului exponențial

Metoda de rezolvare în Excel Solver: *GRG Nonlinear* (Generalized Reduced Gradient)

Ecuția trendului exponențial

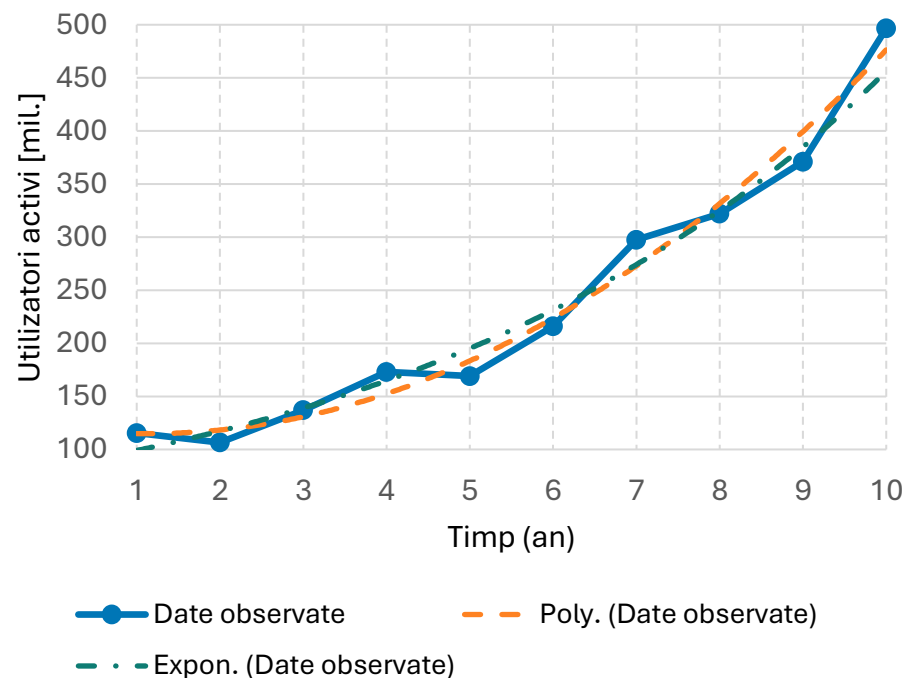
$$T_t = 77.12 \cdot (1.2)^t$$

Eroarea medie pătratică

$$MSE = 307.80$$

b0	77.12			
b1	1.20		MSE	307.80
An	Utilizatori [mil.]	Valoare estimată	Eroare de prognoză	Eroare pătratică
1	115.50	92.51	22.99	528.61
2	106.50	110.96	-4.46	19.92
3	137.00	133.10	3.90	15.22
4	173.00	159.65	13.35	178.22
5	169.00	191.50	-22.50	506.19
6	216.00	229.70	-13.70	187.70
7	297.50	275.52	21.98	482.99
8	322.00	330.49	-8.49	72.02
9	371.00	396.41	-25.41	645.91
10	496.50	475.49	21.01	441.22
Total				3077.99

Date observate și modele de trend



Dacă tendința observată în ultimii 10 ani se menține, modelul exponențial poate fi utilizat pentru prognozarea valorilor viitoare

Pentru anul 11: utilizând coeficienții cu precizia completă furnizați de Solver, rezultă $T_{11} = 570.35$

➤ Se prognozează aproximativ **570.35 mil de utilizatori activi** pentru anul 11

Compararea modelelor de trend neliniar

Criteriu	MODEL CUADRATIC	MODEL EXPONENȚIAL
Ecuatie	$T_t = b_0 + b_1t + b_2t^2$	$T_t = b_0(b_1)^t$
Tip de evoluție	Ritmul de creștere sau de scădere se modifică liniar în timp	Rata procentuală de creștere este aproximativ constantă
Flexibilitate	Poate surprinde curburi și schimbări ale ritmului	Potrivit pentru creștere sau scădere proporțională
MSE pentru seria analizată	276.62	307.80
Proгноza pentru anul 11	562.64 milioane utilizatori	570.35 milioane utilizatori
Limitare	Poate conduce la schimbarea sensului evoluției și la extrapolări nerealiste pe termen lung	Presupune o rată procentuală constantă și poate genera creșteri foarte rapide pe termen lung

Atenție la extrapolare

Pentru seria analizată, modelul quadratic oferă o ajustare istorică mai bună, având un MSE mai mic

Alegerea modelului trebuie să țină cont atât de eroarea istorică, cât și de plauzibilitatea evoluției viitoare

Optimizarea modelului exponențial

Estimarea parametrilor modelului exponențial conduce la o problemă de optimizare neliniară. Soluția furnizată de Solver poate depinde de valorile inițiale. Se recomandă:

- testarea mai multor puncte de pornire, pentru a reduce riscul convergenței la un optim local
- compararea valorilor finale ale SSE sau MSE
- impunerea unor domenii plauzibile pentru parametri

Cuprins

Proгноza utilizând metode de trend

Proгноza utilizând metode de sezonality

- Sezonality fără trend
- Sezonality și trend
- Metode bazate pe date lunare

Rezumat

Prognoza seriilor cu sezonabilitate

Atunci când o serie temporală prezintă **sezonalitate**, componenta sezonieră trebuie inclusă în model pentru a obține **prognoze mai precise**

Sezonalitatea constă în variații care se repetă la intervale regulate, de exemplu: luni, trimestre, anotimpuri sau zile ale săptămânii

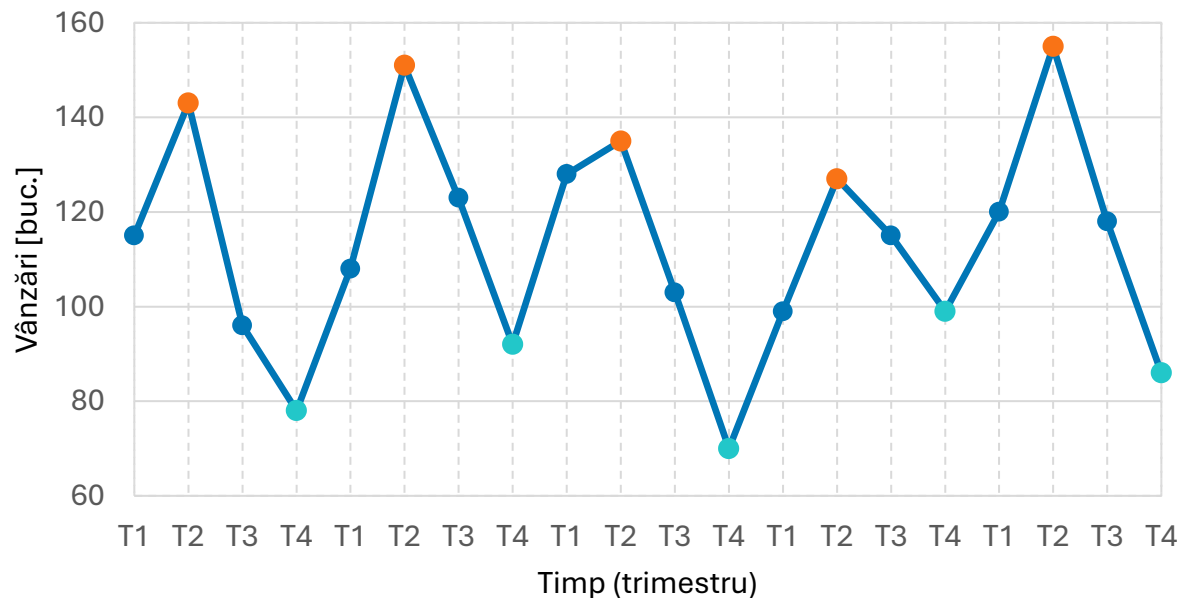


Sezonalitate fără trend

- Revenim la seria temporală privind vânzările trimestriale de produse pentru grădinărit, analizată în cursul anterior

An	1				2				3				4				5			
Trimestru	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Vânzări [buc.]	115	143	96	78	108	151	123	92	128	135	103	70	99	127	115	99	120	155	118	86

Graficul seriei temporale



Observație

Graficul nu evidențiază o tendință clară pe termen lung
Totuși, anumite creșteri și scăderi se repetă în aceleași trimestre ale fiecărui an

Concluzie

Dacă analizăm doar nivelul general, seria ar putea părea aproximativ orizontală
O examinare mai atentă arată însă un tipar trimestrial repetitiv: valorile tind să fie mai ridicate în trimestrul 2 și mai scăzute în trimestrul 4

Model candidat

Model sezonier trimestrial fără trend

Sezonalitate fără trend: model cu variabile fictive

Trimestrul este o **variabilă categorială** cu patru niveluri
Pentru o variabilă categorială cu k niveluri se utilizează $k - 1$ **variabile fictive**

Variabilă categorială = variabilă calitativă ale cărei valori descriu categorii

Variabilă fictivă (*dummy variable*) = variabilă binară care ia valorile 0 sau 1 și indică apartenența unei observații la o anumită categorie

Trimestru	D_{1t}	D_{2t}	D_{3t}
T1	1	0	0
T2	0	1	0
T3	0	0	1
T4	0	0	0

Model cu variabile fictive

$$F_t = b_0 + b_1 \cdot D_{1t} + b_2 \cdot D_{2t} + b_3 \cdot D_{3t}$$

- F_t – valoarea prognozată pentru perioada t ($t = \overline{1,20}$)
- D_{it} – variabila fictivă asociată trimestrului i , pentru perioada t
- b_0 – nivelul estimat pentru trimestrul 4, categoria de referință
- b_1 – diferența estimată dintre trimestrele 1 și 4
- b_2 – diferența estimată dintre trimestrele 2 și 4
- b_3 – diferența estimată dintre trimestrele 3 și 4

$$D_{it} = \begin{cases} 1, & \text{dacă observația din perioada } t \text{ aparține trimestrului } i \\ 0, & \text{în caz contrar} \end{cases}$$

Trimestrul 4 este categoria de referință

Funcția obiectiv pentru seria analizată

$$\text{Min } (115 - F_1)^2 + (143 - F_2)^2 + (96 - F_3)^2 + (78 - F_4)^2 + \dots + (86 - F_{20})^2$$

Calculul valorilor prognozate

$$\begin{aligned} F_1 &= b_0 + 1 \cdot b_1 + 0 \cdot b_2 + 0 \cdot b_3 & \dots \\ F_2 &= b_0 + 0 \cdot b_1 + 1 \cdot b_2 + 0 \cdot b_3 & F_{17} &= b_0 + 1 \cdot b_1 + 0 \cdot b_2 + 0 \cdot b_3 \\ F_3 &= b_0 + 0 \cdot b_1 + 0 \cdot b_2 + 1 \cdot b_3 & F_{18} &= b_0 + 0 \cdot b_1 + 1 \cdot b_2 + 0 \cdot b_3 \\ F_4 &= b_0 + 0 \cdot b_1 + 0 \cdot b_2 + 0 \cdot b_3 & F_{19} &= b_0 + 0 \cdot b_1 + 0 \cdot b_2 + 1 \cdot b_3 \\ & \dots & F_{20} &= b_0 + 0 \cdot b_1 + 0 \cdot b_2 + 0 \cdot b_3 \end{aligned}$$

Sezonalitate fără trend: soluția modelului

Metoda de rezolvare în Excel Solver: *GRG Nonlinear*

Modelul este linear în parametrii b_0, b_1, b_2, b_3 și poate fi estimat și prin regresie liniară cu variabile fictive. Utilizarea *GRG Nonlinear* este posibilă, dar nu este obligatorie

Modelul sezonier estimat

$$F_t = 85 + 29D_{1t} + 57.2D_{2t} + 26D_{3t}$$

Eroarea medie pătratică

$$MSE = 101.84$$

Pentru anul 6, prognozele sunt:

- Trimestrul 1: **114** produse de grădinărit vândute
- Trimestrul 2: **142.2** produse de grădinărit vândute
- Trimestrul 3: **111** produse de grădinărit vândute
- Trimestrul 4: **85** produse de grădinărit vândute

Concluzie

În absența unui trend, prognozele sezoniere se repetă de la un an la altul

b0	85.00		b2	57.20				
b1	29.00		b3	26.00		MSE	101.84	
Perioada	An	Trimestru	D1t	D2t	D3t	Vânzări	Prognoză	Eroare pătratică
1	1	1	1	0	0	115	114.00	1.00
2		2	0	1	0	143	142.20	0.64
3		3	0	0	1	96	111.00	225.00
4		4	0	0	0	78	85.00	49.00
5	2	1	1	0	0	108	114.00	36.00
6		2	0	1	0	151	142.20	77.44
7		3	0	0	1	123	111.00	144.00
8		4	0	0	0	92	85.00	49.00
9	3	1	1	0	0	128	114.00	196.00
10		2	0	1	0	135	142.20	51.84
11		3	0	0	1	103	111.00	64.00
12		4	0	0	0	70	85.00	225.00
13	4	1	1	0	0	99	114.00	225.00
14		2	0	1	0	127	142.20	231.04
15		3	0	0	1	115	111.00	16.00
16		4	0	0	0	99	85.00	196.00
17	5	1	1	0	0	120	114.00	36.00
18		2	0	1	0	155	142.20	163.84
19		3	0	0	1	118	111.00	49.00
20		4	0	0	0	86	85.00	1.00
		(valorile variabilelor fictive)					Total	2036.80

Sezonalitate și trend: identificarea tiparului

- Modelul cu variabile fictive poate fi extins pentru a descrie seriile temporale care prezintă simultan sezonalitate și trend liniar
 - Pentru reprezentarea trendului se introduce variabila temporală t
- Revenim la seria temporală a vânzărilor trimestriale de tablete, analizată în cursul anterior

Observație

Graficul evidențiază simultan un **trend ascendent** și un **tipar sezonier trimestrial**

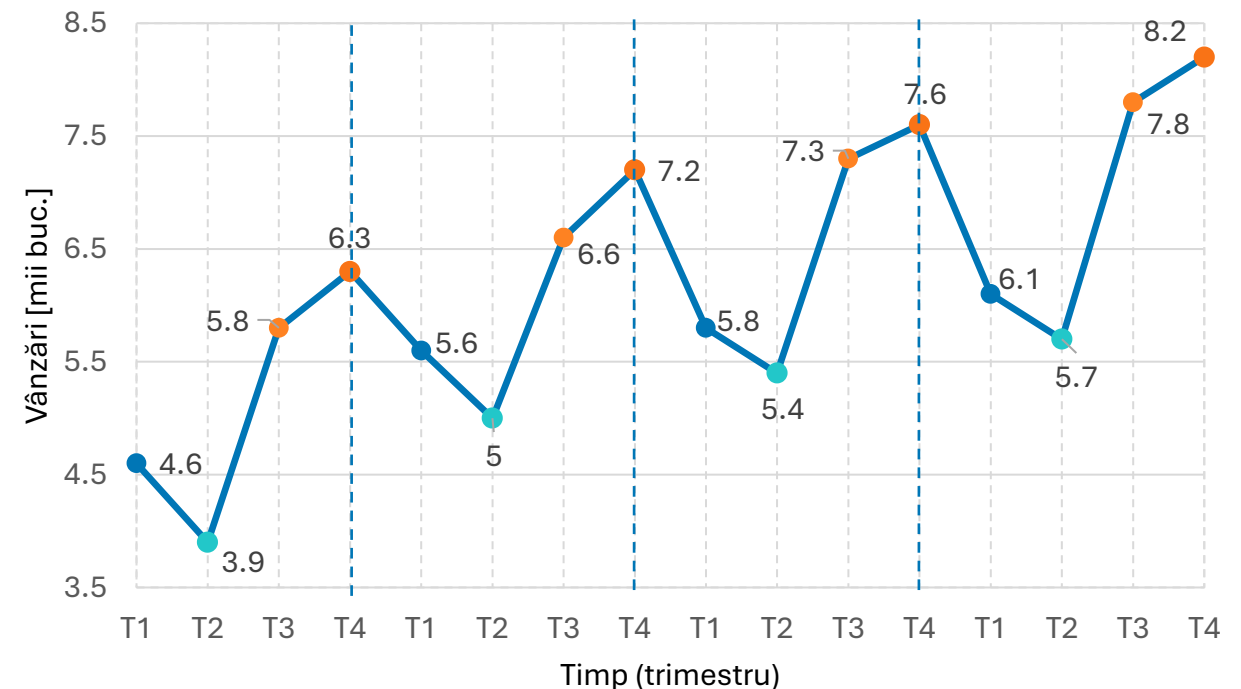
Concluzie

În fiecare an, vânzările sunt minime în **trimestrul 2**, apoi cresc în **trimestrele 3 și 4**, atingând nivelul maxim în **trimestrul 4**

Model candidat

Model sezonier trimestrial cu trend liniar

Graficul seriei temporale



Sezonalitate și trend: model cu variabile fictive

- Pentru prognozarea unei serii care prezintă simultan sezonalitate și trend liniar, variabilele fictive sezoniere se combină cu variabila temporală t

Model cu variabile fictive și trend liniar

$$F_t = b_0 + b_1 \cdot D_{1t} + b_2 \cdot D_{2t} + b_3 \cdot D_{3t} + b_4 t$$

- F_t – valoarea prognozată pentru perioada t ($t = \overline{1,16}$)
- D_{it} – variabila fictivă asociată trimestrului i , pentru perioada t
- b_0 – interceptul modelului pentru trimestrul 4, categoria de referință, la $t = 0$
- b_1, b_2, b_3 – diferența estimată dintre trimestrele 1, 2, 3 și trimestrul 4
- b_4 – modificarea medie a valorii prognozate la trecerea unei perioade, după controlul efectului sezonier

Modelul este **liniar în parametri**, chiar dacă descrie simultan trendul și sezonalitatea

$$D_{it} = \begin{cases} 1, & \text{dacă observația din perioada } t \text{ aparține trimestrului } i \\ 0, & \text{în caz contrar} \end{cases}$$

Obiectiv: determinarea parametrilor b_0, b_1, b_2, b_3 și b_4 care minimizează suma erorilor pătratice (SSE)

$$\min_{b_0, \dots, b_4} \sum_{t=1}^{16} (A_t - F_t)^2$$

Minimizarea SSE este echivalentă cu minimizarea MSE, deoarece n este fix

Trimestru	D_{1t}	D_{2t}	D_{3t}
T1	1	0	0
T2	0	1	0
T3	0	0	1
T4	0	0	0

Trimestrul 4 este categoria de referință

Sezonalitate și trend: estimarea modelului și prognoza

Metoda de rezolvare în Excel Solver: *GRG Nonlinear*

Modelul este liniar în parametri și poate fi estimat și prin regresie liniară cu variabile fictive. Utilizarea *GRG Nonlinear* este posibilă, dar nu este obligatorie

Modelul sezonier cu trend estimat

$$F_t = 5868.74 - 1363.12D_{1t} - 2033.75D_{2t} - 304.38D_{3t} + 145.63t$$

Eroarea medie pătratică

$$MSE = 32273.44$$

Coeficientul $b_4 = 145.63$ indică o creștere medie estimată de aproximativ **146 de tablete pe trimestru**, după controlul efectelor sezoniere

În Solver, valorile sunt exprimate în bucăți; pe graficul anterior au fost prezentate în mii de bucăți

b0 5868.74			b3 -304.38					
b1 -1363.12			b4 145.63				MSE	32273.44
b2 -2033.75								
Perioada	An	Trimestru	D1t	D2t	D3t	Vânzări [buc]	Prognoză	Eroare pătratică
1	1	1	1	0	0	4600	4651.24	2626.01
2		2	0	1	0	3900	4126.25	51187.17
3		3	0	0	1	5800	6001.24	40498.48
4		4	0	0	0	6300	6451.24	22874.64
5	2	1	1	0	0	5600	5233.75	134141.03
6		2	0	1	0	5000	4708.75	84827.43
7		3	0	0	1	6600	6583.75	264.22
8		4	0	0	0	7200	7033.75	27640.28
9	3	1	1	0	0	5800	5816.25	264.06
10		2	0	1	0	5400	5291.25	11826.30
11		3	0	0	1	7300	7166.25	17889.67
12		4	0	0	0	7600	7616.25	264.03
13	4	1	1	0	0	6100	6398.75	89253.17
14		2	0	1	0	5700	5873.75	30190.42
15		3	0	0	1	7800	7748.75	2626.52
16		4	0	0	0	8200	8198.75	1.56
	(valorile variabilelor fictive)						Total	516375.00
Prognoza pentru anul 5								
17	5	1	1	0	0		6981	Trimestrul 1
18		2	0	1	0		6456	Trimestrul 2
19		3	0	0	1		8331	Trimestrul 3
20		4	0	0	0		8781	Trimestrul 4

Extinderea modelului sezonier la date lunare

De la date trimestriale la date lunare

Luna este o variabilă categorială cu 12 niveluri. Dacă modelul conține un termen constant, sunt necesare 11 variabile fictive. Una dintre luni este aleasă drept **categorie de referință**; de exemplu, luna decembrie

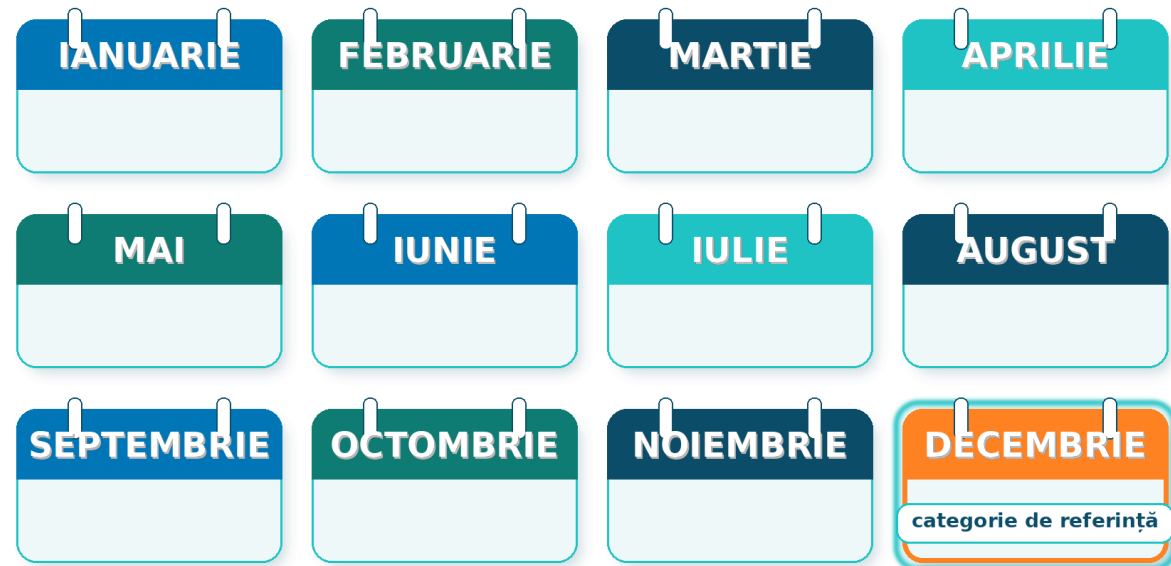
Model sezonier lunar fără trend

$$F_t = b_0 + \sum_{m=1}^{11} b_m D_{mt}$$

Model sezonier lunar cu trend liniar

$$F_t = b_0 + \sum_{m=1}^{11} b_m D_{mt} + b_{12} t$$

$$D_{mt} = \begin{cases} 1, & \text{dacă observația din perioada } t \text{ aparține lunii } m \\ 0, & \text{în caz contrar} \end{cases} \quad m = \overline{1,11}; t = \overline{1,n}$$



Ideea-cheie: principiul de modelare nu se schimbă. Pentru o sezonality cu k niveluri și un model care conține termen constant, se utilizează $k - 1$ variabile fictive

Fără trend, prognozele lunare se repetă anual; prin includerea variabilei t , modelul descrie simultan **sezonalitatea lunară și evoluția în timp**

Rezumat



În cursurile 10 și 11 am prezentat o introducere în analiza și prognoza seriilor temporale

Am arătat că primul pas în alegerea unei metode de prognoză este identificarea tiparului seriei temporale

- Prin reprezentarea grafică a datelor, pot fi observate tipare orizontale, tipare de trend, tipare sezoniere, combinații între trend și sezonalitate sau, în unele cazuri, tipare ciclice

Am discutat apoi despre acuratețea prognozei și despre rolul indicatorilor de eroare în evaluarea metodelor

- Indicatori precum MAE, MSE și MAPE permit compararea prognozelor pe date istorice și sprijină alegerea metodei care oferă cea mai bună ajustare pentru seria analizată

Rezumat



Pentru seriile temporale cu tipar orizontal, am prezentat metode de prognoză bazate pe valorile istorice recente

Media alunecătoare utilizează media ultimelor observații pentru a prognoza perioada următoare, iar media alunecătoare ponderată permite acordarea unor ponderi mai mari observațiilor recente

- Netezirea exponențială simplă extinde această idee, folosind o medie ponderată a valorilor anterioare, în care ponderile scad exponențial odată cu vechimea observațiilor

Am arătat că alegerea ponderilor sau a constantei de netezire influențează compromisul dintre netezirea fluctuațiilor și viteza de adaptare la schimbările din serie

Rezumat



Pentru seriile temporale cu trend, am prezentat utilizarea metodei celor mai mici pătrate pentru estimarea unei ecuații de trend

- Trendul liniar poate fi modelat printr-o dreaptă, iar seriile cu evoluții curbilinii pot fi descrise prin modele neliniare, precum modelul quadratic sau modelul exponențial

Pentru seriile cu sezonabilitate, am arătat cum pot fi utilizate variabilele fictive pentru a reprezenta efectele sezoniere

- Modelul poate fi extins pentru situațiile în care seria prezintă simultan sezonabilitate și trend, prin combinarea variabilelor fictive cu variabila temporală

În cazul datelor lunare, același principiu se aplică prin utilizarea a 11 variabile fictive, una dintre luni fiind aleasă drept categorie de referință

Rezumat

Identificarea tiparului seriei: orizontal, trend, sezonalitate, trend + sezonalitate
Alegerea metodei depinde de tiparul observat.

Tipar orizontal: medie alunecătoare,
medie ponderată, netezire
exponențială

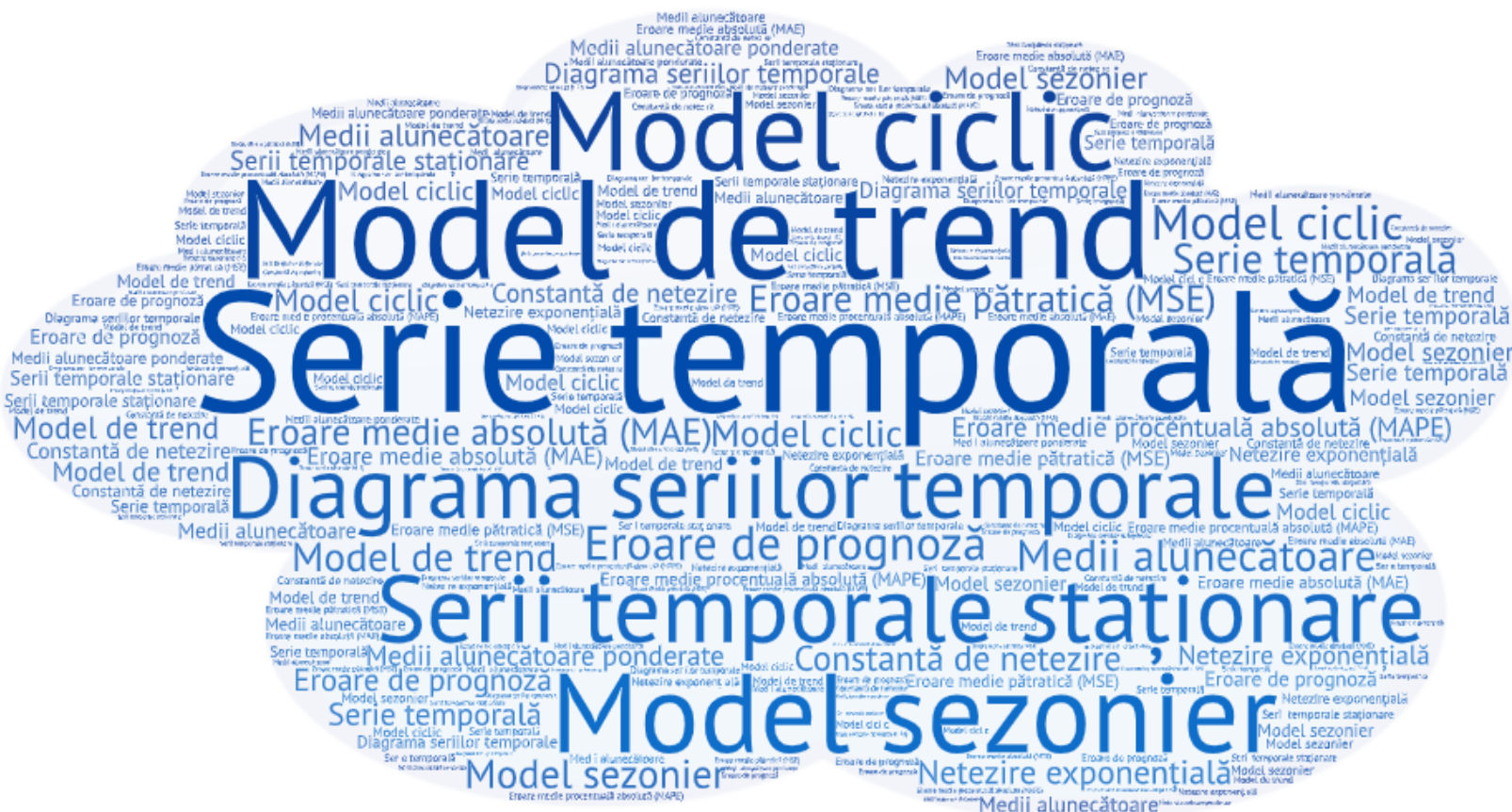
Acuratețea se evaluează cu MAE, MSE, MAPE

Trend linier: model $T_t = b_0 + b_1 t$

Trend nelinier: model quadratic sau exponential.

Sezonalitate: variabile fictive

Sezonalitate + trend: variabile fictive + variabila temporală t



Rezumat

Proгноza utilizând metode de trend

- ✓ Trend liniar
- ✓ Trend neliniar
 - ✓ Model quadratic
 - ✓ Model exponențial

Proгноza utilizând metode de sezonaliitate

- ✓ Sezonaliitate fără trend
- ✓ Sezonaliitate și trend
- ✓ Metode bazate pe date lunare

→ **Modelarea și analiza datelor pentru decizii de management – pe scurt**